2

Introduction to the Methods of Science

Learning Objectives

2.1 Describe naturalistic observation

2.2 Describe the correlational approach

2.3 Explain the experimental method

2.4 Analyze the role of logic and inference in the scientific method

2.5 Distinguish between valid and invalid forms of logic

2.6 Explain the different types of validity

2.7 Explain the effects of perspectives on scientific observation

Most of us remember how different our first few weeks of college were from anything we had known before. Remember how you expected your roommate or your professor to act, and how you reacted when they did not act that way? If you stop and think about your first reactions to college, you can see that there were three aspects of these experiences. First, there was an idea or expectation concerning what was about to happen in college. Second, there was the actual experience of what did happen during those first few weeks—and it was probably quite different, at least in some respects, from what you had expected. Third, there was the resulting reorganization of your ideas about college and the potential impact of this experience on your life.

The methods of science closely parallel these three aspects of your experience. First, scientists begin with an idea or expectation. As we will discuss, a formally stated expectation is called a *hypothesis*. The scientist says, "I expect this to happen under these conditions," and thus states the hypothesis. Second, scientists look to experience to evaluate the accuracy of their ideas or expectations about the world. That is, they try to find or create the situation that will allow them to observe what they are interested in studying. Through observation and experimentation, scientists can begin to evaluate their ideas and expectations about the world. As mentioned in Chapter 1, learning about the world through observation and experimentation is an example of *empiricism,* which means nothing more than the acceptance of sensory information as valid. Third, based on their observations and experiments, scientists seek to draw conclusions or inferences about their ideas and expectations. They reorganize their ideas and consider the impact of the new information on their theoretical conceptualizations.

As mentioned earlier, science is a way of determining what we can infer about the world. In its simplest form, the scientific method consists of asking a question about the world and then experiencing the world to determine the answer. When we begin an inquiry, what we already know about our topic leaves us in one of a number of positions. In some cases, we know little about our topic, or our topic may be very complex. Consequently, our ideas and questions are general. For example, how does our memory work? What causes mental illness? What factors make a fruitful marriage? How can we model the brain?

Sometimes, it is a worldwide problem. With the COVID-19 virus some psychological scientists asked what factors encourage people worldwide to minimize close contact and wear a face mask (Pfattheicher, Nockur, Böhm, Sassenrath, & Petersen, 2020). One answer was to induce empathy for those most vulnerable to the virus. Others wanted to know if offering behavioral health services online was as effective as when psychotherapy was performed in person (Batastini, Paprzycki, Jones, & MacLean, 2021). This type of research suggests that even after COVID-19 is no longer a problem, online therapy is an alternative.

If little is known about a particular phenomenon, it often is useful simply to watch the phenomenon occur naturally and get a general idea of what is involved in the process. Initially, this is accomplished by observing and describing what occurs. This scientific technique is called **naturalistic observation**.

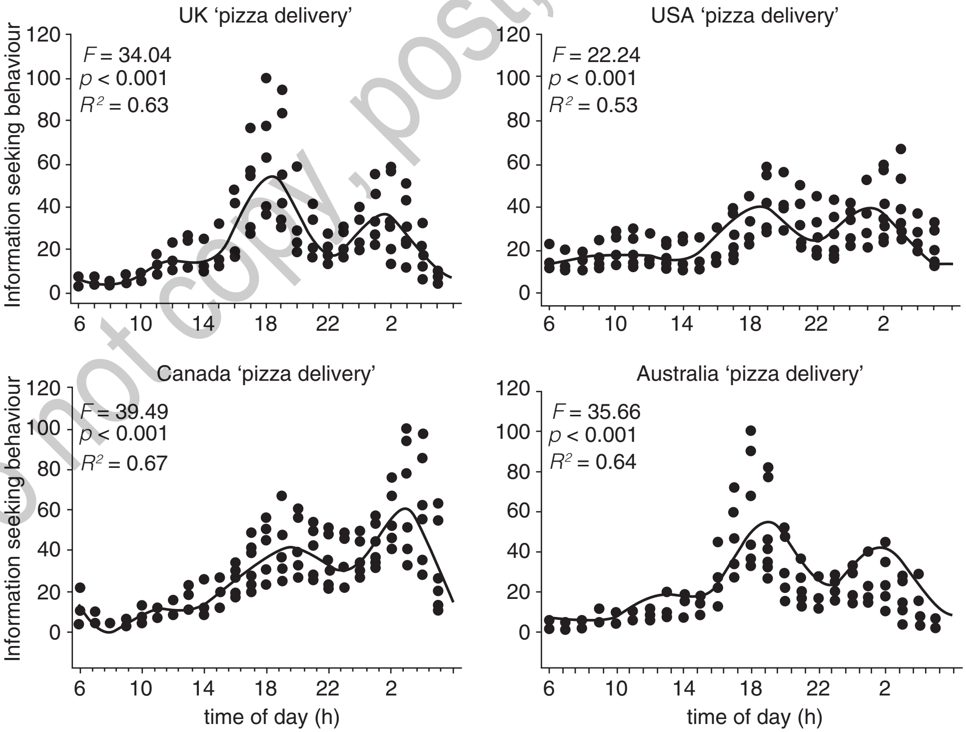
A classic example of this approach is Charles Darwin's observation of animals in the Galápagos Islands, which formed the basis of his theory of evolution. More recently, the study of animals in the wild has led to new insights into animal cognition and social systems. Scientists have continued to return to the Galápagos Islands with more advanced methods (such as observing molecular DNA changes) and still continue to update and confirm Darwin's original ideas (see Weiner, 1994, for a review of this work; Grant & Grant, 2006, for a study of current evolutionary changes; and Lamichhaney et al., 2014, for genetic changes).

Another example of this method has come from observing female dwarf mongooses in Tanzania's Serengeti National Park over several years (Morell, 1996). Before this research was conducted, it was assumed that the dominant female in a pack of animals would be under the least stress and the subordinate females under the most. However, by examining some 14 packs of animals, scientists have found that dominant female mongooses show the highest level of cortisol, a stress hormone, when compared with all other females in the pack. Other researchers (for example, De Waal, 2000, 2016) seek to understand how animals and humans use peacemaking to resolve aggressive episodes. Did you know that chimpanzees, like humans, often kiss their partner on the mouth after a fight?

Although traditional naturalistic observation involves actually watching the participant over a period of time, new technologies allow for an extension of this technique. For example, databases such as Google trends allow researchers to look for patterns of behaviors over a period of time. Alvarado and Stevenson (2018) looked at data from 5 years concerning when individuals looked for pizza companies on the Internet. What they observed was that in four different countries (U.S., UK, Canada, and Australia) individuals sought pizza companies at 2 times of day, 6 and 7 p.m. in the evening and 2 a.m. in the morning (see Figure 2.1).

At other times, we may want to understand certain aspects of a complex system with the goal of better describing how one aspect of the system may be associated with another aspect. For example, we may want to know whether people who have friends have fewer health-related problems than people who do not or whether eating certain foods is associated with not having cancer. How would you go about answering such questions?

FIGURE 2.1 ● Searches for "Pizza Delivery" on the Internet in Four Countries. Across all Countries, There Were Searches Around 6 and 7 PM and 2 AM.



Source: Alvarado and Stevenson (2018).

One way is to examine and note the relationship between a person's health and the number of friends that person has. But how are you to understand these data? Look at an example of an earlier question of the relationship between smoking tobacco and having lung cancer. The first step is to ask whether two events go together. In this example, researchers sought to determine whether, when one event occurred (a person smoked tobacco), the other event also occurred (the person had cancer). Such a scientific approach is called by various names, including the **correlational approach**.

As we will see later in this book, just finding that a relationship exists between two events does not allow us to determine exactly what that relationship is, much less to determine that one event actually caused the other event to happen. If we want to state that one event produced another event, we need to develop a much stronger case for our position. For example, we now know vaping electronic cigarettes is associated with health problems. Knowing this we would want to determine the effects of nicotine as well as other substances in electronic cigarettes. We could see how some single event or substance over which we have control affects the phenomenon we are studying. To do this, we begin to interact with the phenomenon.

We structure our question in this form: "If I do this, what will happen?" Numerous questions can be asked in this way, such as, "Will you learn words better if each word is of the same class (for example, food words) than if they are from different classes (foods, cars, and toys)?" Did you know that leaving your cell phone at home will improve your grade point average (Katz & Lambert, 2016)? However, playing video games at home may actually help you to think better (Stanmore, Stubbs, Vancapfort, de Bruin, & Firth, 2017).

As our knowledge grows, we may even get to the point of formulating specific predictions. In this case, our questions are structured in the form, "If I do this, I expect this will happen." Sometimes our predictions are more global, and we predict that one factor will be stronger than another. We might predict that more people are likely to help a stranger if they perceive the environment to be safe than if they think it is dangerous. Sometimes, however, we may know enough about an area to make a more precise prediction, or *point prediction*. For example, we might predict that 3 months of exercise will lead to a 10-mm Hg decrease in blood pressure. These approaches, in which we interact directly with the phenomenon we are studying, are examples of the **experimental method**.

During the 1980s a different type of scientific approach was developed within psychology, especially in response to cognitive questions. The approach was based less on observation or direct manipulation than on attempting to establish a model (either conceptual or mathematical) capable of performing operations similar to the topic being studied. This approach is called **modeling**. For example, Freeman and colleagues asked whether they could develop a model based on calculus that would mimic the brain's electrical patterns of an animal sensing odors (Freeman, 1991, 1999, 2004, 2015).

Currently, insights from cognitive psychology are being used in artificial intelligence (AI) applications. For example, when you begin to ask questions using your computer or cell phone, AI programs are predicting what type of information you are seeking. Although these techniques are beyond the scope of this text, we want you to know about their existence.

We also want you to know about a different type of statistical modeling than that emphasized in this book, that of Bayesian statistics. Bayesian statistics creates a model of what can be expected from our experiment based on our present knowledge and the accuracy of our observations (Puga, Krzywinski, & Altman, 2015).

Weather forecasting would be one example that could use a Bayesian approach. That is, the more we know about how weather works, the better we can use seasonal records and our current observations to predict what tomorrow's weather will be. If our prediction does not fit the data, we can use this information to refine our predictions in the future. Such a process of drawing conclusions from data is thus based on both information we have learned in the past and new evidence that is collected in our current experiments. This type of modeling has been used in every area of psychological research including perception (Walker, Cotton, Ma, & Tolias, 2020).

Before we continue, we want to emphasize that there is no set number of methods for practicing science. Methods are developed in response to specific questions. Often our area of study determines which methods we use. For example, in sciences such as astronomy and zoology, scientists often use **retrospective, or post hoc (after the fact)**, methods; like Darwin, a zoologist might ask how a certain species developed. Clinical psychologists use similar methods when they speculate on the development of personality or the origin of mental illness. Other areas of psychology may rely on *single-case approaches* when the problem they are studying is rare, such as a specific brain disorder.

For example, Antoine Bechara and colleagues studied three patients with particular lesions or injuries to areas deep within their brains (Bechara et al., 1995). The first patient had damage to the hippocampus, the second had damage to the amygdala, and the third had damage to both areas. Using a classic conditioning procedure in which a loud sound (unconditioned stimulus) is paired with a color or sound (conditioned stimulus), these researchers were able to show different effects depending on the type of damage to the brain.

In particular, damage to the amygdala resulted in a lack of conditioning, but the person was able to describe the color or tone paired with the unconditioned stimulus of a loud sound. However, damage to the hippocampus resulted in the opposite effect: The person did show conditioning but could not describe the conditioned stimulus. The person with damage to both areas of the brain neither knew about the stimulus nor showed the effects of conditioning. We discuss single-case experiments as well as case study and other approaches to research later in this book.

We also want you to know about another approach, called **qualitative methods** (Camic, Rhodes, & Yardley, 2003; Colorafi & Evans, 2016; Creswell & Poth, 2018). Actually, you already know about qualitative methods, since you often ask your friends what they think about a certain movie or how they like particular pieces of music. Qualitative methods emphasize the subjective state of the person under study and are particularly useful when we wish to describe the experience of a particular person or group. For example, there are times when it is important to understand how clients experience various forms of psychotherapy, what it is like to go to an emergency room, what the experience of depression is for a particular person, or even how parents experience their children or their parenting roles. The focus group, in which individuals talk about a particular topic, consumer product, or political candidate in a group, is one type of qualitative research.

Qualitative research is useful when you initially study ongoing processes or seek to understand little-researched processes. For example, Martin, Sadlom, and Stew (2006) asked individuals to describe their experience of boredom, which helped these authors suggest how future research in this area should be conducted. In another qualitative study, Deutsch and Saxon (1998) interviewed mothers and fathers concerning how much praise or criticism each participant received from others concerning the amount of time spent in parenting and at work. Qualitative research can be especially useful for understanding the experience of individuals in their everyday life (Coloeafi & Evans, 2016). This approach is also used for understanding how a particular group of people who share similar cultural experiences approach a particular aspect of life.

Thus, one aspect of the qualitative tradition is a reliance on the phenomenological or subjective aspect of experience. There can be a real richness and depth in the description of subjective experiences. This approach has a long history in philosophy and sociology and has recently gained recognition in psychology. For example, qualitative approaches reveal that some parents of young children with conduct disorders often feel that they are being held hostage by the situation and lose contact with their partners. This work by Webster-Stratton and Spitzer (1996) describes the progression from coping efforts, to intense feelings of inadequacy, to feelings of helplessness in such parents. Such insight into the subjective feelings of parents of oppositional and aggressive children may help therapists develop better approaches for both the parents and the children in such situations.

An interesting and important approach in qualitative research is called the *action-project method* (Young, Valach, & Domene, 2005). In this approach, individuals are recorded discussing a particular topic. The individuals then watch the video separately, and the video is stopped at natural points approximately every minute. Each person comments on how he or she thought or felt during specific segments. These reactions form the data for the qualitative analysis. By repeating the process over a period of time, this procedure could be used to study various relationships, for example, between adolescents and their parents.

An intriguing use of this technique is with suicide attempters (Valach, Michel, Dey, & Young, 2002). Often studies of people who have attempted suicide rely on interviews with mental health workers (for example, Valach, Michel, Young, & Dey, 2006). For a variety of reasons, people who have recently attempted suicide often hold back or withdraw in these interviews. However, having these people watch the recorded interviews and comment on their thoughts and feelings supplies the qualitative researcher with additional, and sometimes more informative, data than the original interview itself.

A variety of sources are available for understanding qualitative research, including the journal *Qualitative Research in Psychology* and various books (for example, Camic, Rhodes, & Yardley, 2003; Sullivan & Forrester, 2018) and chapters in books (for example, Babbie, 2016, which discusses qualitative methods in relation to field research). For an overview in relation to clinical psychology, the interested reader should consult Kazdin (2016). As with any approach, the question arises of how to view the trustworthiness of the information presented, which is a topic of debate within the field of qualitative methods (Leong & Austin, 1996; Morrow, 2005).

In contrast to qualitative methods, the methods we emphasize in this book often are called **quantitative methods**. Using quantitative methods, there is generally an emphasis on behavior as opposed to experience as well as an attempt to describe constructs in terms of numbers and find laws or patterns that describe behavioral processes. In this chapter, we focus on three main approaches: *naturalistic observation, correlational approaches,* and *experimental manipulation.*

Let's consider the relationship between the scientist and the research participant in each method. With the naturalistic method, the scientist is passive and observes carefully the activity of the research participant. In this method, the scientist does not try to change the environment of the research participant. The research participant simply goes about normal activity and the scientist watches, preferably without influencing the participant's behavior. In this way, the scientist can make a detailed description of some aspect of the research participant's natural behavior.

In contrast, when using the experimental method, the scientist is more active, and the research participant's activities are restricted. The scientist intentionally structures the situation so that he or she can study the effect of a particular factor on the research participant's behavior. In between these two approaches are correlational methods, which may range from simple observation and correlation of factors to a more active manipulation, although without the same degree of manipulation and control characteristic of the experimental method.

As an analogy, we can view each method as an extension of the way a child explores his or her world, although the scientist and the child have very different goals in mind. A common way children and scientists begin to explore a new phenomenon is simply to watch it occur naturally. In the process of watching, the child and the scientist may describe what they see. However, the scientist usually goes further and uses mathematical terms to specify what is being seen. As a further step, both children and scientists extend their observations by means of limited interaction with the phenomenon.

Once we as scientists begin to see the relationship involved in a particular phenomenon, we can study it more precisely with more controlled manipulations. From this understanding, we may even move from science to technology and use our understanding profitably in our everyday lives. This general approach works for a child, and it works for scientists, too, although on a more complex level.

To provide you with a more accurate conception of how the scientific method is an extension of our everyday activities, we examine the three fundamental scientific strategies mentioned previously: the naturalistic observation technique, which is akin to a child observing a phenomenon; the correlational approach, which, depending on the situation, may be more or less like either of the other two methods; and the experimental method, which is akin to a child interacting with the phenomenon to learn more about it. Let us now turn to a more detailed discussion of each method.

Naturalistic Observation

Imagine that it is 20,000 years in the future and you have been sent to a strange part of the galaxy to study members of a particular species that have been described by astronauts as "cultural apes." Assume that you could arrive at the appointed place and remain undetected. Because you know virtually nothing about these cultural apes, the method of naturalistic observation would be an efficient way to get a general idea of these beings.

Like many scientists who have studied animals in the wild on 19th, 20th, and 21st century earth, you might set up a blind so that you would not be detected and observe the behavior of these apes. Often with animals in the wild, scientists try to find a place where the animals come together, such as a watering hole, and set up an observation post near this place. After some preliminary observations, you find that these cultural apes meet every morning in large structures. Consequently, you set up your observation post within one of these structures. Assuming that you remain undetected, what would you do next?

The answer is deceptively simple yet difficult to accomplish. *You just watch*. "Just watching" might be compared to seeing a movie in a foreign language that you do not understand. It is easy to see the interactions between the actors in the movie, yet you can only guess what they mean. In the beginning, the most difficult part of just watching is not to guess. Until you have observed a given interaction repeatedly, you can easily distort what you are seeing by your expectation that it occurs in a certain way. After much observation, you may begin to notice certain patterns of behavior by the apes. For example, they may say "Hello" each time they meet and "Goodbye" each time they leave each other. One hallmark of naturalistic observation is the discovery of patterns in the behaviors of different organisms.

An important part of the naturalistic technique is to record what you observe. At one time, the only method of recording was to reduce the observations to written notes or drawn pictures, much as Darwin did when he went to the Galápagos Islands. Today, however, we can record the observed behavior digitally so that we can listen or watch the behavior at a later time. Of course, scientists are still an important part of the process because they select what will be recorded and thus determine the observations for later analysis. So once we have observed many instances of the typical behavior of this species, we can withdraw from our observation post and begin to analyze our recorded observations. We now can begin to make summary statements about the natural ongoing behavior of these species.

Coming down to earth, let's consider the work of one scientist who has used the naturalistic method. Nikolaas Tinbergen (who received the Nobel Prize for his research) became interested in children who experienced autism spectrum disorder. Because little was known about the overall behavior of children with autism, Tinbergen began his work by using the naturalistic observation method and simply observed children who experienced autism spectrum disorder (Tinbergen & Tinbergen, 1972). Individuals with autism have difficulty in three separate areas. The first is social interactions. Children with autism do not connect with other children or adults in the manner that other children do. They do not look others in the eye or may appear to ignore others while being more interested in other aspects of their environment. The second area is communication. The communication patterns of those with autism spectrum disorder do not usually show the give-and-take of most conversations. The third area is behavioral processes. Individuals with autism spectrum disorder often display stereotypical behaviors and the desire to engage in the same behavior in a repetitive manner. As Tinbergen watched these children, he observed that there was a pattern to their atypical behavior in that this behavior appeared most often when they were in an unfamiliar social situation. Even a smile from a stranger might be followed by an attempt by the child with autism to withdraw from the situation.

Tinbergen also was interested in how autistic children were different from children without autism. To understand these differences, Tinbergen observed children without autism and also children with varying degrees of autism. He found that some facial expressions displayed by the children with autism differed from those children without autism. Thus, the naturalistic method offered a starting point for describing differences between children with and without a pattern of autism. In the over 50 years since Tinbergen did his original research, scientists have learned a great deal in how to diagnosis and treat autism spectrum disorder, as it is now called (Maxwell, Merckelbach, Lilienfeld, & Lynn, 2018; Müller & Fishman, 2018).

Another scientist who used the naturalistic method of observation is Konrad Lorenz. (Lorenz also received the Nobel Prize for his behavior research.) Lorenz (1952) described the behavioral interactions in a colony of jackdaws:

> A jackdaw sits feeding at the communal dish, a second bird approaches ponderously, in an attitude of self-display, with head proudly erected, whereupon the first visitor moves slightly to one side, but otherwise does not allow himself to be disturbed. Now comes a third bird, in a much more modest attitude which, surprisingly enough, puts the first bird to flight; the second, on the other hand, assumes a threatening pose, with his back feathers ruffled, attacks the latest comer and drives him from the spot. (p. 149)

At times, field observations may bring to light unknown patterns of behavior that in turn lead to new theories concerning these behaviors. For example, Nelson, Badura, and Goldman (1990) reviewed field studies showing a seasonal shift in the activity patterns of rodents. These animals tend to show the greatest amount of activity at night during the summer; in the winter, they show more activity during the daylight hours. Once a scientist knows an animal's patterns in the wild, ideas can be developed concerning the function such patterns serve as well as the mechanism that controls these patterns.

Rowsemitt (1986), for example, suggested that the winter/summer pattern of activity level in the rodents allows for better use of energy because being active during the winter day allows the animals to avoid extreme cold, and being active during the summer night allows them to avoid the extreme heat of day. To understand the mechanism that mediates the seasonal shift in activity patterns, a scientist may want to use the experimental method, which we discuss later in this chapter. Rowsemitt thought the change in the activity level of the rodent to be controlled by one particular hormone, testosterone, and centered his research in this area. However, one could also study a variety of other factors.

In summary, the naturalistic observation method has four characteristics:

1. Noninterference is of prime importance. Scientists using this method must not disrupt the process or flow of events. In this way we can see things as they really are, without influencing the ongoing phenomenon.

2. This method emphasizes the invariants or patterns that exist in the world. For example, if you could observe yourself in a noninterfering manner, you might conclude that your moods vary with the time of day, particular weather patterns, or even particular thoughts.
3. This method is most useful when we know little about the subject of our investigation. It is most useful for understanding the "big picture" by observing a series of events rather than isolated happenings.
4. The naturalistic method may not shed light on the factors that directly influence the behavior observed. The method provides a description of a phenomenon; it does not answer the question of why it happened.

To better understand how one variable is related to another, we use the correlational method. This method emphasizes scientists' ability to describe whether and to what degree two variables are associated with each other. Through this approach, scientists are better able to understand and describe our world.

The Correlational Approach

At times you may want to know whether a relationship exists between two events that cannot be manipulated easily. For example, you may want to know whether playing sports or drinking alcohol in high school is more associated with developing heart problems later in life. Barefoot, Dahlstrom, and Williams (1983) sought to determine whether one's emotionality (especially hostility) during medical school was associated with coronary heart disease later in life. It was. It is also the case that hostility in tweets (twitter responses) is related to heart disease (Eichstaedt et al., 2015).

Coronary heart disease is associated with hostility and a variety of studies since this initial study has shown this to be the case (see Busch, Pössel, & Valentine, 2017; Chida & Steptoe, 2009, for overviews). The opposite is also the case; that is, subjective well-being is associated with decreased mortality (Martin-María et al., 2017). In many studies of this type, it would be unethical or impractical to manipulate the events actively (for example, drinking alcohol in high school or provoking emotionality in graduate school). What you can do instead is collect information on the particular events under study without attempting to manipulate these events. Formally, we ask whether the frequency or magnitude of one event is related to the frequency or magnitude of the other event, but we do not attempt to establish how one event influences the other. This type of research is called correlational research, or natural association research (Ray, 1989).

Correlation is a measure of association that we present statistically in Chapter 5. For now, we can introduce some of the basic ideas. In correlational studies, the researcher is interested in asking whether there is an association between two variables, but he or she does not attempt to establish how one variable influences the other. Establishing that such an association exists may be the first step in dealing with a complex problem.

For example, a physiological psychologist might ask whether a person's pulse rate is related to the age of the person. To answer this question, she could simply measure various people's heart rates and correlate these measures with their ages. What would this tell us? First, it would tell us how heart rate and age are related. If they were to increase together—that is, if low heart rates were associated with young ages, medium heart rates with middle ages, and high heart rates with older ages—then this relationship would be called a *positive correlation*. However, if low heart rates were associated with older ages and high heart rates with younger ages, then this relationship would be called a *negative correlation*. Of course, as you will see in Chapter 5, few relationships are perfectly related to one another. Thus, we use a mathematical technique, the correlation statistic, to reflect the *degree* of an association between two variables.

What we cannot know from correlational research is whether either variable influences the other. That is, if two variables are related, what might the reason be for the relationship? As you begin to suggest factors that might have produced a high degree of relationship, you realize that a third, unspecified variable actually may have influenced the two variables in the correlational study. Thus, the nature of a correlational study is to suggest relationships but not to suggest which variable influences which other variable.

It is often said that correlation does not imply causality. For example, a researcher might want to know whether a relationship exists between the type of food one eats and the likelihood of one having a heart attack. An approach would be to examine the diets of people who have had heart attacks and of those who have not. What if there were a high association between eating steak, for example, and heart attacks? You could conclude little other than that there was an association or correlation between the two variables. The association of two factors does not in itself imply that one influences the other. However, if there is a *low* correlation between the events, you can infer that one event does *not* cause the other. A high degree of association is always necessary for establishing that one variable influences another; a correlational study is often the first step for providing the needed support for later experimental research, especially in complex areas.

✓ CONCEPT CHECK 2.1

A physician reports data from a study that compares the amount of television a person watches with his or her health. The point is made that people who watch more television are sicker; therefore, watching TV is bad for your health. Someone in the audience says, "No! That is not the case. Sick people have nothing to do, so they watch TV." Whom do you believe?

The Experimental Method

As we suggested, you already know a great deal about the experimental method. Indeed, all of us have used it in one form or another to explore our world since we were small children. Like the child, the scientist begins to interact with the phenomenon that he or she is studying and asks the question, "If I do this, what will happen?"

From these interactions, the scientist gains increased understanding of the phenomenon under study. To test this understanding further, the scientist asks, "Was what happened really a result of what I did?"

To give you a more accurate understanding of how scientists learn from interacting with the environment, let's consider the following line of fictitious research. Before we actually describe the study, we want to suggest that you, both as a scientist and as an informed consumer, consider many of the "scientific" claims that you hear on television or read in magazines and look for alternative explanations to the claims being made. Thinking scientifically is not something you do only when you design experiments; rather, it is a way of approaching all information.

Assume that the makers of a brand of children's cereal, Roasty-Toasties, claim that their breakfast cereal helps children to grow. In its enthusiasm to demonstrate the claim and add "scientific evidence" to its television commercials, the company designed the following experiment. A group of children were given daily a bowl of Roasty-Toasties with cream, bananas, and sugar. After several months, each child was weighed. It was found that they gained an average of 8 pounds each. The company concluded that the weight increase was due to the nourishing breakfast, and consequently the company recommended this breakfast for all children. When a thoughtful scientist heard the results, he admitted to their appeal but added that several things bothered him. One thing was that the children also ate lunch and dinner. Consequently, the weight gain might be due to the food eaten at these other meals.

Dismayed that it had not thought of that, the company designed a new experiment. This time it used two groups of children. The average age and average weight were the same for each group. For breakfast, one group received the recommended cereal with cream, bananas, and sugar; the other was given scrambled eggs. The two groups ate approximately the same foods for lunch and dinner. After several months, each child was weighed. It was found that there was an average gain of 5 pounds in the group that received the recommended breakfast cereal and an average gain of only 1 pound in the group that was given eggs for breakfast.

The company was excited and assailed our thoughtful scientist with the new findings, which seemed to confirm the earlier results. The scientist pointed out that he was even more impressed than before. However, he grew silent again, looked up, and asked, "Could the weight gain be caused by the cream, sugar, and bananas and not by the cereal?" Although the company was confident that the results were due to the cereal, logically the scientist was right. The entire effect could have been due to the cream, sugar, and bananas, and not the cereal.

Crushed by the scientist's keen insight, the company's researchers returned to the laboratory. After much debate, they decided to do the following experiment. As before, one group received the cereal with cream, sugar, and bananas for breakfast, but now another group received equal amounts of cream, sugar, and bananas (but no cereal) each morning. Once again, lunch and dinner were approximately the same for both groups, and the children's weights at the onset of the study were about the same. The company researchers were confident of replicating the earlier findings. After several months, they weighed each child. Much to their dismay, they found that children in both groups gained an average of 5 pounds. The group that received cereal did not gain more weight than the other group.

Definitions in the Experimental Method

The goal of the fictitious study just described was to determine whether eating cereal affected the growth of a child. The **hypothesis**, or idea being tested, was that eating the new cereal influenced growth. To test its hypothesis, in the final study the company gave its cereal to one group of children and not to a second group. The group that received the cereal is called the **experimental group**. The group that did not receive the cereal is called the **control group**. A control group is a group that is treated exactly like the experimental group except for the factor being studied. In this case, the factor being studied was the breakfast cereal. The control group was subject to the effects of all the same factors as the first group except for the cereal. The study is characterized in Table 2.1.

In any experiment we must define the terms in the hypothesis so that the hypothesis can be tested. To minimize possible confusion, the crucial terms in the hypothesis are defined clearly in reference to concrete operations. This definition is called an **operational definition**, and it forms a crucial link between our ideas and the world. Kerlinger (1973, 1986; Kerlinger & Lee, 2000) suggests that there are two types of operational definitions: *measured* and *experimental*. The first type relates to measurement and may specify both *how* observations are to be made and *what* is to be observed and measured. For example, in a study that measures anxiety during certain types of tasks, it would be necessary to define operationally how anxiety was measured. That is, were the anxiety scores derived from self-report measures, physiological measures such as heart rate, or observation of a specific behavior? The second type of operational definition refers to experimentation. This type of operational definition describes how experimental procedures are to be followed.

For example, in a study that examines the effects of praise on improvement in psychotherapy, it would be necessary to define operationally both what praise is and under what conditions it is to be given and withheld. In one sense, operational definitions function like a recipe for a cook. In the same way that it would be difficult to follow a recipe that said only “Heat eggs, milk, and flour” (without specifying the amounts of the ingredients, the temperature at which the mixture is to be heated, and so on), it would be impossible to test a hypothesis that said only “Anxiety hurts performance.” For a complete understanding, it would be necessary to specify (that is, operationally define) how anxiety is to be measured and how performance is to be assessed. Thus, one of the first tasks in developing a research study is to specify the operational definitions related to measurement and experimentation.

TABLE 2.1 ● Design of the Final Roasty-Toasties Study

	Pretreatment weight	Treatment	Posttreatment weight
Group 1 (experimental)	50 lb.	Cereal with cream, bananas, and sugar	*55 lb.*
Group 2 (control)	50 lb.	Cream, bananas, and sugar *without cereal*	*55 lb.*

It should be pointed out that we are using the term *operational definition* in the more popular and less technical sense. We are not speaking of the total definition of a construct, a point we discuss in more detail in Chapter 3. (For a more complete discussion of operational definitions and the related concept *reduction sentences,* see Suppe, 1977.)

In the imaginary cereal experiment, the researchers had to define operationally both what was meant by the construct *growth* and how the eating of the cereal was to be manipulated experimentally. The variable that an experimenter manipulates in an experiment is called an **independent variable**. A variable is said to be independent when its levels are established by the experimenter before the experiment begins and are thus independent of anything that happens during the experiment. In this manner the independent variable precedes and potentially influences the measurements that we take in an experiment.

The aspect of the world that the experimenter expects will be affected by the independent variable is the **dependent variable**. The dependent variable is so called because if a relationship does exist, its value *depends* on the independent variable. Our experience suggests that some people confuse the concept of independent and dependent variables. Thus, it is important to remember that (1) it is the independent variable that you as experimenter control and (2) it is the dependent variable that you as experimenter measure. Because the independent variable is the variable you manipulate, some people remember this by saying "**I** manipulate the **I**ndependent variable."

In the cereal experiment, the researchers hoped that growth would be enhanced through ingestion of the new cereal. However, because there are many aspects of growth (physical maturation, height, weight, intellectual ability, emotional maturation), the task of deciding which aspect to measure is difficult. Notice that their final decision to define growth operationally in terms of weight is quite arbitrary and ignores other aspects of growth that might be evaluated.

The difference in the magnitude of the dependent variable for the control and experimental groups is called the **treatment effect**. Ideally, the only difference between the experimental and control groups should be the independent variable. If we are certain that this is the case, then we can infer that any difference in the magnitude of the dependent variable is due to the independent variable. If there were more than one difference between the two groups, then we would not know which of these differences was responsible for any treatment effects that we might observe. (Note, however, that more complex experimental designs, called *factorial designs,* allow us to investigate the effects of two or more independent variables in the same experiment. This class of designs is discussed in Chapter 9.)

If we suspect that some unintended factor may also be operating, then the truth or validity of the experiment is seriously threatened, and the entire experiment must be questioned. In the second cereal experiment, the fact that the control group did not receive cream, bananas, and sugar constitutes an alternative explanation for that group's lower weight gain. Whenever two or more independent variables are operating, the unintended independent variables (those not chosen by the experimenter) are called **confounding variables**. In the second experiment, the cream, sugar, and bananas represent this type of confounding variable.

Other confounding variables may covary with the independent variable and be more difficult to notice. For example, assume that a researcher compared a new medication against a problem-solving approach for the treatment of anxiety. If she found the problem-solving approach to show a greater reduction in anxiety, could she conclude that problem solving produced the reduction? Although that is one possibility, it also may have been the case that spending time with a professional produced the reduction in anxiety. That is, because giving medications requires less time with a patient than discussing problem-solving techniques, the results found may not have been due to the independent variable as planned in the study but rather to a confounding variable: time with the patient. We discuss confounds and their influence in more detail in Chapters 7 and 8.

✓ CONCEPT CHECK 2.2

A study was conducted to determine whether using videotapes rather than audiotapes to learn German during the semester influenced the grade received on a final exam in German. Name the independent and the dependent variables in this study.

Causation

Before we continue, we would like to clear up some confusion experienced by a number of people when the word *caused* is used. In psychology, when we speak about an independent variable causing a change in a dependent variable, we mean that these two variables reflect a consistent association. That is, with every change in the independent variable, there comes a related and predictable change in the dependent variable. The idea of causality in science is generally a conclusion concerning the relationship between the independent and dependent variables. If research shows that each time we change one aspect of a situation, a predictable change follows in another aspect, we usually say that the first aspect caused a change in the second. As we repeat an experiment in varying situations and under different conditions, if the same relationship between the independent variable and the dependent variable continues to hold true, then we have more confidence in our conclusion. Many philosophers of science see causation as something we ascribe to the situation, and they remind us that what we are really doing is making inferences about the world (that is, epistemology) rather than making statements about what really exists (that is, ontology). As we discuss later in this chapter, modern physics is now regarded as the study of *observations of reality* rather than the study of *reality* itself.

Another way to discuss causation is to consider what conditions are required for an event to occur. In particular, we discuss *necessary* and *sufficient* conditions (compare Copi, 1986). A necessary condition is the condition that *must* exist in order for the particular event to occur. For example, it is necessary for a human to be a woman in order to become pregnant. Although it is a necessary condition, just being a woman in itself will not make one pregnant. For this to occur, there must also exist a sufficient condition. A sufficient condition for the occurrence of an event is a situation that, when it happens, produces the event. For example, we say that fertilization is a sufficient condition for producing an embryo.

To consider another example, we can say that the presence of oxygen alone is a necessary but not a sufficient condition for combustion to occur. However, the presence of oxygen at a certain critical range of temperatures is a sufficient condition for combustion. Likewise, we would not say that feeding milk to a child is a necessary condition for producing growth because a variety of substances fed to a child will produce growth. In the scientific literature, the word *cause* sometimes has been used in the sense of a necessary condition and at other times in the sense of a sufficient condition. This has led many researchers to suggest that we avoid the word altogether. Although we find it difficult to avoid the word completely, we have sympathy with the search for exact definitions. In this book, when we do use the word *cause,* we do not mean the one and only cause but rather the case in which two events (the independent variable and the dependent variable) are systematically connected in a variety of situations.

Exploratory Research

Psychologists often use the experimental method either more or less rigidly, depending on how much they already know about the phenomena they are studying and the types of questions they want to ask. In some cases, following an extensive library search for relevant information, a scientist may realize that almost nothing is known about a particular phenomenon and simply wonder what effect a given treatment will have on a person's experience. Given this situation, the scientist can use the experimental method in either of two ways.

First, when we have no idea what the effect of the independent variables will be, we are sometimes content to give the experimental treatment to a single group of research participants and then informally observe the research participants to get some idea of what aspects of behavior are affected by the independent variable. This initial exploratory use of the experimental method often is used in the initial stages of various experiments, including psychotherapy research and drug evaluation studies. Strictly speaking, this way of gaining information is not an experiment because it does not involve a control group or test specific hypotheses. It is no more than a

HELPFUL HINTS AND EXERCISES

DETERMINING THE TYPE OF RESEARCH USED

In reading about psychological research, knowing what type of procedure was used is important.

One of the first questions to ask is, Was anything done to influence the participant?
If the answer is no, you are dealing with a naturalistic observation design.
If the answer is yes, you are dealing with an experimental design.
In both cases you can then ask what the research was trying to accomplish.

simple demonstration that may provide either clues to fruitful independent variables for more refined analysis or potential attributes of behavior that should be reflected in the future selection of dependent variables.

The second way we use the experimental method as an exploratory tool occurs when we have some idea about which aspects of behavior will be affected by the independent variable and, consequently, have a reasonable idea about what types of dependent variables and control groups we should use, yet we do not understand the phenomenon well enough to make a specific prediction. In this way, our understanding of the phenomenon is refined progressively by a more detailed search for influential factors (independent variables that produce treatment effects) and by a more accurate estimate of their influence (measured by the dependent variables) on the phenomenon under study.

At other times, when we know a great deal about a particular topic, we can move beyond the exploratory uses of the experimental method. In these cases we are able to formulate specific predictions that reflect a more detailed theoretical understanding of the phenomenon. Because we have a clearer understanding, we can refine our independent and dependent variables and our use of control groups so that we can isolate more precisely the important relationships involved in the phenomenon being studied.

Whether we use the method of naturalistic observation, the correlational approach, or the experimental method, the task before us is to make inferences about the research participants' experience from the behaviors we observe. These inferences generally are related to our hypothesis or some larger theory that we want to evaluate. Thus, after we look at the world through these methods, we are faced with the task of deciding how to evaluate new information that we receive in light of both the methods used and our theoretical perspective. To accomplish this task, we use reason and logic. In particular, we ask whether the results of our methods as well as our conclusions are valid. To aid you in evaluating your own research and that of others, we focus on the question of validity and differentiate among some common types of validity in the next section. We also take a brief look at propositional logic.

Logic and Inference: The Detective Work of Science

Perhaps you have heard the story of our friend from Boston who got up every morning, went outside his house, walked around in a circle three times, and yelled at the top of his voice. His neighbor, being somewhat curious after days of this ritual, asked for the purpose behind his strange behavior. The man answered that the purpose was to keep away tigers. "But," the neighbor replied, "there are no tigers within thousands of miles of here." To which our friend replied, "Works quite well, doesn't it?"

How could we demonstrate to our friend that his yelling is not causally related to the absence of tigers? One strategy might be to point out that the absence of tigers might have come about for other reasons, including the fact that there are no tigers roaming in the greater Boston area. In technical terms, we would say that yelling could be a *necessary* condition but not a *sufficient* condition for the absence of tigers. Our friend's reasoning was incorrect because it overlooked many other plausible explanations for the obvious absence of tigers. Although our friend sought to infer a relationship between his yelling and the absence of tigers, his inference was weak.

Logic is particularly important in science as an aid to answering the question, "What question should my experimental study answer to test my ideas about the world?" That is, logic can help us to answer questions of *inference*. Inference is the process by which we look at the evidence available to us and then use our powers of reasoning to reach a conclusion. Like Sherlock Holmes engaged in solving a mystery, we attempt to solve a problem based on the available evidence. Did the butler do it? No, the butler could not have done it because there was blond hair on the knife and the butler had black hair. But perhaps the butler left the blond hair there to fool us. Like a detective, scientists try to determine other factors that may be responsible for the outcome of their experiments or to piece together available information and draw general conclusions about the world. Also like the detective, the scientist is constantly asking, "Given these clues, what inference can I make, and is the inference valid?" Logic is one method for answering these questions.

Validity

Logical procedures are also important for helping us to understand the accuracy or validity of our ideas and research. *Valid* means true and capable of being supported. Historically, we have discussed various types of validity in psychology, which arise from differing contexts. These contexts range from developing types of tests to running experiments. The overall question is, "Does a certain procedure, whether it is an intelligence test or an experiment, do what it was intended to do?" In the next chapter, we will emphasize validity associated with issues of measurement including the ability of our instruments and experimental apparatus to give us true results. In this chapter we look at psychological research and focus on two general types of validity and the logic of our conclusions (Campbell & Stanley, 1963).

The first is **internal validity**. The word *internal* refers to the experiment itself. Internal validity asks the question, "Is there another reason that might explain the outcome of our experimental procedures?" Students are particularly sensitive to questions of internal validity—for example, when it is time for final exams; they can make a number of alternative suggestions about what the exam actually measures and why it does not measure their knowledge of a particular subject. Like students, scientists look for reasons (threats to internal validity) why a particular piece of research may not measure what it claims to measure. In the case of our friend from Boston, the absence of tigers near his house could have reflected a long-standing absence of tigers in his part of the world rather than the effectiveness of his yelling. We discuss specific threats to internal validity in Chapter 7.

The second type of validity is **external validity**. The word *external* refers to the world outside the setting in which the experiment was performed. External validity often is called **generalizability**. Remember the story of Semmelweis from Chapter 1. His finding that the deaths of the mothers who had just given birth were the result of physicians touching them after handling diseased tissue was true not only for his hospital but also for all other hospitals. Thus, in addressing the question of external validity of Semmelweis's work, we would infer that his answers could be generalized to other hospitals with other women and not just to his own original setting. Now consider the story of Galen. We would not fault his research concerning why the woman did not sleep, but we would say that it lacked external validity, or generalizability. Although the insomnia of one particular woman was attributed to her love of a particular dancer, it is not true that all women who suffer from insomnia are in love with dancers. In summary, internal validity refers to the internal consistency or logic of the experiment that allows the results to be meaningful. External validity, however, refers to the possibility of applying the results from an internally valid experiment to other situations and other research participants.

We logically design our research to rule out as many alternative interpretations of our findings as possible and to have any new facts be applicable to as wide a variety of other situations as possible. In many real-life situations in which external validity is high, however, it is impossible to rule out alternative interpretations of our findings. In a similar way, in laboratory settings in which internal validity is high, the setting is often artificial, and in many cases our findings cannot be generalized beyond the laboratory. Consequently, designing and conducting research is always a trade-off between internal and external validity. Which one we emphasize depends on the particular research questions being asked.

Before we continue we would like to clear up one misconception that some students have. It is the idea of designing "the one perfect study." Although we strive to design good research, there are always alternative explanations and conditions not studied in any single study. It is for this reason that Donald Campbell, who introduced scientists to the idea of internal and external validity, also emphasized the importance of replicating studies. If the same study is performed a number of times with similar results, then we can have more assurance that the results were valid. Even better, if the study is performed in a variety of settings around the world, we have even more confidence in our results. Currently, there is a strong emphasis and encouragement to perform replications of studies (Koole & Lakens, 2012; National Academy of Sciences, 2019; Shrout & Rodgers, 2019). The American Psychological Society (APS) is supporting such groups as the Center for Open Science (https://cos.io/about/mission/) in getting researchers to register their replication study before it is performed. In this way, one can see after a number of the replication studies are completed which ones found data consistent with the original study and which did not.

In the next section, we examine propositional logic. However, before you begin that section, we suggest that you try to solve the problem presented in the box, Understanding How We Reason. (*Hint:* In problem solving, as well as in science, it is often more important to show what is not true rather than just focus on what is true. The solution to the problem relies on one type of logical argument, called *modus tollens*, that you will learn about in the next section.) You can also read research based on 228 experiments in which individuals sought to solve the "think square" task in terms of how the task was approached (Ragni, Kola, & Johnson-Laird, 2018).

✓ CONCEPT CHECK 2.3

A tabloid newspaper recently described a diet in which sleep caused a person to lose weight. All the person needed to do was to exercise directly before going to bed and not eat for 6 hours before this time. The diet worked not only with the original group of research participants but also with a variety of other groups. Discuss the internal and external validity of this study.

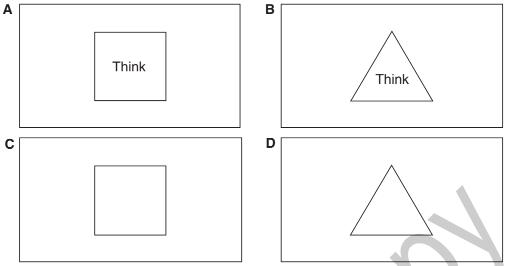
Propositional Logic

In the previous section, we introduced the terms *internal validity* and *external validity* and emphasized the scientist's attempt to rule out alternative explanations. In this section, we emphasize the way in which a scientist relies on the rules of formal logic to both *deduce* and *induce* valid conclusions. As a starting point, keep in mind that **deduction** (to deduce) is the process by which one moves from a general theory to particular statements concerning data, whereas **induction** (to induce) is the process by which one moves from a particular set of data to a general theory or concept.

When we begin with a statement and arrive at its logical consequences, this is called *deductive* reasoning. For example, we use deductive reasoning when saying, "If it is true that schizophrenia is genetically determined, then we should find greater similarity in the presence or absence of the disorder between twins than between

UNDERSTANDING HOW WE REASON

"THINK" SQUARES



The illustration in this box represents four cards: A, B, C, and D. Assume that each card has a square with or without the word "Think" on one side, and a triangle with or without "Think" on the other side. Which of the cards would you have to turn over to determine whether every card that has a "Think" square on one side has a triangle without "Think" on the other?

We will make a prediction, based on previous experience, that you chose cards A and D. You may be in good company—many people choose these cards—but you were wrong. However, you were right when you chose card A. As you correctly reasoned, card A must have a triangle without "Think" on the other side for the statement to be true. However, if you turned over card D, which we did when we first tried the problem, you missed the point of the statement. The statement talks about what is opposite a "Think" square and says nothing about what is opposite a triangle. The other card that we need to turn over to solve the problem is card B. We need to demonstrate that the negative of the statement is not true; that is, we need to demonstrate that there is no card that has a "Think" square on one side and a "Think" triangle on the other side.

We begin with the hypothesis that all cards with a "Think" square on one side have an empty triangle on the other side, which we want to test. We then move to the real situation in which there are four cards. One of these four cards, A, has a "Think" square, and we can test the truth of the statement just by turning over card A. Now comes the problem. We have tested all the visible "Think" squares (card A), so what do we do next? If it is true that all "Think" squares have empty triangles on the other side, then it must also be true that no card with a "Think" triangle has a "Think" square on the other side. Once we realize this, we know that we must look at card B to test this assumption. Thus, the correct answer is cards A and B.

Source: Based on Wason (1977).

strangers." On the other hand, when we begin with an observation and figure out a general rule that covers it, this is called *inductive* reasoning. For example, inductive reasoning might be of the form, "I just saw a monkey use sign language to ask me for food; therefore, it is true that monkeys can communicate with humans." In summary, deductive reasoning goes from theory (the premise) to data (the conclusion), whereas inductive reasoning goes from data to theory.

Suppose a friend said to you, "You know, all experimental textbooks are really dull." You might respond, "That's not true; I am reading one right now that is really interesting." (Well, what did you expect that we would have you say?) This is a logical way to disprove the statement, "All experimental textbooks are really dull." By finding an exception to a statement, you can show it to be false.

To begin an introduction to deductive logic, let's consider this procedure in a formal way. You begin with the statement, "If this is an experimental textbook, then it will be really dull." This statement is presented in a certain "If . . . then" form. Each part has a particular name. The first part, "If this is an experimental textbook," is called the **antecedent**. The second part of the statement, "then it will be really dull," is called the **consequent**. The antecedent is referred to by the letter *p* and the consequent by the letter *q*. We now will discuss four types of propositions of the form "If *p*, then *q*." These four types are presented in Table 2.2. We should point out that other forms besides "If *p*, then *q*" are studied in propositional logic (compare Copi, 1986; Kourany, 1987).

Let's return to our original statement. Several logical consequences may follow from that statement. Suppose this is indeed an experimental textbook; then it

TABLE 2.2 ● Forms of Propositional Logic

	Modus Ponens (confirmatory) If *p*, then *q*. *p*. Therefore, *q*.	Modus Tollens (disconfirmatory) If *p*, then *q*. Not *q*. Therefore, not p.
Valid Arguments	If anxiety is increased, then heart rate will be increased.	If anxiety is increased, then heart rate will be increased.
	Anxiety is increased. Therefore, heart rate will be increased.	Heart rate is not increased. Therefore, anxiety is not increased.
	Affirming the Consequent If *p*, then *q*. *q*. Therefore, *p*.	Denying the Antecedent If *p*, then *q*. Not *p*. Therefore, not *q*.
Invalid Arguments	If anxiety is increased, then heart rate will be increased.	If anxiety is increased, then heart rate will be increased.
	Heart rate is increased. Therefore, anxiety is increased.	Anxiety is not increased. Therefore, heart rate will not be increased.

follows that it will be really dull. In propositional logic, the reasoning is written out as follows:

If this is an experimental textbook, then it will be really dull.

It is an experimental textbook.

Therefore, it will be really dull.

Technically, this is an argument of the form "If *p,* then *q;* given *p* (it is an experimental textbook), then we can conclude *q* (it will be really dull)." This form of argument is called confirmatory reasoning or **modus ponens**. Confirmatory reasoning is logically valid. We use it often in both our everyday lives and science.

However, not all forms of argument are logically valid. Suppose we have an instance in which *q* is true; what does this say about *p*? That is, if I can find a book that is dull, does that demonstrate that it is an experimental textbook? Of course not. There may be other kinds of books that are dull. Yet people make the following type of argument:

If anxiety is increased, then heart rate will be increased.

Heart rate is increased.

Therefore, anxiety is increased.

In such a situation, we can think of many reasons why heart rate could be increased (for example, running up the stairs, doing mental arithmetic) without anxiety being increased. This form of argument is called **affirming the consequent**. Although it is a logically invalid form of argument, one sees it used almost daily. The cartoon on the next page offers one example of this form of reasoning. Technically, the argument could be stated as follows:

If I were a cat, I would have four legs.

I do have four legs.

Therefore, I am a cat.

Let's now take the situation in which *p* is shown not to be true. What can we logically conclude? Again, we begin with a statement of the form "If *p,* then *q.*"

If I were a cat, I would have four legs.

I am not a cat.

Therefore, I do not have four legs.

Of course, this is not a valid argument. We know that other animals besides cats have four legs. The dog in the cartoon is not a cat but has four legs. Thus, if *p* is not true, it does not follow that *q* is also not true. This invalid form of argument is called **denying the antecedent**. Consider the following:

If Freud's theories are correct, then his therapy will be effective.

Freud's theories are not correct.

Therefore, his therapy will not be effective.

This is logically invalid. However, we still hear people using this form of invalid reasoning by suggesting that because Sigmund Freud's theories have been shown to be incorrect, his therapy cannot be effective. The effectiveness of a particular therapy is determined empirically; such a determination has little to do with theoretical formulations.

This brings us to the fourth situation. If the consequent is not true, then what does this tell us about the antecedent? (If *q* is not true, then *p* is . . .?) If *q* is not true, then *p* is also not true. Returning to a previous example and modifying it appropriately, we have the following argument:

Source: ScienceCartoonsPlus.com

If I were a cat, I would have four legs.

I do not have four legs.

Therefore, I am not a cat.

This is a valid form of reasoning. It is called *disconfirmatory reasoning,* or **modus tollens**. In the Freud example, we could make the following valid argument using modus tollens:

If Freud's theories are valid, then his therapy will be effective.

His therapy is not effective.

Therefore, Freud's theories are not valid.

Karl Popper (see box, Philosophy of Science) suggests that modus tollens arguments are at the heart of the scientific testing of theories. He points out that our theories should lead to some prediction. If the prediction is found to be untrue, then we conclude that our theory is incorrect. This is valid reasoning. However, the opposite is not the case. Finding that a prediction is true does not logically lead one to conclude that the theory is true, any more than having four legs made the dog in the cartoon into a cat. Research results may fail to refute a theory, but they cannot prove the theory to be correct. Thus, Popper suggests that our efforts in science should be concentrated on attempting to disprove hypotheses; he calls this procedure *falsification*.

PHILOSOPHY OF SCIENCE

SIR KARL POPPER (FALSIFICATION APPROACH)

Sir Karl Popper has devoted much of his career to answering the questions "What is science?" and "How is science performed?" Although these questions may at first seem easy to you, consider such areas as astrology and Marxism. Could these approaches be considered scientific? Why not?

Falsificationism is the name given to Popper's description of how science is performed. Falsificationism suggests that science should be concerned with disproving or falsifying theories through logic based on observation. How is this accomplished? First, a scientist must create a consistent, falsifiable hypothesis. A falsifiable hypothesis is one that can be shown to be false. For example, the hypothesis "It will rain in Tuscaloosa, Alabama, on Tuesday, December 23, 2021" is falsifiable: If it does not rain on that day, the hypothesis will be shown to be false. Likewise, the hypothesis "All objects regardless of weight will fall to earth at approximately the same speed" is falsifiable; it can be tested by experiment. However, a hypothesis such as "ESP (extrasensory perception) exists" is not falsifiable. Even the hypothesis "Gravity exists" is untestable. It may be true that both ESP and gravity exist, yet until the hypothesis is stated in a form that can be falsified, the hypothesis is not testable. Second, once a scientist has a falsifiable hypothesis, the task is to develop a test of the hypothesis. Third, the hypothesis is tested. Fourth, if the hypothesis is shown to be false, a new hypothesis is developed.

Using this model, Popper emphasizes science as a process for the elimination of false theories. Furthermore, Popper suggests that the elimination of false theories is aided by placing one theory in competition with other theories. In this sense, science is a form of natural selection with a more fit theory being the result (Popper, 1959). If you accept the suggestion that all psychological research is inductive in nature, you must then conclude with Popper that the major role of science is the falsification of incorrect theories. This line of reasoning also leads one to conclude that science, particularly psychology, never *proves* a hypothesis. Science, according to Popper, only shows that the hypothesis has not been proved false.

Note: Although Popper usually is discussed in terms of the falsification position, his later writings emphasized research programs rather than single theories. Even for Popper, science is more than just eliminating false theories.

We want to emphasize that deductive logic is one of many tools of the scientist. Like the experimental method, logic is an approach to knowledge designed to help us to evaluate and direct our research questions. Because of the complexity of the world in which we live and the limits of our own minds in perceiving this complexity precisely, we find ourselves as scientists using a combination of both inductive and deductive approaches to knowledge in our science as well as in our lives. We often use deductive and inductive approaches as a means of gaining information, which becomes a clue as we attempt to interact with and understand the world in which we live. Logic offers us a means of evaluating the inferences we draw from these clues. Logic helps us to understand the limits on our claims to certitude. Quite often logic helps us to see that we do not know enough to make any claim at all. In this manner, logic tends to make the scientific process conservative in its claims. However, it does not follow that our research topics, our ideas, or our theories also must be conservative.

Scientific Observation: The Raw Data of Science

Have you ever heard the question, "If a tree fell in the middle of the forest without anyone around, would there be any sound?" This question reflects a philosophical problem in science that was solved in physics at the beginning of the 20th century. Until that time, the notion of physicists was that they study *events* in the world. In this earlier worldview, the job of the scientist was to be a passive observer and accurately watch events that take place either in the real world or in experiments. There was no thought that the process of observing might influence the perception of the very events being observed. From the earlier perspective, it was meaningful to ask whether the sound existed. However, this has changed.

According to modern physics, scientists do not record events. Instead, scientists record their *observations of events*. They record their experience of the world and base their science on these perceptions. This development amounts to a simple acceptance of the fact that in science we can get no closer to the world than our observations of it.

Let's return for a moment to the child who is discovering the world for the first time. Imagine that you are a small child who is still crawling and cannot yet stand. As you move around your world, what do you see? What do you know about events that take place and objects that are more than 3 feet above the ground? Some events you may know by their sound, such as a passing car or your father's electric razor. Other events you may recognize only by their smell, such as the cooking of bacon or bleach being added to the wash. Other events you may know only from the sensation involved, such as your father or mother picking you up and throwing you in the air and catching you. Suppose someone could talk with you at this age and ask you to describe what the world was like. What would you say? How would the adult you were talking to react to your description? Would he or she say it was a true, accurate, and acceptable view of the world?

As you begin to answer these questions, you see that your description as a child was from your own perspective. You also may realize that it is difficult to say whether this description was true or false. From your viewpoint as an adult, it was incomplete. In the same way that the view of a child's world is relative to where and when the child lives and observes the world, the view of the scientist and, consequently, the *facts of science* are *relative* to the current notions of working scientists and the instruments that they use to make observations.

PHILOSOPHY OF SCIENCE

THOMAS KUHN

When Newton said, "I stand on the shoulders of giants," he was referring to the people who came before him and on whose work he was able to build his scientific system. Many of us have similar ideas when it comes to the progression of science. We think that each new discovery is simply added to old discoveries, with the results being a gradual accumulation of knowledge.

Thomas Kuhn (1970) suggests that this view is wrong. Kuhn proposes that science actually

(*Continued*)

(Continued)

goes through a series of revolutions. Following each revolution, a new system or method for performing science is instituted. The new system or worldview is called a paradigm, or set of assumptions, which guides scientific activity until a new revolution and paradigm shift take place. The stable period between revolutions is called *normal science*. Normal science is the process of problem solving, which most of us think of when someone uses the term *science*. Normal science for Kuhn is always science performed in relation to a particular paradigm.

As an example of the role of paradigms, assume that you were a mapmaker in a culture that believed the world was flat. You would draw your maps as if the world were flat because that was the accepted belief. As a mapmaker, you would never think to question this belief; it was a given in your task of drawing maps. Then there came a mapmaker's version of a scientific revolution. The paradigm shifted to that of a world that was round. As a mapmaker, you would now draw the world as if it were round and you would continue with this system until a new revolution came along. This, of course, was the replacing of the Earth as the center of the solar system with the sun as the center. In the same way that mapmakers work in relation to present-day assumptions and beliefs about the world, Kuhn suggests that scientists also work in relation to a set of beliefs or paradigms until these are replaced by a revolution.

The current notion concerning science and accepted methods, which encompasses a philosophical way of seeing the world, is called a **paradigm**. Philosopher of science Thomas Kuhn has elaborated a particular view of how science progresses, based on the concept of paradigm (see box, Philosophy of Science). Although there is much debate about the exact meaning of the word *paradigm,* most scientists understand it to mean shared beliefs, which include topics to be studied and the types of answers that will be given. For example, the current scientific paradigm in psychology emphasizes the importance of quantitative measurement on a number of levels. Thus, scientific psychology, as you will learn it, directs you toward topics that can be measured quantitatively in terms of biological, psychological, and cultural factors.

Not only are the results and conclusions of our research relative to our current notions of science, but they also may relate to our own psychology. Consider the role of the experimenter in the psychological experiment. Do you think your own state (hungry, sad, tired, excited, and so on) could influence the data of an experiment? That is just one of the factors that we consider later in this book. The important point now is to realize that the state of the experimenter is important. Because the scientist is not *passive* but is *actively* searching for answers, he or she can actually influence the event being recorded by the very manner in which the observation is being made. The scientist can change the world and our understanding of it. Thus, the scientist is more than a passive observer; he or she is a real actor in the drama of science.

Evaluating Scientific Research

Regardless of the amount of work involved in scientific research, an extremely important aspect of *any* research endeavor is whether the final product is worthy of being reported to the scientific community. We must ask whether our conclusions are *accurate,* capable of being *replicated,* and *relevant* to others. In this book we emphasize four ways to ensure the high quality of our research.

The first is through impartial, systematic observation using logically sound experimental design. The experimental method based on random sampling and assignment (as we describe later) is the most powerful class of research design currently available. We emphasize this technique in the initial chapters of this book and later discuss some other scientific approaches.

The second way to ensure meaningful research is through statistical description and inference. We show how statistics can help us to decide whether our results are due to some causal agent or merely to chance.

The third method of quality control is through reason and logic. In discussing logic we emphasize types of validity as well as types of propositional logic and how they help us to evaluate research.

The fourth and final way is by emphasizing perspective and context. In particular, we suggest that conclusions be viewed from the perspectives of the scientist, the research participant, and the informed consumer.

Although this book emphasizes the perspective of the scientist, it is important to remember the experiences of the research participant and the perspective of the informed consumer if our conclusions are to have meaning. We believe that through these four ways of evaluating research a person can use science and maintain the high level of excellence that a science of behavior and experience requires.

Communication in Science

Unlike the child who is busy learning about the world, scientists must share what they learn about the world with other people, especially other scientists. More than 2,000 years ago, Aristotle emphasized this when he taught that science has two parts: inquiry and argument. In modern terms, inquiry is represented by the research that answers our questions about the world, and argument refers in part to the scientist's responsibility to inform others of the findings. Consequently, we design our research, record our observations, and summarize our findings in a manner that others can understand. For scientists to answer a question in terms that only he or she can understand would not be complete science because it is not shared knowledge. The final product of mature science is a communication that summarizes a conclusion about the world and is directed to both scientists and nonscientists.

Learning to communicate in science may be compared to learning a foreign language. One of your first tasks is to learn the vocabulary of science. You need to understand what a scientist means by certain words. You initially may say that is easy because many scientists speak English anyway. That may be true, but it can also be a problem because English words can have slightly different or even totally different meanings when used in the context of a scientific statement.

For example, suppose you were reading a newspaper article concerning a new discovery in subatomic physics. The article is about particles with "color and charm." If you were to talk to a physicist, you would find that in this context *color* and *charm* have nothing to do with colors or with the particles being appealing. These words have special meaning for the physicist.

Likewise in psychology, common words may be used in a special or technical way. For example, B. F. Skinner discussed *negative reinforcements* as applied to people, yet negative reinforcements have nothing to do with punishment, as many people think. Likewise, Carl Jung invented the words *extravert* and *introvert,* which have a technical meaning different from their uses in newspapers and magazines. Even as common a word as *sex* was given a scientific meaning (by Freud) as distinctive as the physicists' terms *color* and *charm.*

At first, the language of science may seem strange. Yet, as with any language, once you learn some words and phrases, you can begin to understand what is going on. This understanding will be useful not only to those of you who pursue careers as scientists, but also to all of us in our daily interactions with the world as we try to understand what we read about science and strive to become more educated consumers. You have a twofold task. First, you must seek to understand how words are used in research in a technical way. You cannot just assume that because you have heard a word, you already know its meaning. Second, in writing your own reports, you must seek to define your words and ideas as precisely as possible so that others can understand and follow what you are saying.

Key Terms

Term	Page
affirming the consequent	50
antecedent	49
confounding variables	42
consequent	49
control group	41
correlational approach	32
deduction	48
denying the antecedent	51
dependent variable	42
experimental group	41
experimental method	32
external validity	46
falsificationism	52
generalizability	46
hypothesis	41
independent variable	42
induction	48
internal validity	46
modeling	32
modus ponens (confirmatory reasoning)	50
modus tollens (disconfirmatory reasoning)	51
naturalistic observation	30
operational definition	41
paradigm	54
qualitative methods	33
quantitative methods	35
retrospective method (post hoc method)	33
treatment effect	42

Concepts

1. Overview of science
    A. Hypothesis
    B. Observation, correlation, and experimentation
    C. Inference and conclusion
    D. Modeling
    E. Qualitative methods
    F. Quantitative methods
2. Types of questions
    A. "If I do this, what will happen?"
    B. "If I do this, I expect this will happen."

3. Role of scientist
    A. In naturalistic observation
    B. In experimental method

4. Naturalistic observation
    A. Four characteristics
        - Noninterference
        - Determining patterns
        - Useful for "big picture"
        - Descriptive

5. Correlational approach
    A. Positive and negative correlations
    B. Association versus causality

6. Experimental method
    A. Key question
        - "Was what happened really a result of what I did?"
    B. Definitions
        - Hypothesis
        - Experimental group
        - Control group
        - Operational definition, two types
        - Independent variable (also called treatment variable)
        - Dependent variable
        - Treatment effect
        - Confounding variables

7. Causation in science
    A. Necessary conditions
    B. Sufficient conditions
    C. Correlation

8. Validity
    A. Internal validity
    B. External validity and generalizability

9. Forms of propositional logic
    A. Deduction and induction
    B. Correct reasoning
        - Confirmatory (modus ponens)
        - Disconfirmatory (modus tollens)
    C. Incorrect reasoning
        - Denying the antecedent
        - Affirming the consequent

10. Falsificationism
    A. Karl Popper

11. Paradigm
    A. Thomas Kuhn

12. The language of science

Summary

1. Science is one way of learning about the world that involves articulating an idea or hypothesis, using experience developed in research to evaluate the idea, and drawing conclusions or inferences from experimentation and observation about the idea or hypothesis.

2. Observation is an important part of science. The naturalistic observation procedure emphasizes observation and has four characteristics: (a) noninterference, (b) observations of patterns and invariants, (c) development of the "big picture" or learning about an unknown process, and (d) provision of descriptions rather than pinpointing specific factors that influence one another.

3. Experimentation is also an important part of science. It offers a means of creating control and determining the manner in which one factor influences another. This determination is aided by the use of a control group, which allows the researcher to evaluate the effects of the independent variable on the dependent variable.

4. Another type of research is called correlational study. The purpose of this research is to determine the association between two variables but not the manner in which one variable affects another.
5. A very important part of science is the use of logic and inference. The particular task is to draw conclusions and rule out alternative hypotheses. The study of propositional logic points to both the logical and the illogical conclusions that may be drawn from general statements.
6. A researcher must question the validity of conclusions drawn from research. Two major types of validity are discussed. Internal validity refers to the experiment itself and asks whether there are alternative explanations (such as confounding variables) that would invalidate the reported relationship between the independent and dependent variables. External validity poses the question of generalizability and asks to what other groups or situations a particular set of findings might be applicable.
7. Science reflects a history of observations of events. As a recorder of observations, it is important for researchers to be sensitive to factors that can influence the records they make and to understand that any record is always presented from a certain perspective, recently called a *paradigm*. It is likewise important that communications in science be clear and be stated in such a manner that they can be evaluated.

Review Questions

1. What is one difference between *qualitative* and *quantitative* methods?
2. What are four characteristics of naturalistic observation?
3. What is an *experimental group*, and what is a *control group*?
4. What is an *operational definition*, and is there more than one type?
5. In the final cereal experiment, identify the independent variable and the dependent variable.
6. Distinguish between a *sufficient* and a *necessary* cause.
7. What is meant by the terms *internal validity* and *external validity*?
8. What does the term *falsificationism* mean in science?
9. How would modern physics answer the question, “If a tree fell in the forest without anyone to hear it, would there be any sound?”
10. What are the two invalid forms of propositional logic discussed in this chapter? Give an example of each.
11. What are the two valid forms of propositional logic discussed in this chapter? Give an example of each.

Discussion Questions and Projects

1. Use the library as the site for a naturalistic observation study. Go to the library, find a place from which you can observe, and record what you see. One focus might be the pattern of interactions among people in the library. If you were an outsider looking at these data, what might you conclude about the function of the library for students?
2. Give the "Think" squares problem in Box 2.1 to some of your friends. Using naturalistic observation, record what they do as they go about solving the problem. You might time them, record what they say (if anything), notice facial expressions, and so on.
3. Have another group verbalize what they are thinking as they try to solve the "Think" squares problem. You might decide whether there are similarities in the verbalizations of the different people. If you record the verbalizations, it will make the task easier. How do you go about deciding whether the verbalizations of two different people are similar? What categories do you look for?
4. Put people who are knowledgeable about a particular sport in one group and people who know little about it in another group. A knowledgeable person should describe in detail some particular play or move from the sport; you then ask both groups to recall what was said. Notice whether there is any difference between the two groups in the amount of recall. What other differences are there between the two groups?
5. Discuss how you might turn the observations in Questions 3 and 4 into experiments. What would be the independent and the dependent variables?
6. Assume that you followed the directions in Question 4 and found that people who knew about the sport remembered more. Discuss the following conclusion: "This experiment demonstrates that playing sports helps to increase your ability to remember, so sports should be required in all schools."
7. Discuss the statement, "Scientists do not record events but only their observations of events."
8. An experimenter was interested in creativity. In particular, she wanted to know whether a person is more creative at one time of day than another. At the time, she was teaching two sections of an introductory psychology course. One class met from 8:00 to 9:00 a.m. and the other from 4:30 to 5:30 p.m. She used a well-known creativity test and gave the test to each of her classes. When she scored the test, she found that those who took the test in the morning did better than those who took it in the afternoon. The experimenter concluded that, in general, college students are more creative in the morning than in the afternoon. Discuss this conclusion. Are there other ways in which these data might be interpreted?

✓ Answers to Concept Checks

2.1 Neither, we hope. Although either or both could be correct, a correlation will not give you that information. A correlation will only describe a relationship but not the direction of the relationship. That is, you cannot state whether one variable influenced another or even whether there was a third variable that influenced both.

2.2 The independent variable is the variable that the experimenter manipulates. The type of tape, video versus audio, is the independent variable. The dependent variable is the score or measurement influenced by the independent variable. In this case, the final exam grade is the dependent variable.

2.3 Because the results were obtained in a variety of settings, we may assume there are few problems with the external validity of the study. However, there may be a problem with the internal validity because the experiment does not actually show that sleep is a necessary condition for weight loss to take place. It may be that the exercise and not eating for 6 hours will result in a weight loss even without sleep.