

Unit II

Research Methods: Thinking Critically With Psychological Science

Modules

- 4** The Need for Psychological Science
- 5** The Scientific Method and Description
- 6** Correlation and Experimentation
- 7** Statistical Reasoning in Everyday Life
- 8** Frequently Asked Questions About Psychology

In a difficult moment—after an argument with a loved one, a social embarrassment, or a bad grade—to whom do you turn? For advice and comfort, we often turn to friends and family, or search online. Psychology can also shed insight. Psychologists start with the questions that intrigue all of us: How can we be happier, healthier, and more successful? What can we do to improve our relationships? Why do people act and think as they do? But psychological *science* takes it a step further and uses careful research to separate uninformed opinions from examined conclusions.



Module 4

The Need for Psychological Science

Module Learning Objectives

- 4-1** Describe how hindsight bias, overconfidence, and the tendency to perceive order in random events illustrate why science-based answers are more valid than those based on intuition and common sense.
- 4-2** Identify how the three main components of the scientific attitude relate to critical thinking.



Top Photo Group/Thinkstock/Getty Images

- 4-1** How do hindsight bias, overconfidence, and the tendency to perceive order in random events illustrate why science-based answers are more valid than those based on intuition and common sense?

Some people suppose that psychology merely documents and dresses in jargon what people already know: "So what else is new—you get paid for using fancy methods to prove what everyone knows?" Others place their faith in human intuition: "Buried deep within each and every one of us, there is an instinctive, heart-felt awareness that provides—if we allow it to—the most reliable guide," offered Prince Charles (2000).

Prince Charles has much company, judging from the long list of pop psychology books on "intuitive managing," "intuitive trading," and "intuitive healing." Today's psychological science does document a vast intuitive mind. As we will see, our thinking, memory, and attitudes operate on two levels—conscious and unconscious—with the larger part operating automatically, off-screen. Like jumbo jets, we fly mostly on autopilot.

So, are we smart to listen to the whispers of our inner wisdom, to simply trust "the force within"? Or should we more often be subjecting our intuitive hunches to skeptical scrutiny?

This much seems certain: We often underestimate intuition's perils. My geographical intuition tells me that Reno is east of Los Angeles, that Rome is south of New York, that Atlanta is east of Detroit. But I am wrong, wrong, and wrong.

Modules to come will show that experiments have found people greatly overestimating their lie detection accuracy, their eyewitness recollections, their interviewee assessments, their risk predictions, and their stock-picking talents. As a Nobel Prize-winning physicist explained, "The first principle is that you must not fool yourself—and you are the easiest person to fool" (Feynman, 1997).



Indeed, observed novelist Madeleine L'Engle, "The naked intellect is an extraordinarily inaccurate instrument" (1973). Three phenomena—*hindsight bias, judgmental overconfidence, and our tendency to perceive patterns in random events*—illustrate why we cannot rely solely on intuition and common sense.

"Those who trust in their own wits are fools." -PROVERBS 28:26

Did We Know It All Along? Hindsight Bias

Consider how easy it is to draw the bull's eye *after* the arrow strikes. After the stock market drops, people say it was "due for a correction." After the football game, we credit the coach if a "gutsy play" wins the game, and fault the coach for the "stupid play" if it doesn't. After a war or an election, its outcome usually seems obvious. Although history may therefore seem like a series of inevitable events, the actual future is seldom foreseen. No one's diary recorded, "Today the Hundred Years War began."

This **hindsight bias** (also known as the *I-knew-it-all-along phenomenon*) is easy to demonstrate: Give half the members of a group some purported psychological finding, and give the other half an opposite result. Tell the first group, "Psychologists have found that separation weakens romantic attraction. As the saying goes, 'Out of sight, out of mind.'" Ask them to imagine why this might be true. Most people can, and nearly all will then view this true finding as unsurprising.

Tell the second group the opposite, "Psychologists have found that separation strengthens romantic attraction. As the saying goes, 'Absence makes the heart grow fonder.'" People given this untrue result can also easily imagine it, and most will also see it as unsurprising. When two opposite findings both seem like common sense, there is a problem.

Such errors in our recollections and explanations show why we need psychological research. Just asking people how and why they felt or acted as they did can sometimes be misleading—not because common sense is usually wrong, but because common sense more easily describes what *has* happened than what *will* happen. As physicist Niels Bohr reportedly said, "Prediction is very difficult, especially about the future."

Some 100 studies have observed hindsight bias in various countries and among both children and adults (Blank et al., 2007). Nevertheless, our intuition is often right. As Yogi Berra once said, "You can observe a lot by watching." (We have Berra to thank for other gems, such as "Nobody ever comes here—it's too crowded," and "If the people don't want to come out to the ballpark, nobody's gonna stop 'em.") Because we're all behavior watchers, it would be

"Life is lived forwards, but understood backwards."
-PHILOSOPHER SØREN KIERKEGAARD,
1813-1855

hindsight bias the tendency to believe, after learning an outcome, that one would have foreseen it. (Also known as the *I-knew-it-all-along phenomenon*.)

"Anything seems commonplace, once explained." -DR. WATSON TO SHERLOCK HOLMES



Hindsight bias When drilling the Deepwater Horizon oil well in 2010, BP employees took some shortcuts and ignored some warning signs, without intending to harm the environment or their company's reputation. After the resulting Gulf oil spill, with the benefit of 20/20 hindsight, the foolishness of those judgments became obvious.

AP® Exam Tip

It is quite common for multiple-choice questions on the AP® exam to test your knowledge of “media myths.” Pay particular attention when psychological findings run counter to “common sense.”

surprising if many of psychology’s findings had *not* been foreseen. Many people believe that love breeds happiness, and they are right (we have what Module 40 calls a deep “need to belong”). Indeed, note Daniel Gilbert, Brett Pelham, and Douglas Krull (2003), “good ideas in psychology usually have an oddly familiar quality, and the moment we encounter them we feel certain that we once came close to thinking the same thing ourselves and simply failed to write it down.” Good ideas are like good inventions; once created, they seem obvious. (Why did it take so long for someone to invent suitcases on wheels and Post-it Notes?)

But sometimes our intuition, informed by countless casual observations, has it wrong. In later modules we will see how research has overturned popular ideas—that familiarity breeds contempt, that dreams predict the future, and that most of us use only 10 percent of our brain. (See also **TABLE 4.1**.) We will also see how it has surprised us with discoveries about how the brain’s chemical messengers control our moods and memories, about other animals’ abilities, and about the effects of stress on our capacity to fight disease.

Table 4.1 True or False?

Psychological research discussed in modules to come will either confirm or refute each of these statements (adapted, in part, from Furnham et al., 2003). Can you predict which of these popular ideas have been confirmed and which refuted? (Check your answers at the bottom of this table.)

1. If you want to teach a habit that persists, reward the desired behavior every time, not just intermittently (see Module 27).
2. Patients whose brains are surgically split down the middle survive and function much as they did before the surgery (see Module 13).
3. Traumatic experiences, such as sexual abuse or surviving the Holocaust, are typically “repressed” from memory (see Module 33).
4. Most abused children do *not* become abusive adults (see Module 50).
5. Most infants recognize their own reflection in a mirror by the end of their first year (see Module 47).
6. Adopted siblings usually do not develop similar personalities, even though they are reared by the same parents (see Module 14).
7. Fears of harmless objects, such as flowers, are just as easy to acquire as fears of potentially dangerous objects, such as snakes (see Module 15).
8. Lie detection tests often lie (see Module 41).
9. The brain remains active during sleep (see Modules 22–23).

Answers: 1. F, 2. T, 3. F, 4. T, 5. F, 6. T, 7. F, 8. T, 9. T

Overconfidence

We humans tend to think we know more than we do. Asked how sure we are of our answers to factual questions (*Is Boston north or south of Paris?*), we tend to be more confident than correct.¹ Or consider these three anagrams, which Richard Goranson (1978) asked people to unscramble:

WREAT → WATER

ETRYN → ENTRY

GRABE → BARGE

About how many seconds do you think it would have taken you to unscramble each of these? Did hindsight influence you? Knowing the answers tends to make us overconfident—surely the solution would take only 10 seconds or so, when in reality the average problem solver spends 3 minutes, as you also might, given a similar anagram without the solution: OCHSA.²

Are we any better at predicting social behavior? University of Pennsylvania psychologist Philip Tetlock (1998, 2005) collected more than 27,000 expert predictions of world events, such as the future of South Africa or whether Quebec would separate from Canada. His repeated finding: These predictions, which experts made with 80 percent confidence on average, were right less than 40 percent of the time. Nevertheless, even those who erred maintained their confidence by noting they were “almost right.” “The Québécois separationists *almost* won the secessionist referendum.”

Perceiving Order in Random Events

In our natural eagerness to make sense of our world—what poet Wallace Stevens called our “rage for order”—we are prone to perceive patterns. People see a face on the moon, hear Satanic messages in music, perceive the Virgin Mary’s image on a grilled cheese sandwich. Even in random data we often find order, because—here’s a curious fact of life—*random sequences often don’t look random* (Falk et al., 2009; Nickerson, 2002, 2005). Consider a random coin flip: If someone flipped a coin six times, which of the following sequences of heads (H) and tails (T) would be most likely: HHHTTT or HTTHHT or HHHHHH?

Daniel Kahneman and Amos Tversky (1972) found that most people believe HTTHHT would be the most likely random sequence. Actually, all three are equally likely (or, you might say, equally unlikely). A poker hand of 10 through ace, all of hearts, would seem extraordinary; actually, it would be no more or less likely than any other specific hand of cards (**FIGURE 4.1**).

In actual random sequences, patterns and streaks (such as repeating digits) occur more often than people expect (Oskarsson et al., 2009). To demonstrate this phenomenon for myself, I flipped a coin 51 times, with these results:

1. H	10. T	19. H	28. T	37. T	46. H
2. T	11. T	20. H	29. H	38. T	47. H
3. T	12. H	21. T	30. T	39. H	48. T
4. T	13. H	22. T	31. T	40. T	49. T
5. H	14. T	23. H	32. T	41. H	50. T
6. H	15. T	24. T	33. T	42. H	51. T
7. H	16. H	25. T	34. T	43. H	
8. T	17. T	26. T	35. T	44. H	
9. T	18. T	27. H	36. H	45. T	

Looking over the sequence, patterns jump out: Tosses 10 to 22 provided an almost perfect pattern of pairs of tails followed by pairs of heads. On tosses 30 to 38 I had a “cold hand,” with only one head in nine tosses. But my fortunes immediately reversed with a “hot hand”—seven heads out of the next nine tosses. Similar streaks happen, about as often as one would expect in random sequences, in basketball shooting, baseball hitting, and mutual fund stock pickers’ selections (Gilovich et al., 1985; Malkiel, 2007; Myers, 2002). These sequences often don’t look random and so are overinterpreted. (“When you’re hot, you’re hot!”)

Overconfidence in history:

“We don’t like their sound. Groups of guitars are on their way out.” -DECCA RECORDS, IN TURNING DOWN A RECORDING CONTRACT WITH THE BEATLES IN 1962

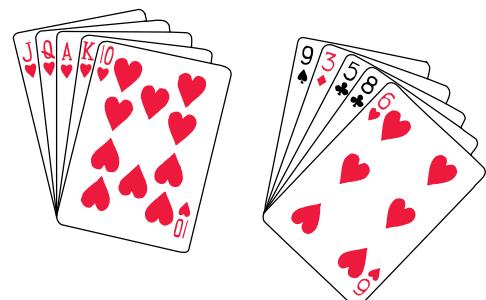
“Computers in the future may weigh no more than 1.5 tons.” -POPULAR MECHANICS, 1949

“They couldn’t hit an elephant at this distance.” -GENERAL JOHN SEDGWICK JUST BEFORE BEING KILLED DURING A U.S. CIVIL WAR BATTLE, 1864

“The telephone may be appropriate for our American cousins, but not here, because we have an adequate supply of messenger boys.” -BRITISH EXPERT GROUP EVALUATING THE INVENTION OF THE TELEPHONE

Figure 4.1

Two random sequences Your chances of being dealt either of these hands are precisely the same: 1 in 2,598,960.



¹ Boston is south of Paris.

² The anagram solution: CHAOS.

What explains these streaky patterns? Was I exercising some sort of paranormal control over my coin? Did I snap out of my tails funk and get in a heads groove? No such explanations are needed, for these are the sorts of streaks found in any random data. Comparing each toss to the next, 23 of the 50 comparisons yielded a changed result—just the sort of near 50-50 result we expect from coin tossing. Despite seeming patterns, the outcome of one toss gives no clue to the outcome of the next.

However, some happenings seem so extraordinary that we struggle to conceive an ordinary, chance-related explanation (as applies to our coin tosses). In such cases, statisticians often are less mystified. When Evelyn Marie Adams won the New Jersey lottery *twice*, newspapers reported the odds of her feat as 1 in 17 trillion. Bizarre? Actually, 1 in 17 trillion are indeed the odds that a given person who buys a single ticket for two New Jersey lotteries will win both times. And given the millions of people who buy U.S. state lottery tickets, statisticians Stephen Samuels and George McCabe (1989) reported, it was “practically a sure thing” that someday, somewhere, someone would hit a state jackpot twice. Indeed, said fellow statisticians Persi Diaconis and Frederick Mosteller (1989), “with a large enough sample, any outrageous thing is likely to happen.” An event that happens to but 1 in 1 billion people every day occurs about 7 times a day, 2500 times a year.

The point to remember: Hindsight bias, overconfidence, and our tendency to perceive patterns in random events often lead us to overestimate our intuition. But scientific inquiry can help us sift reality from illusion.

“The really unusual day would be one where nothing unusual happens.” -STATISTICIAN PERSI DIACONIS (2002)

The Amazing Randi The magician James Randi exemplifies skepticism. He has tested and debunked supposed psychic phenomena.



AP Photo/Alan Diaz

The Scientific Attitude: Curious, Skeptical, and Humble

4-2 How do the scientific attitude’s three main components relate to critical thinking?

Underlying all science is, first, a hard-headed *curiosity*, a passion to explore and understand without misleading or being misled. Some questions (*Is there life after death?*) are beyond science. Answering them in any way requires a leap of faith. With many other ideas (*Can some people demonstrate ESP?*), the proof is in the pudding. Let the facts speak for themselves.

Magician James Randi has used this *empirical approach* when testing those claiming to see auras around people’s bodies:

Randi: Do you see an aura around my head?

Aura-seer: Yes, indeed.

Randi: Can you still see the aura if I put this magazine in front of my face?

Aura-seer: Of course.

Randi: Then if I were to step behind a wall barely taller than I am, you could determine my location from the aura visible above my head, right?

Randi told me that no aura-seer has agreed to take this simple test.

No matter how sensible-seeming or wild an idea, the smart thinker asks: *Does it work?* When put to the test, can its predictions be confirmed? Subjected to such scrutiny, crazy-sounding ideas sometimes find support. During the 1700s, scientists scoffed at the notion that meteorites had extraterrestrial origins. When two Yale scientists challenged the conventional opinion, Thomas Jefferson jeered, “Gentlemen, I would rather believe that those two Yankee professors would lie than to believe that stones fell from Heaven.” Sometimes scientific inquiry turns jeers into cheers.

More often, science becomes society’s garbage disposal, sending crazy-sounding ideas to the waste heap, atop previous claims of perpetual motion machines, miracle cancer cures, and out-of-body travels into centuries past. To sift reality from fantasy, sense from nonsense, therefore requires a scientific attitude: being skeptical but not cynical, open but not gullible.

"To believe with certainty," says a Polish proverb, "we must begin by doubting." As scientists, psychologists approach the world of behavior with a *curious skepticism*, persistently asking two questions: *What do you mean? How do you know?*

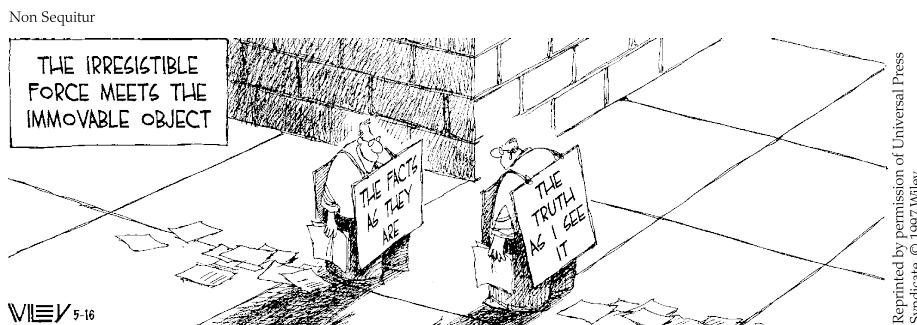
When ideas compete, skeptical testing can reveal which ones best match the facts. Do parental behaviors determine children's sexual orientation? Can astrologers predict your future based on the position of the planets at your birth? Is electroconvulsive therapy (delivering an electric shock to the brain) an effective treatment for severe depression? As we will see, putting such claims to the test has led psychological scientists to answer *No* to the first two questions and *Yes* to the third.

Putting a scientific attitude into practice requires not only curiosity and skepticism but also *humility*—an awareness of our own vulnerability to error and an openness to surprises and new perspectives. In the last analysis, what matters is not my opinion or yours, but the truths nature reveals in response to our questioning. If people or other animals don't behave as our ideas predict, then so much the worse for our ideas. This humble attitude was expressed in one of psychology's early mottos: "The rat is always right."

Historians of science tell us that these three attitudes—curiosity, skepticism, and humility—helped make modern science possible. Some deeply religious people today may view science, including psychological science, as a threat. Yet, many of the leaders of the scientific revolution, including Copernicus and Newton, were deeply religious people acting on the idea that "in order to love and honor God, it is necessary to fully appreciate the wonders of his handiwork" (Stark, 2003a,b).

"I'm a skeptic not because I do not want to believe but because I want to *know*. I believe that the truth is out there. But how can we tell the difference between what we would like to be true and what is actually true? The answer is science." -MICHAEL SHERMER,
"I WANT TO BELIEVE," SCIENTIFIC AMERICAN, 2009

"My deeply held belief is that if a god anything like the traditional sort exists, our curiosity and intelligence are provided by such a god. We would be unappreciative of those gifts . . . if we suppressed our passion to explore the universe and ourselves." -CARL SAGAN,
BROCA'S BRAIN, 1979



Of course, scientists, like anyone else, can have big egos and may cling to their preconceptions. Nevertheless, the ideal of curious, skeptical, humble scrutiny of competing ideas unifies psychologists as a community as they check and recheck one another's findings and conclusions.

Critical Thinking

The scientific attitude prepares us to think smarter. Smart thinking, called **critical thinking**, examines assumptions, assesses the source, discerns hidden values, confirms evidence, and assesses conclusions. Whether reading a news report or listening to a conversation, critical thinkers ask questions. Like scientists, they wonder: How do they know that? What is this person's agenda? Is the conclusion based on anecdote and gut feelings, or on evidence? Does the evidence justify a cause-effect conclusion? What alternative explanations are possible?

Critical thinking, informed by science, helps clear the colored lenses of our biases. Consider: Does climate change threaten our future, and, if so, is it human-caused? In 2009, climate-action advocates interpreted an Australian heat wave and dust storms as evidence of climate change. In 2010, climate-change skeptics perceived North American bitter cold and East Coast blizzards as discounting global warming. Rather than having their understanding

critical thinking thinking that does not blindly accept arguments and conclusions. Rather, it examines assumptions, assesses the source, discerns hidden values, evaluates evidence, and assesses conclusions.

"The real purpose of the scientific method is to make sure Nature hasn't misled you into thinking you know something you don't actually know." -ROBERT M. PIRSIG,
ZEN AND THE ART OF MOTORCYCLE MAINTENANCE, 1974

of climate change swayed by today's weather, or by their own political views, critical thinkers say, "Show me the evidence." Over time, is the Earth actually warming? Are the polar ice caps melting? Are vegetation patterns changing? And is human activity spewing gases that would lead us to expect such changes? When contemplating such issues, critical thinkers will consider the credibility of sources. They will look at the evidence ("Do the facts support them, or are they just makin' stuff up?"). They will recognize multiple perspectives. And they will expose themselves to news sources that challenge their preconceived ideas.

Has psychology's critical inquiry been open to surprising findings? The answer, as ensuing modules illustrate, is plainly Yes. Believe it or not, massive losses of brain tissue early in life may have minimal long-term effects (see Module 12). Within days, newborns can recognize their mother's odor and voice (see Module 45). After brain damage, a person may be able to learn new skills yet be unaware of such learning (see Modules 31–33). Diverse groups—men and women, old and young, rich and middle class, those with disabilities and without—report roughly comparable levels of personal happiness (see Module 83).

And has critical inquiry convincingly debunked popular presumptions? The answer, as ensuing modules also illustrate, is again Yes. The evidence indicates that sleepwalkers are *not* acting out their dreams (see Module 24). Our past experiences are *not* all recorded verbatim in our brains; with brain stimulation or hypnosis, one *cannot* simply "hit the replay button" and relive long-buried or repressed memories (see Module 33). Most people do *not* suffer from unrealistically low self-esteem, and high self-esteem is not all good (see Module 59). Opposites do *not* generally attract (see Module 79). In each of these instances and more, what has been learned is not what is widely believed.

Before You Move On

► ASK YOURSELF

How might critical thinking help us assess someone's interpretations of people's dreams or their claims to communicate with the dead?

► TEST YOURSELF

How does the scientific attitude contribute to critical thinking?

Answers to the Test Yourself questions can be found in Appendix E at the end of the book.

Module 4 Review

4-1

How do hindsight bias, overconfidence, and the tendency to perceive order in random events illustrate why science-based answers are more valid than those based on intuition and common sense?

- *Hindsight bias* (also called the "I-knew-it-all-along phenomenon") is the tendency to believe, after learning an outcome, that we would have foreseen it.
- Overconfidence in our judgments results partly from our bias to seek information that confirms them.
- These tendencies, plus our eagerness to perceive patterns in random events, lead us to overestimate our intuition.

Although limited by the testable questions it can address, scientific inquiry can help us overcome our intuition's biases and shortcomings.

4-2

How do the scientific attitude's three main components relate to critical thinking?

- The scientific attitude equips us to be curious, skeptical, and humble in scrutinizing competing ideas or our own observations.
- This attitude carries into everyday life as *critical thinking*, which puts ideas to the test by examining assumptions, assessing the source, discerning hidden values, evaluating evidence, and assessing conclusions.

Multiple-Choice Questions

- 1.** After the student council election, a friend tells you he has known for weeks who would be elected president. What does this seem to illustrate?
- Skepticism
 - Critical thinking
 - Hindsight bias
 - Overconfidence
 - Perceiving order in random events
- 2.** While taking a standardized test with randomly scrambled answers, you notice that your last four answers have been "c." Which of the following is true concerning the probability of the next answer being "c"?
- It is higher. Once a streak begins, it is likely to last for a while.
 - It is lower. Since answers are distributed randomly, "c" answers become less common.
 - It is unaffected by previous answers. It is as likely to be "c" as any other answer.
 - You should check your previous answers. Four "c's" in a row is impossible.
 - It is higher. Test constructors trick students by keeping the same answer many times in a row.
- 3.** What do we call the tendency to exaggerate the correctness or accuracy of our beliefs and predictions prior to testing?
- Hindsight bias
 - Overconfidence
 - Critical thinking
 - Skepticism
 - Reliability
- 4.** Which of the following is an example of hindsight bias?
- Tom is certain that electric cars will represent 80 percent of vehicles in twenty years and only reads research studies that support his hypothesis.
 - Liza underestimates how much time it will take her to finish writing her college application essays and as a result fails to meet an important deadline.
 - Experts predicting world events with 80 percent confidence turned out to be correct less than 40 percent of the time.
 - Marcy cannot recognize a definition on a flashcard. After turning the card over and viewing the term, she tells herself she knew what the answer was all along.
 - Dr. Grace overestimates how effectively her new treatment method works because she fails to seek out any evidence refuting her theory.

Practice FRQs

- 1.** Name the three components of the scientific attitude. Provide an example to show how each component contributes to the investigation of competing ideas in psychology.

Answer

1 point: Curiosity, or passion to explore, leads us to questions we want to investigate. Any examples of such questions will serve (For example, Does more money make us happier? Is schizophrenia inherited?).

1 point: Skepticism keeps us from accepting ideas without sound support. The work of The Amazing Randi would be a good example here.

1 point: Humility keeps us open to the possibility of changing our ideas when they are not supported by the data. For example, "the rat is always right."

- 2.** Aziz has read that handwriting reveals important details about personality. Explain how each component of the scientific attitude can help Aziz investigate the accuracy of the information he has read about handwriting analysis.

(3 points)

Module 5

The Scientific Method and Description

Module Learning Objectives

5-1 Describe how theories advance psychological science.

5-2 Describe how psychologists use case studies, naturalistic observation, and surveys to observe and describe behavior, and explain the importance of random sampling.



Creative Crop/Jupiterimages

AP® Exam Tip

As you read this module, keep in mind that the scientific method is a set of principles and procedures, not a list of facts. You will be expected to understand how the science of psychology is done, not just what it has discovered.

Psychologists arm their scientific attitude with the *scientific method*—a self-correcting process for evaluating ideas with observation and analysis. In its attempt to describe and explain human nature, psychological science welcomes hunches and plausible-sounding theories. And it puts them to the test. If a theory works—if the data support its predictions—so much the better for that theory. If the predictions fail, the theory will be revised or rejected.

The Scientific Method

5-1 How do theories advance psychological science?

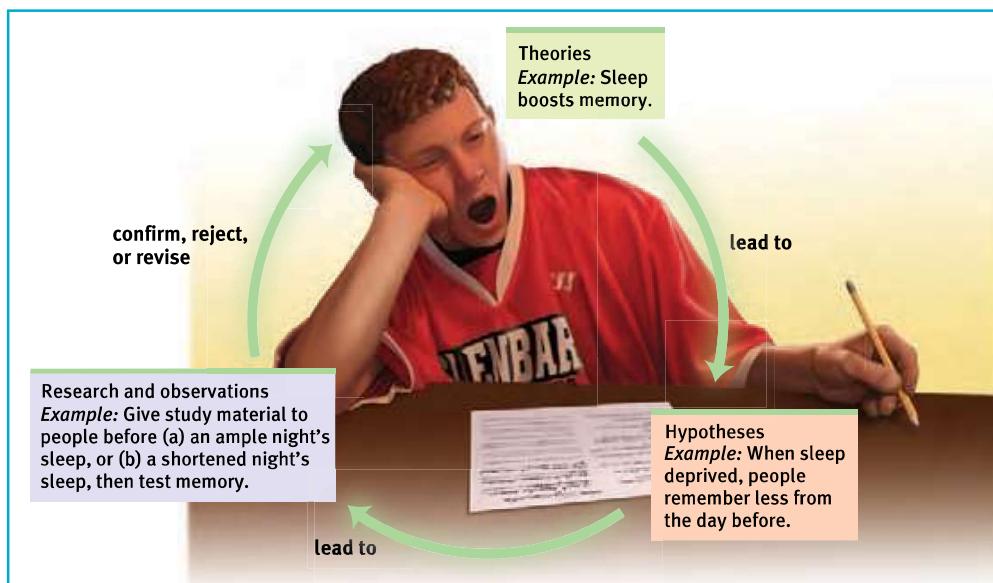
Chatting with friends and family, we often use theory to mean “mere hunch.” In science, a **theory** *explains* behaviors or events by offering ideas that *organize* what we have observed. By organizing isolated facts, a theory simplifies. By linking facts with deeper principles, a theory offers a useful summary. As we connect the observed dots, a coherent picture emerges.

A theory about the effects of sleep on memory, for example, helps us organize countless sleep-related observations into a short list of principles. Imagine that we observe over and over that people with good sleep habits tend to answer questions correctly in class, and they do well at test time. We might therefore theorize that sleep improves memory. So far so good: Our principle neatly summarizes a list of facts about the effects of a good night’s sleep on memory.

Yet no matter how reasonable a theory may sound—and it does seem reasonable to suggest that sleep could improve memory—we must put it to the test. A good theory produces testable predictions, called **hypotheses**. Such predictions specify what results (what behaviors or events) would support the theory and what results would cast doubt on the theory. To test our theory about the effects of sleep on memory, our hypothesis might be that when sleep deprived, people will remember less from the day before. To test that hypothesis, we might assess how well people remember course materials they studied before a good night’s sleep, or before a shortened night’s sleep (**FIGURE 5.1**). The results will either confirm our theory or lead us to revise or reject it.

theory an explanation using an integrated set of principles that organizes observations and predicts behaviors or events.

hypothesis a testable prediction, often implied by a theory.

**Figure 5.1**

The scientific method A self-correcting process for asking questions and observing nature's answers.

Our theories can bias our observations. Having theorized that better memory springs from more sleep, we may see what we expect: We may perceive sleepy people's comments as less insightful. Perhaps you are aware of students who, because they have developed an excellent reputation, can now do no wrong in the eyes of teachers. If they're in the hall during class, nobody worries. Other students can do no good. Because they have behaved badly in the past, even their positive behaviors are viewed suspiciously.

As a check on their biases, psychologists use **operational definitions** when they report their studies. "Sleep deprived," for example, may be defined as "X hours less" than the person's natural sleep. Unlike dictionary definitions, operational definitions describe concepts with precise procedures or measures. These exact descriptions will allow anyone to **replicate** (repeat) the research. Other people can then re-create the study with different participants and in different situations. If they get similar results, we can be confident that the findings are reliable.

Let's summarize. A good theory:

- effectively organizes a range of self-reports and observations.
- leads to clear *hypotheses* (predictions) that anyone can use to check the theory.
- often stimulates research that leads to a revised theory which better organizes and predicts what we know. Or, our research may be replicated and supported by similar findings. (This has been the case for sleep and memory studies, as you will see in Module 24.)

We can test our hypotheses and refine our theories in several ways.

- *Descriptive* methods describe behaviors, often by using case studies, surveys, or naturalistic observations.
- *Correlational* methods associate different factors, or *variables*. (You'll see the word *variable* often in descriptions of research. It refers to anything that contributes to a result.)
- *Experimental* methods manipulate variables to discover their effects.

To think critically about popular psychology claims, we need to understand the strengths and weaknesses of these methods.

operational definition a carefully worded statement of the exact procedures (operations) used in a research study. For example, *human intelligence* may be operationally defined as what an intelligence test measures.

replication repeating the essence of a research study, usually with different participants in different situations, to see whether the basic finding extends to other participants and circumstances.

Description

5-2

How do psychologists use case studies, naturalistic observation, and surveys to observe and describe behavior, and why is random sampling important?

The starting point of any science is description. In everyday life, we all observe and describe people, often drawing conclusions about why they act as they do. Professional psychologists do much the same, though more objectively and systematically, through

- *case studies* (analyses of special individuals).
- *naturalistic observation* (watching and recording the natural behavior of many individuals).
- *surveys* and interviews (by asking people questions).

case study a descriptive technique in which one individual or group is studied in depth in the hope of revealing universal principles.

naturalistic observation observing and recording behavior in naturally occurring situations without trying to manipulate and control the situation.

"Well my dear," said Miss Marple, "human nature is very much the same everywhere, and of course, one has opportunities of observing it at closer quarters in a village." —AGATHA CHRISTIE, *THE TUESDAY CLUB MURDERS*, 1933

The Case Study

Psychologists use the **case study**, which is among the oldest research methods, to examine one individual or group in depth in the hope of revealing things true of all of us. Some examples: Much of our early knowledge about the brain came from case studies of individuals who suffered a particular impairment after damage to a certain brain region. Jean Piaget taught us about children's thinking through case studies in which he carefully observed and questioned individual children. Studies of only a few chimpanzees have revealed their capacity for understanding and language. Intensive case studies are sometimes very revealing. They show us what *can* happen, and they often suggest directions for further study.

But individual cases may mislead us if the individual is atypical. Unrepresentative information can lead to mistaken judgments and false conclusions. Indeed, anytime a researcher mentions a finding ("*Smokers die younger: ninety-five percent of men over 85 are nonsmokers*") someone is sure to offer a contradictory anecdote ("*Well, I have an uncle who smoked two packs a day and lived to 89*"). Dramatic stories and personal experiences (even psychological

case examples) command our attention and are easily remembered. Journalists

understand that, and so begin an article about bank foreclosures with the sad story of one family put out of their house, not with foreclosure statistics. Stories move us. But stories can mislead. Which of the following do you find more memorable? (1) "In one study of 1300 dream reports concerning a kidnapped child, only 5 percent correctly envisioned the child as dead" (Murray & Wheeler, 1937). (2) "I know a man who

dreamed his sister was in a car accident, and two days later she died in a head-on collision!" Numbers can be numbing, but the plural of *anecdote* is not *evidence*. As psychologist Gordon Allport (1954, p. 9) said, "Given a thimbleful of [dramatic] facts we rush to make generalizations as large as a tub."

The point to remember: Individual cases can suggest fruitful ideas. What's true of all of us can be glimpsed in any one of us. But to discern the general truths that cover individual cases, we must answer questions with other research methods.

Jupiterimages/Bridgeman Images/Alamy



Freud and Little Hans Sigmund Freud's case study of 5-year-old Hans' extreme fear of horses led Freud to his theory of childhood sexuality. He conjectured that Hans felt unconscious desire for his mother, feared castration by his rival father, and then transferred this fear into his phobia about being bitten by a horse. As Module 56 will explain, today's psychological science discounts Freud's theory of childhood sexuality but acknowledges that much of the human mind operates outside our conscious awareness.

Naturalistic Observation

A second descriptive method records behavior in natural environments. These **naturalistic observations** range from watching chimpanzee societies in the jungle, to unobtrusively videotaping (and later systematically analyzing) parent-child interactions in different cultures, to recording racial differences in students' self-seating patterns in a school cafeteria.

Like the case study, naturalistic observation does not *explain* behavior. It *describes* it. Nevertheless, descriptions can be revealing. We once thought, for example, that only humans use tools. Then naturalistic observation revealed that chimpanzees sometimes insert a stick in a termite mound and withdraw it, eating the stick's load of termites.

Such unobtrusive naturalistic observations paved the way for later studies of animal thinking, language, and emotion, which further expanded our understanding of our fellow animals. "Observations, made in the natural habitat, helped to show that the societies and behavior of animals are far more complex than previously supposed," chimpanzee observer Jane Goodall noted (1998). Thanks to researchers' observations, we know that chimpanzees and baboons use deception. Psychologists Andrew Whiten and Richard Byrne (1988) repeatedly saw one young baboon pretending to have been attacked by another as a tactic to get its mother to drive the other baboon away from its food. The more developed a primate species' brain, the more likely it is that the animals will display deceptive behaviors (Byrne & Corp, 2004).

Naturalistic observations also illuminate human behavior. Here are four findings you might enjoy.

- *A funny finding.* We humans laugh 30 times more often in social situations than in solitary situations. (Have you noticed how seldom you laugh when alone?) As we laugh, 17 muscles contort our mouth and squeeze our eyes, and we emit a series of 75-millisecond vowel-like sounds, spaced about one-fifth of a second apart (Provine, 2001).
- *Sounding out students.* What, really, are college psychology students saying and doing during their everyday lives? To find out, researchers equipped 52 such students from the University of Texas with electronic recorders (Mehl & Pennebaker, 2003). For up to four days, the recorders captured 30 seconds of the students' waking hours every 12.5 minutes, thus enabling the researchers to eavesdrop on more than 10,000 half-minute life slices by the end of the study. On what percentage of the slices do you suppose they found the students talking with someone? What percentage captured the students at a computer? The answers: 28 and 9 percent. (What percentage of *your* waking hours are spent in these activities?)
- *What's on your mind?* To find out what was on the mind of their University of Nevada, Las Vegas, students, researchers gave them beepers (Heavey & Hurlburt, 2008). On a half-dozen occasions, a beep interrupted students' daily activities, signaling them to pull out a notebook and record their inner experience at that moment. When the researchers later coded the reports in categories, they found five common forms of inner experience (**TABLE 5.1** on the next page).
- *Culture, climate, and the pace of life.* Naturalistic observation also enabled researchers to compare the pace of life in 31 countries (Levine & Norenzayan, 1999). (Their operational definition of *pace of life* included walking speed, the speed with which postal clerks completed a simple request, and the accuracy of public clocks.) Their conclusion: Life is fastest paced in Japan and Western Europe, and slower paced in economically less-developed countries. People in colder climates also tend to live at a faster pace (and are more prone to die from heart disease).



A natural observer Chimpanzee researcher Frans de Waal (2005) reported, "I am a born observer. . . . When picking a seat in a restaurant I want to face as many tables as possible. I enjoy following the social dynamics—love, tension, boredom, antipathy—around me based on body language, which I consider more informative than the spoken word. Since keeping track of others is something I do automatically, becoming a fly on the wall of an ape colony came naturally to me."



Courtesy of Matthias Mehl

Naturalistic observation

Researchers at the University of Texas used electronic recorders to sample naturally occurring slices of daily life.

Table 5.1 A Penny for Their Thoughts: The Inner Experience of University Students*

Inner Experience	Example	Frequency
Inner speech	Susan was saying to herself, "I've got to get to class."	26%
Inner seeing	Paul was imagining the face of a best friend, including her neck and head.	34%
Unsymbolized thinking	Alphonse was wondering whether the workers would drop the bricks.	22%
Feeling	Courtney was experiencing anger and its physical symptoms.	26%
Sensory awareness	Fiona was feeling the cold breeze on her cheek and her hair moving.	22%

* More than one experience could occur at once.

Naturalistic observation offers interesting snapshots of everyday life, but it does so without controlling for all the variables that may influence behavior. It's one thing to observe the pace of life in various places, but another to understand what makes some people walk faster than others.

The Survey

survey a technique for ascertaining the self-reported attitudes or behaviors of a particular group, usually by questioning a representative, random sample of the group.

A **survey** looks at many cases in less depth. Researchers do surveys when wanting to estimate, from a representative sample of people, the attitudes or reported behaviors of a whole population. Questions about everything from cell-phone use to political opinions are put to the public. In recent surveys,

- half of all Americans reported experiencing more happiness and enjoyment than worry and stress on the previous day (Gallup, 2010).
- online Canadians reported using new forms of electronic communication and thus receiving 35 percent fewer e-mails in 2010 than 2008 (Ipsos, 2010a).
- 1 in 5 people across 22 countries reported believing that alien beings have come to Earth and now walk among us disguised as humans (Ipsos, 2010b).
- 68 percent of all humans—some 4.6 billion people—say that religion is important in their daily lives (Diener et al., 2011).

But asking questions is tricky, and the answers often depend on the ways questions are worded and respondents are chosen.

WORDING EFFECTS

As we will see in Module 35, even subtle changes in the order or wording of questions—the way we *frame* a question—can have major effects. People are much more approving of “aid to the needy” than of “welfare,” of “affirmative action” than of “preferential treatment,” of “not allowing” televised cigarette ads and pornography than of “censoring” them, and of “revenue enhancers” than of “taxes.” In 2009, three in four Americans in one national survey approved of giving people “a choice” of public, government-run, or private health insurance. Yet in another survey, most Americans were not in favor of “creating a public health care plan administered by the federal government that would compete directly with private health insurance companies” (Stein, 2009). Because wording is such a delicate matter, critical thinkers will reflect on how the phrasing of a question might affect people’s expressed opinions.

RANDOM SAMPLING

In everyday thinking, we tend to generalize from samples we observe, especially vivid cases. Given (a) a statistical summary of auto owners' evaluations of their car make and (b) the vivid comments of a biased sample—two frustrated owners—our impression may be influenced as much by the two unhappy owners as by the many more evaluations in the statistical summary. The temptation to ignore the **sampling bias** and to generalize from a few vivid but unrepresentative cases is nearly irresistible.

The point to remember: The best basis for generalizing is from a *representative sample*.

But it's not always possible to survey everyone in a group. So how do you obtain a representative sample—say, of the students at your high school? How could you choose a group that would represent the total student **population**, the whole group you want to study and describe? Typically, you would seek a **random sample**, in which every person in the entire group has an equal chance of participating. You might number the names in the general student listing and then use a random number generator to pick your survey participants. (Sending each student a questionnaire wouldn't work because the conscientious people who returned it would not be a random sample.) Large representative samples are better than small ones, but a small representative sample of 100 is better than an unrepresentative sample of 500.

Political pollsters sample voters in national election surveys just this way. Using only 1500 randomly sampled people, drawn from all areas of a country, they can provide a remarkably accurate snapshot of the nation's opinions. Without random sampling (also called *random selection*), large samples—including call-in phone samples and TV or website polls (think of *American Idol* fans voting)—often merely give misleading results.

The point to remember: Before accepting survey findings, think critically: Consider the sample. You cannot compensate for an unrepresentative sample by simply adding more people.

Before You Move On

► ASK YOURSELF

Can you recall examples of misleading surveys you have experienced or read about? What survey principles did they violate?

► TEST YOURSELF

What are some strengths and weaknesses of the three different methods psychologists use to describe behavior—case studies, naturalistic observation, and surveys?

Answers to the Test Yourself questions can be found in Appendix E at the end of the book.

sampling bias a flawed sampling process that produces an unrepresentative sample.

population all those in a group being studied, from which samples may be drawn. (*Note:* Except for national studies, this does *not* refer to a country's whole population.)

random sample a sample that fairly represents a population because each member has an equal chance of inclusion.

FYI

With very large samples, estimates become quite reliable. *E* is estimated to represent 12.7 percent of the letters in written English. *E*, in fact, is 12.3 percent of the 925,141 letters in Melville's *Moby Dick*, 12.4 percent of the 586,747 letters in Dickens' *A Tale of Two Cities*, and 12.1 percent of the 3,901,021 letters in 12 of Mark Twain's works (*Chance News*, 1997).

Module 5 Review

5-1 How do theories advance psychological science?

- Psychological *theories* are explanations that apply an integrated set of principles to organize observations and generate *hypotheses*—predictions that can be used to check the theory or produce practical applications of it. By testing their hypotheses, researchers can confirm, reject, or revise their theories.
- To enable other researchers to *replicate* the studies, researchers report them using precise *operational definitions* of their procedures and concepts. If others achieve similar results, confidence in the conclusion will be greater.

5-2 How do psychologists use case studies, naturalistic observation, and surveys to observe and describe behavior, and why is random sampling important?

- Description methods, which include *case studies*, *naturalistic observations*, and *surveys*, show us what can happen, and they may offer ideas for further study.
- The best basis for generalizing about a *population* is a representative sample; in a *random sample*, every person in the entire population being studied has an equal chance of participating.
- Descriptive methods describe but do not *explain* behavior, because these methods do not control for the many variables that can affect behavior.

Multiple-Choice Questions

- Why is an operational definition necessary when reporting research findings?
 - An operational definition allows others to replicate the procedure.
 - An operational definition provides more context and includes many examples of the concept described.
 - An operational definition is easier to translate into multiple languages than a dictionary definition.
 - An operational definition uses more scientific language than a dictionary definition.
 - An operational definition is not necessary since a dictionary definition will work as well for replication.
- A researcher looking for gender differences in 3-year-olds observes a preschool class and records how many minutes children of each gender play with dolls. She then compares the two sets of numbers. What type of descriptive research is she conducting?
 - Case study
 - National study
 - Random sample method
 - Naturalistic observation
 - Survey
- Which of the following questions is best investigated by means of a survey?
 - Is IQ related to grades?
 - Are violent criminals genetically different from nonviolent criminals?
 - Does extra sleep improve memory?
 - What is the best study technique for AP® tests?
 - Are students more likely to be politically liberal or conservative?
- A testable prediction that drives research is known as a(n)
 - theory.
 - hypothesis.
 - operational definition.
 - guess.
 - random sample.
- Researchers are interested in finding out if winning Congressional candidates display more positive facial expressions than losing candidates. The researchers attend political debates and record how frequently each candidate displays positive facial expressions. Which research method are the researchers using?
 - Random sample
 - Case study
 - Naturalistic observation
 - Survey
 - Interview

- 6.** An individual with an exceptional memory is identified. She is capable of recalling major events, the weather, and what she did on any given date. What research method is being used if a psychologist conducts an in-depth investigation of this individual including questionnaires, brain scans, and memory tests?
- Naturalistic observation
 - Survey
 - Interview
 - Case study
 - Correlational method

- 7.** Which of the following is most important when conducting survey research?
- Choosing a representative sample
 - Choosing a large sample
 - Choosing a biased sample
 - Choosing a sample that includes every member of the population
 - Choosing a sample whose answers will likely support your hypothesis

Practice FRQs

- 1.** A teacher wants to know if nightmares are more common than dreams. He asks volunteers from his second-period class to report how many dreams they had last week. He asks volunteers from his third-period class to report the number of nightmares they had last week. Describe two things wrong with the design of this study.

Answer (2 of the following)

1 point: There is no hypothesis stated.

1 point: In asking for volunteers, the teacher is taking a nonrandom sample that is probably not representative of the population of interest.

1 point: Neither dreams nor nightmares are operationally defined, so they might be interpreted differently by later researchers.

1 point: The research is not blind. The teacher could influence the results by the way he asked questions.

- 2.** Naturalistic observation is a research method used by psychologists to investigate human and animal behavior. Identify three weaknesses of naturalistic observation.

(3 points)

Module 6

Correlation and Experimentation

Module Learning Objectives

- 6-1** Describe positive and negative correlations, and explain how correlational measures can aid the process of prediction but not provide evidence of cause-effect relationships.
- 6-2** Explain illusory correlations.
- 6-3** Describe the characteristics of experimentation that make it possible to isolate cause and effect.



imagebroker/Alamy

Study finds that increased parental support for college results in lower grades" (Jaschik, 2013). "People with mental illness more likely to be smokers" (Belluck, 2013). What should we make of such news headlines—telling us that students whose parents pay the college bill tend to underachieve, and that smoking is associated with mental illness? Do these correlations indicate that students would achieve more if their parents became less supportive and that stopping smoking could produce better mental health? No. Read on.

Correlation

correlation a measure of the extent to which two variables change together, and thus of how well either variable predicts the other.

correlation coefficient a statistical index of the relationship between two variables (from -1.0 to $+1.0$).

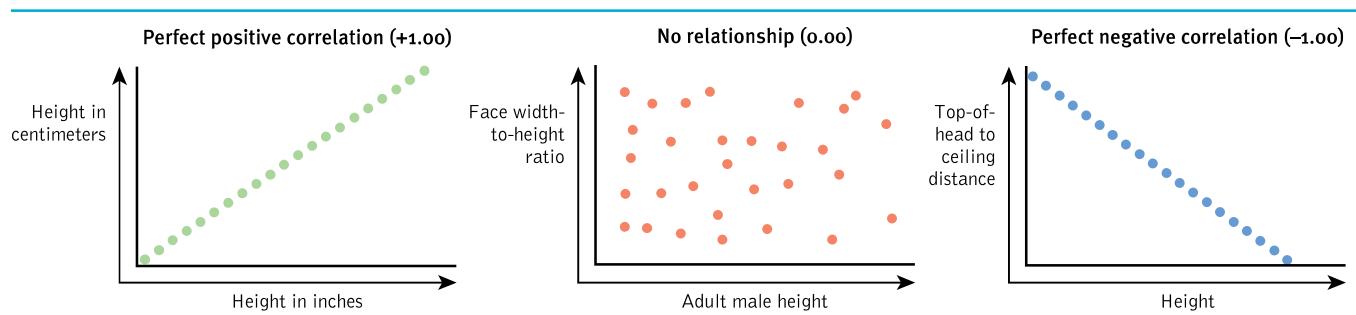
scatterplot a graphed cluster of dots, each of which represents the values of two variables. The slope of the points suggests the direction of the relationship between the two variables. The amount of scatter suggests the strength of the correlation (little scatter indicates high correlation).

- 6-1** What are positive and negative correlations, and why do they enable prediction but not cause-effect explanation?

Describing behavior is a first step toward predicting it. Naturalistic observations and surveys often show us that one trait or behavior is related to another. In such cases, we say the two **correlate**. A statistical measure (the **correlation coefficient**) helps us figure how closely two things vary together, and thus how well either one predicts the other. Knowing how much aptitude test scores correlate with school success tells us how well the scores predict school success.

Throughout this book we will often ask how strongly two things are related: For example, how closely related are the personality scores of identical twins? How well do intelligence test scores predict career achievement? How closely is stress related to disease? In such cases, **scatterplots** can be very revealing.

Each dot in a scatterplot represents the values of two variables. The three scatterplots in **FIGURE 6.1** illustrate the range of possible correlations from a perfect positive to a perfect negative. (Perfect correlations rarely occur in the "real world.") A correlation is positive if two sets of scores, such as height and weight, tend to rise or fall together.

**Figure 6.1**

Scatterplots, showing patterns of correlation Correlations can range from +1.00 (scores on one measure increase in direct proportion to scores on another) to -1.00 (scores on one measure decrease precisely as scores rise on the other).

Saying that a correlation is “negative” says nothing about its strength or weakness. A correlation is negative if two sets of scores relate inversely, one set going up as the other goes down. The study of Nevada university students’ inner speech discussed in Module 5 also included a correlational component. Students’ reports of inner speech correlated negatively ($-.36$) with their scores on another measure: psychological distress. Those who reported more inner speech tended to report slightly *less* psychological distress.

Statistics can help us see what the naked eye sometimes misses. To demonstrate this for yourself, try an imaginary project. Wondering if tall men are more or less easygoing, you collect two sets of scores: men’s heights and men’s temperaments. You measure the heights of 20 men, and you have someone else independently assess their temperaments (from zero for extremely calm to 100 for highly reactive).

With all the relevant data right in front of you (**TABLE 6.1**), can you tell whether the correlation between height and reactive temperament is positive, negative, or close to zero?

Comparing the columns in Table 6.1, most people detect very little relationship between height and temperament. In fact, the correlation in this imaginary example is positive, +0.63, as we can see if we display the data as a scatterplot. In **FIGURE 6.2** on the next page, moving from left to right, the upward, oval-shaped slope of the cluster of points shows that our two imaginary sets of scores (height and temperament) tend to rise together.

If we fail to see a relationship when data are presented as systematically as in Table 6.1, how much less likely are we to notice them in everyday life? To see what is right in front of us, we sometimes need statistical illumination. We can easily see evidence of gender discrimination when given statistically summarized information about job level, seniority, performance,

Table 6.1 Height and Temperamental Reactivity of 20 Men

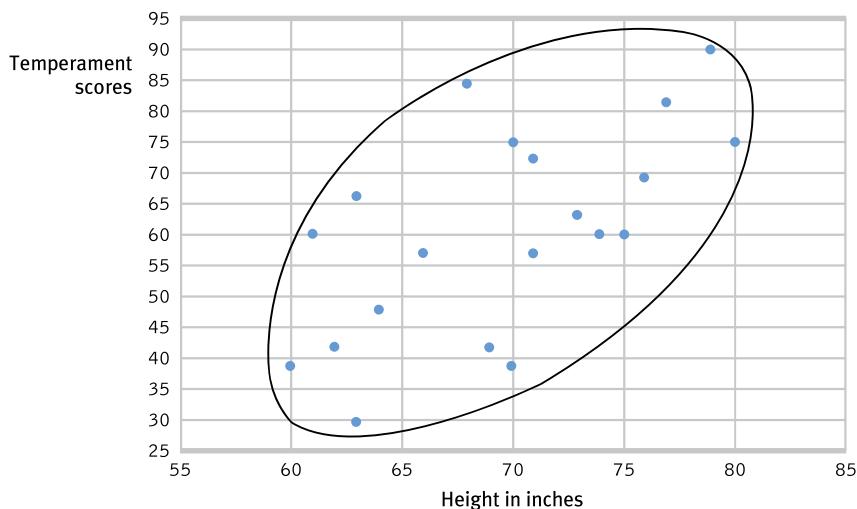
Person	Height in Inches	Temperament
1	80	75
2	63	66
3	61	60
4	79	90
5	74	60
6	69	42
7	62	42
8	75	60
9	77	81
10	60	39
11	64	48
12	76	69
13	71	72
14	66	57
15	73	63
16	70	75
17	63	30
18	71	57
19	68	84
20	70	39

AP® Exam Tip

This is the first of several times in your psychology course that you will see something labeled as being positive or negative. We often think that if something is positive it is good and if it’s negative it’s bad. That is rarely the case in this course. Here, positive and negative refer only to the direction of the correlation. They say nothing about whether the relationship is desirable or not.

Figure 6.2

Scatterplot for height and reactive temperament This display of data from 20 imagined people (each represented by a data point) reveals an upward slope, indicating a positive correlation. The considerable scatter of the data indicates the correlation is much lower than +1.0.



gender, and salary. But we often see no discrimination when the same information dribbles in, case by case (Twiss et al., 1989). See **TABLE 6.2** to test your understanding further.

The point to remember: A correlation coefficient, which can range from -1.0 to $+1.0$, reveals the extent to which two things relate. The closer the score gets to -1 or $+1$, the stronger the correlation.

Table 6.2

Test your understanding of correlation. Which of the following news reports are examples of a *positive* correlation, and which are examples of a *negative* correlation? (Check your answers below.)

1. The more children and youth used various media, the less happy they were with their lives (Kaiser, 2010). _____
2. The less sexual content teens saw on TV, the less likely they were to have sex (Collins et al., 2004). _____
3. The longer children were breast-fed, the greater their later academic achievement (Horwood & Ferguson, 1998). _____
4. The more income rose among a sample of poor families, the fewer psychiatric symptoms their children experienced (Costello et al., 2003). _____

ANSWERS: 1. negative, 2. positive, 3. positive, 4. negative

Correlation and Causation

Correlations help us predict. The *New York Times* reports that U.S. counties with high gun ownership rates tend to have high murder rates (Luo, 2011). Gun ownership predicts homicide. What might explain this guns-homicide correlation?

I can almost hear someone thinking, “Well, of course, guns kill people, often in moments of passion.” If so, that could be an example of A (guns) causes B (murder). But I can hear other readers saying, “Not so fast. Maybe people in dangerous places buy more guns for self-protection—maybe B causes A.” Or maybe some third variable C causes both A and B.

Another example: Self-esteem correlates negatively with (and therefore predicts) depression. (The lower people's self-esteem, the more they are at risk for depression.) So, does low self-esteem *cause* depression? If, based on the correlational evidence, you assume that it does, you have much company. A nearly irresistible thinking error is assuming that an association, sometimes presented as a correlation coefficient, proves causation. But no matter how strong the relationship, it does not.

As options 2 and 3 in **FIGURE 6.3** show, we'd get the same negative correlation between self-esteem and depression if depression caused people to be down on themselves, or if some third variable—such as heredity or brain chemistry—caused both low self-esteem and depression.

This point is so important—so basic to thinking smarter with psychology—that it merits one more example. A survey of over 12,000 adolescents found that the more teens feel loved by their parents, the less likely they are to behave in unhealthy ways—having early sex, smoking, abusing alcohol and drugs, exhibiting violence (Resnick et al., 1997). “Adults have a powerful effect on their children’s behavior right through the high school years,” gushed an Associated Press (AP) story reporting the finding. But this correlation comes with no built-in cause-effect arrow. The AP could as well have reported, “Well-behaved teens feel their parents’ love and approval; out-of-bounds teens more often think their parents are disapproving jerks.”

The point to remember (turn the volume up here): *Association does not prove causation.*¹ Correlation indicates the *possibility* of a cause-effect relationship *but does not prove such*. Remember this principle and you will be wiser as you read and hear news of scientific studies.



© Nancy Brown/Getty Images

Correlation need not mean causation Length of marriage correlates with hair loss in men. Does this mean that marriage *causes* men to lose their hair (or that balding men make better husbands)? In this case, as in many others, a third variable probably explains the correlation: Golden anniversaries and baldness both accompany aging.

FYI

A New York Times writer reported a massive survey showing that “adolescents whose parents smoked were 50 percent more likely than children of nonsmokers to report having had sex.” He concluded (would you agree?) that the survey indicated a causal effect—that “to reduce the chances that their children will become sexually active at an early age” parents might “quit smoking” (O’Neil, 2002).

AP® Exam Tip

Take note of how much emphasis is put on this idea. Correlation and association do not prove a cause-effect relationship.

¹Because many associations are stated as correlations, the famously worded principle is “Correlation does not prove causation.” That’s true, but it’s also true of associations verified by other nonexperimental statistics (Hatfield et al., 2006).

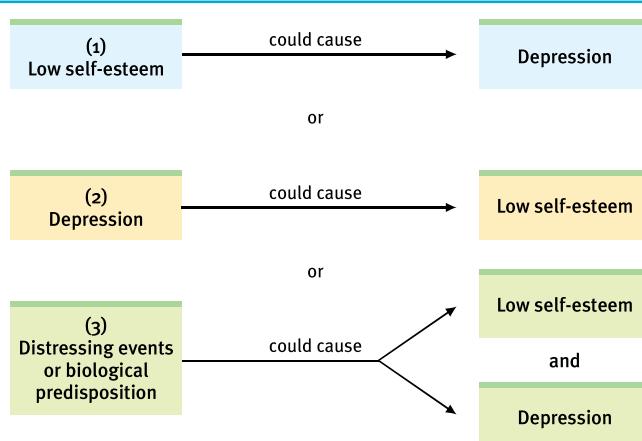


Figure 6.3
Three possible cause-effect relationships People low in self-esteem are more likely to report depression than are those high in self-esteem. One possible explanation of this negative correlation is that a bad self-image causes depressed feelings. But, as the diagram indicates, other cause-effect relationships are possible.

Illusory Correlations

6-2

What are illusory correlations?

illusory correlation the perception of a relationship where none exists.

FYI

A study reported in the *British Medical Journal* found that youths who identify with the goth subculture attempt, more often than other young people, to harm or kill themselves (Young et al., 2006). Can you imagine multiple possible explanations for this association?

Correlation coefficients make visible the relationships we might otherwise miss. They also restrain our “seeing” relationships that actually do not exist. A perceived but nonexistent correlation is an **illusory correlation**. When we *believe* there is a relationship between two things, we are likely to *notice* and *recall* instances that confirm our belief (Trolier & Hamilton, 1986).

Because we are sensitive to dramatic or unusual events, we are especially likely to notice and remember the occurrence of two such events in sequence—say, a premonition of an unlikely phone call followed by the call. When the call does not follow the premonition, we are less likely to note and remember the nonevent. Illusory correlations help explain many superstitious beliefs, such as the presumption that infertile couples who adopt become more likely to conceive (Gilovich, 1991). Couples who conceive after adopting capture our attention. We’re less likely to notice those who adopt and never conceive, or those who conceive without adopting. In other words, illusory correlations occur when we over-rely on the top left cell of **FIGURE 6.4**, ignoring equally essential information in the other cells.

Such illusory thinking helps explain why for so many years people believed (and many still do) that sugar makes children hyperactive, that getting chilled and wet causes people to catch a cold, and that changes in the weather trigger arthritis pain. We are, it seems, prone to perceiving patterns, whether they’re there or not.

The point to remember: When we notice random coincidences, we may forget that they are random and instead see them as correlated. Thus, we can easily deceive ourselves by seeing what is not there.

Figure 6.4

Illusory correlation in everyday life

Many people believe infertile couples become more likely to conceive a child after adopting a baby. This belief arises from their attention being drawn to such cases. The many couples who adopt without conceiving or conceive without adopting grab less attention. To determine whether there actually is a correlation between adoption and conception, we need data from all four cells in this figure. (From Gilovich, 1991.)

		Conceive	Do not conceive	
Adopt	confirming evidence	disconfirming evidence		
	disconfirming evidence	confirming evidence		



Michael Newman Jr./Photodisc

Experimentation

6-3

What are the characteristics of experimentation that make it possible to isolate cause and effect?

Happy are they, remarked the Roman poet Virgil, “who have been able to perceive the causes of things.” How might psychologists perceive causes in correlational studies, such as the correlation between breast feeding and intelligence?

Researchers have found that the intelligence scores of children who were breast-fed as infants are somewhat higher than the scores of children who were bottle-fed (Angelsen et al., 2001; Mortensen et al., 2002; Quinn et al., 2001). In Britain, breast-fed babies have also been more likely than their bottle-fed counterparts to eventually move into a higher social class (Martin et al., 2007). The “breast is best” intelligence effect shrinks when researchers compare breast-fed and bottle-fed children from the same families (Der et al., 2006).

What do such findings mean? Do smarter mothers (who in modern countries more often breast feed) have smarter children? Or, as some researchers believe, do the nutrients of mother's milk contribute to brain development? To find answers to such questions—to isolate cause and effect—researchers can **experiment**. Experiments enable researchers to isolate the effects of one or more variables by (1) manipulating the variables of interest and (2) holding constant ("controlling") other variables. To do so, they often create an **experimental group**, in which people receive the treatment, and a contrasting **control group** that does not receive the treatment.

Earlier we mentioned the place of *random sampling* in a well-done survey. Consider now the equally important place of *random assignment* in a well-done experiment. To minimize any preexisting differences between the two groups, researchers **randomly assign** people to the two conditions. Random assignment effectively equalizes the two groups. If one-third of the volunteers for an experiment can wiggle their ears, then about one-third of the people in each group will be ear wiggler. So, too, with ages, attitudes, and other characteristics, which will be similar in the experimental and control groups. Thus, if the groups differ at the experiment's end, we can surmise that the treatment had an effect.

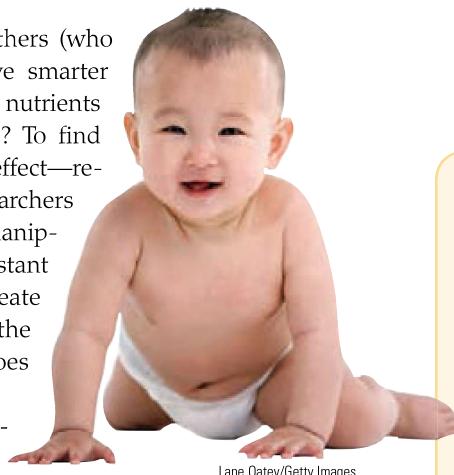
To experiment with breast feeding, one research team randomly assigned some 17,000 Belarus newborns and their mothers either to a breast-feeding promotion group or to a normal pediatric care program (Kramer et al., 2008). At 3 months of age, 43 percent of the infants in the experimental group were being exclusively breast-fed, as were 6 percent in the control group. At age 6, when nearly 14,000 of the children were restudied, those who had been in the breast-feeding promotion group had intelligence test scores averaging six points higher than their control condition counterparts.

No single experiment is conclusive, of course. But randomly assigning participants to one feeding group or the other effectively eliminated all variables except nutrition. This supported the conclusion that breast is indeed best for developing intelligence: If a behavior (such as test performance) changes when we vary an experimental variable (such as infant nutrition), then we infer the variable is having an effect.

The point to remember: Unlike correlational studies, which uncover naturally occurring relationships, an experiment manipulates a variable to determine its effect.

Consider, then, how we might assess therapeutic interventions. Our tendency to seek new remedies when we are ill or emotionally down can produce misleading testimonies. If three days into a cold we start taking vitamin C tablets and find our cold symptoms lessening, we may credit the pills rather than the cold naturally subsiding. In the 1700s, blood-letting seemed effective. People sometimes improved after the treatment; when they didn't, the practitioner inferred the disease was too advanced to be reversed. So, whether or not a remedy is truly effective, enthusiastic users will probably endorse it. To determine its effect, we must control for other variables.

And that is precisely how investigators evaluate new drug treatments and new methods of psychological therapy (Modules 72–73). They randomly assign participants in these studies to research groups. One group receives a treatment (such as a medication). The other group receives a pseudotreatment—an inert placebo (perhaps a pill with no drug in it). The participants are often blind (uninformed) about what treatment, if any, they are receiving. If the study is using a **double-blind procedure**, neither the participants nor the research assistants who administer the drug and collect the data will know which group is receiving the treatment.



Lane Oatey/Getty Images

experiment a research method in which an investigator manipulates one or more factors (independent variables) to observe the effect on some behavior or mental process (the dependent variable). By *random assignment* of participants, the experimenter aims to control other relevant variables.

experimental group in an experiment, the group exposed to the treatment, that is, to one version of the independent variable.

control group in an experiment, the group *not* exposed to the treatment; contrasts with the experimental group and serves as a comparison for evaluating the effect of the treatment.

random assignment assigning participants to experimental and control groups by chance, thus minimizing preexisting differences between the different groups.

double-blind procedure an experimental procedure in which both the research participants and the research staff are ignorant (blind) about whether the research participants have received the treatment or a placebo. Commonly used in drug-evaluation studies.

DORA: removed acronyms from AP Exam Tip--per note from Christine. --JO'N

In such studies, researchers can check a treatment's actual effects apart from the participants' and the staff's belief in its healing powers. Just thinking you are getting a treatment can boost your spirits, relax your body, and relieve your symptoms. This **placebo effect** is well documented in reducing pain, depression, and anxiety (Kirsch, 2010). And the more expensive the placebo, the more "real" it seems to us—a fake pill that costs \$2.50 works better than one costing 10 cents (Waber et al., 2008). To know how effective a therapy really is, researchers must control for a possible placebo effect.

placebo [pluh-SEE-bo; Latin for "I shall please"] **effect** experimental results caused by expectations alone; any effect on behavior caused by the administration of an inert substance or condition, which the recipient assumes is an active agent.

independent variable the experimental factor that is manipulated; the variable whose effect is being studied.

confounding variable a factor other than the independent variable that might produce an effect in an experiment.

dependent variable the outcome factor; the variable that may change in response to manipulations of the independent variable.

AP® Exam Tip

The identification of independent and dependent variables is the single most likely concept to be tested on the AP® exam. Experiments are critical to psychology, and independent and dependent variables are critical to experiments.

Independent and Dependent Variables

Here is a practical experiment: In a not yet published study, Victor Benassi and his colleagues gave college psychology students frequent in-class quizzes. Some items served merely as *review*—students were given questions with answers. Other *self-testing* items required students to actively produce the answers. When tested weeks later on a final exam, students did far better on material on which they had been tested (75 percent correct) rather than merely reviewed (51 percent correct). By a wide margin, testing beat restudy.

This simple experiment manipulated just one factor: the study procedure (reading answers versus self-testing). We call this experimental factor the **independent variable** because we can vary it *independently* of other factors, such as the students' memories, intelligence, and age. These other factors, which can potentially influence the results of the experiment, are called **confounding variables**. Random assignment controls for possible confounding variables.

Experiments examine the effect of one or more independent variables on some measurable behavior, called the **dependent variable** because it can vary *depending* on what takes place during the experiment. Both variables are given precise *operational definitions*, which specify the procedures that manipulate the independent variable (the review versus self-testing study method in this analysis) or measure the dependent variable (final exam performance). These definitions answer the "What do you mean?" question with a level of precision that enables others to repeat the study. (See **FIGURE 6.5** for the previously mentioned breast-milk experiment's design.)

Let's pause to check your understanding using a simple psychology experiment: To test the effect of perceived ethnicity on the availability of a rental house, researchers sent identically worded e-mail inquiries to 1115 Los Angeles-area landlords (Carpusor & Loges, 2006). The researchers varied the ethnic connotation of the sender's name and tracked the percentage of positive replies (invitations to view the apartment in person). "Patrick McDougall," "Said Al-Rahman," and "Tyrell Jackson" received, respectively, 89 percent, 66 percent, and 56 percent invitations. (In this experiment, what was the independent variable? The dependent variable?²²)

Figure 6.5

Experimentation To discern causation, psychologists may randomly assign some participants to an experimental group, others to a control group. Measuring the dependent variable (intelligence score in later childhood) will determine the effect of the independent variable (whether breast feeding was promoted).

Random assignment
(controlling for other confounding variables,
such as parental intelligence and environment)



© Radius Images/Alamy

Group	Independent variable	Dependent variable
Experimental	Promoted breast feeding	Intelligence score, age 6
Control	Did not promote breast feeding	Intelligence score, age 6

A key goal of experimental design is **validity**, which means the experiment will test what it is supposed to test. In the rental housing experiment, we might ask, “Did the e-mail inquiries test the effect of perceived ethnicity? Did the landlords’ response actually vary with the ethnicity of the name?”

Experiments can also help us evaluate social programs. Do early childhood education programs boost impoverished children’s chances for success? What are the effects of different antismoking campaigns? Do school sex-education programs reduce teen pregnancies? To answer such questions, we can experiment: If an intervention is welcomed but resources are scarce, we could use a lottery to randomly assign some people (or regions) to experience the new program and others to a control condition. If later the two groups differ, the intervention’s effect will be supported (Passell, 1993).

Let’s recap. A *variable* is anything that can vary (infant nutrition, intelligence, TV exposure—anything within the bounds of what is feasible and ethical). Experiments aim to *manipulate* an *independent variable*, *measure* the *dependent variable*, and allow random assignment to *control* all other variables. An experiment has at least two different conditions: an *experimental condition* and a *comparison or control condition*. *Random assignment* works to equate the groups before any treatment effects occur. In this way, an experiment tests the effect of at least one independent variable (what we manipulate) on at least one dependent variable (the outcome we measure). **TABLE 6.3** compares the features of psychology’s research methods.

²The independent variable, which the researchers manipulated, was the ethnicity-related names. The dependent variable, which they measured, was the positive response rate.

validity the extent to which a test or experiment measures or predicts what it is supposed to.

AP® Exam Tip

Almost 15 pages of text are summarized in this one table. Spend some time with it, as it is information you will likely encounter on the AP® exam.

Table 6.3 Comparing Research Methods

Research Method	Basic Purpose	How Conducted	What Is Manipulated	Strengths	Weaknesses
Descriptive	To observe and record behavior	Do case studies, naturalistic observations, or surveys	Nothing	Case studies require only one participant; naturalistic observations may be done when it is not ethical to manipulate variables; surveys may be done quickly and inexpensively (compared with experiments)	Uncontrolled variables mean cause and effect cannot be determined; single cases may be misleading
Correlational	To detect naturally occurring relationships; to assess how well one variable predicts another	Collect data on two or more variables; no manipulation	Nothing	Works with large groups of data, and may be used in situations where an experiment would not be ethical or possible	Does not specify cause and effect
Experimental	To explore cause and effect	Manipulate one or more variables; use random assignment	The independent variable(s)	Specifies cause and effect, and variables are controlled	Sometimes not feasible; results may not generalize to other contexts; not ethical to manipulate certain variables



"If I don't think it's going to work, will it still work?"

Before You Move On

► ASK YOURSELF

If you were to become a research psychologist, what questions would you like to explore with experiments?

► TEST YOURSELF

Why, when testing a new drug to control blood pressure, would we learn more about its effectiveness from giving it to half of the participants in a group of 1000 than to all 1000 participants?

Answers to the Test Yourself questions can be found in Appendix E at the end of the book.

Module 6 Review

6-1

What are positive and negative correlations, and why do they enable prediction but not cause-effect explanation?

- In a positive correlation, two variables rise or fall together. In a negative correlation, one item rises as the other falls.
- *Scatterplots* can help us to see *correlations*.
- A *correlation coefficient* can describe the strength and direction of a relationship between two variables, from +1.00 (a perfect positive correlation) through zero (no correlation at all) to -1.00 (a perfect negative correlation).

6-2

What are illusory correlations?

- *Illusory correlations* are random events that we notice and falsely assume are related.
- Patterns or sequences occur naturally in sets of random data, but we tend to interpret these patterns as meaningful connections, perhaps in an attempt to make sense of the world around us.

6-3

What are the characteristics of experimentation that make it possible to isolate cause and effect?

- To discover cause-effect relationships, psychologists conduct *experiments*, manipulating one or more variables of interest and controlling other variables.
- Using *random assignment*, they can minimize *confounding variables*, such as preexisting differences between the *experimental group* (exposed to the treatment) and the *control group* (given a placebo or different version of the treatment).
- The *independent variable* is the factor the experimenter manipulates to study its effect; the *dependent variable* is the factor the experimenter measures to discover any changes occurring in response to the manipulation of the independent variable.
- Studies may use a *double-blind procedure* to avoid the *placebo effect* and researcher's bias.
- An experiment has *validity* if it tests what it is supposed to test.

Multiple-Choice Questions

1. Which of the following is an example of negative correlation?
 - a. People who spend more time exercising tend to weigh less.
 - b. Teenage females tend to have fewer speeding tickets than teenage males.
 - c. Students with low IQ scores tend to have lower grades.
 - d. As hours studying for a test decrease, so do grades on that test.
 - e. Students' shoe sizes are not related to their grades.

- 2.** Which of the following is used only in correlation studies?
- Double blind
 - Placebo
 - Random assignment
 - Scatterplot
 - Random sample
- 3.** Researchers have discovered that individuals with lower income levels report having fewer hours of total sleep. Therefore,
- income and sleep levels are positively correlated.
 - income and sleep levels are negatively correlated.
 - income and sleep levels are inversely correlated.
 - income and sleep levels are not correlated.
 - lower income levels cause individuals to have fewer hours of sleep.
- 4.** Which of the following correlation coefficients represents the strongest relationship between two variables?
- +.30
 - .75
 - +1.3
 - .85
 - 1.2
- 5.** The purpose of random assignment is to
- allow participants in both the experimental and control groups to be exposed to the independent variable.
 - ensure that every member of the population had an equal chance of being selected to participate in the research.
 - eliminate the placebo effect.
 - reduce potential confounding variables.
 - generate operational definitions for the independent and dependent variables.

Practice FRQs

- 1.** Students with higher scores on anxiety scales were found to have lower scores on standardized tests. What research method would show this relationship? Why can no cause-effect conclusion be drawn from the results?

Answer

1 point: This research method is a correlation study.

1 point: There are three possibilities for causation: Anxiety could cause low test scores, low test scores could cause anxiety, or a third factor could cause both anxiety and low test scores. No conclusions can be drawn about causation because this is not an experiment.

- 2.** Ms. Ledbetter wants to determine if the new review activity she developed will improve student performance on unit exams. She randomly separates 160 students into two groups. Group A reviews for the unit exam in the traditional manner they have always used. Group B participates in the new review activity. After reviewing both groups are given the same unit exam and their scores are compared. Identify the independent and dependent variables for this experiment.

(2 points)