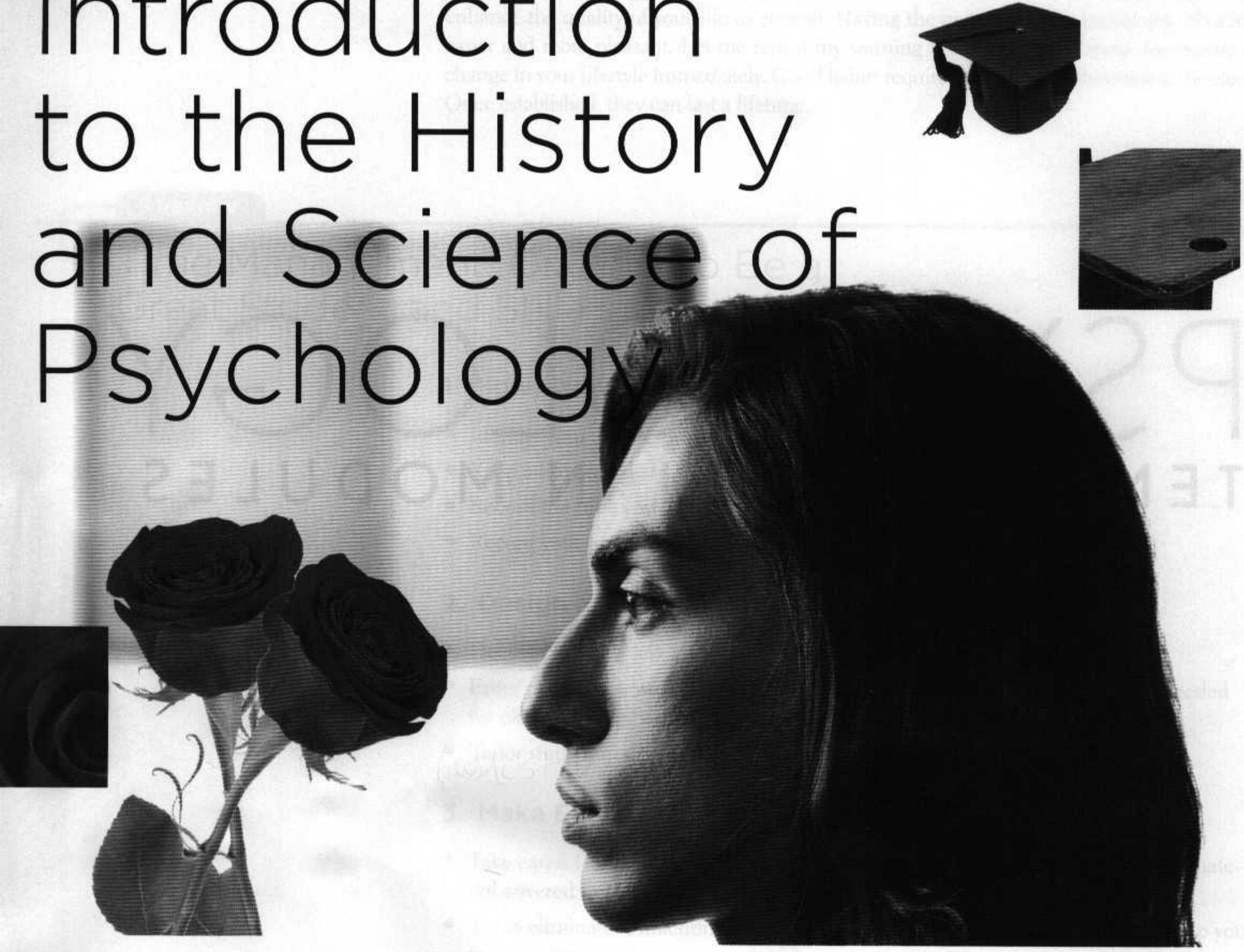
Introduction to the History and Science of Psychology



Harvard astronomer Owen Gingerich (2006) reports that there are more than 100 billion galaxies. One of these, our own relative speck of a galaxy, has some 200 billion stars, many of which, like our Sun-star, are circled by planets. On the scale of outer space, we are less than a single grain of sand on all the oceans' beaches, and our lifetime but a relative nanosecond.

Yet there is nothing more awe inspiring and absorbing than our own inner space. Our brain, adds Gingerich, "is by far the most complex physical object known to us in the entire cosmos" (p. 29). Our consciousness—our mind somehow arising from matter—remains a profound mystery. Our thinking, emotions, and actions (and their interplay with others' thinking, emotions, and actions) fascinate us. Outer space staggers us with its enormity. But inner space enthralls us. Enter psychological science.

For people whose exposure to psychology comes from popular media, psychologists seem to analyze personality, offer counseling, and dispense child-rearing advice. Do they? *Yes*, and much more. Consider some of psychology's questions that you may wonder about:

- Have you ever found yourself reacting to something as one of your biological parents would—perhaps in a way you vowed you never would—and then wondered how much of your personality you inherited? *To what extent do genes predispose our person-to-person differences in personality? To what extent do home and community environments shape us?*

MODULES

1 The Story of Psychology

2 Thinking Critically With Psychological Science

3 Research Strategies: How Psychologists Ask and Answer Questions

- Have you ever worried about how to act among people of a different culture, race, gender, or sexual orientation? *In what ways are we alike as members of the human family? How do we differ?*
- Have you ever awakened from a nightmare and, with a wave of relief, wondered why you had such a crazy dream? *How often, and why, do we dream?*
- Have you ever played peekaboo with a 6-month-old and wondered why the baby finds the game so delightful? *What do babies actually perceive and think?*
- Have you ever wondered what fosters school and work success? *Are some people just born smarter? And why do some people get richer, think more creatively, or relate more sensitively?*
- Have you ever become depressed or anxious and wondered whether you'll ever feel "normal"? *What triggers our bad moods? What's the line between a normal mood swing and a* psychological disorder *for which someone should seek help?*
- Have you ever wondered how the Internet, video games, and electronic social networks affect people? *How do today's electronic media influence how we think and how we relate?*

Psychology is a science that seeks to answer such questions about us all—how and why we think, feel, and act as we do. In Module 1, we trace psychology's roots and survey its scope. In Module 2, we consider how psychological science can help you to think critically and to understand the dangers of relying on intuition and common sense. In Module 3, we survey psychology's methods—how psychologists ask and answer questions.

MODULE 2

Thinking Critically With Psychological Science

- The Need for Psychological Science
- Frequently Asked Questions About Psychology

Hoping to satisfy their curiosity about people and to remedy their own woes, millions turn to "psychology." They listen to talk-radio counseling. They read articles on psychic powers. They attend stop-smoking hypnosis seminars. They immerse themselves in self-help websites and books on the meaning of dreams, the path to ecstatic love, and the roots of personal happiness.

Others, intrigued by claims of psychological truth, wonder: Do mothers and infants bond in the first hours after birth? How—and how much—does parenting shape children's personalities and abilities? Are first-born children more driven to achieve? Does psychotherapy heal?

In working with such questions, how can we separate uninformed opinions from examined conclusions? *How can we best use psychology to understand why people think, feel, and act as they do?*

The Need for Psychological Science

2-1 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 my grandmother knew?" 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). Former President George W. Bush described the feeling to journalist Bob Woodward (2002) in explaining his decision to launch the Iraq war: "I'm a gut player. I rely on my instincts."

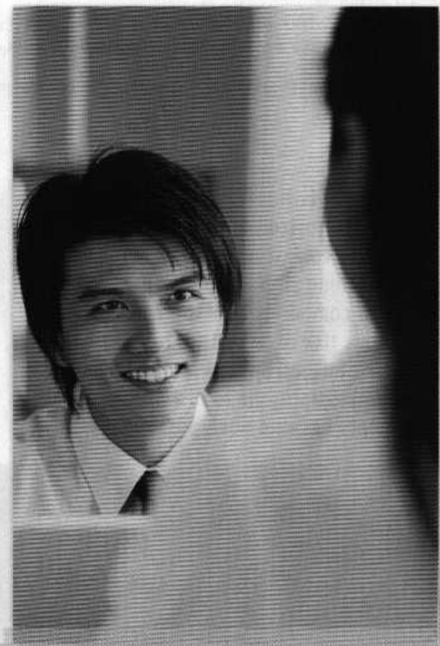
Prince Charles and former President Bush have much company, judging from the long list of pop psychology books on "intuitive managing," "intuitive trading," and "intuitive healing." Today's psychological science does document a vast intuitive mind. As we will see, our thinking, memory, and attitudes 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.

Experiments have found people greatly overestimating their lie detection accuracy, their eyewitness recollections, their interviewee assessments, their risk predictions, and their stock-picking talents. "The first principle," said Richard Feynman (1997), "is that you must not fool yourself—and you are the easiest person to fool."

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.



Asia Images Group Pte Ltd / Alamy

The limits of intuition Personnel interviewers tend to be overconfident of their gut feelings about job applicants. Their confidence stems partly from their recalling cases where their favorable impression proved right, and partly from their ignorance about rejected applicants who succeeded elsewhere.

"Those who trust in their own wits are fools."

Proverbs 28:26

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.*)

"Life is lived forwards, but understood backwards."

Philosopher Søren Kierkegaard, 1813–1855

"Anything seems commonplace, once explained."

Dr. Watson to Sherlock Holmes

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.

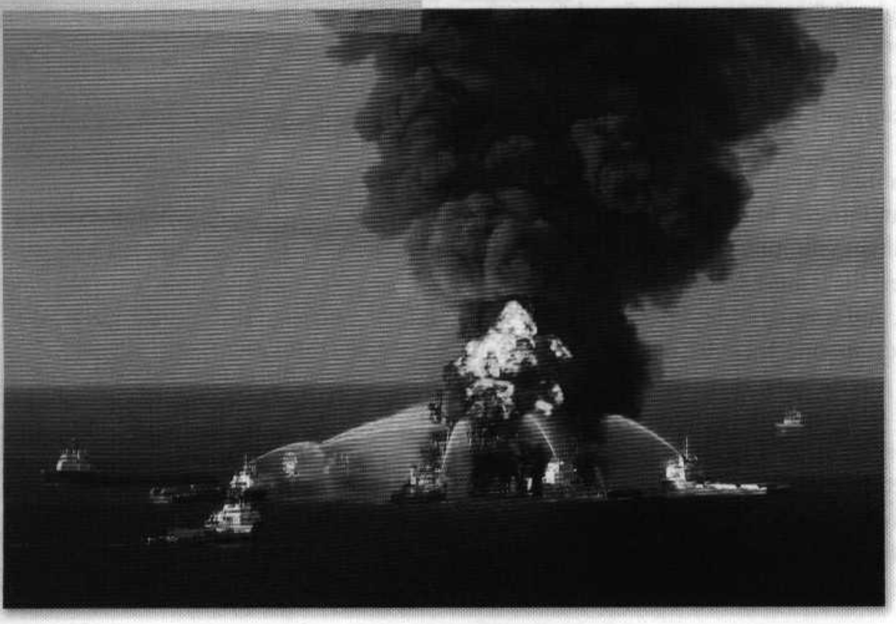
Tell the second group the opposite, "Psychologists have found that separation strengthens romantic attraction. As the saying goes, 'Absence makes the heart grow fonder.'" People given this untrue result can also easily imagine it, and most will also see it as unsurprising. When two opposite findings 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, Grandma's 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 surprising if many of psychology's findings had *not* been foreseen. Many people believe that love breeds happiness, and they are right, according to researchers who have found that we have 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 Grandma's intuition, informed by countless casual observations, has it wrong. Psychological 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. It has also 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.

Hindsight bias When drilling its Deepwater Horizon oil well in 2010, BP employees took some shortcuts and ignored some warning signs, without intending to put their company and the environment at serious risk of devastation. *After* the resulting Gulf oil spill, with the benefit of 20/20 hindsight, the foolishness of those judgments became obvious.



REUTERS/U.S. Coast Guard/Handout

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 separatists *almost* won the secessionist referendum."

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

RETRIEVAL PRACTICE

- Why, after friends start dating, do we often feel that we *knew* they were meant to be together?

ANSWER: We often suffer from hindsight bias—after we've learned a situation's outcome, that outcome seems familiar and therefore obvious.

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

Maciej Oleksy /Shutterstock

[1] Boston is south of Paris.

[2] The anagram solution: CHAOS.

Bizarre-looking, perhaps. But actually no more unlikely than any other number sequence.

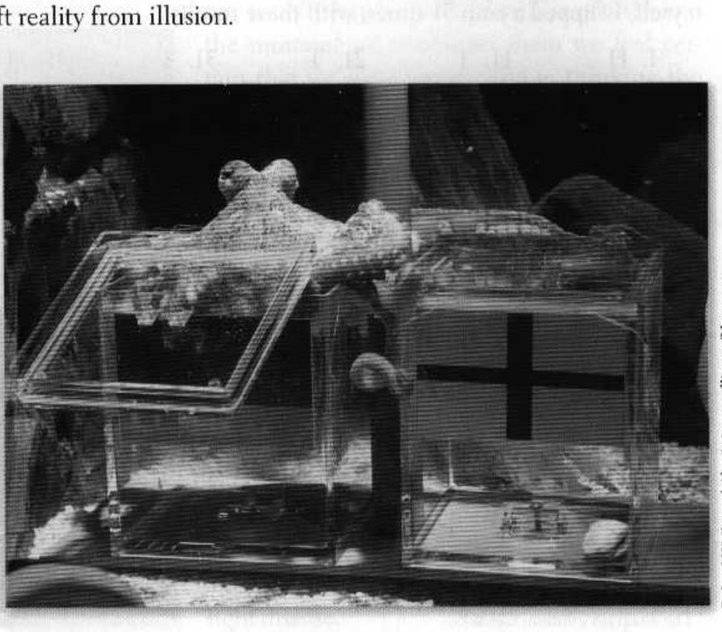
Looking over the sequence, patterns jump out: Tosses 10 to 22 provided an almost perfect pattern of pairs of tails followed by pairs of heads. On tosses 30 to 38 I had a "cold hand," with only one head in nine tosses. But my fortunes immediately reversed with a "hot hand"—seven heads out of the next 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!")

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 really unusual day would be one where nothing unusual happens."
Statistician Persi Diaconis (2002)

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.



Given enough random events, some weird-seeming streaks will occur During the 2010 World Cup, a German octopus—Paul, "the oracle of Oberhausen"—was offered two boxes, each with mussels and with a national flag on one side. Paul selected the right box eight out of eight times in predicting the outcome of Germany's seven matches and Spain's triumph in the final.

The Scientific Attitude: Curious, Skeptical, and Humble

2-2 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.

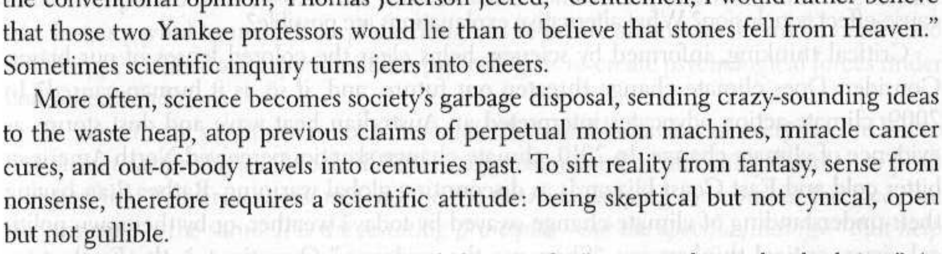
No matter how sensible-seeming or wild an idea, the smart thinker asks: *Does it work?* When put to the test, can its predictions be confirmed? Subjected to such scrutiny, crazy-sounding ideas sometimes find support. During the 1700s, scientists scoffed at the notion that meteorites had extraterrestrial 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.

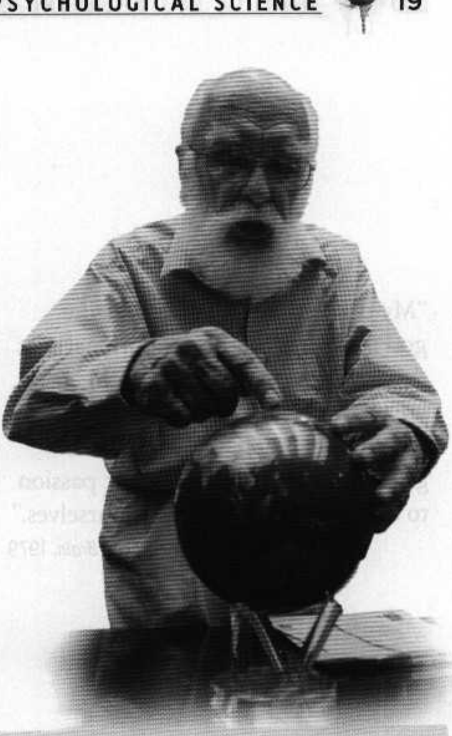
More often, science becomes society's garbage disposal, sending crazy-sounding ideas to the waste heap, atop previous claims of perpetual motion machines, miracle cancer cures, and out-of-body travels into centuries past. To sift reality from fantasy, sense from nonsense, therefore requires a scientific 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.

Non Sequitur





The Amazing Randi The magician James Randi exemplifies skepticism. He has tested and debunked a variety of psychic phenomena.
AP Photo/Alan Diaz

"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

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."

"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

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).

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, discerns hidden values, evaluates 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 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.

"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

Has psychology's critical inquiry been open to surprising findings? The answer is plainly *Yes*. Some examples: massive losses of brain tissue early in life may have minimal long-term effects. Within days, newborns can recognize their mother's odor and voice. After brain damage, a person may be able to learn new skills yet be unaware of such learning. Diverse groups—men and women, old and young, rich and middle class, those with disabilities and without—report roughly comparable levels of personal happiness.

And has critical inquiry convincingly debunked popular presumptions? The answer is again *Yes*. The evidence indicates that sleepwalkers are *not* acting out their dreams. 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. Most people do *not* suffer from unrealistically low self-esteem, and high self-esteem is not all good. Opposites do *not* generally attract. In each of these instances and more, what has been learned is not what is widely believed.

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

RETRIEVAL PRACTICE

- How does the scientific attitude contribute to critical thinking?

ANSWER: The scientific attitude combines (1) *curiosity* about the world around us, (2) *skepticism* toward various claims and ideas, and (3) *humility* about one's own understanding. Evaluating evidence, assessing conclusions, and examining our own assumptions are essential parts of critical thinking.

Frequently Asked Questions About Psychology

You are now prepared to think about psychological matters. Yet, even knowing this much, you may still be approaching psychology with a mixture of curiosity and apprehension. So before we plunge in, let's entertain some frequently asked questions.

2-3 Can laboratory experiments illuminate everyday life?

When you see or hear about psychological research, do you ever wonder whether people's behavior in the lab will predict their behavior in real life? For example, does detecting the blink of a faint red light in a dark room have anything useful to say about flying a plane at night? After viewing a violent, sexually explicit film, does an aroused man's increased willingness to push buttons that he thinks will electrically shock a woman really say anything about whether violent pornography makes a man more likely to abuse a woman?

Before you answer, consider: The experimenter *intends* the laboratory environment to be a simplified reality—one that simulates and controls important features of everyday life. Just as a wind tunnel lets airplane designers re-create airflow forces under controlled conditions, a laboratory experiment lets psychologists re-create psychological forces under controlled conditions.

An experiment's purpose is not to re-create the exact behaviors of everyday life but to test *theoretical principles* (Mook, 1983). In aggression studies, deciding whether to push a button that delivers a shock may not be the same as slapping someone in the face, but the principle is the same. *It is the resulting principles—not the specific findings—that help explain everyday behaviors.*

When psychologists apply laboratory research on aggression to actual violence, they are applying theoretical principles of aggressive behavior, principles they have refined through many experiments. Similarly, it is the principles of the visual system, developed from experiments in artificial settings (such as looking at red lights in the dark), that researchers apply to more complex behaviors such as night flying. And many investigations show that principles derived in the laboratory do typically generalize to the everyday world (Anderson et al., 1999).

The point to remember: Psychological science focuses less on particular behaviors than on seeking general principles that help explain many behaviors.

2-4 Does behavior depend on one's culture and gender?

What can psychological studies done in one time and place—often with people from what Joseph Henrich, Steven Heine, and Ara Norenzayan (2010) call the WEIRD cultures (*W*estern, *E*ducated, *I*ndustrialized, *R*ich, and *D*emocratic cultures that contribute most study participants but are only 12 percent of humanity)—really tell us about people in general? As we will see time and again, **culture**—shared ideas and behaviors that one generation passes on to the next—matters. Our culture shapes our behavior. It influences our standards of promptness and frankness, our attitudes toward premarital sex and varying body shapes, our tendency to be casual or formal, our willingness to make eye contact, our conversational distance, and much, much more. Being aware of such differences, we can

culture the enduring behaviors, ideas, attitudes, values, and traditions shared by a group of people and transmitted from one generation to the next.

A cultured greeting Because culture shapes people's understanding of social behavior, actions that seem ordinary to us may seem quite odd to visitors from far away. Yet underlying these differences are powerful similarities. Supporters of newly elected leaders everywhere typically greet them with pleased deference, though not necessarily with bows and folded hands, as in India. Here influential and popular politician Sonia Gandhi greets some of her constituents shortly after her election.

Ami Vitale/Getty Images

restrain our assumptions that others will think and act as we do. Given the growing mixing and clashing of cultures, our need for such awareness is urgent.

It is also true, however, that our shared biological heritage unites us as a universal human family. The same underlying processes guide people everywhere.

- People diagnosed with dyslexia, a reading disorder, exhibit the same brain malfunction whether they are Italian, French, or British (Paulesu et al., 2001).
- Variation in languages may impede communication across cultures. Yet all languages share deep principles of grammar, and people from opposite hemispheres can communicate with a smile or a frown.
- People in different cultures vary in feelings of loneliness. But across cultures, loneliness is magnified by shyness, low self-esteem, and being unmarried (Jones et al., 1985; Rokach et al., 2002).

We are each in certain respects like all others, like some others, and like no other. Studying people of all races and cultures helps us discern our similarities and our differences, our human kinship and our diversity.

You will see throughout this book that *gender* matters, too. Researchers report gender differences in what we dream, in how we express and detect emotions, and in our risk for alcohol dependence, depression, and eating disorders. Gender differences fascinate us, and studying them is potentially beneficial. For example, many researchers believe that women carry on conversations more readily to build relationships, while men talk more to give information and advice (Tannen, 2001). Knowing this difference can help us prevent conflicts and misunderstandings in everyday relationships.

But again, psychologically as well as biologically, women and men are overwhelmingly similar. Whether female or male, we learn to walk at about the same age. We experience the same sensations of light and sound. We feel the same pangs of hunger, desire, and fear. We exhibit similar overall intelligence and well-being.

The point to remember: Even when specific attitudes and behaviors vary by gender or across cultures, as they often do, the underlying processes are much the same.

"All people are the same; only their habits differ."

Confucius, 551–479 B.C.E

2-5 Why do psychologists study animals, and what ethical guidelines safeguard human and animal research participants?

Many psychologists study animals because they find them fascinating. They want to understand how different species learn, think, and behave. Psychologists also study animals to learn about people. We humans are not *like* animals; we *are* animals, sharing a common biology. Animal experiments have therefore led to treatments for human diseases—insulin for diabetes, vaccines to prevent polio and rabies, transplants to replace defective organs.

Humans are complex. But the same processes by which we learn are present in rats, monkeys, and even sea slugs. The simplicity of the sea slug's nervous system is precisely what makes it so revealing of the neural mechanisms of learning. Sharing such similarities, should we not respect our animal relatives? "We cannot defend our scientific work with animals on the basis of the similarities between them and ourselves and then defend it morally on the basis of differences," noted Roger Ulrich (1991). The animal protection movement protests the use of animals in psychological, biological, and medical research.

Out of this heated debate, two issues emerge. The basic one is whether it is right to place the well-being of humans above that of animals. In experiments on stress and cancer, is it right that mice get tumors in the hope that people might not? Should some monkeys be exposed to an HIV-like virus in the search for an AIDS vaccine? Is our use and consumption of other animals as natural as the behavior of carnivorous hawks, cats, and whales? The answers to such questions vary by culture. In Gallup surveys in Canada and the United States, about 60 percent of adults deem medical testing on animals "morally acceptable." In Britain, only 37 percent do (Mason, 2003).

If we give human life first priority, what safeguards should protect the well-being of animals in research? One survey of animal researchers gave an answer. Some 98 percent supported government regulations protecting primates, dogs, and cats, and 74 percent supported regulations providing for the humane care of rats and mice (Plous & Herzog, 2000). Many professional associations and funding agencies already have such guidelines. British Psychological Society guidelines call for housing animals under reasonably natural living conditions, with companions for social animals (Lea, 2000). American Psychological Association guidelines state that researchers must ensure the "comfort, health, and humane treatment" of animals and minimize "infection, illness, and pain" (APA, 2002). The European Parliament now mandates standards for animal care and housing (Vogel, 2010).

Animals have themselves benefited from animal research. One Ohio team of research psychologists measured stress hormone levels in samples of millions of dogs brought each year to animal shelters. They devised handling and stroking methods to reduce stress and ease the dogs' transition to adoptive homes (Tuber et al., 1999). Other studies have helped improve care and management in animals' natural habitats. By revealing our behavioral kinship with animals and the remarkable intelligence of chimpanzees, gorillas, and other animals, experiments have also led to increased empathy and protection for them. At its best, a psychology concerned for humans and sensitive to animals serves the welfare of both.

What about human participants? Does the image of white-coated scientists delivering electric shocks trouble you? If so, you'll be relieved to know that most psychological studies are free of such stress. With people, blinking lights, flashing words, and pleasant social interactions are more common. Moreover, psychology's experiments are mild compared with the stress and humiliation often inflicted by reality TV shows. In one episode of *The Bachelor*, a man dumped his new fiancée—on camera, at the producers' request—for the woman who earlier had finished second (Collins, 2009).

Occasionally, though, researchers do temporarily stress or deceive people, but only when they believe it is essential to a justifiable end, such as understanding and controlling violent behavior or studying mood swings. Some experiments won't work if participants know everything beforehand. (Wanting to be helpful, the participants might try to confirm the researcher's predictions.)

"Rats are very similar to humans except that they are not stupid enough to purchase lottery tickets."

Dave Barry, July 2, 2002

"Please do not forget those of us who suffer from incurable diseases or disabilities who hope for a cure through research that requires the use of animals."

Psychologist Dennis Feeney (1987)

"The greatness of a nation can be judged by the way its animals are treated."

Mahatma Gandhi, 1869–1948

Animal research benefiting animals Thanks partly to research on the benefits of novelty, control, and stimulation, these gorillas are enjoying an improved quality of life in New York's Bronx Zoo.

Ami Vitale/Getty Images

informed consent an ethical principle that research participants be told enough to enable them to choose whether they wish to participate.

debriefing the postexperimental explanation of a study, including its purpose and any deceptions, to its participants.

The American Psychological Association's ethics code urges researchers to (1) obtain potential participants' **informed consent**, (2) protect them from harm and discomfort, (3) keep information about individual participants confidential, and (4) fully **debrief** people (explain the research afterward). Moreover, university ethics committees screen research proposals and safeguard participants' well-being.

2-6 Is psychology free of value judgments?

"It is doubtless impossible to approach any human problem with a mind free from bias."

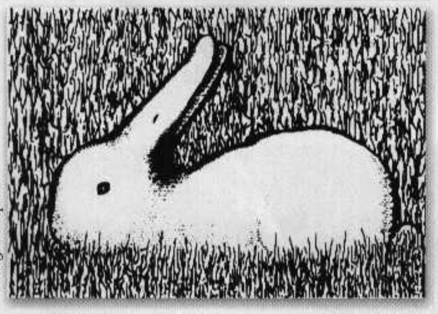
Simone de Beauvoir, *The Second Sex*, 1953

Psychology is definitely not value free. Values affect what we study, how we study it, and how we interpret results. Researchers' values influence their choice of topics. Should we study worker productivity or worker morale? Sex discrimination or gender differences? Conformity or independence? Values can also color "the facts." As we noted earlier, our preconceptions can bias our observations and interpretations; sometimes we see what we want or expect to see (**FIGURE 2.1**).

Mike Kemp/Getty Images

Even the words we use to describe something can reflect our values. Are the sex acts we do not practice "perversions" or "sexual variations"? In psychology and in everyday speech, labels describe and labels evaluate: One person's *rigidity* is another's *consistency*. One person's *faith* is another's *fanaticism*. One country's *enhanced interrogation techniques*, such as cold-water immersion, become *torture* when practiced by its enemies. Our labeling someone as *firm* or *stubborn*, *careful* or *picky*, *discreet* or *secretive* reveals our own attitudes.

FIGURE 2.1
What do you see? Our expectations influence what we perceive. Did you see a duck or a rabbit? Show some friends this image with the rabbit photo above covered up and see if they are more likely to perceive a duck lying on its back instead. (From Shepard, 1990).



© Roger Shepard

Popular applications of psychology also contain hidden values. If you defer to "professional" guidance about how to live—how to raise children, how to achieve self-fulfillment, what to do with sexual feelings, how to get ahead at work—you are accepting value-laden advice. A science of behavior and mental processes can help us reach our goals. But it cannot decide what those goals should be.

If some people see psychology as merely common sense, others have a different concern—that it is becoming dangerously powerful. Is it an accident that astronomy is the oldest science and psychology the youngest? To some, exploring the external universe seems far safer than exploring our own inner universe. Might psychology, they ask, be used to manipulate people?

Knowledge, like all power, can be used for good or evil. Nuclear power has been used to light up cities—and to demolish them. Persuasive power has been used to educate people—and to deceive them. Although psychology does indeed have the power to deceive, its purpose is to enlighten. Every day, psychologists are exploring ways to enhance learning, creativity, and compassion. Psychology speaks to many of our world's great problems—war, overpopulation, prejudice, family crises, crime—all of which involve attitudes and behaviors. Psychology also speaks to our deepest longings—for nourishment, for love, for happiness. Psychology cannot address all of life's great questions, but it speaks to some mighty important ones.

Psychology speaks In making its historic 1954 school desegregation decision, the U.S. Supreme Court cited the expert testimony and research of psychologists Kenneth Clark and Mamie Phipps Clark (1947). The Clarks reported that, when given a choice between Black and White dolls, most African-American children chose the White doll, which seemingly indicated internalized anti-Black prejudice.

Office of Public Affairs at Columbia University

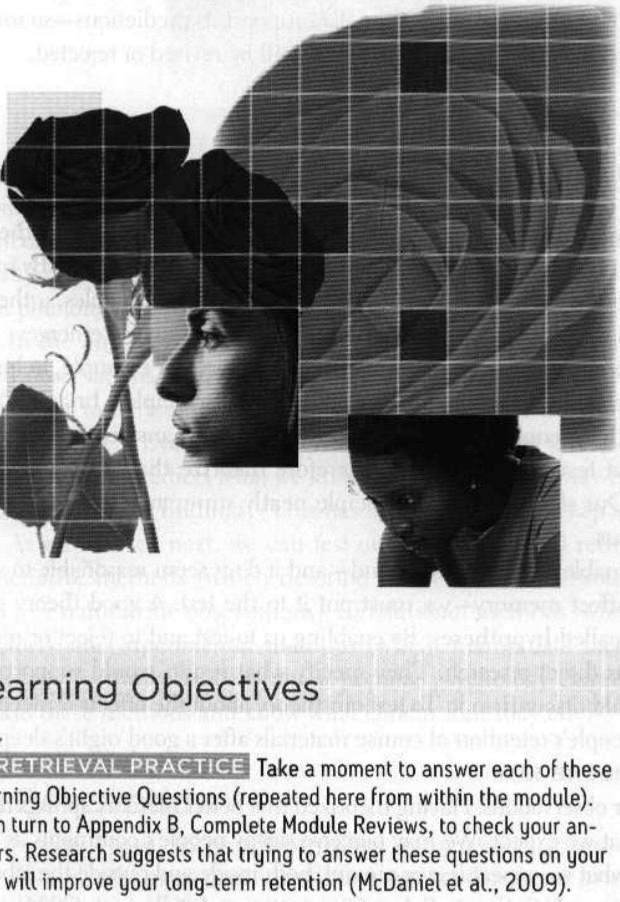
RETRIEVAL PRACTICE

- How are human and animal research participants protected?

ANSWER: Animal protection legislation, laboratory regulation and inspection, and local ethics committees serve to protect human and animal welfare. University ethics committees screen research proposals. Ethical principles developed by the American Psychological Association and other organizations urge researchers using human participants to obtain *informed consent*, to protect them from harm and discomfort, to treat their personal information confidentially, and to fully *debrief* all participants.

MODULE REVIEW

Thinking Critically With Psychological Science



Learning Objectives

RETRIEVAL PRACTICE Take a moment to answer each of these Learning Objective Questions (repeated here from within the module). Then turn to Appendix B, Complete Module Reviews, to check your answers. Research suggests that trying to answer these questions on your own will improve your long-term retention (McDaniel et al., 2009).

2-1: 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?

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

2-3: Can laboratory experiments illuminate everyday life?

2-4: Does behavior depend on one's culture and gender?

2-5: Why do psychologists study animals, and what ethical guidelines safeguard human and animal research participants?

2-6: Is psychology free of value judgments?

Terms and Concepts to Remember

RETRIEVAL PRACTICE Test yourself on these terms by trying to write down the definition before flipping back to the referenced page to check your answer.

hindsight bias, p. 16
critical thinking, p. 20
culture, p. 21
informed consent, p. 24
debriefing, p. 24

RETRIEVAL PRACTICE Gain an advantage, and benefit from immediate feedback, with the interactive self-testing resources at www.worthpublishers.com/myers.

- How *Do* Psychologists Ask and Answer Questions?
- Statistical Reasoning in Everyday Life

MODULE 3

Research Strategies: How Psychologists Ask and Answer Questions

How *Do* Psychologists Ask and Answer Questions?

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

3-1 How do theories advance psychological science?

In everyday conversation, we often use *theory* to mean "mere hunch." In science, a **theory** *explains* with principles that *organize* observations and *predict* behaviors or events. 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 good theory about the effects of sleep deprivation 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 poor sleep habits cannot answer questions in class, and they do poorly at test time. We might therefore theorize that sleep improves memory. So far so good: Our sleep-retention principle neatly summarizes a list of facts about the effects of sleep loss.

Yet no matter how reasonable a theory may sound—and it does seem reasonable to suggest that sleep loss could affect memory—we must put it to the test. A good theory produces testable predictions, called **hypotheses.** By enabling us to test and to reject or revise our theory, such predictions direct research. They specify what results would support the theory and what results would disconfirm it. To test our theory about the effects of sleep on memory, we might assess people's retention of course materials after a good night's sleep, or a shortened night's sleep (**FIGURE 3.1**).

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. The urge to see what we expect is ever-present, both inside and outside the laboratory. According to the bipartisan U.S. Senate Select Committee on Intelligence (2004), preconceived expectations that Iraq had weapons of mass destruction led intelligence analysts to wrongly interpret ambiguous observations as confirming that theory (much as people's views of climate change may influence their interpretation of local weather events). This theory-driven conclusion then led to the preemptive U.S. invasion of Iraq.

As a check on their biases, psychologists report their research with precise **operational definitions** of procedures and concepts. *Hunger,* for example, might be defined as "hours without eating," *generosity* as "money contributed." Using these carefully worded statements, others can **replicate** (repeat) the original observations with different participants,

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.

operational definition a statement of the procedures (operations) used to define research variables. 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.

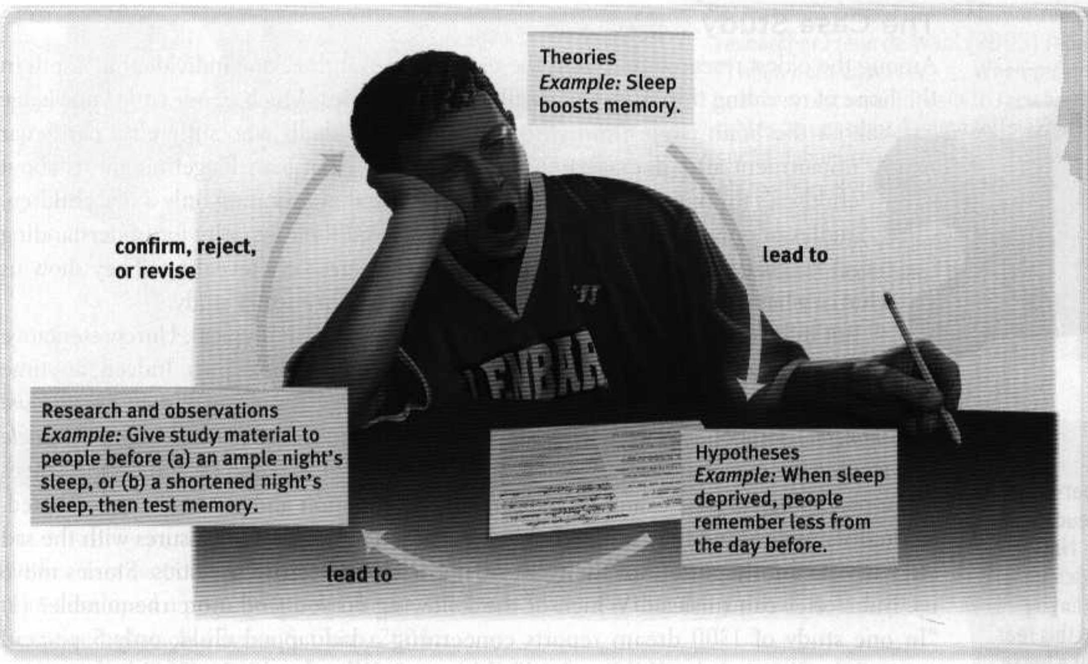


FIGURE 3.1
The scientific method A self-correcting process for asking questions and observing nature's answers.

materials, and circumstances. If they get similar results, confidence in the finding's reliability grows. The first study of hindsight bias (the tendency to believe, after learning an outcome, that one would have foreseen it) aroused psychologists' curiosity. Now, after many successful replications with differing people and questions, we feel sure of the phenomenon's power.

In the end, our theory will be useful if it (1) *organizes* a range of self-reports and observations, and (2) implies *predictions* that anyone can use to check the theory or to derive practical applications. (If people sleep more, will their retention improve?) Eventually, our research may lead to a revised theory that 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 we will see next, we can test our hypotheses and refine our theories using *descriptive* methods (which describe behaviors, often through case studies, surveys, or naturalistic observations), *correlational* methods (which associate different factors), and *experimental* methods (which manipulate factors to discover their effects). To think critically about popular psychology claims, we need to understand these methods and know what conclusions they allow.

RETRIEVAL PRACTICE

- What does a good theory do?

ANSWER: 1. It *organizes* observed facts. 2. It implies hypotheses that offer testable *predictions* and, sometimes, practical applications.

- Why is replication important?

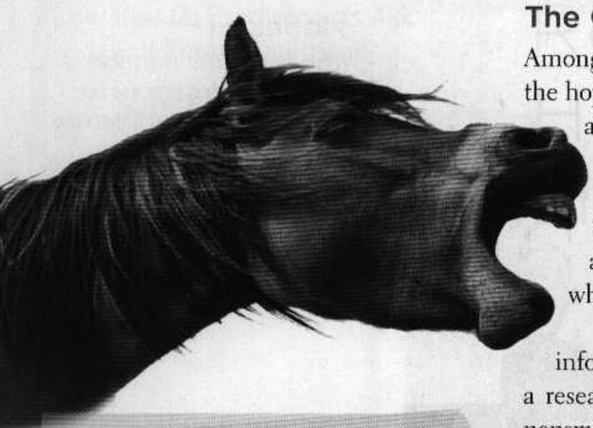
ANSWER: Psychologists watch eagerly for new findings, but they also proceed with caution—by awaiting other investigators' repeating the experiment to see if the finding will be confirmed (the result replicated).

Description

3-2 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).



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 an 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. Today's psychological science discounts Freud's theory of childhood sexuality but acknowledges that much of the human mind operates outside our conscious awareness.
Skye Hohmann/Alamy

The Case Study

Among the oldest research methods, the **case study** examines one individual in depth in the hope of revealing things true of us all. 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 after carefully observing and questioning only a few 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: 95 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.

"'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

RETRIEVAL PRACTICE

- Case studies do not allow us to learn about general principles that apply to all of us. Why not?

ANSWER: Case studies involve only one individual, so we can't know for sure whether the principles observed would apply to a larger population.

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 lunchroom.

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

case study an observation technique in which one person 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.

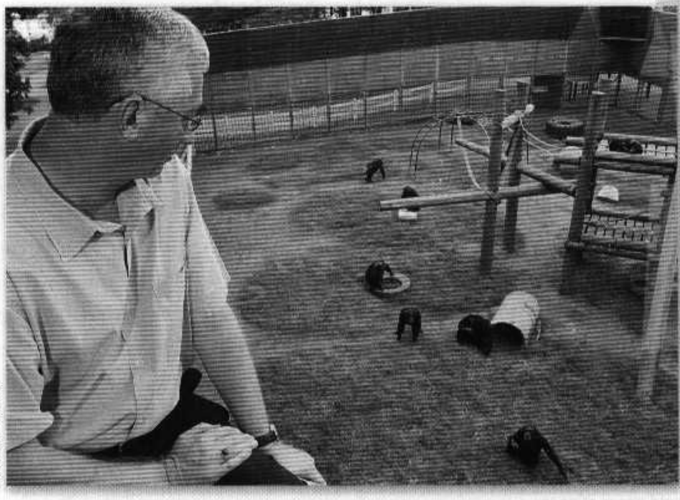


Photo by Jack Kearse, Emory University for Yerkes National Primate Research Center

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."

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 introductory psychology students saying and doing during their everyday lives? To find out, Matthias Mehl and James Pennebaker (2003) equipped 52 such students from the University of Texas with belt-worn Electronically Activated Recorders (EARs). For up to four days, the EAR 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 minds of their University of Nevada, Las Vegas, students, Christopher Heavey and Russell Hurlburt (2008) gave them beepers. 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 3.1** on the next page).
- *Culture, climate, and the pace of life.* Naturalistic observation also enabled Robert Levine and Ara Norenzayan (1999) to compare the pace of life in 31 countries. (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).

Naturalistic observation offers interesting snapshots of everyday life, but it does so without controlling for all the factors 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.

RETRIEVAL PRACTICE

An EAR for naturalistic observation Psychologists Matthias Mehl and James Pennebaker have used Electronically Activated Recorders (EARs) to sample naturally occurring slices of daily life.

- What are the advantages and disadvantages of naturalistic observation, such as Mehl and Pennebaker used in this study?

ANSWER: The study by Mehl and Pennebaker carefully observed and recorded naturally occurring behaviors—outside the artificiality of the lab. Because this was not an experiment, the study did not reveal the factors that influence everyday speech.

Courtesy of Matthias Mehl

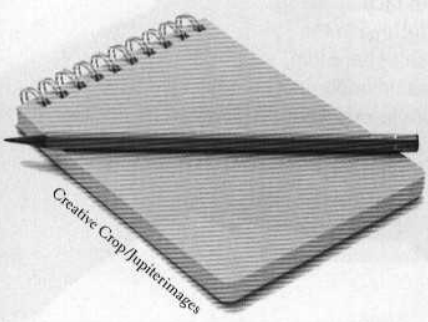

Creative Crop/Jupiterimages

TABLE 3.1

A Penny for Your 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.

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.

population all the cases 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.

The Survey

A **survey** looks at many cases in less depth. A survey asks people to report their behavior or opinions. Questions about everything from sexual practices 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 Even subtle changes in the order or wording of questions 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.

This Modern World by Tom Tomorrow © 1991.

Random Sampling In everyday thinking, we tend to generalize from samples we observe, especially vivid cases. Given (a) a statistical summary of a professor's student evaluations and (b) the vivid comments of a biased sample—two irate students—an administrator's impression of the professor may be influenced as much by the two unhappy students as by the many favorable 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 college or university? 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, large samples—including call-in phone samples and TV or website polls—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.

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).

RETRIEVAL PRACTICE

- What is *sampling bias,* and how do researchers avoid it?

ANSWER: *Random sampling* helps researchers avoid sampling bias, which occurs when a survey group is not representative of the population being studied.

Correlation

3-3 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 vocational 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 3.2** 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.

correlation a measure of the extent to which two factors vary together, and thus of how well either factor predicts the other.

correlation coefficient a statistical index of the relationship between two things (from −1 to +1).

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).

FIGURE 3.2
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).

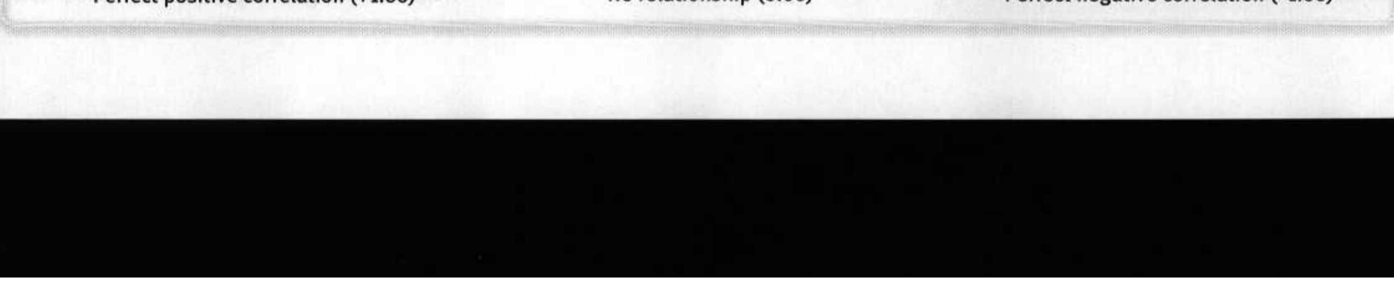


TABLE 3.2

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

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 earlier in this module 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 3.2**), can you tell whether the correlation between height and reactive temperament is positive, negative, or close to zero?

Comparing the columns in Table 3.2, 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 3.3**, 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 3.2, 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, gender,

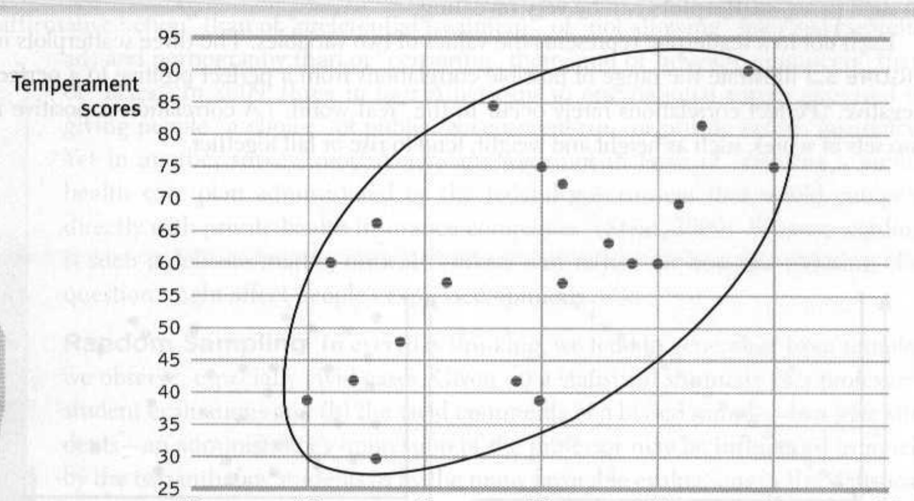


FIGURE 3.3
Scatterplot for height and reactive temperament This display of data from 20 imagined people (each represented by a data point) reveals an upward slope, indicating a positive correlation. The considerable scatter of the data indicates the correlation is much lower than +1.0.

and salary. But we often see no discrimination when the same information dribbles in, case by case (Twiss et al., 1989).

The point to remember: A correlation coefficient helps us see the world more clearly by revealing the extent to which two things relate.

RETRIEVAL PRACTICE

- Indicate whether each association is a positive correlation or a negative correlation.

1. The more children and youth use various media, the less happy they are with their lives (Kaiser, 2010). ________
2. The more sexual content teens see on TV, the more likely they are to have sex (Collins et al., 2004). ________
3. The longer children are 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

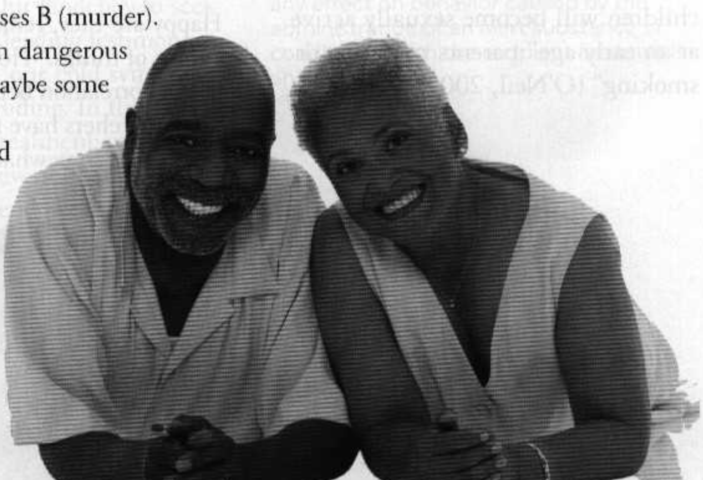
Correlations help us predict. The *New York Times* reports that U.S. counties with high gun ownership rates tend to have high murder rates (Luo, 2011). Gun ownership predicts homicide. What might explain this guns-homicide correlation?

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 factor 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 3.4** on the next page 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 factor—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."



© Nancy Brown/Getty Images

RETRIEVAL PRACTICE

Correlation and causation

- Length of marriage positively correlates with hair loss in men. Does this mean that marriage *causes* men to lose their hair (or that balding men make better husbands)?

ANSWER: In this case, as in many others, a third factor obviously explains the correlation: Golden anniversaries and baldness both accompany aging.

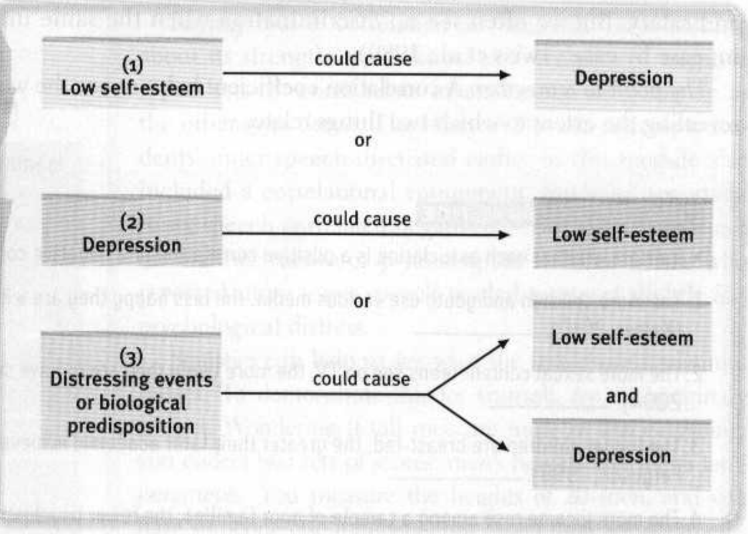


FIGURE 3.4
Three possible cause-effect relationships People low in self-esteem are more likely to report depression than are those high in self-esteem. One possible explanation of this negative correlation is that a bad self-image causes depressed feelings. But, as the diagram indicates, other cause-effect relationships are possible.

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).

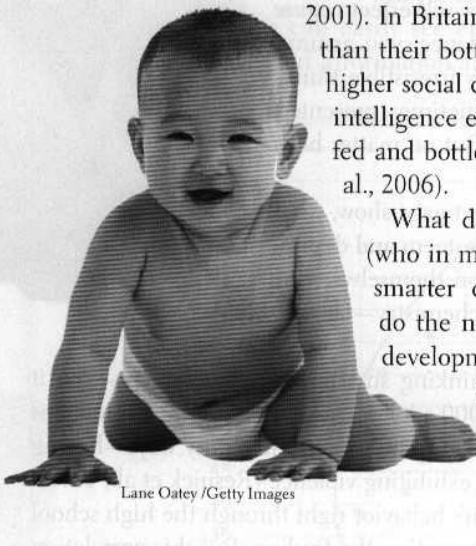
The point to remember (turn the volume up here): *Association does not prove causation.*[1] Correlation indicates the *possibility* of a cause-effect relationship *but does not prove such.* Remember this principle and you will be wiser as you read and hear news of scientific studies.

Experimentation

3-4 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 with cow's milk (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).


Lane Oatey /Getty Images

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

[1] Because 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).

factors by (1) *manipulating the factor(s) of interest* and (2) *holding constant ("controlling") other factors*. 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. 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 wigglers. 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 three 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 factors 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 factor (such as infant nutrition), then we infer the factor is having an effect.

The point to remember: Unlike correlational studies, which uncover naturally occurring relationships, an experiment manipulates a factor 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, bloodletting *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 factors.

And that is precisely how investigators evaluate new drug treatments and new methods of psychological therapy. 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.

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.

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 factors.

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

control group in an experiment, the group that is *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 those assigned to 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.

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.

RETRIEVAL PRACTICE

- What measure do researchers use to prevent the *placebo effect* from confusing their results?

ANSWER: Use of a *control group*, which is given a placebo and not the real treatment, allows results to be compared to the group that *is* given the real treatment, thus demonstrating whether the real treatment produces better results than *belief* in that treatment.

Independent and Dependent Variables

Here is an even more potent example: The drug Viagra was approved for use after 21 clinical trials. One trial was an experiment in which researchers randomly assigned 329 men with erectile dysfunction to either an experimental group (Viagra takers) or a control group (placebo takers). It was a double-blind procedure—neither the men nor the person giving them the pills knew what they were receiving. The result: At peak doses, 69 percent of Viagra-assisted attempts at intercourse were successful, compared with 22 percent for men receiving the placebo (Goldstein et al., 1998). Viagra worked.

This simple experiment manipulated just one factor: the drug dosage (none versus peak dose). We call this experimental factor the **independent variable** because we can vary it *independently* of other factors, such as the men's age, weight, and personality. 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 precise drug dosage and timing in this study) or measure the dependent variable (the questions that assessed the men's responses). These definitions provide a level of precision that enables others to repeat the study. (See **FIGURE 3.5** for the breast milk experiment's design.)

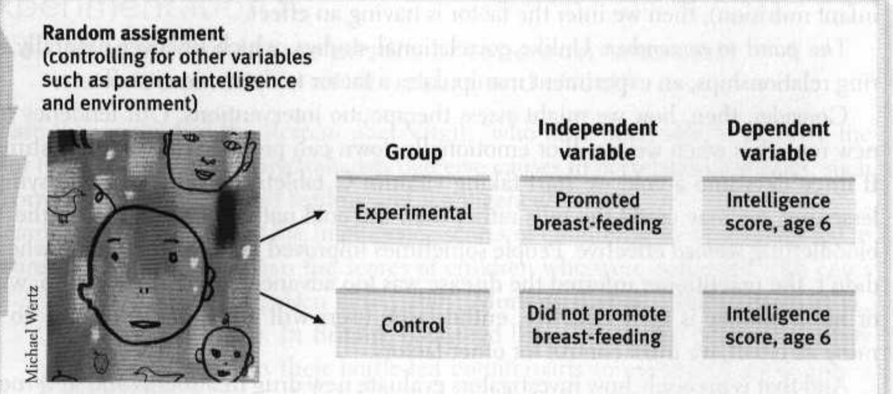


FIGURE 3.5
Experimentation To discern causation, psychologists may randomly assign some participants to an experimental group, others to a control group. Measuring the dependent variable (intelligence score in later childhood) will determine the effect of the independent variable (type of milk).

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, Adrian Carpusor and William Loges (2006) sent identically worded e-mail inquiries to 1115 Los Angeles–area landlords. 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. (Retrieval Practice: In this experiment, what was the independent variable? The dependent variable?[2])

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 anti-smoking campaigns? Do school sex-education programs reduce teen pregnancies? To answer such questions, we can experiment: If an intervention is welcomed

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.

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

TABLE 3.3

Comparing Research Methods

Research Method	Basic Purpose	How Conducted	What Is Manipulated	Weaknesses
Descriptive	To observe and record behavior	Do case studies, naturalistic observations, or surveys	Nothing	No control of variables; 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	Does not specify cause and effect
Experimental	To explore cause and effect	Manipulate one or more factors; use random assignment	The independent variable(s)	Sometimes not feasible; results may not generalize to other contexts; not ethical to manipulate certain variables

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* a *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 3.3** compares the features of psychology's research methods.

✓ RETRIEVAL PRACTICE

- By using *random assignment,* researchers are able to control for ________, which are other factors besides the independent variable(s) that may influence research results.

ANSWER: confounding variables

- Match the term on the left with the description on the right.

1. double-blind procedure
2. random sampling
3. random assignment

a. helps researchers generalize from a small set of survey responses to a larger population

b. helps minimize preexisting differences between experimental and control groups

c. controls for the placebo effect; neither researchers nor participants know who receives the real treatment

ANSWERS: 1. c, 2. a, 3. b

- 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?

ANSWER: To determine the drug's effectiveness, we must compare its effect on those randomly assigned to receive it (the experimental group) with the other half of the participants (control group), who receive a placebo. If we gave the drug to all 1000 participants, we would have no way of knowing if the drug is serving as a placebo or if it is actually medically effective.

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

Asked about the *ideal* wealth distribution in America, Democrats and Republicans were surprisingly similar. In the Democrats' ideal world, the richest 20 percent would possess 30 percent of the wealth. Republicans preferred a similar 35 percent (Norton & Ariely, 2011).

Statistical Reasoning in Everyday Life

In descriptive, correlational, and experimental research, statistics are tools that help us see and interpret what the unaided eye might miss. Sometimes the unaided eye misses badly. Researchers Michael Norton and Dan Ariely (2011) invited 5522 Americans to estimate the percent of wealth possessed by the richest 20 percent in their country. Their average person's guess—58 percent—"dramatically underestimated" the actual wealth inequality. (The wealthiest 20 percent possess 84 percent of the wealth.)

Accurate statistical understanding benefits everyone. To be an educated person today is to be able to apply simple statistical principles to everyday reasoning. One needn't memorize complicated formulas to think more clearly and critically about data.

Off-the-top-of-the-head estimates often misread reality and then mislead the public. Someone throws out a big, round number. Others echo it, and before long the big, round number becomes public misinformation. A few examples:

- *Ten percent of people are lesbians or gay men.* Or is it 2 to 3 percent, as suggested by various national surveys?
- *We ordinarily use but 10 percent of our brain.* Or is it closer to 100 percent?
- *The human brain has 100 billion nerve cells.* Or is it more like 40 billion, as suggested by extrapolation from sample counts?

The point to remember: Doubt big, round, undocumented numbers.

Statistical illiteracy also feeds needless health scares (Gigerenzer et al., 2008, 2009, 2010). In the 1990s, the British press reported a study showing that women taking a particular contraceptive pill had a 100 percent increased risk of blood clots that could produce strokes. This caused thousands of women to stop taking the pill, leading to a wave of unwanted pregnancies and an estimated 13,000 additional abortions (which also are associated with increased blood clot risk). And what did the study find? A 100 percent increased risk, indeed—but only from 1 in 7000 to 2 in 7000. Such false alarms underscore the need to teach statistical reasoning and to present statistical information more transparently.

©Patrick Hardin

"Figures can be misleading—so I've written a song which I think expresses the real story of the firm's performance this quarter."

Describing Data

3-5 How can we describe data with measures of central tendency and variation?

Once researchers have gathered their data, they must organize them in some meaningful way. One way to do this is to convert the data into a simple *bar graph*, as in **FIGURE 3.6**, which displays a distribution of different brands of trucks still on the road after a decade. When reading statistical graphs such as this, take care. It's easy to design a graph to make a difference look big (Figure 3.6a) or small (Figure 3.6b). The secret lies in how you label the vertical scale (the *y-axis*).

The point to remember: Think smart. When viewing figures in magazines and on television, read the scale labels and note their range.

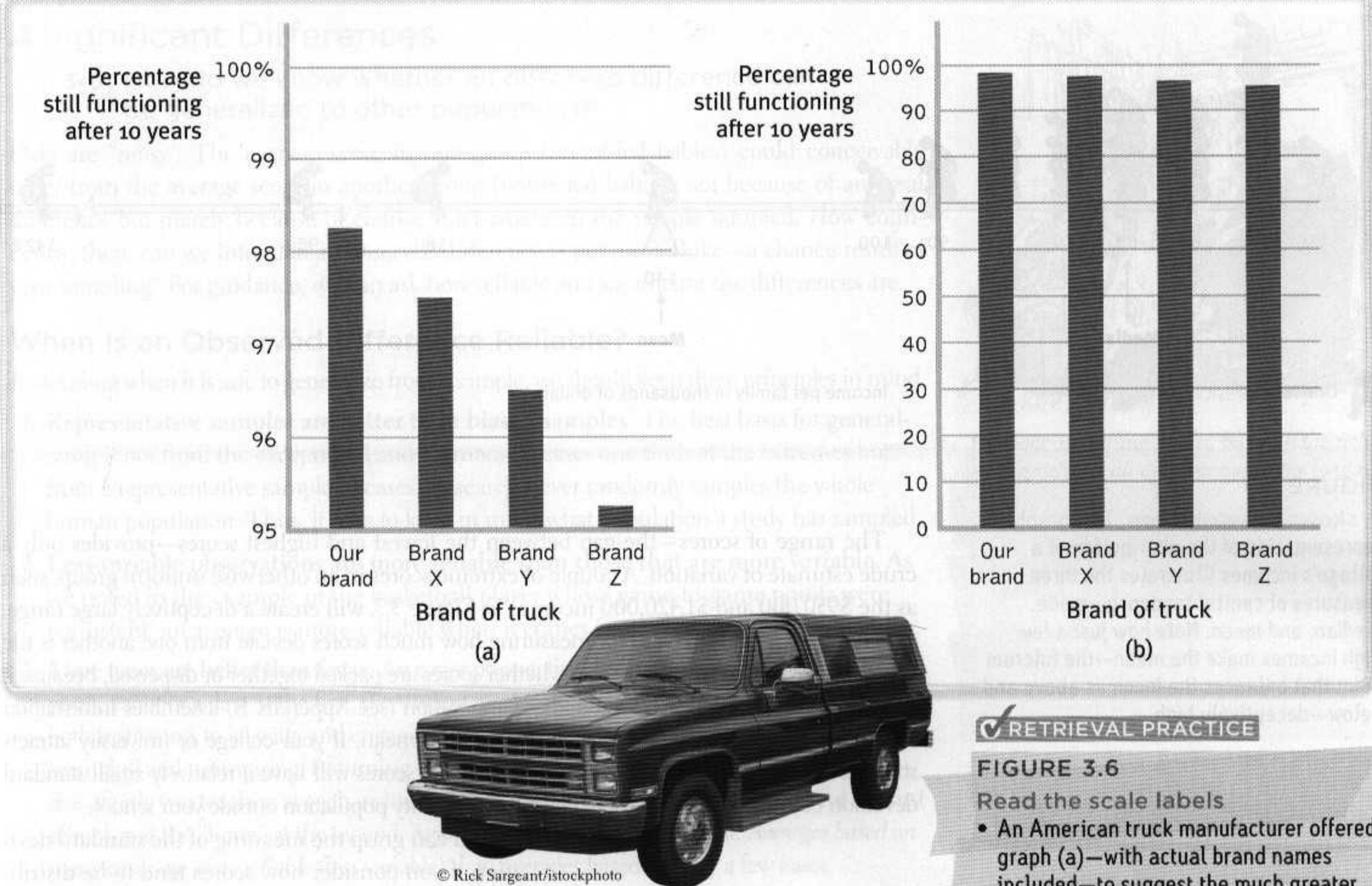
Measures of Central Tendency

The next step is to summarize the data using some *measure of central tendency,* a single score that represents a whole set of scores. The simplest measure is the **mode,** the most frequently occurring score or scores. The most commonly reported is the **mean,** or arithmetic average—the total sum of all the scores divided by the number of scores. On a divided highway, the median is the middle. So, too, with data: The **median** is the midpoint—the 50th percentile. If you arrange all the scores in order from the highest to the lowest, half will be above the median and half will be below it.

mode the most frequently occurring score(s) in a distribution.

mean the arithmetic average of a distribution, obtained by adding the scores and then dividing by the number of scores.

median the middle score in a distribution; half the scores are above it and half are below it.



RETRIEVAL PRACTICE

FIGURE 3.6

Read the scale labels

- **An American truck manufacturer offered graph (a)—with actual brand names included—to suggest the much greater durability of its trucks. What does graph (b) make clear about the varying durability, and how is this accomplished?**

ANSWER: Note how the y-axis of each graph is labeled. The range for the y-axis labels in graph (a) is only from 95 to 100. The range for graph (b) is from 0 to 100. All the trucks rank as 95% and up, so almost all of them are "still functioning," which graph (b) makes clear.

Measures of central tendency neatly summarize data. But consider what happens to the mean when a distribution is lopsided, or *skewed*, by a few way-out scores. With income data, for example, the mode, median, and mean often tell very different stories (**FIGURE 3.7** on the next page). This happens because the mean is biased by a few extreme scores. When Microsoft co-founder Bill Gates sits down in an intimate café, its average (mean) customer instantly becomes a billionaire. But the customer's median wealth remains unchanged. Understanding this, you can see how a British newspaper could accurately run the headline "Income for 62% Is Below Average" (Waterhouse, 1993). Because the bottom *half* of British income earners receive only a *quarter* of the national income cake, most British people, like most people everywhere, make less than the mean. Mean and median tell different true stories.

The point to remember: Always note which measure of central tendency is reported. If it is a mean, consider whether a few atypical scores could be distorting it.

The average person has one ovary and one testicle.

Measures of Variation

Knowing the value of an appropriate measure of central tendency can tell us a great deal. But the single number omits other information. It helps to know something about the amount of *variation* in the data—how similar or diverse the scores are. Averages derived from scores with low variability are more reliable than averages based on scores with high variability. Consider a basketball player who scored between 13 and 17 points in each of her first 10 games in a season. Knowing this, we would be more confident that she would score near 15 points in her next game than if her scores had varied from 5 to 25 points.

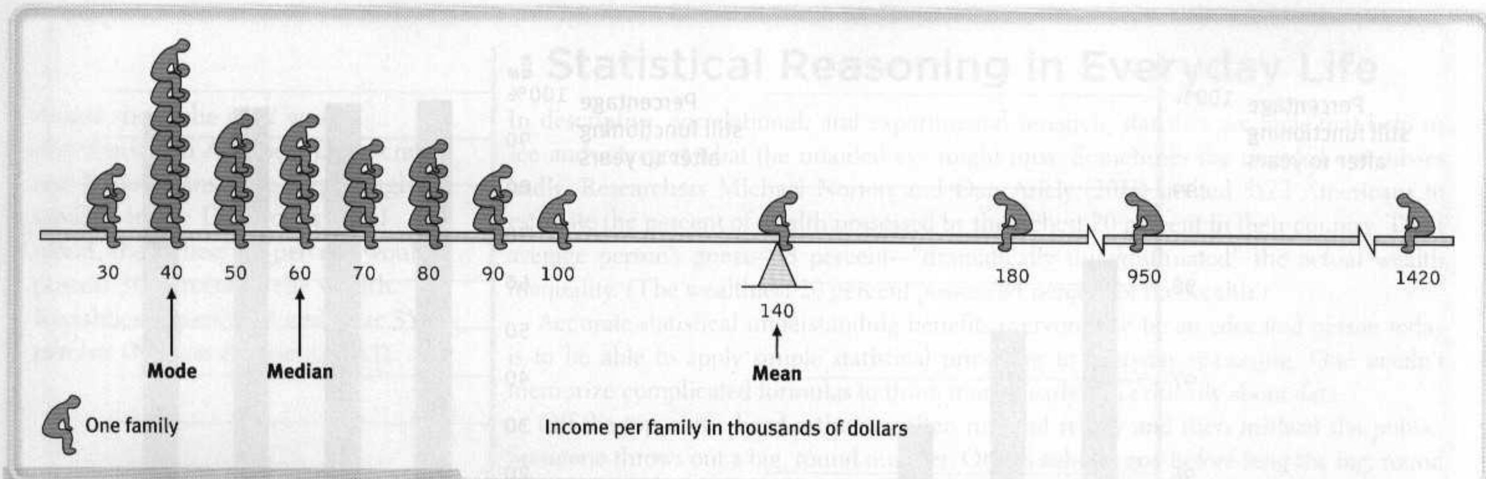


FIGURE 3.7
A skewed distribution This graphic representation of the distribution of a village's incomes illustrates the three measures of central tendency—mode, median, and mean. Note how just a few high incomes make the mean—the fulcrum point that balances the incomes above and below—deceptively high.

The **range** of scores—the gap between the lowest and highest scores—provides only a crude estimate of variation. A couple of extreme scores in an otherwise uniform group, such as the $950,000 and $1,420,000 incomes in Figure 3.7, will create a deceptively large range.

The more useful standard for measuring how much scores deviate from one another is the **standard deviation.** It better gauges whether scores are packed together or dispersed, because it uses information from each score. The computation (see Appendix B) assembles information about how much individual scores differ from the mean. If your college or university attracts students of a certain ability level, their intelligence scores will have a relatively small standard deviation compared with the more diverse community population outside your school.

You can grasp the meaning of the standard deviation if you consider how scores tend to be distributed in nature. Large numbers of data—heights, weights, intelligence scores, grades (though not incomes)—often form a symmetrical, *bell-shaped* distribution. Most cases fall near the mean, and fewer cases fall near either extreme. This bell-shaped distribution is so typical that we call the curve it forms the **normal curve.**

As **FIGURE 3.8** shows, a useful property of the normal curve is that roughly 68 percent of the cases fall within one standard deviation on either side of the mean. About 95 percent of cases fall within two standard deviations. Thus, about 68 percent of people taking an intelligence test will score within ±15 points of 100. About 95 percent will score within ±30 points.

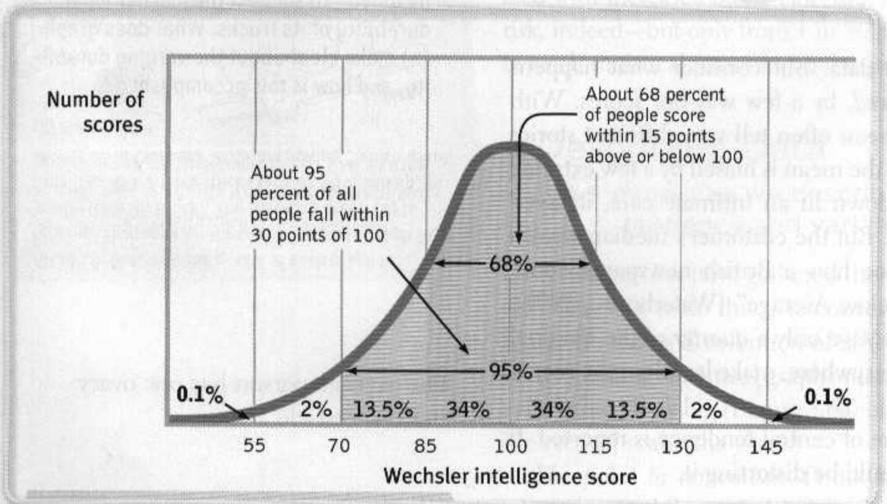


FIGURE 3.8
The normal curve Scores on aptitude tests tend to form a normal, or bell-shaped, curve. For example, the Wechsler Adult Intelligence Scale calls the average score 100.

RETRIEVAL PRACTICE

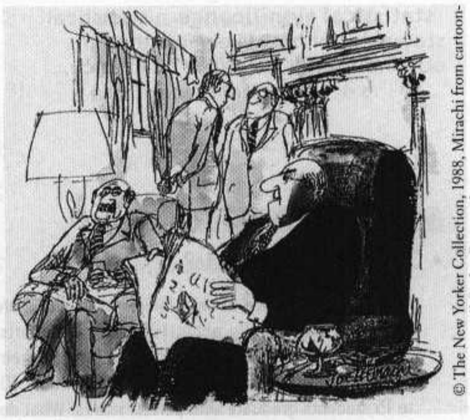
- The average of a distribution of scores is the ________. The score that shows up most often is the ________. The score right in the middle of a distribution (half the scores above it; half below) is the ________. We determine how much scores vary around the average in a way that includes information about the ________ of scores (difference between highest and lowest) by using the ________ formula.

ANSWERS: mean; mode; median; range; standard deviation

Significant Differences

3-6 How do we know whether an observed difference can be generalized to other populations?

Data are "noisy." The average score in one group (breast-fed babies) could conceivably differ from the average score in another group (bottle-fed babies) not because of any real difference but merely because of chance fluctuations in the people sampled. How confidently, then, can we infer that an observed difference is not just a fluke—a chance result of your sampling? For guidance, we can ask how reliable and significant the differences are.



"The poor are getting poorer, but with the rich getting richer it all averages out in the long run."

When Is an Observed Difference Reliable?

In deciding when it is safe to generalize from a sample, we should keep three principles in mind.

1. **Representative samples are better than biased samples.** The best basis for generalizing is not from the exceptional and memorable cases one finds at the extremes but from a representative sample of cases. Research never randomly samples the whole human population. Thus, it pays to keep in mind what population a study has sampled.
2. **Less-variable observations are more reliable than those that are more variable.** As we noted in the example of the basketball player whose game-to-game points were consistent, an average is more reliable when it comes from scores with low variability.
3. **More cases are better than fewer.** An eager prospective student visits two university campuses, each for a day. At the first, the student randomly attends two classes and discovers both instructors to be witty and engaging. At the next campus, the two sampled instructors seem dull and uninspiring. Returning home, the student (discounting the small sample size of only two teachers at each institution) tells friends about the "great teachers" at the first school, and the "bores" at the second. Again, we know it but we ignore it: *Averages based on many cases are more reliable* (less variable) than averages based on only a few cases.

The point to remember: Smart thinkers are not overly impressed by a few anecdotes. Generalizations based on a few unrepresentative cases are unreliable.

When Is a Difference Significant?

Perhaps you've compared men's and women's scores on a laboratory test of aggression, and found a gender difference. But individuals differ. How likely is it that the difference you found was just a fluke? Statistical testing can estimate the probability of the result occurring by chance.

Here is the underlying logic: When averages from two samples are each reliable measures of their respective populations (as when each is based on many observations that have small variability), then their *difference* is likely to be reliable as well. (Example: The less the variability in women's and in men's aggression scores, the more confidence we would have that any observed gender difference is reliable.) And when the difference between the sample averages is *large*, we have even more confidence that the difference between them reflects a real difference in their populations.

PEANUTS

range the difference between the highest and lowest scores in a distribution.

standard deviation a computed measure of how much scores vary around the mean score.

normal curve *(normal distribution)* a symmetrical, bell-shaped curve that describes the distribution of many types of data; most scores fall near the mean (about 68 percent fall within one standard deviation of it) and fewer and fewer near the extremes.

statistical significance a statistical statement of how likely it is that an obtained result occurred by chance.

In short, when sample averages are reliable, and when the difference between them is relatively large, we say the difference has **statistical significance.** This means that the observed difference is probably not due to chance variation between the samples.

In judging statistical significance, psychologists are conservative. They are like juries who must presume innocence until guilt is proven. For most psychologists, proof beyond a reasonable doubt means not making much of a finding unless the odds of its occurring by chance, if no real effect exists, are less than 5 percent.

When reading about research, you should remember that, given large enough or homogeneous enough samples, a difference between them may be "statistically significant" yet have little practical significance. For example, comparisons of intelligence test scores among hundreds of thousands of first-born and later-born individuals indicate a highly significant tendency for first-born individuals to have higher average scores than their later-born siblings (Kristensen & Bjerkedal, 2007; Zajonc & Markus, 1975). But because the scores differ by only one to three points, the difference has little practical importance.

The point to remember: Statistical significance indicates the *likelihood* that a result will happen by chance. But this does not say anything about the *importance* of the result.

RETRIEVAL PRACTICE

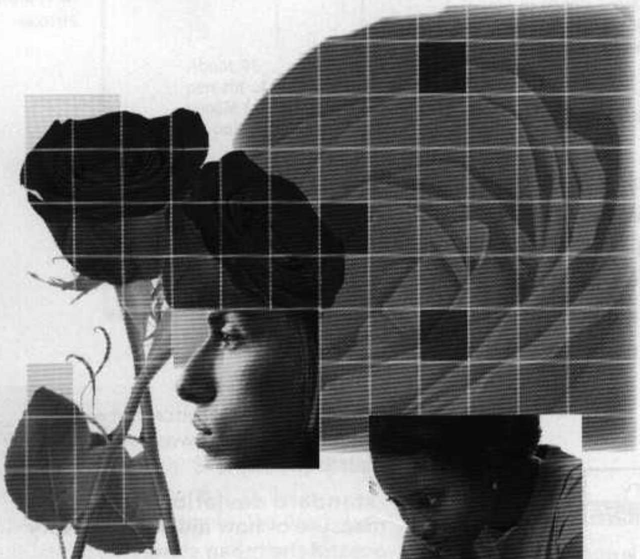
- Can you solve this puzzle?

The registrar's office at the University of Michigan has found that usually about 100 students in Arts and Sciences have perfect marks at the end of their first term at the University. However, only about 10 to 15 students graduate with perfect marks. What do you think is the most likely explanation for the fact that there are more perfect marks after one term than at graduation (Jepson et al., 1983)?

ANSWER: Averages based on fewer courses are more variable, which guarantees a greater number of extremely low and high marks at the end of the first term.

MODULE REVIEW

Research Strategies: How Psychologists Ask and Answer Questions



Learning Objectives

RETRIEVAL PRACTICE Take a moment to answer each of these Learning Objective Questions (repeated here from within the module). Then turn to Appendix B, Complete Module Reviews, to check your answers. Research suggests that trying to answer these questions on your own will improve your long-term retention (McDaniel et al., 2009).

3-1: How do theories advance psychological science?

3-2: How do psychologists use case studies, naturalistic observation, and surveys to observe and describe behavior, and why is random sampling important?

3-3: What are positive and negative correlations, and why do they enable prediction but not cause-effect explanation?

3-4: What are the characteristics of experimentation that make it possible to isolate cause and effect?

3-5: How can we describe data with measures of central tendency and variation?

3-6: How do we know whether an observed difference can be generalized to other populations?

Terms and Concepts to Remember

RETRIEVAL PRACTICE **Test yourself on these terms by trying to write down the definition before flipping back to the referenced page to check your answer.**

theory, p. 26
hypothesis, p. 26
operational definition, p. 26
replication, p. 26
case study, p. 28
naturalistic observation, p. 28
survey, p. 30
population, p. 31
random sample, p. 31
correlation, p. 31
correlation coefficient, p. 31
scatterplot, p. 31
experiment, p. 34
experimental group, p. 35
control group, p. 35
random assignment, p. 35
double-blind procedure, p. 35
placebo effect, p. 35
independent variable, p. 36
confounding variable, p. 36
dependent variable, p. 36
mode, p. 38
mean, p. 38
median, p. 38
range, p. 40
standard deviation, p. 40
normal curve, p. 40
statistical significance, p. 42

RETRIEVAL PRACTICE **Gain an advantage, and benefit from immediate feedback, with the interactive self-testing resources at www.worthpublishers.com/myers.**