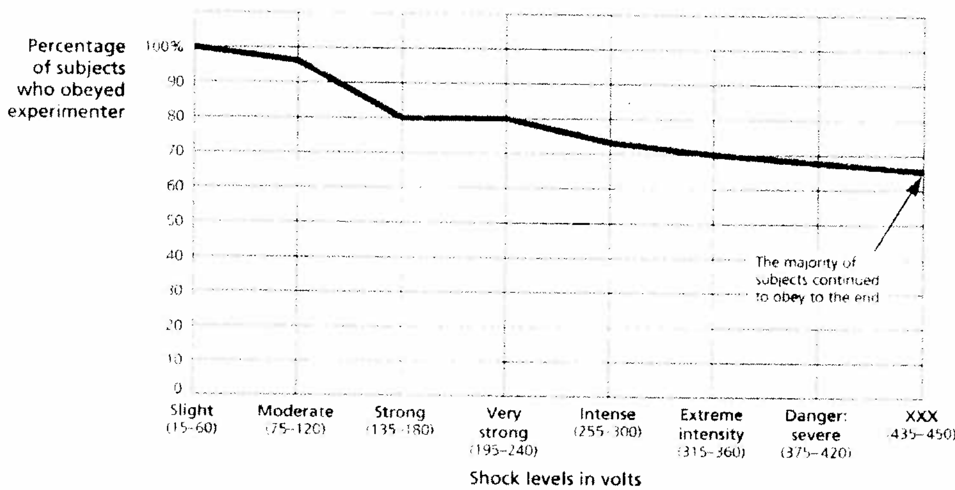
Ch 1: Thinking Critically with Psychological Science



T-104

The Need for Psychological Science
The Limits of Intuition and Common Sense
The Scientific Attitude
The Scientific Method

Description
The Case Study
The Survey
Naturalistic Observation

Correlation
Correlation and Causation
Illusory Correlations
Perceiving Order in Random Events
Thinking Critically About: Hot and Cold Streaks in Basketball and the Stock Market

Experimentation
Evaluating Therapies
Can Subliminal Tapes Improve Your Life?

Statistical Reasoning
Describing Data
Measures of Central Tendency
Measures of Variation
Making Inferences

Frequently Asked Questions About Psychology
Thinking Critically About: The Death Penalty—When Beliefs Collide with Psychological Science

What good fortune for those in power that people do not think.

Adolf Hitler, 1889–1945

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

Others, intrigued by claims of psychological truth, wonder: Do mothers and infants bond in the first hours after birth? Should we trust childhood sexual abuse memories that get "recovered" in adulthood—and prosecute the alleged predators? Are first-born children more driven to achieve? Does handwriting offer clues to personality? 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

Preview: As we familiarize ourselves with psychological science's strategies and incorporate its underlying principles into our daily thinking, our thinking becomes smarter. Two phenomena—hindsight bias and judgmental overconfidence—illustrate why we cannot rely solely on intuition and common sense. The critical inquiry that flows from a scientific approach—undergirded by curiosity, skepticism, and humility—helps winnow sense from nonsense. Psychologists, like all scientists, use the scientific method to construct theories that organize observations and imply testable hypotheses.

The Limits of Intuition and Common Sense

In sifting reality from illusion, won't intuition and plain common sense suffice for everyday life? Some say psychology merely documents what people already know and dresses it in jargon: "So what else is new—you get paid for using fancy methods to prove what my grandmother knew?"

Bob Daemmrich/The Image Works

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 from their ignorance about rejected applicants who succeeded elsewhere.

"... naked intellect is an extraordinarily inaccurate instrument."

Madeleine L'Engle, *A Wind in the Door*, 1973

Others scorn a scientific approach because of their faith in human intuition. Advocates of "intuitive management" urge us to distrust statistical predictors and tune into our hunches when hiring, firing, and investing. Like *Star Wars'* Luke Skywalker, should we trust the force within?

Actually, our intuition can lead us astray. Consider two examples:

- Imagine (or ask someone to imagine) folding a sheet of paper on itself 100 times. Roughly how thick would it then be?
- Given our year with 365 days, a group needs 366 people to ensure that at least two people share the same birthday; how big should a group be to have a 50 percent chance of finding a birthday match? (See page 22 for the answers.)

Our notions of common sense similarly err. We're all after-the-fact pundits, presuming we could have foreseen what we know happened.

Did We Know It All Along? The Hindsight Bias

How easy it is to seem astute when drawing the bull's eye after the arrow has struck. *After* each stock market downswing—after the bursting of the dot-com bubble, for example—investment gurus say "the market was obviously overdue for a correction." After the first World Trade Center tower was hit on 9/11, people in the second tower *should* have immediately evacuated (it became obvious only later that it was not an accident). And after physicians receive case information *plus* an autopsy report, they find the cause of death to be self-evident—something they easily could have foreseen, knowing the symptoms. But *before* the arrow strikes, the stock market drops, the terrorists attack, and death occurs, these results are anything but obvious. Causes of death, for example, are not so clear to doctors told the same symptoms without the autopsy report (Dawson & others, 1988). Finding out that something has happened makes it seem inevitable. Psychologists call this 20/20 hindsight vision the **hindsight bias**, also known as the *I-knew-it-all-along phenomenon*.

"... is lived forwards, but understood backwards."

Philosopher Søren Kierkegaard, 1813–1855

Psychologists Paul Slovic and Baruch Fischhoff (1977) and Gordon Wood (1979) have shown how unanticipated scientific results and historical happenings can indeed *seem* like common sense. This phenomenon is easy to demonstrate: Give half the members of a group some purported psychological finding, and 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 regard this true finding as unsurprising.

Hindsight bias

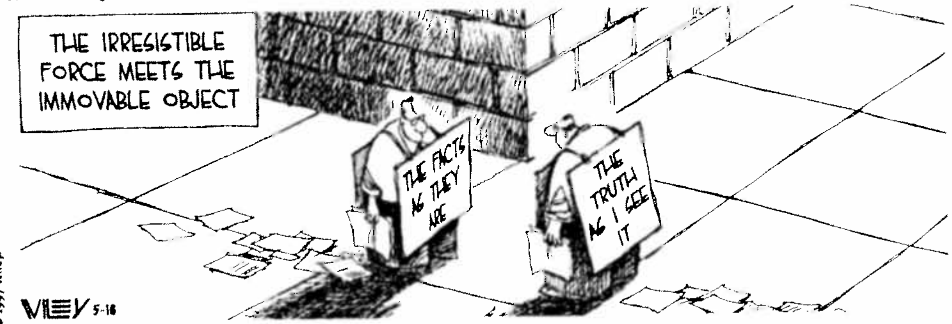
After the horror of 9/11 it seemed obvious that the American intelligence services should have taken advance warnings more seriously, that airport security should have anticipated box-cutter–wielding terrorists, that occupants of the second tower should have known to play it safe and leave. With 20/20 hindsight, everything seems obvious.

Tim Boyle/Getty

Tell the second group just the opposite—that "psychologists have found that separation strengthens romantic attraction. As the saying goes, 'Absence makes the heart grow fonder.'" People given this result can also easily explain it, and they overwhelmingly see it as unsurprising common sense. Obviously, when both a supposed finding and its opposite seem like common sense, there is a problem.

Consider hindsight bias in a police context. When viewing a police lineup, eyewitnesses often feel uncertain: "I'm not sure. . . . I think it's one of those two, maybe the shorter guy on the left." If told they have chosen the actual suspect, they may later, when testifying in court, recall identifying the person easily. "There was no maybe about it," recalled one formerly uncertain eyewitness. Gary Wells and Amy Bradfield (1998) demonstrated hindsight bias after showing 352 Iowa State University students a grainy security video of a man entering a store just before murdering a security guard. When shown a photospread from the actual case, minus the actual gunman's photo, all 352 students made a false identification. Those told "Good. You identified the actual sus-

NON SEQUITUR



pect" were now understandably more confident in their identification—but were also *four times* more likely to recall having felt great confidence when earlier making their identification. Were they aware of how the experimenter's off-hand comment had influenced their recollections? No, most denied being influenced by the casual feedback.

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.***)

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 sometimes can be misleading—*not* because common sense is usually wrong, but because it is after the fact. Common sense describes what has happened more easily than it predicts what will happen.

"Anything seems commonplace, once explained."

Dr. Watson to Sherlock Holmes

Nevertheless, Grandmother 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 (we have what Chapter 12 calls a deep "need to belong").

But some research findings *do* jolt our common sense. Sometimes Grandmother's intuition has it wrong. Informed by countless casual observations, our intuition may tell us that familiarity breeds contempt, that dreams predict the future, and that emotional reactions coincide with menstrual phase. As we will see in later chapters, the available evidence suggests that these commonsense ideas are wrong, wrong, and wrong. Throughout this book we will see how research has both inspired and overturned popular ideas—about aging, about sleep and dreams, about personality. And we will also see how it has surprised us with discoveries about how the brain's chemical messengers control our moods and memories, about animal abilities, and about the effects of stress on our capacity to fight disease.

Overconfidence

Our everyday thinking is limited not only by our after-the-fact common sense but also by our human tendency to be overly confident. As Chapter 10 explains, we 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

Reflect for a moment: About how many seconds do you think it would have taken you to unscramble each of these?

Fun anagram solutions from Wordsmith.org:
Elvis = lives
Dormitory = dirty room
The Morse code = here come the dots
Slot machines = cash lost in 'em

[1]Boston is south of Paris.

Answers to questions on page 20: Given a -millimeter-thick sheet, the thickness after 100 folds would be 800 trillion times the distance between the Earth and the Sun (Gilovich, 1991). Only 23 people are needed to give better than even odds of any two people having the same birthday.

"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

"The telephone may be appropriate for our American cousins, but not here, because we have an adequate supply of messenger boys."

British expert group evaluating the invention of the telephone

"They couldn't hit an elephant at this dist—."

General John Sedgwick's last words, uttered during a U.S. Civil War battle, 1864

"The scientist . . . must be free to ask any question, to doubt any assertion, to seek for any evidence, to correct any errors."

Physicist J. Robert Oppenheimer, *Life*, October 10, 1949

Once people know the target word, hindsight makes it seem obvious—so much so that they become overconfident. They think they would have seen the solution in only 10 seconds or so, when in reality the average subject spent 3 minutes, as you also might, given a similar anagram without the solution: OCHSA (see page 24 to check your answer).

Are we any better at predicting our social behavior? To find out, Robert Vallone and his associates (1990) had students predict at the beginning of the school year whether they would drop a course, vote in an upcoming election, call their parents more than twice a month, and so forth. On average, the students felt 84 percent confident in making these self-predictions. Later quizzes about their actual behavior showed their predictions were correct only 71 percent of the time. Even when they were 100 percent sure of themselves, their self-predictions erred 15 percent of the time.

It's not just collegians. For a dozen years, Ohio State University psychologist Philip Tetlock (1998) collected experts' predictions of political, economic, and military situations. In the late 1980s, for example, he invited expert professors, think-tank analysts, government experts, and journalists to project the governance of the Soviet Union or of South Africa five years later, and to rate how confident they felt. Others did the same for the future of Canada in 1992. After the five years had elapsed (and communism had collapsed in the Soviet Union, South Africa had become a multiracial democracy, and the Canadian constitution continued), Tetlock invited the experts to recall and reflect on their predictions—which, as in laboratory studies, were far more confident than correct. Experts who had felt more than 80 percent confident were right less than 40 percent of the time.

Despite their lackluster predictions, those who erred were nearly as likely as those who got it right to convince themselves that their initial analysis was *still basically right.* I was "almost right," many of them felt. "The hardliners almost succeeded in their coup attempt against Gorbachev." "The Quebecois separatists almost won the secessionist referendum." "But for the coincidence of de Klerk and Mandela, the transition to black majority rule in South Africa would have been a lot bloodier." The overconfidence of political experts (and stock market forecasters and sports prognosticators) is therefore hard to dislodge, no matter what the outcome.

The Scientific Attitude

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. To answer them in any way requires a leap of faith. With many other ideas (Can some people demonstrate ESP?), the proof is in the pudding. No matter how sensible or crazy-sounding an idea, the hard-headed question is, Does it work? When put to the test, can its predictions be confirmed?

This scientific approach has a long history. As ancient a figure as Moses used such an approach. How do you evaluate a self-proclaimed prophet? His answer: Put the prophet to the test. If the predicted event "does not take place or prove true," then so much the worse for the prophet (*Deuteronomy* 18:22). Magician James Randi uses Moses' 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 tells me that no aura-seer has yet agreed to take this simple test.

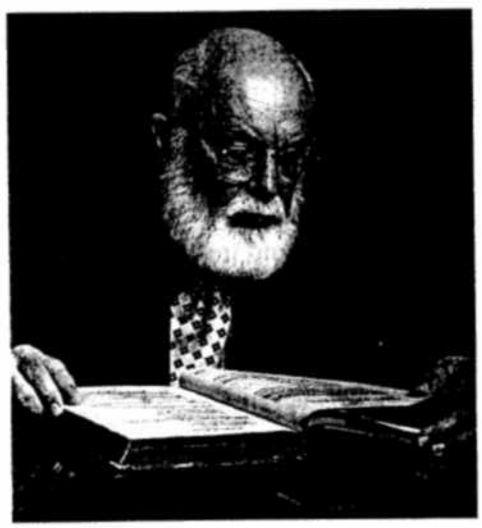
When subjected to such scrutiny, crazy-sounding ideas sometimes find support. During the 1700s, scientists scoffed at the notion that meteorites had extraterrestrial origins. When two Yale scientists dared to deviate from the conventional opinion, Thomas Jefferson jeered, "Gentlemen, I would rather believe that those two Yankee Professors would lie than to believe that stones fell from heaven." Sometimes scientific inquiry refutes skeptics.

More often, science relegates crazy-sounding ideas to the mountain of forgotten 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.

As scientists, psychologists, too, approach the world of behavior with a *curious skepticism*. They persistently ask two questions: What do you mean? How do you know? In business, the motto is "show me the money." In science, it is "show me the evidence."

Consider some familiar claims: that parental behaviors determine their children's sexual orientation; that lie detectors tell the truth; that astrologers can analyze your character and predict your future based on the position of the planets at your birth. As you will see in the chapters that follow, putting such claims to the test has led most psychologists to doubt them. In the arena of competing ideas, skeptical testing can reveal which ones best match the facts. "To believe with certainty," says a Polish proverb, "we must begin by doubting."

Putting a scientific attitude into practice requires not only skepticism but also *humility*, because we may have to reject our own ideas. In the last analysis, what matters is not my opinion or yours, but the truths nature reveals in response to our questioning. If people don't behave as our ideas predict, then so much the worse for our ideas. This is the humble attitude expressed in one of psychology's early mottos: "The rat is always right."



Rob Kinmonth

The amazing Randi
The magician James Randi exemplifies skepticism. He has tested and debunked a variety of psychic phenomena.

Historians of science tell us that these attitudes of curiosity, skepticism, and humility helped make modern science possible. Many of its founders were people whose religious convictions made them humble before nature and skeptical of mere human authority (Hooykaas, 1972; Merton, 1938). Of course, scientists, like anyone else, can have big egos and may cling to their preconceptions. We all view nature through the spectacles of our preconceived ideas. Yet the ideal that unifies psychologists with all scientists is the curious, skeptical, humble scrutiny of competing ideas. As a community, scientists check and recheck one another's findings and conclusions.

This 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's 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? Carried to an extreme, healthy skepticism can degenerate into a negative cynicism that scorns any unproven idea.

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

"A skeptic is one who is willing to question any truth claim, asking for clarity in definition, consistency in logic, and adequacy of evidence."

Philosopher Paul Kurtz, *The Skeptical Inquirer*, 1994

"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

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

hypothesis a testable prediction, often implied by a theory.

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

"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

Better to have a critical attitude that produces humility—an awareness of our own vulnerability to error and an openness to surprises and new perspectives.

Has psychology's critical inquiry been open to surprising findings? The answer, as ensuing chapters illustrate, is plainly yes. Believe it or not . . .

- massive losses of brain tissue early in life may have minimal long-term effects (see page 85).
- within days, newborns can recognize their mother's odor and voice (see page 138).
- brain damage can leave a person able to learn new skills, yet be unaware of such (see pages 86–88).
- diverse groups—men and women, old and young, rich and working class, those with disabilities and without—report roughly comparable levels of personal happiness (see pages 523–525).
- electroconvulsive ("shock") therapy is often a very effective treatment for severe depression (see pages 689–690).

And has critical inquiry convincingly debunked popular presumptions? The answer, as ensuing chapters also illustrate, is again yes. The evidence indicates that . . .

- as part of their passage to middle adulthood, men in their early forties do *not* typically undergo a midlife crisis (see pages 182–183) and most mothers are *not* depressed for a time after their children grow up and leave home (see page 185).
- sleepwalkers are *not* acting out their dreams, and sleeptalkers are *not* verbalizing their dreams (see Chapter 7).
- our past experiences are *not* all recorded verbatim in our brains; with brain stimulation or hypnosis, one *cannot* simply "play the tape" and relive long-buried or repressed memories (see Chapter 9).
- most people do *not* suffer from unrealistically low self-esteem (see page 609).
- opposites do *not* generally attract (see page 729).

The Scientific Method

Psychologists arm their scientific attitude with the *scientific method*: They make observations, form theories, and then refine their theories in the light of new observations. In everyday conversation, we tend to use *theory* to mean "mere hunch." In science, *theory* is linked with observation. A scientific **theory** *explains* through an integrated set of principles that *organizes* and *predicts* behaviors or events. By organizing isolated facts, a theory simplifies things. There are too many facts about behavior to remember them all. By linking facts and bridging them to deeper principles, a theory offers a useful summary. When we connect the observed dots, we may discover a coherent picture.

A good theory of depression, for example, helps us organize countless observations concerning depression into a much shorter list of principles. Say we observe over and over that people with depression describe their past, present, and future in gloomy terms. We might therefore theorize that low self-esteem contributes to depression. So far so good: Our self-esteem principle neatly summarizes a long list of facts about people with depression.

Yet no matter how reasonable a theory may sound—and low self-esteem seems a reasonable explanation of depression—we must put it to the test. A good theory doesn't just sound appealing. It must imply testable predictions, called **hypotheses**. By enabling us to test and reject or revise the theory, such predictions give direction to research. They specify what results would support the theory and what results would disconfirm it. To test our self-esteem theory of depression, we might give people a test of self-esteem on which they respond to statements such as "I have good ideas." Then we could see whether, as we hypothesized, people who report poorer self-images also score higher on a depression scale (**FIGURE 1.1**).

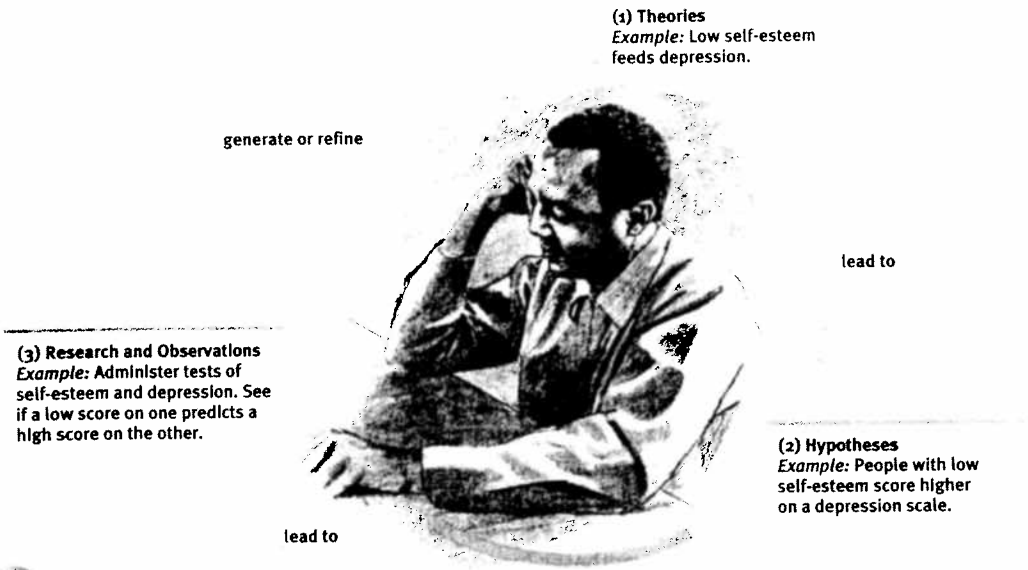


FIGURE 1.1
The scientific method
A self-correcting process for asking questions and observing nature's answer.

In testing our theory, we should be aware that it can bias subjective observations. Having theorized that depression springs from low self-esteem, we may see what we expect. We may perceive depressed people's neutral comments as self-disparaging.

As a check on their biases, psychologists report their research precisely enough—with clear **operational definitions** of concepts—to allow others to **replicate** (repeat) their observations. If other researchers re-create a study with different participants and materials and get similar results, then our confidence in the finding's reliability grows. The first study of hindsight bias 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) effectively *organizes* a range of self-reports and observations and (2) implies clear *predictions* that anyone can use to check the theory or to derive practical applications. (If we boost people's self-esteem, will their depression lift?) Eventually, our research will probably lead to a revised theory (such as the one on pages 642–643) that better organizes and predicts what we know about depression.

...and theories explain by organizing and linking observed facts. ...implying hypotheses that offer testable predictions and, sometimes, practical applications.

Our research strategies include descriptive, correlational, and experimental methods. We test hypotheses and refine our theories by making *observations* that *describe* behavior, detecting *correlations* that help *predict* behavior, and doing *experiments* that help *explain* behavior. To think critically about popular psychology claims, we need to recognize these designs and to know what conclusions they allow.

REVIEW AND REFLECT

The Need for Psychological Science

The Limits of Intuition and Common Sense

Although in some ways we outsmart the smartest computers, our intuition often goes awry. To err is human. Without scientific inquiry and critical thinking we readily succumb to *hindsight bias*, also called the I-knew-it-all-along phenomenon. Learning the outcome of a study (or of an everyday happening) can make it seem like obvious common sense. We also are routinely *overconfident* of our judgments, thanks partly to our bias to seek information that confirms them. Such biases lead us to overestimate our unaided intuition.

Enter psychological science. Science, with its procedures for gathering and sifting evidence, restrains error. Although limited by the testable questions it can address, a scientific approach helps us sift reality from illusion, taking us beyond the limits of our intuition and common sense.

The Scientific Attitude

Scientific inquiry begins with an attitude—a *curious* eagerness to *skeptically* scrutinize competing ideas and an open-minded *humility* before nature. Putting ideas, even crazy-sounding ideas, to the test helps us winnow sense from nonsense. The curiosity that drives us to test ideas and to expose their underlying assumptions carries into everyday life as *critical thinking*.

The Scientific Method

Research stimulates the construction of *theories*, which organize *observations* and imply predictive *hypotheses*. These hypotheses (predictions) are then tested to validate and refine the theory and to suggest practical applications.

CHECK YOURSELF: What is the scientific attitude and why is it important for critical thinking?

ASK YOURSELF: How might the scientific method help us understand the roots of terrorism?

Answers to the Check Yourself questions can be found in the yellow appendix at the end of the book.

Description

Preview: Psychologists describe behavior using case studies, surveys, and naturalistic observations.

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

The Case Study

"'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

Among the oldest research methods is the **case study**, in which psychologists study one individual in great 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. Sigmund Freud constructed his theory of personality from a handful of case studies. Developmental psychologist Jean Piaget taught us about children's thinking after carefully observing and questioning but a few children. Studies of only a few chimpanzees have revealed their capacity for understanding and language. Intensive case studies are sometimes very revealing.

Although case studies can also suggest hypotheses for further study, they sometimes mislead us: An individual may be 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 case ("Well, I have an uncle who smoked two packs a day and lived to be 89"). Anecdotal cases—dramatic stories, personal experiences, even psychological case examples—have a way of overwhelming general truths. Highly publicized school shootings can raise alarm about school violence even while school violence rates are subsiding. Numbers can be numbing (in one study of 1300 dream reports concerning a kidnapped child, only 5 percent correctly envisioned the child as dead—see page 259). Anecdotes are often more startling. ("But I know a man who dreamed his sister was in a car accident, and two days later she was badly injured.")

Susan Kuklin/Photo Researchers

The case of the conversational chimpanzee
In intensive case studies of chimpanzees, psychologists have explored the intriguing question of whether language is uniquely human. Here Nim Chimpsky signs *hug* as his trainer, psychologist Herbert Terrace, shows him the puppet Ernie. But is Nim really capable of using language? We'll explore that issue in Chapter 10.

After 2-year-old James Bulger was abducted from a Liverpool shopping mall and bludgeoned to death, and after a murderous rampage at Colorado's Columbine High School, children and parents in both countries became noticeably "scared" (as a *Newsweek* cover story put it)—much more scared than they were of car accidents or cancer, which cause child deaths many hundreds of times more often than kidnappings and school assassinations do. The brutal kidnapping was impressed on people's memories, and people intuitively judge various risks based on how easily they remember examples of them. As psychologist Gordon Allport said, "Given a thimbleful of [dramatic] facts we rush to make generalizations as large as a tub."

So, 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 methods.

The Survey

The **survey** method, commonly used in both descriptive and correlational studies, 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 get put to the public. It's hard to think of a significant question that survey researchers have not asked. For example, Harris and Gallup polls have revealed that 72 percent of Americans think there is too much TV violence, 84 percent favor equal job opportunities for homosexual people, 89 percent say they face high stress, 95 percent believe in God, and 96 percent would like to change something about their appearance.

Wording Effects

Asking questions is tricky. Even subtle changes in the order or wording of questions can have major effects. Should cigarette ads or pornography be allowed on television? People are much more likely to approve "not allowing" such things than "forbidding" or "censoring" them. In a national survey, only 27 percent of Americans approved of "government censorship" of media sex and violence, though 66 percent approved of "more restrictions on what is shown on television" (Lacayo, 1995). People are similarly much more approving of "aid to the needy" than of "welfare," of "affirmative action" than of "preferential treatment," and of "revenue enhancers" than of "taxes." Because wording questions is such a delicate matter, critical thinkers will reflect on how the phrasing of a question might have affected the opinions respondents expressed.

- **case study** an observation technique in which one person is studied in depth in the hope of revealing universal principles.
- **survey** a technique for ascertaining the self-reported attitudes or behaviors of people, usually by questioning a representative, random sample of them.

false consensus effect the tendency to overestimate the extent to which others share our beliefs and behaviors.

population all the cases in a group, from which samples may be drawn for a study. (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.

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

Sampling

In our everyday experience we spend most of our time with a biased sample of people—mostly those who share our attitudes and habits. Thus, when we wonder how many people hold a particular belief, those who think as we do come to mind most readily. This tendency to overestimate others' agreement with us is the **false consensus effect** (Ross & others, 1977). Vegetarians will think more people are vegetarians than will meat-eaters, and conservatives will perceive more support for conservative views than will liberals. To restrain this bias, researchers aim to gather a representative sample of people.

Most surveys sample a target group. If you wished to survey the students at your college or university you could question them all, but probably there are too many. Instead, you could survey a representative sample of the total student **population**—the whole group you wanted to study and describe. How could you make your sample representative of this population? Typically by making it a **random sample**, one in which every person in the entire group has an equal chance of participating.

To sample the students at your institution randomly, you would *not* send them all a questionnaire. (The conscientious people who return it would not be a random sample.) Rather, you would aim for a representative sample by, say, using a table of random numbers to pick participants from a student listing and then making sure you involve as many as possible. Large representative samples are better than small ones, but a small representative sample of 100 is better than an unrepresentative sample of 500.

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).

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

You can forecast the weather by taking an arbitrary sample—by looking at the clouds and holding your finger in the wind—or you can look at weather maps based on comprehensive reporting. You can describe human experience using common sense, dramatic anecdotes, personal experience, and arbitrary samples. But for an accurate picture of the experiences and attitudes of a whole population, there's only one game in town—the representative sample.

We can extend this point to everyday thinking, as we generalize from samples we observe. We meet a few students and attend a few classes during a visit to a college and infer from those instances how friendly the campus is and how good the teaching is. We observe the weather during a three-day visit to Copenhagen and then tell our friends about the climate there.

This Modern World by Tom Tomorrow © 1991.

Overgeneralizing from such select samples is tempting, especially when they are vivid cases. Given (a) a statistical summary of a professor's student evaluations and (b) the vivid comments of 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. Standing in the checkout line at the supermarket, George sees the woman in front of him pay with government-provided food stamps and then watches with dismay as she drives away in a fancy car. In both situations, the temptation 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 of cases.

The random-sampling principle also works in national surveys. Imagine that you had a giant barrel containing 60 million white beans mixed with 40 million red beans. A scoop that randomly sampled 1500 of them would contain about 60 percent white and 40 percent red beans, give or take 2 or 3 percent. Sampling voters in a national election survey is like sampling the beans;

1500 randomly sampled people, drawn from all areas of a country, provide a remarkably accurate snapshot of the opinions of a nation (**FIGURE 1.2**).

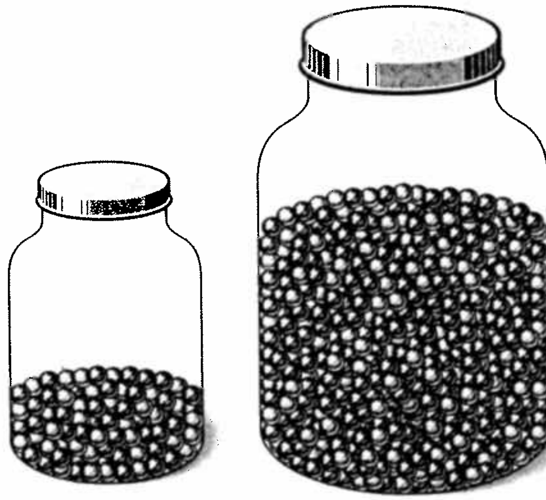


FIGURE 1.2
World in a jar
If marbles of two colors are mixed well in the large jar, the fastest way to know their ratio is to blindly transfer a few into a smaller jar and count them. This approach is called random sampling.

Because gathering a random sample can be a huge task, some don't make the effort. Shere Hite's book *Women and Love* reported survey findings based on only a 4.5 percent response rate from mailings to an unrepresentative sample of 100,000 women. The response was doubly unrepresentative because not only did she have a modest, self-selected return, but the women initially contacted were members of women's organizations. Nonetheless, "It's 4500 people. That's enough for me," reported Hite. And it was apparently enough for *Time* magazine, which made a cover story of her findings—that 70 percent of women married five or more years were having affairs, and that 95 percent of women felt emotionally harassed by the men they love (Wallis, 1987). Evidently it didn't matter that on less publicized surveys, *randomly* sampled American women expressed much higher levels of satisfaction. And only 1 in 7 reported having had an affair during their current marriage—a level of faithfulness replicated in British, French, and Danish surveys (Greeley, 1991, 1994). Without random sampling, large samples like Hite's—including call-in phone samples and TV web site polls—often merely give misleading results.

"How would you like me to answer that question? As a member of my ethnic group, educational class, income group, or religious category?"

Naturalistic Observation

A third descriptive research method involves watching and recording the behavior of organisms in their natural environment. These **naturalistic observations** range from watching chimpanzee societies in the jungle, to using unobtrusive measures of parent-child interactions in different cultures, to recording students' self-seating patterns in the lunchrooms of multiracial schools.

Like the case study and survey methods, 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 naturalistic observations, recalls chimpanzee observer Jane Goodall (1998), paved the way for later studies of animal thinking, language, and emotion. "Observations, made in the natural habitat, helped to show that the societies and behavior of animals are far more complex than previously supposed," and is expanding our understanding of our fellow animals. We later learned that chimps and baboons also use deception to achieve their aims.

Naturalistic observation
Some psychologists study human and animal behavior in natural environments. As University of St. Andrews psychologist Richard Byrne observes an adult gorilla, recording its behavior on a hand-held computer, a curious infant approaches and investigates his camera lens cap.

Psychologists Andrew Whiten and Richard Byrne (1988) repeatedly saw one young baboon pretending to have been attacked by another as a tactic to get its mother to drive the other baboon away from its food.

Naturalistic observations are also done with humans. Here's one 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?) And when we do laugh, 17 muscles contort our mouth and squeeze our eyes, and we emit a series of 75-millisecond vowellike sounds that are spaced about one-fifth of a second apart (Provine, 2001).

Naturalistic observation also enabled Robert Levine and Ara Norenzayan (1999) to compare the pace of life in 31 countries. By operationally defining *pace of life* as walking speed, the speed with which postal clerks completed a simple request, and the accuracy of public clocks, they concluded that 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 is often used to describe behavior. But this study, showing how pace of life is associated with culture and climate, illustrates how naturalistic observation can also be used with correlational research, our next topic.

REVIEW AND REFLECT

Description

The Case Study, the Survey, and Naturalistic Observation

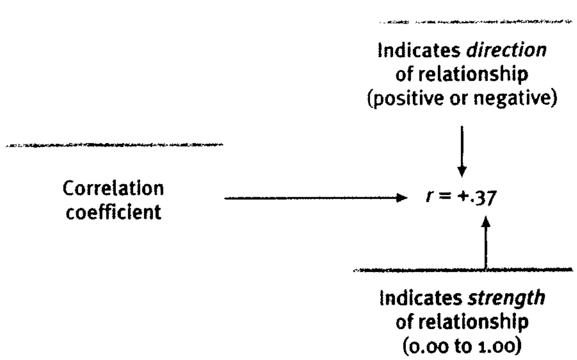
Through individual case studies, surveys among random samples of a population, and naturalistic observations, psychologists observe and describe behavior and mental processes. In generalizing from observations, remember: Representative samples are a better guide than vivid examples.

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

ASK YOURSELF: Can you recall examples of misleading surveys you have experienced or read about? What principles for a good survey did they violate?

Answers to the Check Yourself questions can be found in the yellow appendix at the end of the book.

FIGURE 1.3
How to read a correlation coefficient



Correlation

Preview: Psychologists use numbers to describe the strength of a relationship expressed as a correlation. But they caution against illusory correlations and incorrectly inferring cause and effect.

Describing behavior is a first step toward predicting it. When surveys and naturalistic observations reveal that one trait or behavior accompanies another, we say the two *correlate*. The **correlation coefficient** is a statistical measure of relationship (**FIGURE 1.3**): It reveals 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.

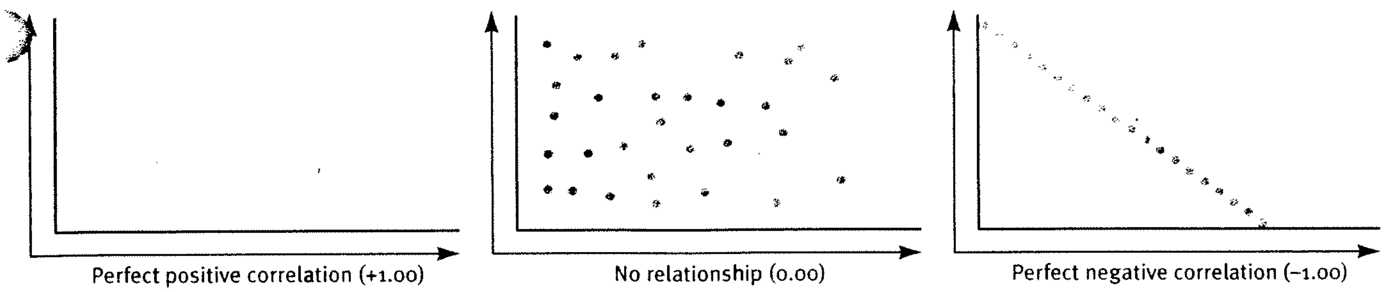


FIGURE 1.4
Scatterplots, showing patterns of correlation
Correlations can range from +1.00 (scores on one measure increase in direct proportion to scores on another) to −1.00 (scores on one measure decrease precisely as scores rise on the other).

Throughout this book we will often ask how strongly two things are related: How closely related are the personality scores of identical twins? How well do intelligence test scores predict achievement? How often does stress lead to disease?

FIGURE 1.4 illustrates perfect positive and negative correlations, which rarely occur in the "real world." These graphs are called **scatterplots**, because each point *plots* the value of two variables. A correlation's being negative has nothing to do with its strength or weakness; a negative correlation means two things relate inversely (one set of scores goes up as the other goes down). As toothbrushing goes up from zero, tooth decay goes down. A weak correlation, indicating little or no relationship, is one that has a coefficient near zero. A positive correlation means that one set of scores increases in direct proportion to the other set of scores' increase.

Statistics can help us see what the naked eye sometimes misses. To demonstrate this for yourself, try an imaginary project. Wondering if tall people 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 have someone else independently assess their temperaments (from zero for extremely calm to 100 for highly reactive).

With all the relevant data (Table 1.1) right in front of you, can you tell whether there is (1) a positive correlation between height and reactive temperament, (2) very little or no correlation, or (3) a negative correlation?

Comparing the columns in Table 1.1, most people detect very little relationship between height and temperament. In fact, the correlation in this imaginary example is moderately positive, +0.63, as we can see if we display the data as a scatterplot. In **FIGURE 1.5** (page 32) the upward, oval-shaped slope of the cluster of points as one moves to the right shows that our two imaginary sets of scores (height and reactivity) tend to rise together.

If we fail to see a relationship when data are presented as systematically as in Table 1.1, how much less likely are we to notice them in everyday life? To see what is

TABLE 1.1

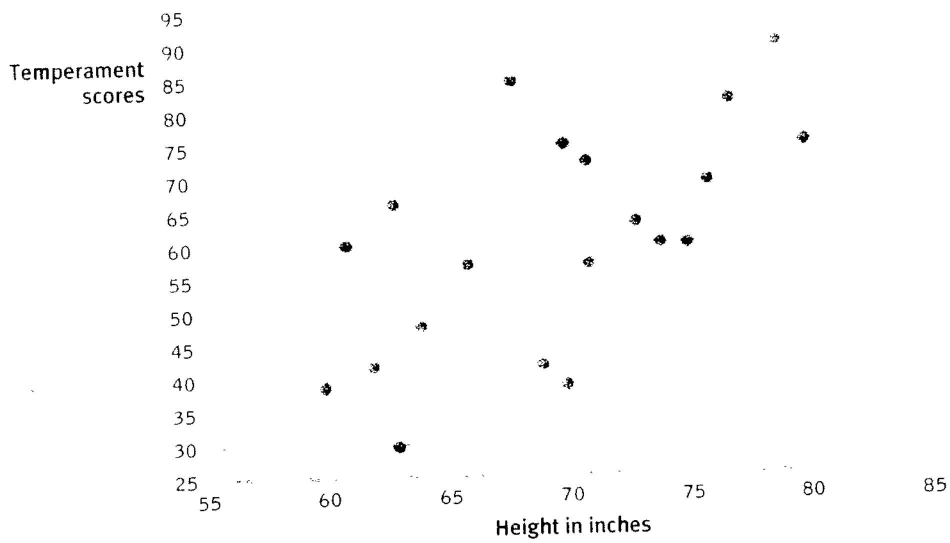
HEIGHT AND TEMPERAMENT OF 20 MEN

Subject	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

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

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). (Also called a *scattergram* or *scatter diagram*.)

FIGURE 1.5
Scatterplot for height and 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.



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

Though informative, psychology's correlations usually leave most of the variation among individuals unpredicted. As we will see, there is a correlation between parents' abusiveness and their children's later abusiveness when they become parents. But this does not mean that most abused children become abusive. The correlation simply indicates a statistical relationship: Although most abused children do not grow into abusers, nonabused children are even less likely to become abusive.

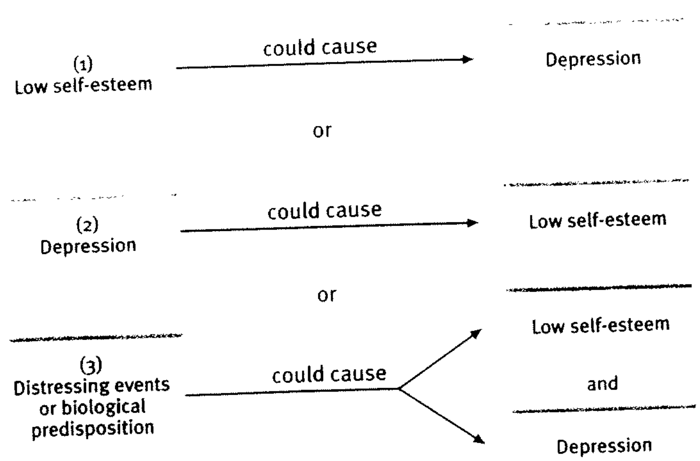
The point to remember: Although the correlation coefficient tells us nothing about cause and effect, it *can* help us see the world more clearly by revealing the actual extent to which two things relate.

Correlation and Causation

We have seen that correlations, however imperfect, do help us predict and restrain the illusions of our flawed intuition. Watching violence correlates with (and therefore predicts) aggression. But does that mean it *causes* aggression? Does low self-esteem *cause* depression? If, based on the correlational evidence, you assume that they do, you have much company. Among the most irresistible thinking errors made both by laypeople and by professional psychologists is assuming that correlation proves causation. But no matter how strong the relationship, it does not!

For example, what about the negative correlation between self-esteem and depression? Perhaps low self-esteem does cause depression. But as **FIGURE 1.6** suggests, we'd get the same correlation of low self-esteem and depression if depression caused people to be down on themselves, or if something else—a third factor such as heredity or brain chemistry—caused both low self-esteem and depression. Among men, length of marriage correlates positively with hair loss—because both are associated with a third factor,

FIGURE 1.6
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.



age. And people who wear hats are *more* likely to suffer skin cancer—because both are associated with fair-skinned people (who are vulnerable to skin cancer and more likely to wear protective hats).

This point is so important—so basic to thinking smarter with psychology—that it merits one more example, from a survey of 12,118 adolescents: 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 & others, 1997). "Adults have a powerful effect on their children's behavior right through the high school years," gushed an Associated Press story on the study. But the correlation comes with no built-in cause-effect arrow. Thus, the AP could as well have said, "Well-behaved teens feel their parents' love and approval; out-of-bounds teens more often think their parents are disapproving jerks."

The point to remember: Correlation indicates the *possibility* of a cause-effect relationship, *but it does not prove causation.* Knowing that two events are correlated need not tell us anything about causation. Remember this principle and you will be wiser as you see reports of scientific studies in the news and in this book.

illusory correlation the perception of a relationship where none exists.

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).

Illusory Correlations

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

Illusory correlations help explain many a superstitious belief, such as the presumption that more babies are born when the moon is full or that infertile couples who adopt become more likely to conceive (Gilovich, 1991). Those who conceive after adopting capture our attention. We're less likely to notice those who adopt and never conceive, or those who conceive without adopting. In other words, illusory correlations occur when we over-rely on the top left cell of **FIGURE 1.7**, ignoring equally essential information in the other cells.

Such illusory thinking helps explain why for so many years people believed (and many still do) that sugar made children hyperactive, that getting cold and wet caused one to catch a cold, and that weather changes trigger arthritis pain. Physician Donald Redelmeier, working with Amos Tversky (1996; Kolata, 1996), a psychologist who specialized in "debugging human intuition," followed 18 arthritis patients for 15 months. The researchers recorded both the patients' pain reports and the daily weather—temperature, humidity, and barometric pressure. Despite patients' beliefs, the weather was uncorrelated with their discomfort, either on the same day or up to two days earlier or later. Shown columns of random numbers labeled "arthritis pain"

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

FIGURE 1.7
Illusory correlation in everyday life
Many people believe infertile couples become more likely to conceive a child after adopting a baby. This belief arises from their attention being drawn to such cases. The many couples who adopt without conceiving or conceive without adopting grab less attention. To determine whether there actually is a correlation between adoption and conception, we need data from all four cells in this figure. (From Gilovich, 1991)

and "barometric pressure," even college students saw a correlation where there was none. We are, it seems, very, very good at detecting patterns, whether they're there or not, and not so good at testing our hypotheses.

Because we are sensitive to dramatic or unusual events, we are especially likely to notice and remember the occurrence of two such events in sequence—say, a premonition of an unlikely phone call followed by the call. When the call does not follow the premonition, we are less likely to note and remember the nonevent.

Likewise, instances of positive-thinking people being cured of cancer impress those who believe that positive attitudes counter disease. But to assess whether positive thinking actually affects cancer, we need three more types of information. First, we need an estimate of how many positive thinkers were *not* cured. Then we need to know how many people with cancer were and were not cured among those not using positive thinking. Without these comparison figures, the positive examples of a few tell us nothing about the actual correlation between attitudes and disease. (Chapter 14 explores the effects of emotions on health and illness.)

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

Perceiving Order in Random Events

Illusory correlations arise from our natural eagerness to make sense of our world—what poet Wallace Stevens called our "rage for order." Given even random data, we look for order, for meaningful patterns. And we usually find such, because *random sequences often don't look random*. Consider a random coin flip: If someone flipped a coin six times, which of the following sequences of heads (H) and tails (T) would be most likely: HHHTTT or HTTHTH or HHHHHH?

Daniel Kahneman and Amos Tversky (1972) found that most people believe HTTHTH would be the most likely random sequence. Actually, all are equally likely (or, you might say, equally unlikely) to occur. A bridge or poker hand of 10 through Ace, all of hearts, would seem extraordinary; actually, it would be no more or less likely than any other specific hand of cards (**FIGURE 1.8**).

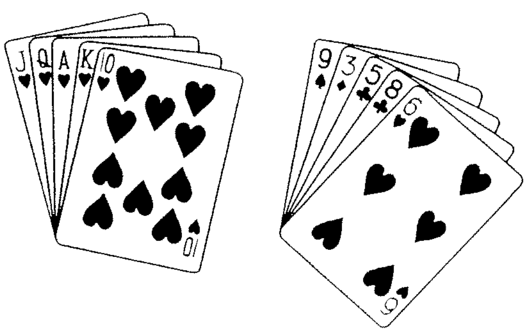


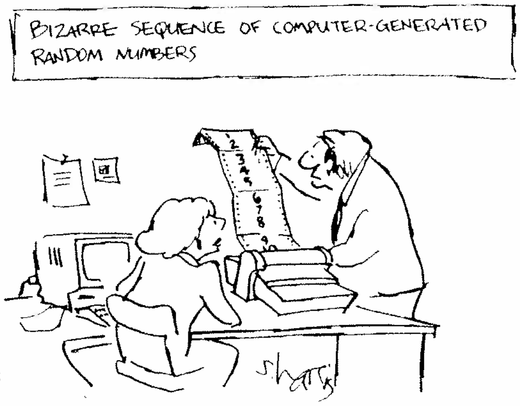
FIGURE 1.8
Two random sequences
Your chances of being dealt either of these hands are precisely the same: 1 in 2,598,960.

Psychologists Thomas Holtgraves and James Skeel (1992) exposed people's perceptions of randomness in their bets placed in Indiana's Pick-3 Lottery. You can play, too: Pick any three-digit number from 0 to 999.

Did your number have a repeated digit (as in 525)? Probably not. Only 14 percent of 2.24 million number strings chosen in July 1991 had a repeated digit. Although repeated digits actually occur in 28 percent of the available numbers, such numbers *look* less random (and people prefer to bet random-looking series). In actual random sequences, seeming patterns and streaks (such as repeating digits) occur more often than people expect. Thus, shown random data, scientists and psychics alike can often "see" an interesting pattern (Guion, 1992). To demonstrate this phenomenon for myself (as you can do), I flipped a coin 51 times, with these results:

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

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



© 1990 by Sidney Harris/American Scientist Magazine.

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

What explains these 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, 24 of the 50 comparisons yielded a changed result—just the sort of near 50-50 result we expect from coin tossing. Despite the seeming patterns in these data, the outcome of one toss gives no clue to the outcome of the next toss.

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 the odds that a given person who buys a single ticket for two New Jersey lotteries will win both times. But statisticians Stephen Samuels and George McCabe (1989) report that, given the millions of people who buy U.S. state lottery tickets, it was "practically a sure thing" that someday, somewhere, someone would hit a state jackpot twice. Indeed, say fellow statisticians Persi Diaconis and Frederick Mosteller (1989), "with a large enough sample, any outrageous thing is likely to happen." "The really unusual day would be one where nothing unusual happens," adds Diaconis (2002).

On the 2002 anniversary of 9/11, New York State's three-pick lottery numbers came up 9-1-1.

We all experience events that make us feel astonished now and then. One day when my daughter bought two pairs of shoes, we later were astounded to discover that the two brand names were her first and last names. Checking out a photocopy counter from our library, I confused the clerk when giving him my six-digit department charge number—which just happened at that moment to be identical to the counter's one-in-a-million number on which the last user had finished. Ron Vachon was astounded while sitting among thousands of fans at a September 1990 baseball game in Boston. Oakland A's outfielder Rickey Henderson hit two foul balls right to him, on successive pitches. That something like that should have happened to Vachon (who dropped them both) was incredibly unlikely. That it sometime would happen to someone was not. An event that happens to but one in 1 billion people every day occurs about six times a day, 2000 times a year. (For two provocative instances of random sequences that don't look random, see Thinking Critically About Hot and Cold Streaks in Basketball and the Stock Market on page 36.)

On March 11, 1998, Utah's Ernie and Lynn Carey gained three new grandchildren when three of their daughters gave birth—on the same day (*Los Angeles Times*, 1998).

UPI/Corbis-Bettmann

Given enough random events, something weird will happen
Evelyn Marie Adams was the beneficiary of one of those extraordinary, chance events when she won the New Jersey lottery a second time.

THINKING CRITICALLY ABOUT:

Hot and Cold Streaks in Basketball and the Stock Market

Misinterpreting random sequences is common in sports and investing. In both arenas, the statistical facts collide with commonsense intuition.

Basketball Players' "Hot Hands"

Every basketball player and every fan intuitively "knows" that players have hot and cold streaks. Players who have "hot hands" can't seem to miss. Those who have "cold" ones can't find the center of the hoop. When Thomas Gilovich, Robert Vallone, and Amos Tversky (1985) interviewed Philadelphia 76ers, the players estimated they were about 25 percent more likely to make a shot after they had just made one than after a miss. In one survey, 9 in 10 basketball fans agreed that a player "has a better chance of making a shot after having just *made* his last two or three shots than he does after having just *missed* his last two or three shots." Believing in shooting streaks, players will feed the ball to a teammate who has just made two or three shots in a row. Many coaches will bench the player who has just missed three in a row.

The only trouble is (believe it or not), it isn't true! When Gilovich and his collaborators studied detailed individual shooting records, they found that the 76ers—and the Boston Celtics, the New Jersey Nets, the New York Knicks, and Cornell University's men's and women's basketball players—were equally likely to score after a miss and after a basket. A typical 50 percent shooter averages 50 percent after just missing three shots, and 50 percent after just making three shots. It works with free throws, too. Celtics star Larry Bird made 88 percent of his free throws after making a free throw and 91 percent after missing. (Did this reduce Larry Bird to a mere puppet, manipulated by statistical laws? No, his skill was reflected in his 90 percent average.)

Why, then, do players and fans alike believe that players are more likely to score after scoring and to miss after missing? (The same phenomenon turns up in baseball, where the individual and team streaks and slumps that fascinate sportswriters are to be expected, given mere random variation [Myers, 2002].) It's because streaks do occur, more than people expect in random sequences. In any series of 20 shots by a 50 percent shooter (or any 20 flips of a coin), there is a 50-50 chance of 4 baskets (or heads) in a row, and it is quite possible that one person out of five will have a streak of 5 or 6. Players and fans notice these random streaks and so form the errant conclusion that "when you're hot, you're hot" (**Figure 1.9**).

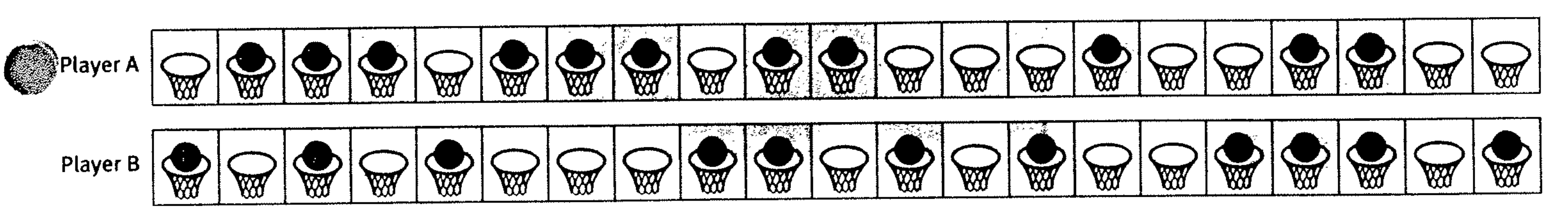
Mutual Funds: Does Past Performance Predict Future Returns?

The same misinterpretation of random sequences occurs when investors believe that a mutual fund that has had a string of good years will likely outperform one that has had a string of bad years. Based on that assumption, investment magazines report mutual funds' performance. But, as economist Burton Malkiel (1989, 1995) documents, past performances of mutual funds do *not* predict their future performance. If on January 1 of each year since 1980 we had bought the previous year's top-performing funds, our hot funds would not have beaten the next year's market average. If we had put our money instead on the *Forbes* "Honor Roll" of funds each year for the two decades following 1975, we would have pulled in 13.5 percent (compared with the market's overall 14.9 percent annual return). Of the top 81 Canadian funds during 1994, 40 performed above average and 41 below average during 1995 (Chalmers, 1995).

When funds have streaks of several good or bad years, we may nevertheless think that past success predicts future success. "Randomness is a difficult notion for people to accept," notes Malkiel. "When events come in clusters and streaks, people look for explanations and patterns. They refuse to believe that such patterns—which frequently occur in random data—could equally well be derived from tossing a coin. So it is in the stock market as well."

The point to remember: When watching basketball, choosing stocks, or flipping coins, remember that our intuition often misleads us. Random sequences frequently don't look random. Expect streaks.

FIGURE 1.9
Who is the chance shooter?
Here are 21 consecutive shots, each scoring either a basket or a miss, by two players who each make 11. Within this sample of shots, which player's sequence looks more like what we would expect in a random sequence? (See page 39; adapted from Barry Ross, *Discover*, 1987.)



REVIEW AND REFLECT

Correlation

Correlation and Causation

The strength of the relationship between one factor and another is expressed as a number in their *correlation coefficient*. Scatterplots and the correlations they reveal help us to see relationships that the naked eye might miss. Knowing how closely two things are positively or negatively correlated tells us how much one predicts the other. But it is crucial to remember that correlation is a measure of relationship; it does not reveal cause and effect.

Illusory Correlations and Perceiving Order in Random Events

Correlations also help us to discount relationships that do not exist. Illusory correlations—random events we notice and assume are related—arise from our search for patterns.

CHECK YOURSELF: Here are some recently reported correlations, with interpretations drawn by journalists. Further research, often including experiments, has clarified cause and effect in each case. Knowing just these correlations, can you come up with other possible explanations for each of these?

a. Alcohol use is associated with violence. (One interpretation: Drinking triggers or unleashes aggressive behavior.)

b. Educated people live longer, on average, than less-educated people. (One interpretation: Education lengthens life and enhances health.)

c. Teens engaged in team sports are less likely than other teens to use drugs, smoke, have sex, carry weapons, and eat junk food less often than teens who do not engage in team sports. (One interpretation: Team sports encourage healthy living.)

d. Adolescents who frequently see smoking in movies are more likely to smoke. (One interpretation: Movie stars' behavior influences impressionable teens.)

ASK YOURSELF: Can you think of an example of correlational research that you recently heard about from a friend or on the news? Was an unwarranted conclusion drawn?

Answers to the Check Yourself questions can be found in the yellow appendix at the end of the book.

R. Sidney/The Image Works

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

Experimentation

Preview: To discern cause and effect, psychologists experiment. In the typical experiment they randomly assign some people to experience a treatment of interest, while others have no such experience. Because the random assignment equalizes the groups at the outset, any later differences were probably caused by the experimental variable being tested.

Happy are they "who have been able to perceive the causes of things," remarked the Roman poet Virgil. We endlessly wonder and debate *why* we act as we do. Why do some people smoke? Have babies while they are still children? Do stupid things when drunk? Become troubled teens and open fire on their classmates? Though psychology cannot answer these questions directly, it has helped us to understand what influences drug use, sexual behaviors, thinking when drinking, and aggression.

Many factors influence our everyday behavior. To isolate cause and effect—say, in looking for possible causes of depression—psychologists sometimes try to statistically control for other factors. For example, many studies have found that breast-fed infants grow up with somewhat higher intelligence scores than those of infants bottle-fed with cow's milk (Angelsen & others, 2001; Gale & Martyn, 1996; Johnson & others, 1996; Lucas & others, 1992; Mortensen & others, 2002; Quinn & others, 2001). Mother's milk correlates modestly but positively with later intelligence. But does this mean that smarter mothers (who more often breast-feed) have smarter children? Or, as some researchers believe, do the nutrients of mother's milk contribute to brain development? To help answer this question, researchers have "controlled for" (statistically removed differences in) maternal age, education, and intelligence. Still, breast-fed infants exhibit slightly higher intelligence as young children.

The clearest and cleanest way to isolate cause and effect is, however, to **experiment**. Experiments enable a researcher to focus on the possible effects of one or more factors by (1) *manipulating the factors of interest* and (2) *holding constant ("controlling") other factors.* Knowing that correlations of infant nutrition and later intelligence can't possibly control for all other possible factors, a British research team led by Alan Lucas (1998) decided to experiment, using 424 hospital preterm infants. With parental permission, the researchers randomly assigned some infants to standard infant formula feedings and others to donated breast milk feedings. When given intelligence tests at age 8, the children nourished with breast milk had significantly higher intelligence scores than their formula-fed counterparts. No single experiment is conclusive of course, but these researchers, by randomly assigning infants to a feeding condition, were able to hold constant all factors except nutrition. This rigorous design helps eliminate alternative explanations and supports the conclusion that, so far as the developing intelligence of preterm infants is concerned, breast is best.

If behavior changes when we vary an experimental factor, such as infant nutrition, then we know that the factor is having an effect. *The important point to remember:* Unlike correlational studies, which uncover naturally occurring relationships, an experiment manipulates a factor to determine its effect. Let's consider some more experiments.

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 experiment controls other relevant factors.

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 effect experimental results caused by expectations alone; any effect on behavior caused by the administration of an inert substance or condition, which is assumed to be an active agent.

experimental condition the condition of an experiment that exposes participants to the treatment, that is, to one version of the independent variable.

control condition the condition of an experiment that contrasts with the experimental condition and serves as a comparison for evaluating the effect of the treatment.

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

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

dependent variable the experimental factor—in psychology, the behavior or mental process—that is being measured; the variable that may change in response to manipulations of the independent variable.

Evaluating Therapies

Our tendencies to seek new remedies when we are ill or emotionally down can produce misleading testimonies. When our health or emotions return to normal, we attribute the return to something we have done. If three days into a cold we start taking vitamin C tablets and find our cold symptoms lessening, the pills may seem more potent than they are (an illusion of control). If, after nearly failing the first exam, we listen to a "peak learning" subliminal tape and then improve on the next exam, we may credit the tape rather than conclude that our performance has returned to our average. In the 1700s, blood-letting *seemed* effective. Sometimes people improved after the treatment; when they didn't, the practitioner inferred the disease was too far advanced to be reversed. So, whether or not a remedy is truly effective, enthusiastic users will probably endorse it. To find out whether it actually is effective, we must experiment.

And that is precisely how new drug treatments and new methods of psychological therapy are evaluated (Chapter 16). In many of these studies, the participants are *blind* (uninformed) about what treatment, if any, they are receiving. One group receives the treatment. Others receive a pseudotreatment—an inert *placebo* (perhaps a pill with no drug in it). Often neither the participant nor the research assistant collecting the data knows whether the participant's group is receiving the treatment. This **double-blind procedure** enables researchers to check a treatment's actual effects apart from the research participants' (and their own) enthusiasm for it and from the healing power of belief. The **placebo effect** is well documented with pain, depression, and anxiety (Kirsch & Sapirstein, 1998). Just *thinking* one is getting a treatment can boost one's spirits, relax one's body, and lead to symptom relief.

The double-blind procedure creates an **experimental condition** in which people receive the treatment and a contrasting **control condition** without the treatment. By **randomly assigning** people to these conditions the two groups should otherwise be identical. Random assignment roughly equalizes the two groups in age, attitudes, and every other characteristic. With random assignment, as occurred with the infants in the breast milk experiment, we can know that any later differences between people in the experimental and control conditions must be the result of the treatment.

Another example: On the advice of their physicians, millions of postmenopausal women turned to hormone replacement therapy after correlational studies found that women on replacement hormones had lower rates of heart disease, stroke, and colon cancer. But women who got the therapy were perhaps more likely to be receiving medical care, exercising, and eating well. So, did the hormones make women healthy or did healthy women take the hormones? In 2002, the National Institutes of Health announced the surprising results of a massive experiment that randomly assigned 16,608 healthy women to either replacement hormones or a placebo: Compared to women in the control condition, women receiving the hormones had *more* health problems (Love, 2002).

Answer to question in Figure 1.9 (page 36): Player B, whose outcomes may look more random, actually has fewer streaks than would be expected by chance. For these players, chance shooting, like chance coin tossing, should produce a change in outcome about 50 percent of the time. But 70 percent of the time (14 times out of 20), Player B's outcome changes on successive shots. Player A's next outcome differs from the last 10 times out of 20.

And an even more potent example: The drug Viagra was approved for use after 21 clinical trials, including an experiment in which researchers randomly assigned 329 men with impotence to either an experimental condition (Viagra) or a control condition (a placebo). It was a double-blind procedure—neither the men nor the person who gave them the pills knew which drug 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 & others, 1998). Viagra worked.

This simple experiment manipulated just one drug factor. 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 (which random assignment controls). 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 answer the "What do you mean?" question with a level of precision that enables others to repeat the study.

Let's recap. A variable is anything (infant nutrition, intelligence, hair color—whatever) that can vary. Experiments aim to *manipulate* an *independent* variable, *measure* the *dependent* variable, and *control* all other variables. An experiment has at least two different conditions: a comparison or control condition and an experimental condition. Random assignment equates the conditions before any treatment effects. In this way, an experiment tests the effect of at least one independent variable (the

Note the distinction between random *sampling* in surveys and random *assignment* in experiments. Random sampling helps us generalize to a larger population. Random assignment controls extraneous influences, which helps us infer cause and effect.

TABLE 1.2

COMPARING RESEARCH METHODS

Research Method	Basic Purpose	How Conducted	What Is Manipulated	Possible Problems
Descriptive	To observe and record behavior	Do case studies, surveys, or naturalistic observations	Nothing	Atypical sample; biased observations
Correlational	To detect naturally occurring relationships; to assess how well one variable predicts another	Compute statistical association, sometimes among survey responses	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

experimental factor) on at least one dependent variable (the measured response). Table 1.2 compares the features of psychology's research methods.

These concepts—experimental and control conditions, independent and dependent variables, random assignment—are important, yet easily confused. So let's put them to work with another intriguing set of experiments.

Can Subliminal Tapes Improve Your Life?

A new generation of entrepreneurs would have you believe so. We are bombarded by mail-order catalogs, cable television ads, and bookstores offering tapes whose imperceptibly faint messages supposedly "reprogram your unconscious mind for success and happiness." While struggling students listen to soothing music, subliminal messages (those below one's hearing threshold) are said to persuade the unconscious that "I am a good student. I love learning." Procrastinators can be similarly reprogrammed: "I set my priorities. I get things done ahead of time!"

Is there anything to these claims? Could positive subliminal messages help us, even a little? Chapter 5 will show that subliminal sensation is for real. We, in fact, do process much information without conscious awareness. And under certain conditions, a stimulus too weak to recognize can affect us, *briefly*.

In this experiment, what was the independent variable? The dependent variable? (See page 42.)

But does this subtle, fleeting effect extend to the powerful, enduring influence claimed by the subliminal tape merchants? Anthony Greenwald and his colleagues (1991) wanted to find out, so they randomly assigned university students to listen daily for five weeks to commercial subliminal tapes claiming to improve either self-esteem or memory. But the researchers had manipulated an experimental factor. On half the tapes they switched the labels. Some students *thought* they were receiving affirmations of self-esteem when they actually were hearing the memory enhancement tape. Others got the self-esteem tape but *thought* their memory was being recharged (**FIGURE 1.10**).

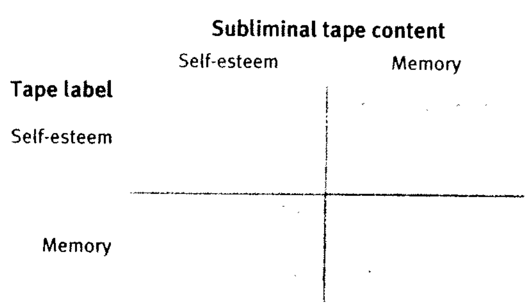


FIGURE 1.10
Design of the subliminal tapes experiment
Students' self-esteem and memory abilities were assessed before and after listening to subliminal tapes purporting to increase either self-esteem or memory. Half the students, however, received deliberately mislabeled tapes.

Were the tapes effective? Their scores on tests for both self-esteem and memory, taken before and after the five weeks, revealed zilch. No effects. None. And yet, those who *thought* they had heard a memory tape *believed* their memories had improved. A similar result occurred for those who thought they had heard a self-esteem tape. The tapes had no effects, yet the students *perceived* themselves receiving the benefits they *expected*. When reading this research, you can hear echoes of the testimonies that ooze from the mail-order tape catalogs. Many customers, having bought what is not supposed to be heard, and having indeed not heard it, actually write things like, "I really know that your tapes were invaluable in reprogramming my mind." Greenwald conducted 16 double-blind experiments evaluating subliminal self-help tapes over one 10-year period. His results were uniform: Not one had any therapeutic effect (Greenwald, 1992).

Unfortunately, the general public is surprisingly uninformed about the importance of controlled experiments such as this. One science literacy survey asked people to imagine testing a new drug to combat high blood pressure (Miller & Pifer, 1996). The survey asked whether it would make more sense to give the drug to 1000 individuals and see what happened, or to give it to half of them and compare their reactions to those who got no drug. One-third said it would make more sense to give the drug to all 1000 people, reasoning that the greater the number tested, the more reliable the finding. Among those who selected the option with the control group, 30 percent did so simply to save lives, saying "If the drug kills people, it kills only half as many." Again, remember: Psychology's most powerful tool for sorting reality from wishful thinking and for evaluating cause and effect is the control group.

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? Does school sex education reduce teen pregnancies? To answer these questions, we can use experiments: If an intervention is welcomed but resources are scarce, we could use a lottery to randomly assign some people (or regions) to experience the new program and others to a control condition. If later the two groups differ, there will be less to argue about (Passell, 1993).

REVIEW AND REFLECT

Experimentation

To discover cause-and-effect relationships, psychologists conduct *experiments*. By constructing a controlled reality, experimenters can manipulate one or more factors and discover how these independent variables affect a particular behavior, the dependent variable.

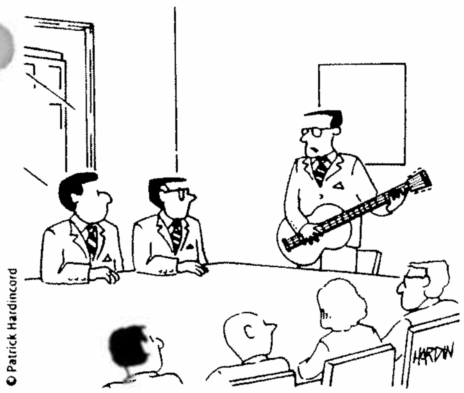
Evaluating Therapies, and Can Subliminal Tapes Improve Your Life?

In many experiments, control is achieved by randomly assigning people either to the experimental condition, the group exposed to the treatment, or to a control condition, a group that experiences no treatment or a different version of the treatment.

CHECK YOURSELF: Why, when testing a new drug for 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?

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

Answers to the Check Yourself questions can be found in the yellow appendix at the end of the book.



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

Statistical Reasoning

Preview: Having gathered data, we must next organize, summarize, and make inferences from it, using statistics. Today's statistics are tools that help us see and interpret what the unaided eye might miss.

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:

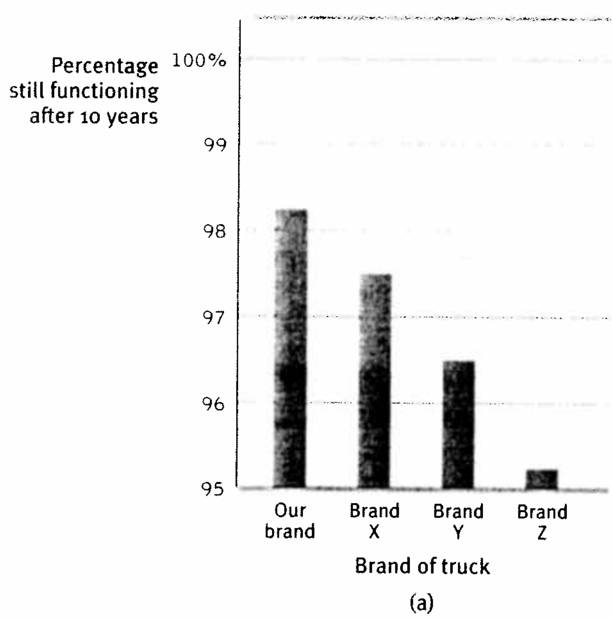
- *One percent of Americans (2.7 million) are homeless.* Or is it 300,000, an earlier estimate by the federal government? Or 600,000, the estimate by the Urban Institute (Crossen, 1994)?
- *Ten percent of people are lesbians or gay men.* Or is it 2 to 3 percent, as suggested by various national surveys (Chapter 12)?
- *We ordinarily use but 10 percent of our brain.* Or is it closer to 100 percent? (Which 90 percent, or even 10 percent, would you be willing to sacrifice?) (Chapter 2)

The point to remember: Doubt big, round, undocumented numbers. Rather than swallow top-of-the-head estimates, focus on thinking smarter by applying simple statistical principles to everyday reasoning.

Answer to question on page 40: In the subliminal tapes experiment, the primary independent variable was the type of subliminal message, self-esteem versus memory. (This experiment actually had a second independent variable as well: people's beliefs about which tape they received.) The primary dependent variable was improvement on the self-esteem and memory measures.

Describing Data

Once researchers have gathered their raw data, their first task is to *organize* it. One way is to use a simple *bar graph*, as in **FIGURE 1.11**, which displays a distribution of trucks of different brands still on the road after a decade. When reading statistical graphs such as this, take care. Depending on what people want to emphasize, they can design the graph to make a difference look small or big. So think smart: When viewing figures in magazines and on television, read the scale labels and note their range.



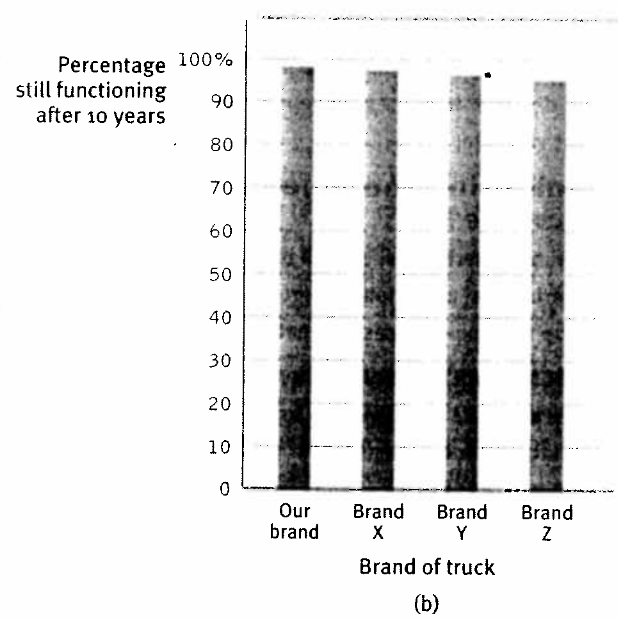


FIGURE 1.11 Read the scale labels An American truck manufacturer offered a graph (a)—with actual brand names included—to suggest the much greater durability of its trucks. Note, however, how the apparent difference hrinks as the vertical scale changes (graph b).

Measures of Central Tendency

The next step is to summarize the data using the three measures of "central tendency." The simplest measure is called the **mode**, the most frequently occurring score. 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 middle score—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.

Measures of central tendency neatly summarize data. But consider what happens to the mean when a distribution is lopsided or *skewed*. With income data, for example, the mode, median, and mean often tell very different stories (**FIGURE 1.12**). This is because the mean is biased by a few extreme scores. When Microsoft founder Bill Gates sits down in an intimate cafe, its average (mean) patron instantly becomes a billionaire. 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. Professional athletes' incomes also form skewed distributions. In 1998, 66 percent of the National Basketball Association's 411 players made less than the average (mean) player salary (DuPree, 1998). The average ($2.24 million) was, of course, inflated by a few superstar salaries, led by Michael Jordan's $33.14 million.

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

mode the most frequently occurring score 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.

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

The average adult has one ovary and one testicle.

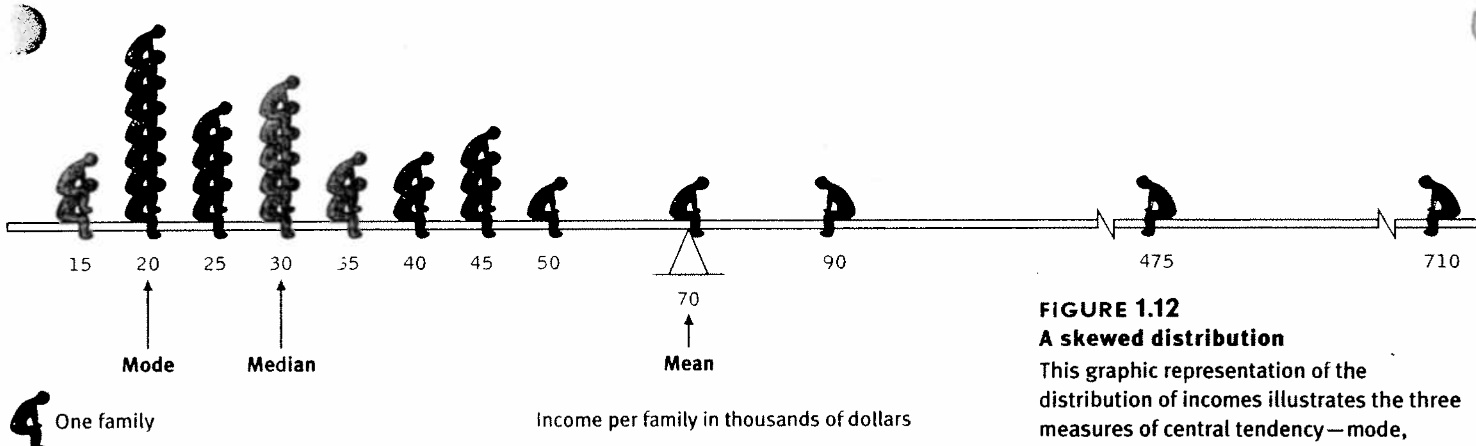


FIGURE 1.12
A skewed distribution
This graphic representation of the distribution of incomes illustrates the three measures of central tendency—mode, median, and mean. Note how just a few high incomes make the mean—the fulcrum point that balances the incomes above and below—deceptively high.

Measures of Variation

Knowing the value of an appropriate measure of central tendency can tell us a great deal. But it also 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.

The **range** of scores—the gap between the lowest and highest scores—provides only a crude estimate of variation because a couple of extreme scores in an otherwise uniform group, such as the $475,000 and $710,000 incomes in Figure 1.12, will create a deceptively large range.

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

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

The more useful measure of 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 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 smaller standard deviation than the one found in the more diverse community population outside your school.

Making Inferences

Data are "noisy." One group's average score (breast-fed babies' intelligence scores) could conceivably differ from another's (the formula-fed babies') not because of any real difference but merely due to chance fluctuation in the people sampled. How confidently, then, can we infer that an observed difference accurately estimates the true difference?

When Is an Observed Difference Reliable?

In deciding when it is safe to generalize from a sample, we should keep three principles in mind. Let's look at each in turn.

1. ***Representative samples are better than biased samples.*** As we have noted, 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. No research involves a representative sample of 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 points scored 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 university student visits two college 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 tells friends about the "great teachers" at the first school, and the "bores" at the second. Again, we know it but we ignore it: Small samples provide less reliable estimates of the average than do large samples. The proportion of heads in samples of 10 coin tosses varies more than in samples of 100 tosses. Said differently, *averages based on many cases are more reliable* (less variable) than averages based on only a few cases.

The point to remember: Don't be overly impressed by a few anecdotes. Generalizations based on a few unrepresentative cases are unreliable.

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

When Is a Difference Significant?

We can justifiably have the most confidence when we generalize from samples that (1) are representative of the population we wish to study, (2) give us consistent rather than highly variable data, and (3) are large rather than small. These principles extend to the inferences we make about differences between groups—as when we generalize from a gender difference in grades in our sample to the whole campus population.

Statistical tests help us determine significance by indicating the reliability of differences. Here is the logic behind them: When *averages* from two samples are each *reliable* measures of their respective populations (as when each is based on many ob-

servations that have small variability), then their difference (sometimes even a very small difference) is likely to be reliable as well. (The less the variability in women's and men's aggression scores, the more confidence we would have that any observed difference is reliable.) But 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.

In short, *when the sample averages are reliable and the difference between them is relatively large*, we say the difference has **statistical significance**. This simply means that the difference we observed 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 are less than 5 percent (an arbitrary criterion).

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 several hundred thousand first-born and later-born individuals indicate that there is a highly significant tendency for first-born individuals within a family to have higher average scores than their later-born siblings (Zajonc & Markus, 1975). But because the scores differ by only one or two points, the difference has little practical importance. Such findings have caused some psychologists to advocate alternatives to significance testing (Hunter, 1997). Better, they say, to use other ways to express a finding's magnitude and reliability.

The point to remember: Statistical significance indicates the *likelihood* that a result will happen by chance. It does not indicate the *importance* of the result.

Using the principles discussed in this chapter will help us to think critically—to see more clearly what we might otherwise miss or misinterpret and to generalize more accurately from our observations. We do think smarter when we understand and use the principles of research methods and statistics (Fong & others, 1986; Lehman & others, 1988; VanderStoep & Shaughnessy, 1997). It requires training and practice, but developing clear and critical thinking abilities is part of your becoming an educated person. The report of the Project on Redefining the Meaning and Purpose of Baccalaureate Degrees (1985) eloquently asserts why there are few higher priorities in a college education:

> If anything is paid attention to in our colleges and universities, thinking must be it. Unfortunately, thinking can be lazy. It can be sloppy. . . . It can be fooled, misled, bullied. . . . Students possess great untrained and untapped capacities for logical thinking, critical analysis, and inquiry, but these are capacities that are not spontaneous: They grow out of wide instruction, experience, encouragement, correction, and constant use.

REVIEW AND REFLECT

Statistical Reasoning

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

From this section's consideration of how we can organize, summarize, and make inferences from data—by constructing distributions and computing measures of central tendency, variation, and statistical significance—we derived five points to remember:

1. Doubt big, round, undocumented numbers.
2. When looking at statistical graphs in books and magazines and on television ads and news broadcasts, think critically: Always read the scale labels and note their range.
3. Always note which measure of central tendency is reported. Then, if it is a mean, consider whether a few atypical scores could be distorting it.
4. Don't be overly impressed by a few anecdotes. Generalizations based on only a few cases are unreliable.
5. Statistical significance indicates the *likelihood* that a result will occur by chance. It does not indicate the importance of the result.

CHECK YOURSELF: Consider a question posed by Christopher Jepson, David Krantz, and Richard Nisbett (1983) to University of Michigan introductory psychology students:

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?

ASK YOURSELF: Find a graph in a popular magazine ad. How has the advertiser used (or abused) statistics to make a point?

Answers to the Check Yourself questions can be found in the yellow appendix at the end of the book.

Frequently Asked Questions About Psychology

Preview: A scientific approach can restrain our flawed intuition while satisfying our curiosity about what predicts or causes behavior. But for many, the idea of applying science to human affairs raises concerns about how well experiments relate to life, whether they apply to all cultures and both genders, how experimenters treat human and animal subjects, and how psychologists' values influence their work and its applications.

We have seen how case studies, surveys, and naturalistic observations help us describe behavior. We have also noted that correlational studies assess the relationship between two factors, which indicates how well, knowing one thing, we can predict another. We have examined the logic that underlies experiments, which use control conditions and random assignment of subjects to isolate the effects of an independent variable on a dependent variable. We have reflected on how a scientific approach, aided by statistics, can restrain biases.

You are now prepared to understand what lies ahead and to think critically 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 confront some typical questions and concerns.

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? Does our tendency to remember best the first and last items in a list of unrelated words tell us anything about why we remember the names of certain people we meet at a party? 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 in which important features of everyday life can be simulated and controlled. Just as an aeronautical wind tunnel enables an engineer to re-create atmospheric forces under controlled conditions, a laboratory experiment enables a psychologist to re-create psychological forces under controlled conditions.

People in the lab are not different creatures from their out-of-lab selves. For example, Cecilia Cheng (2001) observed that Hong Kong adults who flexibly coped with laboratory stresses also coped flexibly with stress in their marriages. 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. And the experiment's purpose, notes Douglas Mook (1983), is not to re-create the exact behaviors of everyday life but to test theoretical principles. *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 we 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 & others, 1999).

The point to remember: As psychologists, our concerns lie less with particular behaviors than with the general principles that help explain many behaviors.

Does Behavior Depend on One's Culture?

If culture shapes behavior, what can psychological studies done in one culture, often with white North Americans, 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 influences our standards of promptness and frankness, our attitudes toward premarital sex and varying body shapes, tendencies to be casual or formal, and much, much more. Being aware of such differences, we can restrain our assumptions that others will think and act as we

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

Mark Wexler/Woodfin Camp & Associates

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. Schoolchildren everywhere greet their teachers with respect, although not necessarily with the formality of this young Japanese schoolchild.

do. Given the growing mixing and clashing of cultures, our need for such awareness is urgent.

Our shared biological heritage does, however, unite 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 & others, 2001).
- Variation in languages—spoken and gestured—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 do vary in feelings of loneliness, but across cultures shyness, low self-esteem, and being unmarried magnify loneliness (Jones & others, 1985; Rokach & others, 2002).
- Most Japanese prefer their fish raw and most North Americans prefer theirs cooked, but the same principles of hunger and taste influence all of us when we sit down to a meal. 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.

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

Confucius 551–479 B.C.

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

Does Behavior Vary with Gender?

At your birth, friends and family immediately wondered which of the two human types you were: male or female. Given how important our gender is to our identity and to others' perceptions of us, do we need a different psychology for women and for men?

You will see throughout this book that gender issues permeate psychology. Researchers report gender differences in what we dream, in how we express and detect emotions, and in our risk for alcoholism, depression, and eating disorders. Not only is studying such differences interesting, it is also potentially beneficial. For example, many researchers believe that women carry on conversations more readily to build relationships; men usually talk to give information and advice (Tannen, 1990). Knowing this difference can help us prevent conflicts and misunderstandings in everyday relationships.

Nevertheless, it's important to remember that 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. We also tend to exhibit and perceive the very behaviors our culture expects of males and females.

So, gender matters. Biology determines our sex, and then culture further bends the genders. But viewing life through the lens of gender can exaggerate differences.

Why Do Psychologists Study Animals?

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, by doing experiments that are permissible only with animals. Rats, critics say, are not long-tailed people. Yet human physiology resembles that of many other animals. Animal experiments have therefore led to treatments for human diseases—insulin for diabetes, vaccines to prevent polio and rabies, transplants to replace defective organs.

Likewise, the same processes by which humans see, exhibit emotion, and become obese are present in rats and monkeys. To discover more about the basics of human learning, researchers even study sea slugs. To understand how a combustion engine works, you would do better to study the engine of a lawn mower than that of a Mercedes. Like Mercedes engines, humans are complex. But it is precisely the simplicity of the sea slug's nervous system that makes it so revealing of the neural mechanisms of learning.

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

Dave Barry, July 2, 2002

Is It Ethical to Experiment on Animals?

If we share important similarities with other animals, then should we not respect them? "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. Researchers remind us that the world's 30 million mammals used each year in research are but a fraction of 1 percent of the billions of animals killed annually for food (which means the average person eats 20 animals a year). While researchers each year conduct experiments on some 200,000 dogs and cats cared for under humane regulations, humane animal shelters are forced to kill 50 times that many (Goodwin & Morrison, 1999).

Mobilization for Animals, a network of animal protection organizations, has nevertheless been concerned. It has declared that animals used in psychological experiments are shocked "until they lose the ability to even scream in pain, . . . [are] deprived of food and water to suffer and die slowly from hunger and thirst, . . . [are] put in total isolation chambers until they are driven insane or even die from despair and terror," and are made "the victims of extreme pain and stress, inflicted upon them out of idle curiosity." However, when psychologists Caroline Coile and Neal Miller (1984) analyzed every animal research article published in the American Psychological Association's journals during the preceding five years, they found no study in which any of these allegations was true. Even when researchers used shock, it was usually of a mild intensity, one that humans can easily endure on their fingers. Only 7 percent of psychology's studies involved animals, 95 percent of which were rats, mice, rabbits, or birds. About 10 percent of these animal studies involved electric shock (Coile & Miller, 1984; Gallup & Suarez, 1985). In British psychology departments, where animal use dropped by two-thirds in the dozen years after 1977, electric shock had been used in only 4 percent of animal studies. All involved rats (Thomas & Blackman, 1991).

Animal protection organizations, such as Psychologists for the Ethical Treatment of Animals, advocate naturalistic observation of animals rather than laboratory manipulation. However, many researchers say this is not the morality of good versus evil but of compassion (for animals) versus compassion (for people). How many of us would have attacked Pasteur's experiments with rabies, which caused some dogs to suffer but led to a vaccine that spared millions of people, and dogs, from agonizing death? And would we really wish to have deprived ourselves of the animal research

"I believe that to prevent, cripple, or needlessly complicate the research that can relieve animal and human suffering is profoundly inhuman, cruel, and immoral."

Psychologist Neal Miller (1983)

that led to effective methods of training children with mental disorders; of understanding aging; of relieving fears and depression; of controlling obesity, alcoholism, and stress-related pain and disease?

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 hopes that people might not? Should some monkeys be exposed to an HIV-like virus in the search for an AIDS vaccine? Is our use of other animals as natural as the behavior of carnivorous hawks, cats, and whales? (Animals themselves do not assign rights to other animals lower on the food chain.) Defenders of research on animals argue that anyone who has eaten a hamburger, worn leather shoes, tolerated hunting and fishing, or supported the extermination of crop-destroying or plague-carrying pests has already agreed that, yes, it is sometimes permissible to sacrifice animals for the sake of human well-being.

"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)

Scott Plous (1993) notes that our compassion for animals varies, as does our compassion for people, based on their perceived similarity to us. As Chapter 18 explains, we feel more attraction, give more help, and act less aggressively toward similar others. Likewise, we value animals according to their perceived kinship with us. Thus, primates and companion pets get top priority. (Western people raise or trap mink and foxes for their fur, but not dogs or cats.) Other mammals occupy the second rung on the privilege ladder, followed by birds, fish, and reptiles on the third rung, with insects at the bottom. In deciding which animals have rights, we each draw our own cut-off line somewhere across the animal kingdom.

"The righteous know the needs of their animals."

Proverbs 12:10

If we give human life first priority, the second issue is the priority given the well-being of the animals in research. What safeguards should protect animals? Most researchers today feel ethically obligated to enhance the well-being of captive animals and protect them from needless suffering. In one survey of animal researchers, 98 percent or more 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 now have guidelines for the humane use of animals. For example, British Psychological Society guidelines now call for housing animals under reasonably natural living conditions, with companions for social animals (Lea, 2000). Humane care also leads to more effective science, because pain and stress would distort the animals' behavior during experiments.

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

Mahatma Gandhi, 1869–1948

Animals have themselves benefitted 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, and studied methods of handling and stroking them that reduced stress and eased their transition to adoptive homes (Tuber & others, 1999). Thanks to animal behavior studies, formerly idle Bronx Zoo animals are now staving off listless boredom by working for their supper as would their counterparts in the wild (Stewart, 2002). Studies have helped improve animal care and management not only in laboratories, shelters, and zoos but also in their natural habitats. By revealing our behavioral kinship with animals and the remarkable intelligence of some animals, experiments have also led to an increase in our empathy for them. At its best, a psychology concerned for humans and sensitive to animals serves the welfare of both.

Is It Ethical to Experiment on People?

If the image of animals or people receiving supposed electric shocks troubles you, you may find it a relief that most psychological research involves no such stress. Blinking lights, flashing words, and pleasant social interactions are the rule.

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

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

controlling violent behavior or studying mood swings. Such experiments wouldn't work if the participants knew all there was to know about the experiment beforehand. Either the procedures would be ineffective or the participants, wanting to be helpful, might try to confirm the researchers' predictions.

Ethical principles developed by the American Psychological Association (1992) and the British Psychological Society (1993) urge investigators to (1) obtain the informed consent of potential participants, (2) protect them from harm and discomfort, (3) treat information about individual participants confidentially, and (4) fully explain the research afterward. Moreover, most universities today screen research proposals through an ethics committee that safeguards the well-being of every participant.

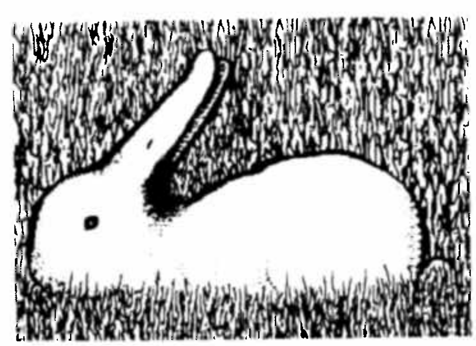
Much research, however, occurs outside of university laboratories, in places where there may be no ethics committees. For example, retail stores routinely survey people, photograph their purchasing behavior, track their buying patterns, and test the effectiveness of advertising. Curiously, such research attracts less attention than the scientific research done to advance human understanding.

Is Psychology Free of Value Judgments?

Psychology is definitely not value-free. Values affect what we study, how we study it, and how we interpret results. Consider: Researchers' values influence their choice of research topics—whether to study worker productivity or worker morale, sex discrimination or gender differences, conformity or independence. Values can even color "the facts." Our preconceptions can bias our observations and interpretations; sometimes we see what we want or expect to see (**FIGURE 1.13**). Even the words we use to describe a phenomenon can reflect our values. Labeling the sex acts we do not practice as "perversions" or as "sexual variations" conveys a value judgment. The same holds true in everyday speech, when one person's "rigidity" is another's "consistency," or one person's "faith" is another's "fanaticism." Our labeling someone as "firm" or "stubborn," "careful" or "picky," "discreet" or "secretive" reveals our feelings. Both in and out of psychology, labels describe and labels evaluate.

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 certainly help us reach our goals, but it cannot decide what those goals should be. (See Thinking Critically About the Death Penalty on pages 52–53.)

FIGURE 1.13
What do you see?
People interpret ambiguous information to fit their preconceptions. Did you see a duck or a rabbit? Before showing some friends this image, ask them if they can see the duck lying on its back (or the bunny in the grass). (From Shepard, 1990.)



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

Simone De Beauvoir, *The Second Sex*, 1953

The Death Penalty—When Beliefs Collide with Psychological Science

An influential modern viewpoint, ironically called *postmodernism*, questions scientific objectivity. Rather than mirroring the real world, say postmodernists, scientific concepts are socially constructed fictions. Like all knowledge, they reflect the culture that formed them. "Intelligence," for instance, is a concept psychologists created and defined. Because personal values guide theory and research, "truth" is actually personal and subjective. (What behaviors shall we call "intelligent"?) In our quest for truth, we cannot help following our hunches, our biases, our cultural bent.

Psychological scientists agree that many important questions lie beyond the reach of science. And they agree that personal beliefs often shape perceptions. But they also believe that there is a real world out there, and that we advance truth by checking our hunches against it. Marie Curie did not just construct the concept of radium, she *discovered* radium. It really exists. In the social sciences, pure objectivity, like pure love, may be unattainable. Yet most would argue that it is better to humble ourselves before reliable evidence than to cling to untested presumptions.

Letting go of presumptions is just what the U.S. Supreme Court justices did after 1950. They considered pertinent social science evidence and decided to disallow five-member juries and to end school desegregation. These very decisions helped inspire hundreds more studies that researchers hoped would inform future judicial decisions. But more recently the Court has joined postmodernists in discounting social science research. In deciding whether the death penalty falls under the Constitution's ban on "cruel and unusual punishment," the Court wrestled with whether society defines execution as cruel and unusual, whether courts inflict the penalty arbitrarily, whether they apply it with racial bias, and whether execution deters crime more than all other available punishments. The social science answers to each of these questions, note psychologists Mark Costanzo (1997) and Craig Haney and Deana Logan (1994), could hardly be clearer. And yet, on two of these issues—the fairness of the death penalty and its effectiveness—the Court has disregarded social science research.

AP/Wide World Photos

Executing the childlike
When granted a stay of execution in November 2000, with three hours to spare, Johnny Paul Penry's immediate concern was whether he would lose his promised last meal of a cheeseburger and fries. Penry, the son of an absent father who taunted him as retarded and a mother who abused and tormented him, has the mental ability of a 7-year-old. In 2002, he was again sentenced to death.

Is Psychology Potentially Dangerous?

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? Exploring the external universe is one thing, but exploring our own inner universe seems even more dangerous and threatening. Might psychology 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 ed-

Is the Death Penalty Applied Fairly?

Should it be permissible to execute a person with mental retardation—someone having the mental age of a 7-year-old, as in the case of Johnny Paul Penry? Attitudes toward capital punishment tend to follow a nation's legal practice. The death penalty is therefore mostly favored by Americans and opposed by those in many other nations (as readers in Canada, Western Europe, Australia, New Zealand, and most of South America will recognize). Nevertheless, public opinion surveys show Americans are overwhelmingly opposed to executing people with mental retardation. Some justices have dismissed such surveys, preferring instead to trust state legislation and jury decisions as indicators of public attitudes. However, studies show that those eligible to serve as jurors in capital punishment cases—those who accept the death penalty—do not represent the greater population. Compared with people excluded by virtue of their qualms about capital punishment, those chosen as jurors are less likely to be minorities and women. They are also more likely to believe the prosecution's arguments, and they are more conviction-prone.

"The state asks that the defendant, although a tadpole, be charged as a frog."

The Court has accepted social science evidence that a 15-year-old is too immature emotionally and too vulnerable to peer pressure for the death penalty to be appropriate. Yet, without explanation, it ignored the very same body of evidence when it decided that a 16-year-old, and even someone with the mental ability of a 7-year-old, could be executed.

Does the Death Penalty Work—Does It Deter Crime?

The evidence is consistent: States with a death penalty do not have lower homicide rates. After instituting the death penalty, these states did not see their rates drop. And homicide has not risen in states that have abandoned the death penalty. A person committing a crime of passion doesn't pause to calculate the consequences (and, if she or he did, would likely consider life in a prison cell an ample deterrent). Yet the Court persists in its belief that "the death penalty undoubtedly is a significant deterrent."

Beliefs guide perceptions. And that, say psychological scientists responding to postmodernists, is why we *need* to think smarter—to restrain our hunches, our biases, and our cultural leanings by checking them against available evidence. Why not put our testable beliefs to the test? If they find support, so much the better for them. If they collide against a wall of observation, so much the worse for them. These ideals of skeptical scrutiny and humility fuel all scientific endeavor.

ucate people—and to deceive them. The power of mind-altering drugs has been used to restore sanity—and to destroy it.

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 also speaks to many of our world's great problems—war, overpopulation, prejudice, family dysfunction, crime—all of which involve attitudes and behaviors. And psychology speaks to our deepest longings—for nourishment, for love, for happiness. True, psychology cannot address all of life's great questions, but it speaks to some mighty important ones.

REVIEW AND REFLECT

Frequently Asked Questions About Psychology

Can Laboratory Experiments Illuminate Everyday Life?

By intentionally creating a controlled, artificial environment in the lab, researchers aim to test theoretical principles. These principles help us to understand, describe, explain, and predict everyday behaviors.

Does Behavior Depend on One's Culture?

Attitudes and behaviors do vary across cultures, but the principles that underlie them vary much less. Cross-cultural psychology explores both our cultural differences and the universal similarities that define our human kinship.

Does Behavior Vary with Gender?

Gender is a basic fact of life. Although gender differences tend to capture attention, it is important to remember our greater gender similarities.

Why Do Psychologists Study Animals?

Some psychologists study animals out of an interest in animal behavior. Others do so because knowledge of the physiological and psychological processes of animals gives them a better understanding of the similar processes operating in humans.

Is It Ethical to Experiment on Animals?

Only about 7 percent of all psychological experiments involve animals, and under ethical and legal guidelines these animals rarely experience pain. Nevertheless, animal rights groups raise an important issue: Even if it leads to the relief of human suffering, is an animal's temporary suffering justified?

Is It Ethical to Experiment on People?

Occasionally researchers temporarily stress or deceive people in order to learn something important. Professional ethical standards provide guidelines concerning the treatment of both human and animal participants.

Is Psychology Free of Value Judgments?

Psychology is not value-free. Psychologists' own values influence their choice of research topics, their theories and observations, their labels for behavior, and their professional advice.

Is Psychology Potentially Dangerous?

Knowledge is power that can be used for good or evil. Applications of psychology's principles have so far been overwhelmingly for the good. Psychology addresses some of humanity's greatest problems and deepest longings.

CHECK YOURSELF: How are human and animal research subjects protected?

ASK YOURSELF: Were any of these Frequently Asked Questions your questions? Do you have other questions or concerns about psychology?

Answers to the Check Yourself questions can be found in the yellow appendix at the end of the book.

TERMS AND CONCEPTS TO REMEMBER

hindsight bias, p. 20
critical thinking, p. 23
theory, p. 24
hypothesis, p. 24
operational definition, p. 25
replication, p. 25
case study, p. 26
survey, p. 27
false consensus effect, p. 28
population, p. 28
random sample, p. 28
naturalistic observation, p. 29
correlation coefficient, p. 30
scatterplot, p. 31
illusory correlation, p. 33
experiment, p. 38
double-blind procedure, p. 39
placebo effect, p. 39
experimental condition, p. 39
control condition, p. 39
random assignment, p. 39
independent variable, p. 39
dependent variable, p. 39
mode, p. 43
mean, p. 43
median, p. 43
range, p. 43
standard deviation, p. 44
statistical significance, p. 45
culture, p. 47

To continue your study and review of Thinking Critically with Psychological Science, visit this book's Web site at www.worthpublishers.com/myers. You will find practice tests, review activities, and many interesting articles and Web links for more information on topics related to Thinking Critically with Psychological Science.