Unit 2

Research Methods: Thinking Critically with Psychological Science

1: Why are the answers that flow from the scientific approach more reliable than those based on intuition and common sense?

SOME PEOPLE SUPPOSE that psychology merely documents and dresses in jargon what people already know: "So what else is new—you get paid for using fancy methods to prove what everyone knows?" Others place their faith in human intuition: "Buried deep within each and every one of us, there is an instinctive, heart-felt awareness that provides—if we allow it to—the most reliable guide," offered Prince Charles (2000). "I know there's no evidence that shows the death penalty has a deterrent effect," George W. Bush (1999) reportedly said as Texas governor, "but I just feel in my gut it must be true." "I'm a gut player. I rely on my instincts," said the former president in explaining to Bob Woodward (2002) his decision to launch the Iraq war.

Prince Charles and former President Bush have much company. A long list of pop psychology books encourage us toward "intuitive managing," "intuitive trading," "intuitive healing," and much more. Today's psychological science does document a vast intuitive mind. As we will see, our thinking, memory, and attitudes operate on two levels, conscious and unconscious, with the larger part operating automatically, offscreen. Like jumbo jets, we fly mostly on autopilot.

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

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

This much seems certain. Intuition is important, but we often underestimate its perils. My geographical intuition tells me that Reno is east of Los Angeles, that Rome is south of New York, that Atlanta is east of Detroit. But I am wrong, wrong, and wrong.

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

Indeed, observed Madeleine L’Engle, “The naked intellect is an extraordinarily inaccurate instrument” (1973). Two phenomena—hindsight bias and judgmental overconfidence— illustrate why we cannot rely solely on intuition and common sense.

2.1.1 Did We Know It All Along? Hindsight Bias

How easy it is to seem astute when drawing the bull’s eye after the arrow has struck. After the North Tower of New York’s World Trade Center was hit on September 11, 2001, commentators said people in the South Tower *should* have immediately evacuated. (It became obvious only later that the strike was not an accident.) After the U.S. occupation of Iraq led to a bloody civil war rather than a peaceful democracy, commentators saw the result as inevitable. *Before* the invasion was launched, these results seemed anything but obvious: In voting to allow the Iraq invasion, most U.S. senators did not anticipate the chaos that would seem so predictable in hindsight. Finding that something has happened makes it seem inevitable, a tendency we call **hindsight bias** (also known as the *I-knew-it-all-along phenomenon*).

“He who trusts in his own heart is a fool.”

Proverbs 28:28

This phenomenon is easy to demonstrate: Give half the members of a group some purported psychological finding, and give the other half an opposite result. Tell the first group, “Psychologists have found that separation weakens romantic attraction. As the saying goes, ‘Out of sight, out of mind.’” Ask them to imagine why this might be true. Most people can, and nearly all will then regard this true finding as unsurprising.

“Life is lived forwards, but understood backwards.”

Philosopher Søren Kierkegaard, 1813–1855

Tell the second group the opposite, “Psychologists have found that separation strengthens romantic attraction. As the saying goes, ‘Absence makes the heart grow fonder.’” People given this untrue result can also easily imagine it, and 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.

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

"Anything seems commonplace, once explained."

Dr. Watson to Sherlock Holmes

Hindsight bias *After* the horror of 9/11, it seemed obvious that the U.S. intelligence analysts should have taken advance warnings more seriously, that airport security should have anticipated box-cutter–wielding terrorists, that occupants of the South Tower of the World Trade Center should have known to play it safe and leave. With 20/20 hindsight, everything seems obvious. Thus we now spend billions to protect ourselves against what the terrorists did last time. AP Photo/*The Roanoke TImes*, Matt Gentry

Hindsight bias is widespread. Some 100 studies have observed it in various countries and among both children and adults (Blank et al., 2007). Nevertheless, our intuition is often right. As Yogi Berra once said, "You can observe a lot by watching." (We have Berra to thank for other gems, such as "Nobody ever comes here—it's too crowded," and "If the people don't want to come out to the ballpark, nobody's gonna stop 'em.") Because we're all behavior watchers, it would be 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 **Unit 8A** calls a deep "need to belong"). Indeed, note Daniel Gilbert, Brett Pelham, and Douglas Krull (2003), "good ideas in psychology usually have an oddly familiar quality, and the moment we encounter them we feel certain that we once came close to thinking the same thing ourselves and simply failed to write it down." Good ideas are like good inventions; once created, they seem obvious.

But sometimes our intuition, informed by countless casual observations, has it wrong. In later units we will see how research has overturned popular ideas—

that familiarity breeds contempt, that dreams predict the future, and that emotional reactions coincide with menstrual phase. (See also **Table 2.1**.) We will also see how it has surprised us with discoveries about how the brain's chemical messengers control our moods and memories, about other animals' abilities, and about the effects of stress on our capacity to fight disease.

2.1.2 Overconfidence

We humans tend to be overconfident. As **Unit 7B** 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

About how many seconds do you think it would have taken you to unscramble each of these?

Once people know the answer, 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 problem solver spends 3 minutes, as you also might, given a similar anagram without the solution: OCHSA. (See margin in the next section to check your answer.)

"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

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 only 71 percent correct. Even when students were 100 percent sure of themselves, their self-predictions erred 15 percent of the time.

"Computers in the future may weigh no more than 1.5 tons."

Popular Mechanics, 1949

It’s not just students. Ohio State University psychologist Philip Tetlock (1998, 2005) has collected more than 27,000 expert predictions of world events, such as the future of South Africa or whether Quebec would separate from Canada. His repeated finding: These predictions, which experts made with 80 percent confidence on average, were right less than 40 percent of the time. Nevertheless, even those who erred maintained their confidence by noting they were “almost right.” “The Québécois separatists *almost* won the secessionist referendum.”

“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

The point to remember: Hindsight bias and overconfidence often lead us to overestimate our intuition. But scientific inquiry can help us sift reality from illusion.

“They couldn’t hit an elephant at this distance.”

General John Sedgwick just before being killed during a U.S. Civil War battle, 1864

2.1.3 The Scientific Attitude

: What are three main components of 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 an idea sounds, the critical thinker’s question is, *Does it work?* When put to the test, can its predictions be confirmed?

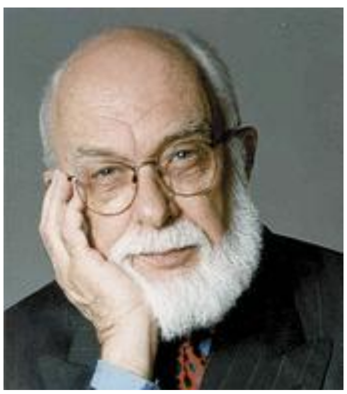
“The scientist…must be free to ask any question, to doubt any assertion, to seek for any evidence, to correct any errors.”

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

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). By letting the facts speak for themselves, Moses was using what we now call an *empirical approach*. Magician James Randi uses this 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 has told me that no aura-seer has agreed to take this simple test.



The Amazing Randi The magician James Randi exemplifies skepticism. He has tested and debunked a variety of psychic phenomena. Courtesy of the James Randi Education Foundation

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 turns jeers into cheers.

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

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

When ideas compete, skeptical testing can reveal which ones best match the facts. Do parental behaviors determine children's sexual orientation? Can astrologers predict your future based on the position of the planets at your birth? As we will see, putting such claims to the test has led psychological scientists to doubt them.

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

"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

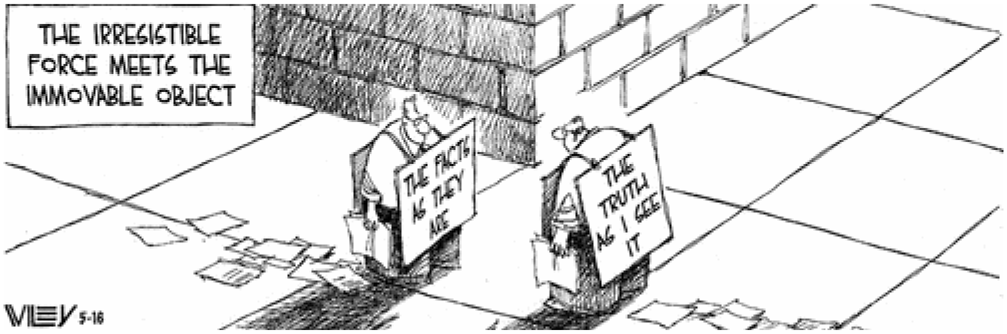
Historians of science tell us that these three attitudes—curiosity, skepticism, and humility—helped make modern science possible. Many of its founders, including Copernicus and Newton, were people whose religious convictions made them humble before nature and skeptical of mere human authority (Hooykaas, 1972; Merton, 1938). Some deeply religious people today may view science, including psychological science, as a threat. Yet, notes sociologist Rodney Stark (2003a,b), the scientific revolution was led mostly by deeply religious people acting on the idea that "in order to love and honor God, it is necessary to fully appreciate the wonders of his handiwork."

"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

Solution to anagram on previous page: CHAOS.

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. Nevertheless, 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.



Non Sequitur Reprinted by permission of Universal Press Syndicate. © 1997 Wiley.

2.1.4 Critical Thinking

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

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

- massive losses of brain tissue early in life may have minimal long-term effects (see **Unit 3B**).
- within days, newborns can recognize their mother's odor and voice (see **Unit 9**).
- brain damage can leave a person able to learn new skills yet unaware of such learning (see **Unit 7B**).
- diverse groups—men and women, old and young, rich and middle class, those with disabilities and those without—report roughly comparable levels of personal happiness (see **Unit 8B**).
- electroconvulsive therapy (delivering an electric shock to the brain) is often a very effective treatment for severe depression (see **Unit 13**).

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

- sleepwalkers are *not* acting out their dreams (see **Unit 5**).
- our past experiences are *not* all recorded verbatim in our brains; with brain stimulation or hypnosis, one *cannot* simply "hit the replay button" and relive long-buried or repressed memories (see **Unit 7A**).

- most people do *not* suffer from unrealistically low self-esteem, and high self-esteem is not all good (see **Unit 10**).
- opposites do *not* generally attract (see **Unit 14**).

In each of these instances and more, what has been learned is not what is widely believed.

"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

BEFORE YOU MOVE ON...

ASK YOURSELF

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

TEST YOURSELF 1

What is the scientific attitude, and why is it important for critical thinking?

2.2 How Do Psychologists Ask and Answer Questions?

PSYCHOLOGISTS ARM THEIR SCIENTIFIC attitude with the *scientific method*. Psychological science evaluates competing ideas with careful observation and rigorous analysis. In its attempt to describe and explain human nature, it welcomes hunches and plausible-sounding theories. And it puts them to the test. If a theory works—if the data support its predictions—so much the better for that theory. If the predictions fail, the theory will be revised or rejected.

2.2.1 The Scientific Method

3: How do theories advance psychological science?

In everyday conversation, we often use *theory* to mean "mere hunch." In science, however, theory is linked with observation. A scientific **theory** *explains* through an integrated set of principles that *organizes* observations and *predicts* behaviors or events. By organizing isolated facts, a theory simplifies. 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. As we connect the observed dots, a coherent picture emerges.

A good theory of depression, for example, helps us organize countless depression-related observations into a short list of principles. Imagine that we observe over and over that people with depression describe their past, present, and future in gloomy terms. We might therefore theorize that at the heart of depression lies low self-esteem. 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 produces testable predictions, called **hypotheses**. By enabling us to test and to 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 selfesteem theory of depression, we might assess people's self-esteem by having them respond to statements such as "I have good ideas" and "I am fun to be with." Then we could see whether, as we hypothesized, people who report poorer self-images also score higher on a depression scale (**Figure 2.1**).

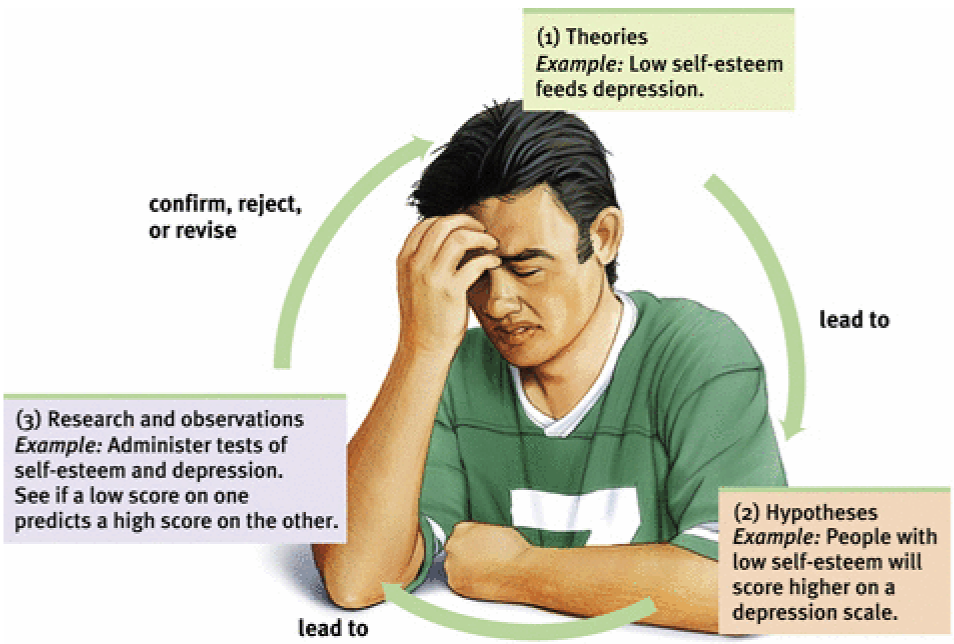


Figure 2.1 The scientific method A self-correcting process for asking questions and observing nature's answers.

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. The urge to see what we expect is an ever-present temptation, in the laboratory and outside of it. Perhaps you are aware of students who, because they have developed an excellent reputation, can now do no wrong in the eyes of teachers. If they're in the hall during class, nobody worries. Other students can do no good. Because they have behaved badly in the past, even their positive behaviors are viewed suspiciously.

As a check on their biases, psychologists report their research with precise **operational definitions** of procedures and concepts. Unlike dictionary definitions, operational definitions describe concepts with precise procedures or measures. *Hunger*, for example, might be defined as "hours without eating," *generosity* as "money contributed." Such carefully worded statements should allow others to **replicate** (repeat) the original 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.

Good theories explain by

(1) organizing and linking observed facts, and

(2) implying hypotheses that offer testable predictions and, sometimes, practical applications.

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 in **Unit 12**) that better organizes and predicts what we know about depression.

As we will see next, we can test our hypotheses and refine our theories using *descriptive* methods (which describe behaviors, often using case studies, surveys, or naturalistic observations), *correlational* methods (which associate different factors), and *experimental* methods (which manipulate factors to discover their effects). To think critically about popular psychology claims, we need to recognize these methods and know what conclusions they allow.

2.2.2 Description

4: How do psychologists observe and describe behavior?

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, though more objectively and systematically.

The Case Study

The case of the conversational chimpanzee In case studies of chimpanzees, psychologists have asked 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 using language? Psychologists debate that issue. Susan Kuklin/Photo Researchers

Among the oldest research methods, the **case study** examines one individual in depth in hopes of revealing things true of us all. Some examples: Much of our early knowledge about the brain came from case studies of individuals who suffered a particular impairment after damage to a certain brain region. Jean Piaget taught us about children's thinking after carefully observing and questioning only a few children. Studies of only a few chimpanzees have revealed their capacity for understanding and language. Intensive case studies are sometimes very revealing.

Case studies often suggest directions for further study, and they show us what *can* happen. But individual cases may mislead us if the individual being studied is atypical. Unrepresentative information can lead to mistaken judgments and false conclusions. Indeed, anytime a researcher mentions a finding ("Smokers die younger: 95 percent of men over 85 are nonsmokers") someone is sure to offer a contradictory anecdote ("Well, I have an uncle who smoked two packs a day and lived to be 89"). Dramatic stories and personal experiences (even psychological case examples) command our attention, and they are easily remembered. Which of the following do you find more memorable? (1) "In one study of 1300 dream reports concerning a kidnapped child, only 5 percent correctly envisioned the child as dead (Murray & Wheeler, 1937)." (2) "I know a man who dreamed his sister was in a car accident, and two days later she died in a head-on collision!" Numbers can be numbing, but the plural of *anecdote* is not *evidence*. As psychologist Gordon Allport (1954, p. 9) said, "Given a thimbleful of [dramatic] facts we rush to make generalizations as large as a tub."

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

"'Well my dear,' said Miss Marple, 'human nature is very much the same everywhere, and of course, one has opportunities of observing it at closer quarters in a village.'"

Agatha Christie, *The Tuesday Club Murders*, 1933

The Survey

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 **survey** method looks at many cases in less depth. Researchers do surveys when wanting to estimate, from a representative sample of people, the attitudes or reported behaviors of a whole population. Questions about everything from cell-phone use to political opinions are put to the public. Harris and Gallup polls have revealed that 72 percent of Americans think there is too much TV violence, 89 percent favor equal job opportunities for homosexual people, 89 percent are facing high stress, and 96 percent would like to change something about their appearance. In Britain, seven in ten 18- to 29-year-olds recently supported gay marriage; among those over 50, about the same ratio opposed it (a generation gap found in many Western countries). But asking questions is tricky, and the answers often depend on the ways questions are worded and respondents are chosen.

This Modern World by Tom Tomorrow © 1991

Wording Effects 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 one 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 is such a delicate matter, critical thinkers will reflect on how the phrasing of a question might affect people's expressed opinions.

Random Sampling We can describe human experience by drawing on memorable anecdotes and personal experience. But for an accurate picture of a whole population's attitudes and experience, 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, especially vivid cases. Given (a) a statistical summary of auto owners' evaluations of their car's make and (b) the vivid comments of two frustrated owners, one's impression may be influenced as much by the two unhappy owners as by the many more evaluations in the statistical summary. The temptation to generalize from a few vivid but unrepresentative cases is nearly irresistible.

"How would you like me to answer that question? As a member of my ethnic group, educational class, income group, or religious category?" Drawing by D. Fradon; © 1969 The New Yorker Magazine, Inc.

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

So how do you obtain a representative sample—say, of the students at your high school? How could you choose a group that would represent the total student **population**, the whole group you want to study and describe? Typically, you would choose a **random sample**, in which every person in the entire group has an equal chance of participating. This means you would not send each student a questionnaire. (The conscientious people who return it would not be a random sample.) Rather, you might number the names in the general student listing and then use a random number generator to pick the participants for your survey. Large representative samples are better than small ones, but a small representative sample of 100 is better than an unrepresentative sample of 500.

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

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

Naturalistic Observation

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

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

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

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

- *A funny finding*. We humans laugh 30 times more often in social situations than in solitary situations. (Have you noticed how seldom you laugh when alone?) As we laugh, 17 muscles contort our mouth and squeeze our eyes, and we emit a series of 75-millisecond vowel-like sounds that are spaced about one-fifth of a second apart (Provine, 2001).

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

- *Sounding out students.* What, really, are introductory psychology students saying and doing during their everyday lives? To find out, Matthias Mehl and James Pennebaker (2003) equipped 52 such students from the University of Texas with electronically activated belt-worn tape recorders. For up to four days, the recorders captured 30 seconds of the students' waking hours every 12.5 minutes, thus enabling the researchers to eavesdrop on more than 10,000 half-minute life slices by the end of the study. On what percentage of the slices do you suppose they found the students talking with someone? What percentage captured the students at a computer? The answers: 28 and 9 percent. (What percentage of *your* waking hours are spent in these activities?)
- *Culture, climate, and the pace of life*. Naturalistic observation also enabled Robert Levine and Ara Norenzayan (1999) to compare the pace of life in 31 countries. (Their operational definition of *pace of life* included walking speed, the speed with which postal clerks completed a simple request, and the accuracy of public clocks.) Their conclusion: Life is fastest paced in Japan and Western Europe, and slower paced in economically less-

developed countries. People in colder climates also tend to live at a faster pace (and are more prone to die from heart disease).

Naturalistic observation offers interesting snapshots of everyday life, but it does so without controlling for all the factors that may influence behavior. It’s one thing to observe the pace of life in various places, but another to understand what makes some people walk faster than others. Yet naturalistic observation, like surveys, can provide data for correlational research, which we consider next.

2.2.3 Correlation

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

Figure 2.2 Scatterplots, showing patterns of correlation Correlations can range from +1.00 (scores on one measure increase in direct proportion to scores on another) to –1.00 (scores on one measure decrease precisely as scores rise on the other).

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

Throughout this book we will often ask how strongly two sets of scores are related: For example, how closely related are the personality scores of identical twins? How well do intelligence test scores predict achievement? How closely is stress related to disease?

Table 2.2

Height and Temperament of 20 Men

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

Figure 2.2 contains three **scatterplots**, illustrating the range of possible correlations from a perfect positive to a perfect negative. (Perfect correlations rarely occur in the "real world.") Each dot in a scatterplot represents the scattered values of two variables. A correlation is positive if two sets of scores, such as height and weight, tend to rise or fall together. Saying that a correlation is "negative" says nothing about its strength or weakness. A correlation is negative if two sets of scores relate inversely, one set going up as the other goes down. Exercise and weight correlate negatively. As exercise goes up from zero, body weight has some tendency to go down. A weak correlation, indicating little relationship, has a coefficient near zero.

Here are four recent news reports of correlational research, some derived from surveys or natural observations. Can you spot which are reporting positive correlations, which negative? (Answers below.)

1. The more young children watch TV, the less they read (Kaiser, 2003).
2. The more sexual content teens see on TV, the more likely they are to have sex (Collins et al., 2004).
3. The longer children are breast-fed, the greater their later academic achievement (Horwood & Fergusson, 1998).
4. The more often adolescents eat breakfast, the lower their body mass (Timlin et al., 2008).

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

With all the relevant data (**Table 2.2**) 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 2.2**, 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 2.3**, moving from left to right, the upward, oval-shaped slope of the cluster of points shows that our two imaginary sets of scores (height and reactivity) tend to rise together.

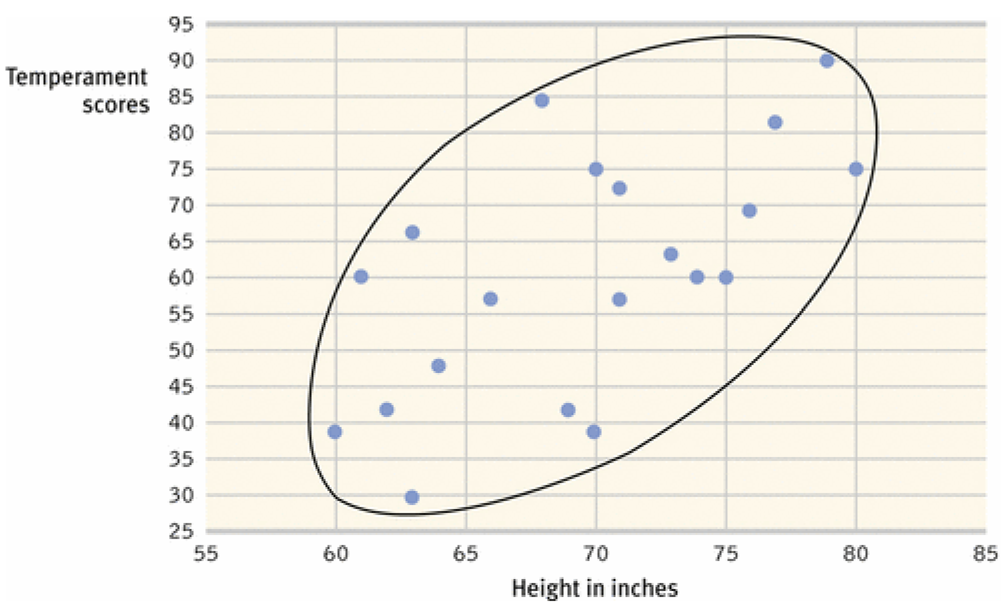


Figure 2.3 Scatterplot for height and temperament This display of data from 20 imagined men (each represented by a data point) reveals an upward slope, indicating a positive correlation. The considerable scatter of the data indicates the correlation is much lower than +1.0.

Answers to correlation question: 1. negative, 2. positive, 3. positive, 4. negative.

If we fail to see a relationship when data are presented as systematically as in **Table 2.2**, how much less likely are we to notice them in everyday life? To see what is right in front of us, we sometimes need statistical illumination. We can easily see evidence of gender discrimination when given statistically summarized information about job level, seniority, performance, gender, and salary. But we often see no discrimination when the same information dribbles in, case by case (Twiss et al., 1989).

The point to remember: A correlation coefficient, which can range from −1.0 to +1.0, reveals the extent to which two things relate.

Correlation and Causation

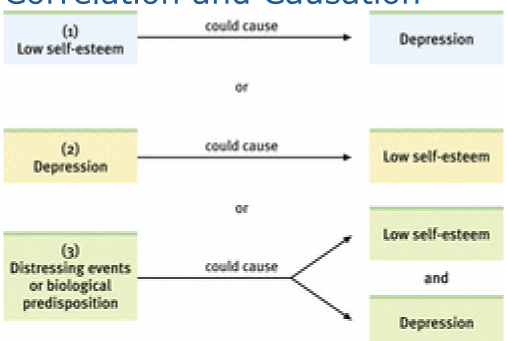
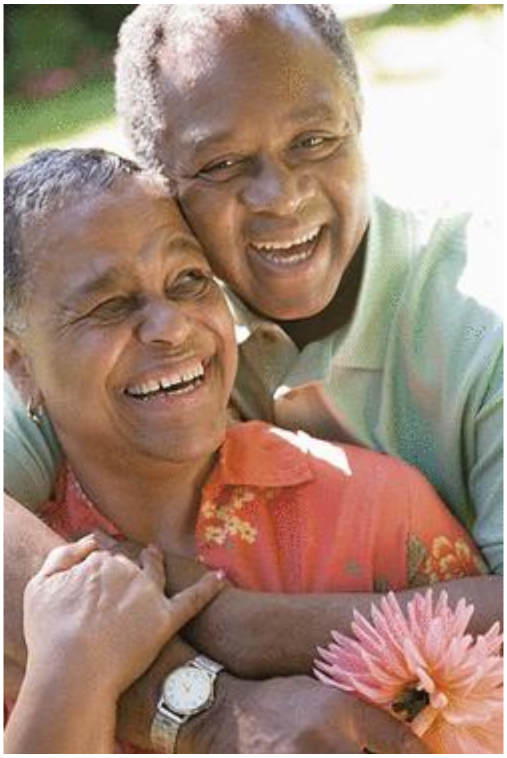


Figure 2.4 **Three possible cause-effect relationships** People low in self-esteem are more likely to report depression than are those high in self-esteem. One possible explanation of this negative correlation is that a bad self-image causes depressed feelings. But, as the diagram indicates, other cause-effect relationships are possible.

Correlations help us predict. Low self-esteem correlates with (and therefore predicts) depression. (This correlation might be indicated by a correlation coefficient, or just by a finding that people who score on the lower half of a self-esteem scale have an elevated depression rate.) So, does low self-esteem *cause* depression? If, based on the correlational evidence, you assume that it does, you have much company. A nearly irresistible thinking error is assuming that an association, sometimes presented as a correlation coefficient, proves causation. But no matter how strong the relationship, it does not *prove* anything!

As options 2 and 3 in **Figure 2.4** show, we'd get the same negative correlation between low self-esteem and depression if depression caused people to be down on themselves, or if some third factor—such as heredity or brain chemistry—caused both low self-esteem and depression. Among men, for example, length of marriage correlates positively with hair loss—because both are associated with a third factor, age.



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. Big Cheese Photo LLC/Alamy

This point is so important—so basic to thinking smarter with psychology—that it merits one more example, from a survey of more than 12,000 adolescents. The study found that the more teens feel loved by their parents, the less likely they are to behave in unhealthy ways—having early sex, smoking, abusing alcohol and drugs, exhibiting violence (Resnick et al., 1997). "Adults have a powerful effect on their children's behavior right through the high school years," gushed an Associated Press (AP) story reporting the finding. But this correlation comes with no built-in cause-effect arrow. Said differently (turn the volume up here), *association does not prove causation*.[2] Thus, the AP could as well have reported, "Well-behaved teens feel their parents' love and approval; out-of-bounds teens more often think their parents are disapproving jerks."

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

Illusory Correlations

6: What are illusory correlations?

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

Because we are sensitive to dramatic or unusual events, we are especially likely to notice and remember the occurrence of two such events in sequence—say, a premonition of an unlikely phone call followed by the call. When the call does not follow the premonition, we are less likely to note and remember the nonevent. Illusory correlations help explain many superstitious beliefs, such as the presumption that infertile couples who adopt become more likely to conceive (Gilovich, 1991). Couples who conceive after adopting capture our attention. We’re less likely to notice those who adopt and never conceive, or those who conceive without adopting. In other words, illusory correlations occur when we over-rely on the top left cell of **Figure 2.5**, ignoring equally essential information in the other cells.

Figure 2.5 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.) Michael Newman Jr./PhotoEdit

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

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

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

Perceiving Order in Random Events

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).

In our natural eagerness to make sense of our world—what poet Wallace Stevens called our "rage for order"—we look for order even in random data. And we usually find it, because—here's a curious fact of life—*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?

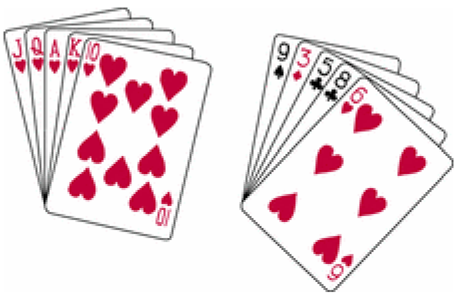


Figure 2.6 Two random sequences Your chances of being dealt either of these hands are precisely the same: 1 in 2,598,960.

Daniel Kahneman and Amos Tversky (1972) found that most people believe HTTHTH would be the most likely random sequence. Actually, all three are equally likely (or, you might say, equally unlikely). A poker hand of 10 through ace, all of hearts, would seem extraordinary; actually, it would be no more or less likely than any other specific hand of cards (**Figure 2.6**).

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

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

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. Similar streaks happen, about as often as one would expect in random sequences, in basketball shooting, baseball hitting, and mutual fund stock pickers' selections (Gilovich et al., 1985; Malkiel, 1989, 1995; Myers, 2002). These sequences often don't look random, and so get overinterpreted ("When you're hot, you're hot!").

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).

What explains these streaky patterns? Was I exercising some sort of paranormal control over my coin? Did I snap out of my tails funk and get in a heads groove? No such explanations are needed, for these are the sorts of streaks found in any random data. Comparing each toss to the next, 24 of the 50 comparisons yielded a changed result—just the sort of near 50-50 result we expect from coin tossing. Despite seeming patterns, the outcome of one toss gives no clue to the outcome of the next.

Bizarre-looking, perhaps. But actually no more unlikely than any other number sequence. © 1990 by Sidney Harris/*American Scientist magazine*.

However, some happenings seem so extraordinary that we struggle to conceive an ordinary, chance-related explanation (as applies to our coin tosses). In such cases, statisticians often are less mystified. When Evelyn Marie Adams won the New Jersey lottery *twice*, newspapers reported the odds of her feat as 1 in 17 trillion. Bizarre? Actually, 1 in 17 trillion are indeed the odds that a given person who buys a single ticket for two New Jersey lotteries will win both times. But statisticians Stephen Samuels and George McCabe (1989) reported 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, said fellow statisticians Persi Diaconis and Frederick Mosteller (1989), "with a large enough sample, any outrageous thing is likely to happen." An event that happens to but 1 in 1 billion people every day occurs about six times a day, 2000 times a year.

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

2.2.4 Experimentation

7: How do experiments, powered by random assignment, clarify cause and effect?

Happy are they, remarked the Roman poet Virgil, "who have been able to perceive the causes of things." To isolate cause and effect, psychologists can statistically control for other factors. For example, researchers have found that breast-fed infants grow up with somewhat higher intelligence scores than do infants bottle-fed with cow's milk (Angelsen et al., 2001; Mortensen et al., 2002; Quinn et al., 2001). They have also found that breast-fed British babies have been more likely than their bottle-fed counterparts to eventually move into a higher social class (Martin et al., 2007). But the "breast is best" intelligence effect shrinks when researchers compare breast-fed and bottle-fed children from the same families (Der et al., 2006).

So, does this mean that smarter mothers (who in modern countries more often breast-feed) have smarter children? Or, as some researchers believe, do the nutrients of mother's milk contribute to brain development? To help answer this question, researchers have "controlled for" (statistically removed differences in) certain other factors, such as maternal age, education, and income. And they have found that in infant nutrition, mother's milk correlates modestly but positively with later intelligence.

Correlational research cannot control for all possible factors. But researchers can isolate cause and effect with an **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*. With parental permission, a British research team randomly assigned 424 hospital preterm infants either to standard infant formula feedings or to donated breast milk feedings (Lucas et al., 1992). On intelligence tests taken at age 8, the children nourished with breast milk had significantly higher intelligence scores than their formula-fed counterparts. Earlier we mentioned the place of *random sampling* in a well-done survey. Consider now the equally important place of *random assignment* in a well-done experiment.

Random Assignment

No single experiment is conclusive, of course. But by **randomly assigning** infants to one feeding group or the other, researchers were able to hold constant all factors except nutrition. This eliminated alternative explanations and supported the conclusion that breast is indeed best for developing intelligence (at least for preterm infants).

If a behavior (such as test performance) changes when we vary an experimental factor (such as infant nutrition), then we infer the factor is having an effect. *The point to remember*: Unlike correlational studies, which uncover naturally occurring relationships, an experiment manipulates a factor to determine its effect.

Consider, too, how we might assess a therapeutic intervention. Our tendency to seek new remedies when we are ill or emotionally down can produce misleading testimonies. If three days into a cold we start taking vitamin C tablets and find our cold symptoms lessening, we may credit the pills rather than the cold naturally subsiding. If, after nearly failing the first test, we listen to a "peak learning" subliminal CD and then improve on the next test, we may credit the CD rather than conclude that our performance has returned to our average. In the 1700s, bloodletting *seemed* effective. Sometimes people improved after the treatment; when they didn't, the practitioner inferred the disease was just too advanced to be reversed. (We, of course, now know that usually bloodletting is a *bad* treatment.) 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 investigators evaluate new drug treatments and new methods of psychological therapy (**Unit 13**). The participants in these studies are randomly assigned to the research groups and are often *blind* (uninformed) about what treatment, if any, they are receiving. One group receives a treatment (such as medication or other therapy). The other group receives a pseudotreatment—an inert *placebo* (perhaps a pill with no drug in it). If the study is using a **double-blind procedure**, neither the participants nor the research assistants collecting the data will know which group is receiving the treatment. In such studies, researchers can check a treatment's actual effects apart from the participants' belief in its healing powers and the staff's enthusiasm for its potential. Just *thinking* you are getting a treatment can boost your spirits, relax your body, and relieve your symptoms. This **placebo effect** is well documented in reducing pain, depression, and anxiety (Kirsch & Sapirstein, 1998). And the more expensive the placebo, the more "real" it seems to us—a fake pill that costs $2.50 works better than one costing 10 cents (Waber et al., 2008). To know how effective a therapy really is, researchers must control for a possible placebo effect.

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

The double-blind procedure is one way to create an **experimental group**, in which people receive the treatment, and a contrasting **control group** that does not receive the treatment. By randomly assigning people to these conditions, researchers can be fairly certain the two groups are otherwise 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 also can conclude that any later differences between people in the experimental and control groups will usually be the result of the treatment.

Independent and Dependent Variables

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

This simple experiment manipulated just one factor: the drug dosage (none versus peak dose). We call this experimental factor the **independent variable** because we can vary it *independently* of other factors, such as the men's age, weight, and personality. These other factors, which can potentially influence the results of the experiment, are called **confounding variables.** Random assignment controls for possible confounding variables.

Experiments examine the effect of one or more independent variables on some measurable behavior, called the **dependent variable** because it can vary *depending* on what takes place during the experiment. Both variables are given precise *operational definitions*, which specify the procedures that manipulate the independent variable (the precise drug dosage and timing in this study) or measure the dependent variable (the questions that assessed the men's responses). These definitions answer the "What do you mean?" question with a level of precision that enables others to repeat the study. (See **Figure 2.7** for the breast milk experiment's design.)

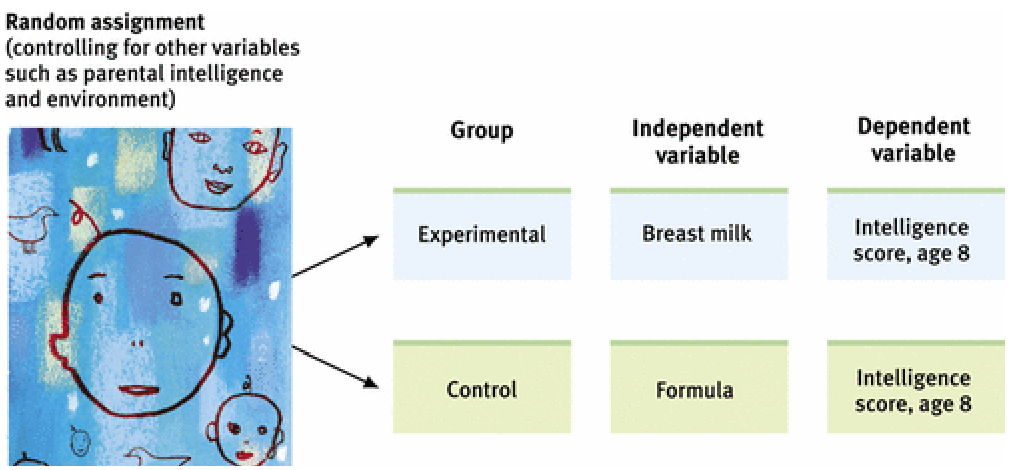


Figure 2.7 Experimentation To discern causation, psychologists may randomly assign some participants to an experimental group, others to a control group. Measuring the dependent variable (intelligence score in later childhood) will determine the effect of the independent variable (type of milk). © Michael Wertz

Note the distinction between random sampling in surveys (discussed earlier in relation to surveys) and random assignment in experiments (depicted in Figure 2.7). *Random sampling* helps us generalize to a larger population. *Random assignment* controls extraneous influences, which helps us infer cause and effect

Let's pause to check your understanding using a simple psychology experiment: To test the effect of perceived ethnicity on the availability of a rental house, Adrian Carpusor and William Loges (2006) sent identically worded e-mail inquiries to 1115 Los Angeles-area landlords. The researchers varied the ethnic connotation of the sender's name and tracked the percentage of positive replies (invitations to view the apartment in person). "Patrick McDougall," "Said Al-

Rahman," and "Tyrell Jackson" received, respectively, 89 percent, 66 percent, and 56 percent invitations. In this experiment, what was the independent variable? The dependent variable?[3]

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

Let's recap. A *variable* is anything that can vary (infant nutrition, intelligence, TV exposure—anything within the bounds of what is feasible and ethical). Experiments aim to *manipulate* an *independent* variable, *measure* the *dependent* variable, and *control* confounding variables. An experiment has at least two different groups: an *experimental group* and a *comparison* or *control group*. *Random assignment* works to equate the groups before any treatment effects. In this way, an experiment tests the effect of at least one independent variable (what we manipulate) on at least one dependent variable (the outcome we measure).

Table 2.3 compares the features of psychology's research methods.

Table 2.3

COMPARING RESEARCH METHODS

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

BEFORE YOU MOVE ON...

ASK YOURSELF

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

TEST YOURSELF 2

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

Show Answer

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

2.3 Statistical Reasoning in Everyday Life

IN DESCRIPTIVE, CORRELATIONAL, and experimental research, statistics are tools that help us see and interpret what the unaided eye might miss. But statistical understanding benefits more than just researchers. To be an educated person today is to be able to apply simple statistical principles to everyday reasoning. One needn't memorize complicated formulas to think more clearly and critically about data.

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

- *Ten percent of people are lesbians or gay men*. Or is it 2 to 3 percent, as suggested by various national surveys (**Unit 8A**)?
- *We ordinarily use but 10 percent of our brain*. Or is it closer to 100 percent (**Unit 3B**)?
- *The human brain has 100 billion nerve cells*. Or is it more like 40 billion, as suggested by extrapolation from sample counts (**Unit 3A**)?

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

2.3.1 Describing Data

8: How can we describe data with measures of central tendency and variation?

Once researchers have gathered their data, they must organize them in some meaningful way. One way to do this is to convert the data into a simple *bar graph*, as in **Figure 2.8** which displays a distribution of different brands of trucks still on the road after a decade. When reading statistical graphs such as this, take care. It's easy to design a graph to make a difference look big (**Figure 2.8a**) or small (**Figure 2.8b**). The secret lies in how you label the vertical scale (the Y-axis).

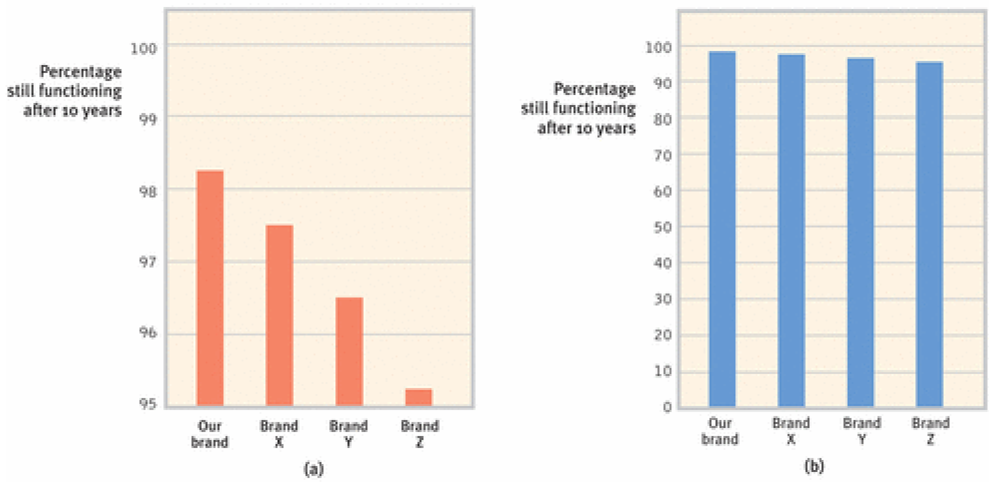


Figure 2.8 **Read the scale labels** An American truck manufacturer offered graph (a)—with actual brand names included—to suggest the much greater durability of its trucks. Note, however, how the apparent difference shrinks as the vertical scale changes in graph (b).

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

Measures of Central Tendency

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

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 2.9**). This happens because the mean is biased by a few extreme scores. When Microsoft co-founder Bill Gates sits down in an intimate café, its average (mean) customer instantly becomes a billionaire. But the customer's median wealth remains unchanged. Understanding this, you can see how a British newspaper could accurately run the headline "Income for 62% Is Below Average" (Waterhouse, 1993). Because the bottom *half* of British income earners receive only a *quarter* of the national income cake, most British people, like most people everywhere, make less than the mean. In the United States, Republicans have tended to tout the economy's solid growth since 2000 using average income; Democrats have lamented the economy's lackluster growth using median income (Paulos, 2006). Mean and median tell different true stories.

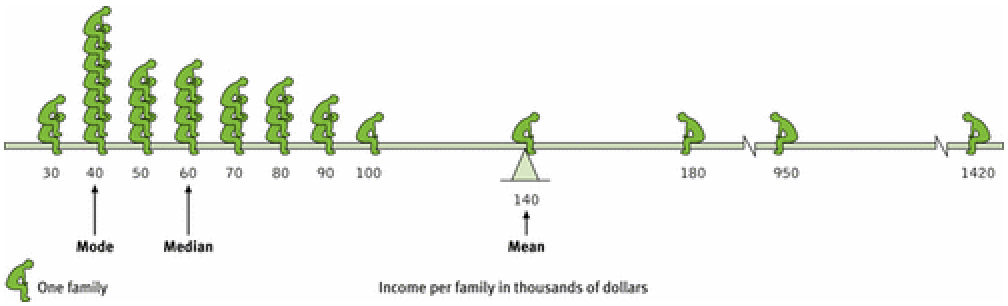


Figure 2.9 **A skewed distribution** This graphic representation of the distribution of a village's incomes illustrates the three measures of central tendency—mode, median, and mean. Note how just a few high incomes make the mean—the fulcrum point that balances the incomes above and below—deceptively high.

The average person has one ovary and one testicle.

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.

Measures of Variation

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

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 $950,000 and $1,420,000 incomes in **Figure 2.9**, will create a deceptively large range.

The more useful standard for measuring how much scores deviate from one another is the **standard deviation**. It better gauges whether scores are packed together or dispersed, because it uses information from each score (**Table 2.4**). The computation assembles information about how much individual scores differ from the mean. If your high school serves a community where most families have similar incomes, family income data will have a relatively small standard deviation compared with a school in a more diverse community population.

Table 2.4

Standard Deviation Is Much More Informative Than Mean Alone

Note that the test scores in Class A and Class B have the same mean (80), but very different standard deviations, which tell us more about how the students in each class are really faring.

Test Scores in Class A			Test Scores in Class B		
Score	Deviation from the Mean	Squared Deviation	Score	Deviation from the Mean	Squared Deviation
72	-8	64	60	-20	400
74	-6	36	60	-20	400
77	-3	9	70	-10	100
79	-1	1	70	-10	100
82	+2	4	90	+10	100
84	+4	16	90	+10	100
85	+5	25	100	+2-	400
87	+7	49	100	+20	400
Total = 640 Mean = 640 ÷ 8 = 80		Sum of (deviations)² = 204	Total = 640 Mean = 640 ÷ 8 = 80	Sum of (deviations)² = 2000	

Standard deviation =

$$\sqrt{\frac{\textit{Sum of (deviations)}^2}{\textit{Number of scores}}} = \sqrt{\frac{204}{8}} = 5.0$$

Standard deviation =

$$\sqrt{\frac{\textit{Sum of (deviations)}^2}{\textit{Number of scores}}} = \sqrt{\frac{2000}{8}} = 15.8$$

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

As **Figure 2.10** shows, a useful property of the normal curve is that roughly 68 percent of the cases fall within one standard deviation on either side of the mean. About 95 percent of cases fall within two standard deviations. Thus, **Unit 11** notes that about 68 percent of people taking an intelligence test will score within ±15 points of 100. About 95 percent will score within ±30 points.

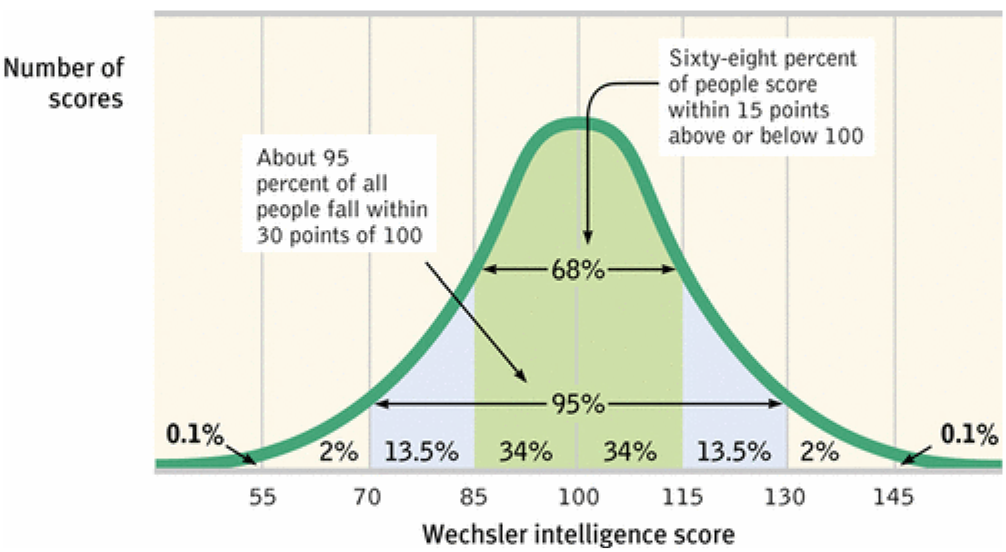


Figure 2.10 The normal curve Scores on aptitude tests tend to form a normal, or bell-shaped, curve. For example, the Wechsler Adult Intelligence Scale calls the average score 100.

2.3.2 Making Inferences

9: What principles can guide our making generalizations from samples and deciding whether differences are significant?

Data are “noisy.” The average score in one group (breast-fed babies) could conceivably differ from the average score in another group (formula-fed babies) not because of any real difference but merely because of chance fluctuations in the people sampled. How confidently, then, can we infer that an observed difference accurately estimates the true difference? For guidance, we can ask how reliable and significant the differences are.

When Is an Observed Difference Reliable?

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

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

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

When Is a Difference Significant?

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

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

In judging statistical significance, psychologists are conservative. They are like juries who must presume innocence until guilt is proven. For most psychologists, proof beyond a reasonable doubt means not making much of a finding unless the odds of its occurring by chance 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 hundreds of thousands of first-born and later-born individuals indicate a highly significant tendency for first-born individuals to have higher average scores than their later-born siblings (Kristensen & Bjerkedal, 2007; Zajonc & Markus, 1975). But because the scores differ by only one to three points, the difference has little practical importance. 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 *effect size*—its magnitude and reliability.

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

BEFORE YOU MOVE ON...

ASK YOURSELF

Find a graph in a popular magazine ad. How does the advertiser use (or abuse) statistics to make a point?

TEST YOURSELF 3

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 grades at the end of their first term at the University. However, only about 10 to 15 students graduate with perfect grades. What do you think is the most likely explanation for the fact that there are more perfect grades after one term than at graduation?

2.4 Frequently Asked Questions About Psychology

WE HAVE REFLECTED ON HOW A scientific approach can restrain biases. We have seen how case studies, surveys, and naturalistic observations help us describe behavior. We have also noted that correlational studies assess the association between two factors, which indicates how well one thing predicts another. We have examined the logic that underlies experiments, which use control conditions and random assignment of participants to isolate the effects of an independent variable on a dependent variable. And we have considered how statistical tools can help us see and interpret the world around us.

Yet, even knowing this much, you may still be approaching psychology with a mixture of curiosity and apprehension. So before we plunge in, let's entertain some frequently asked questions.

2.4.1 Psychology Applied

10: Can laboratory experiments illuminate everyday life?

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

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

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

When psychologists apply laboratory research on aggression to actual violence, they are applying theoretical principles of aggressive behavior, principles they have refined through many experiments. Similarly, it is the principles of the visual system, developed from experiments in artificial settings (such as looking at red lights in the dark), that 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 et al., 1999).

The point to remember: Psychologists' concerns lie less with unique behaviors than with discovering general principles that help explain many behaviors.

11: Does behavior depend on one's culture and gender?

Soccer shoes? Because culture shapes social behavior, actions that seem ordinary to others may seem odd to us. Yet underlying these differences are powerful similarities. Children everywhere love to play sports such as soccer. But most U.S. children would play with athletic shoes on grass, not barefoot on the street, as with these Burkina Faso boys. Alistair Berg/Alamy

What can psychological studies done in one time and place, often with White Europeans or 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 shapes our behavior. It influences our standards of promptness and frankness, our attitudes toward premarital sex and varying body shapes, our tendency to be casual or formal, our willingness to make eye contact, our conversational distance, and much, much more. Being aware of such differences, we can restrain our assumptions that others will think and act as we do. Given the growing mixing and clashing of cultures, our need for such awareness is urgent.

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

- People diagnosed with dyslexia, a reading disorder, exhibit the same brain malfunction whether they are Italian, French, or British (Paulesu et al., 2001).
- Variation in languages may impede communication across cultures. Yet all languages share deep principles of grammar, and people from opposite hemispheres can communicate with a smile or a frown.

- People in different cultures vary in feelings of loneliness. But across cultures, loneliness is magnified by shyness, low self-esteem, and being unmarried (Jones et al., 1985; Rokach et al., 2002).

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

You will see throughout this book that gender matters, too. Researchers report gender differences in what we dream, in how we express and detect emotions, and in our risk for alcohol dependence, depression, and eating disorders. Gender differences fascinate us, and studying them is potentially beneficial. For example, many researchers believe that women carry on conversations more readily to build relationships, while men talk more to give information and advice (Tannen, 1990). Knowing this difference can help us prevent conflicts and misunderstandings in everyday relationships.

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

“All people are the same; only their habits differ.”

Confucius, 551–479 B.C.E.

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

2.4.2 Ethics in Research

12: Why do psychologists study animals, and is it ethical to experiment on 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 permissible only with animals. Human physiology resembles that of many other animals. We humans are not *like* animals; we *are* 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.

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

Dave Barry, July 2, 2002

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 a lawn mower’s engine than a Mercedes’. Like Mercedes engines, human nervous systems are complex. But the simplicity of the sea slug’s nervous system is precisely what makes it so revealing of the neural mechanisms of learning.

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 animals used worldwide each year in research are but a fraction of 1 percent of the billions of animals killed annually for food. And yearly, for every dog or cat used in an experiment and cared for under humane regulations, 50 others are killed in humane animal shelters (Goodwin & Morrison, 1999).

“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

Some animal protection organizations want to replace experiments on animals with naturalistic observation. Many animal researchers respond that this is not a question of good versus evil but of compassion for animals versus compassion for people. How many of us would have attacked Louis 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 that led to effective methods of training children with mental disorders; of understanding aging; and of relieving fears and depression? The answers to such questions vary by culture. In Gallup surveys in Canada and the United States, about 60 percent of adults deem medical testing on animals “morally acceptable.” In Britain, only 37 percent do (Mason, 2003).

“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]

Out of this heated debate, two issues emerge. The basic one is whether it is right to place the well-being of humans above that of animals. In experiments on stress and cancer, is it right that mice get tumors in the hope that people might not? Should some monkeys be exposed to an HIV-like virus in the search for an AIDS vaccine? Is our use and consumption of other animals as natural as the behavior of carnivorous hawks, cats, and whales? 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.

"The righteous know the needs of their animals."

Proverbs 12:10

Scott Plous (1993) notes, however, that our compassion for animals varies, as does our compassion for people—based on their perceived similarity to us. As **Unit 14** 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.

If we give human life first priority, the second issue is the priority we give to the well-being of animals in research. What safeguards should protect them? 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 already have such guidelines. For example, British Psychological Society guidelines call for housing animals under reasonably natural living conditions, with companions for social animals (Lea, 2000). American Psychological Association (2002) guidelines mandate ensuring the "comfort, health, and humane treatment" of animals, and of minimizing "infection, illness, and pain of animal subjects." 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

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

Animals have themselves benefited from animal research. One Ohio team of research psychologists measured stress hormone levels in samples of millions of dogs brought each year to animal shelters. They devised handling and stroking methods to reduce stress and ease the dogs' transition to adoptive homes (Tuber et al., 1999). In New York, formerly listless and idle Bronx Zoo animals now stave off boredom by working for their supper, as they would in the wild (Stewart, 2002). Other studies have helped improve care and management in animals' natural habitats. By revealing our behavioral kinship with animals and the remarkable intelligence of chimpanzees, gorillas, and other animals, experiments have also led to increased empathy and protection for them. At its best, a psychology concerned for humans and sensitive to animals serves the welfare of both.

13: Is it ethical to experiment on people?

If the image of researchers delivering supposed electric shocks troubles you, you may be relieved to know that in most psychological studies, especially those with human participants, blinking lights, flashing words, and pleasant social interactions are more common.

Occasionally, though, researchers do temporarily stress or deceive people—though more mildly than the stresses that people sometimes willingly undergo in network reality TV programs. But stress and deception are used sparingly—only when researchers believe it is essential to a justifiable end, such as understanding and 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. Wanting to be helpful, the participants might try to confirm the researcher's predictions.

Ethical principles developed by the American Psychological Association (1992), by the British Psychological Society (1993), and by psychologists internationally (Pettifor, 2004), 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 **debrief** people: explain the research afterward. Moreover, most universities (where a great deal of research is conducted) now screen research proposals through an ethics committee—an "Institutional Review Board"—that safeguards the well-being of every participant.

The ideal is for a researcher to be sufficiently informative *and* considerate so that participants will leave feeling at least as good about themselves as when they came in. Better yet, they should be repaid by having learned something. If treated respectfully, most participants enjoy or accept their engagement (Epley & Huff, 1998; Kimmel, 1998). Indeed, say psychology's defenders, teachers provoke much greater anxiety by giving and returning class tests than do researchers in a typical experiment.

Much research 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.

14: Is psychology free of value judgments?

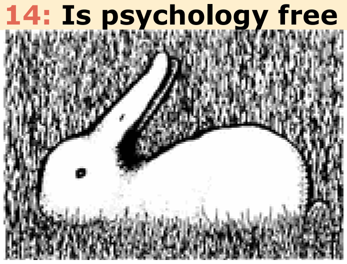


Figure 2.11 What do you see? People interpret ambiguous information to fit their preconceptions. Do you see a duck or a rabbit? Before showing some friends this image, ask them if they can see the duck lying on its back (or the bunny in the grass). (From Shepard, 1990.) © Roger Shepard

Psychology is definitely not value-free. Values affect what we study, how we study it, and how we interpret results. Researchers' values influence their choice of topics. Should we study worker productivity or worker morale? Sex discrimination or gender differences? Conformity or independence? Values can also color "the facts." As we noted earlier, our preconceptions can bias our observations and interpretations; sometimes we see what we want or expect to see (**Figure 2.11**).

Even the words we use to describe something can reflect our values. Are the sex acts that an individual does not practice "perversions" or "sexual variations"? Both in and out of psychology, labels describe and labels evaluate: The same holds true in everyday speech. One person's "rigidity" is another's "consistency." One person's "faith" is another's "fanaticism." Our labeling someone as "firm" or "stubborn," "careful" or "picky," "discreet" or "secretive" reveals our feelings.

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.

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

Simone de Beauvoir, *The Second Sex*, 1953

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

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

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

Unit 2 Objectives – Research Methods

Describe how research design drives the reasonable conclusions that can be drawn (e.g., experiments are useful for determining cause and effect; the use of experimental controls reduces alternative explanations).

Identify independent, dependent, confounding, and control variables in experimental designs.

Distinguish between random assignment of participants to conditions in experiments and random selection of participants, primarily in correlational studies and surveys.

Predict the validity of behavioral explanations based on the quality of research design (e.g., confounding variables limit confidence in research conclusions).

Distinguish the purposes of descriptive statistics and inferential statistics.

Apply basic descriptive statistical concepts, including interpreting and constructing graphs and calculating simple descriptive statistics
(e.g., measures of central tendency, standard deviation).

Discuss the value of reliance on operational definitions and measurement in behavioral research.

Identify how ethical issues inform and constrain research practices.

Describe how ethical and legal guidelines (e.g., those provided by the American Psychological Association, federal regulations, local institutional review boards) protect research participants and promote sound ethical practice.