Module 7

Statistical Reasoning in Everyday Life


Norm Hall/Getty Images

Module Learning Objectives

7-1 Describe the three measures of central tendency, and discuss the relative usefulness of the two measures of variation.

7-2 Explain how we know whether an observed difference can be generalized to other populations.

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

FYI

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

The Need for Statistics

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

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

- *Ten percent of people are lesbians or gay men.* Or is it 2 to 3 percent, as suggested by various national surveys (Module 53)?
- *We ordinarily use but 10 percent of our brain.* Or is it closer to 100 percent (Module 12)?
- *The human brain has 100 billion nerve cells.* Or is it more like 40 billion, as suggested by extrapolation from sample counts (Module 10)?

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

Statistical illiteracy also feeds needless health scares (Gigerenzer et al., 2008, 2009, 2010). In the 1990s, the British press reported a study showing

Patrick Hardin/cartoonstock

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

AP® Exam Tip

Do math and statistics scare you? Take a couple of deep breaths and relax before continuing. You will not be asked to do difficult computations on the AP® exam. Nothing will be beyond the scope of simple mental math. You need to focus on the concepts. Why do these statistics exist? How can they help us understand the real world?

that women taking a particular contraceptive pill had a 100 percent increased risk of blood clots that could produce strokes. This caused thousands of women to stop taking the pill, leading to a wave of unwanted pregnancies and an estimated 13,000 additional abortions (which also are associated with increased blood clot risk). And what did the study find? A 100 percent increased risk, indeed—but only from 1 in 7000 women to 2 in 7000 women. Such false alarms underscore the need to teach statistical reasoning and to present statistical information more transparently.

Descriptive Statistics

7-1 How do we describe data using three measures of central tendency, and what is the relative usefulness of the two measures of variation?

Once researchers have gathered their data, they may use **descriptive statistics** to organize that data meaningfully. One way to do this is to convert the data into a simple *bar graph,* called a **histogram,** as in **FIGURE 7.1,** 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 7.1a) or small (Figure 7.1b). The secret lies in how you label the vertical scale (the *y-axis*).

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

descriptive statistics numerical data used to measure and describe characteristics of groups. Includes measures of central tendency and measures of variation.

histogram a bar graph depicting a frequency distribution.

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

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

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

Measures of Central Tendency

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

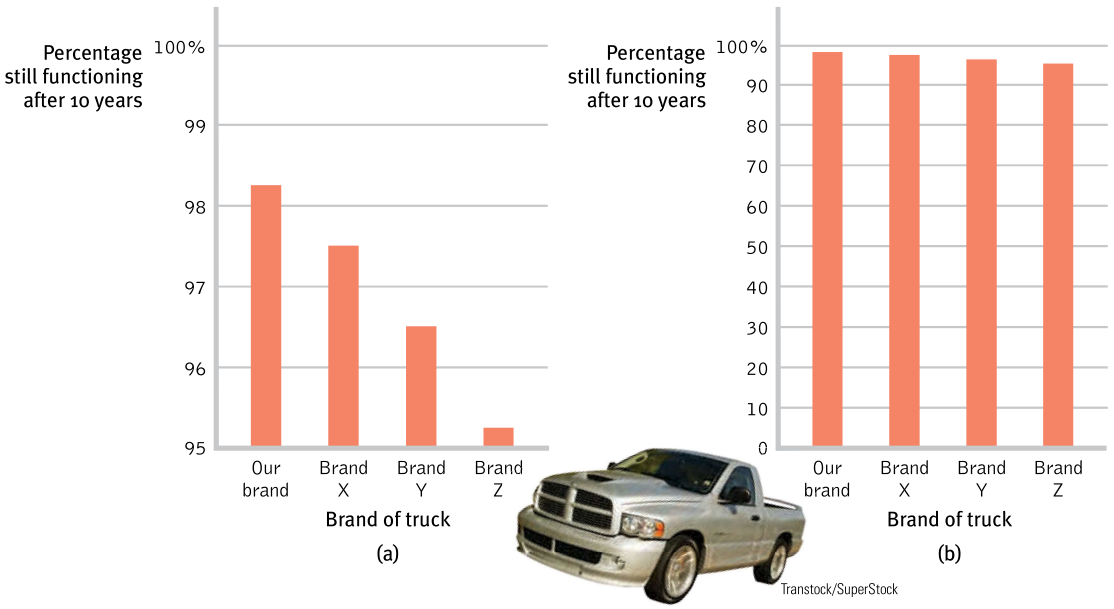


Figure 7.1

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).

Measures of central tendency neatly summarize data. But consider what happens to the mean when a distribution is lopsided, or **skewed,** by a few way-out scores. With income data, for example, the mode, median, and mean often tell very different stories (**FIGURE 7.2**). 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 customers' median wealth remains unchanged. Understanding this, you can see how a British newspaper could accurately run the headline "Income for 62% Is Below Average" (Waterhouse, 1993). Because the bottom *half* of British income earners receive only a *quarter* of the national income cake, most British people, like most people everywhere, make less than the mean. Mean and median tell different true stories.

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

FYI

The average person has one ovary and one testicle.

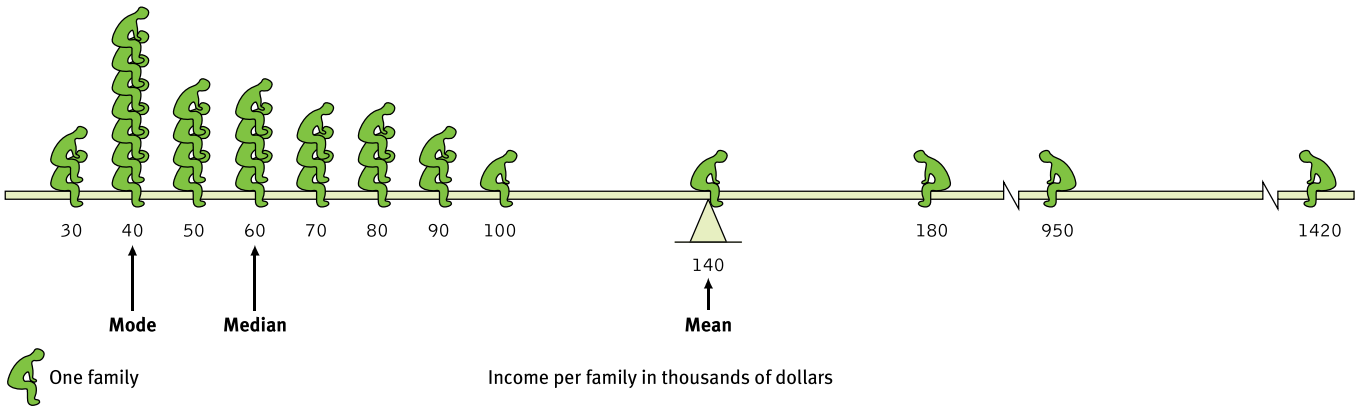


Figure 7.2

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.

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. A couple of extreme scores in an otherwise uniform group, such as the $950,000 and $1,420,000 incomes in Figure 7.2, 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 7.1**). 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 the more diverse community population outside your school.

You can grasp the meaning of the standard deviation if you consider how scores tend to be distributed in nature. Large numbers of data—heights, weights, intelligence scores, grades (though not incomes)—often form a symmetrical, *bell-shaped* distribution.

skewed distribution a representation of scores that lack symmetry around their average value.

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

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

Table 7.1 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	+20	400
87	+7	49	100	+20	400
Total = 640	Sum of (deviations)^2 = 204		Total = 640	Sum of (deviations)^2 = 2000	
Mean = 640 ÷ 8 = 80			**Mean** = 640 ÷ 8 = 80		

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

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

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 7.3** 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, as Module 61 notes, about 68 percent of people taking an intelligence test will score within ±15 points of 100. About 95 percent will score within ±30 points.

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

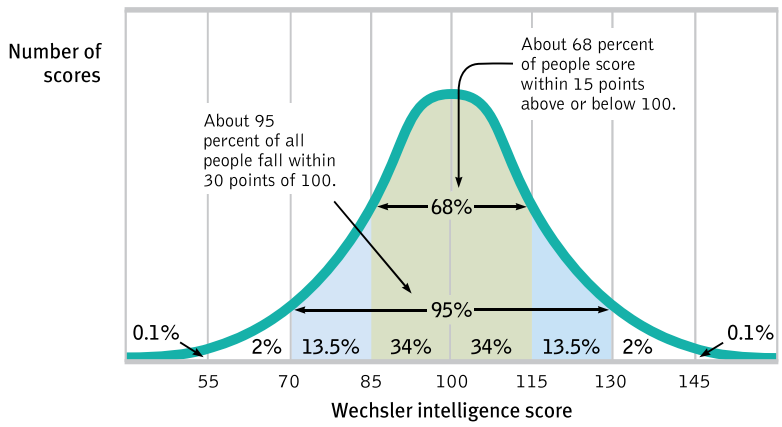


Figure 7.3
The normal curve Scores on aptitude tests tend to form a normal, or bell-shaped, curve. For example, the most commonly used intelligence test, the Wechsler Adult Intelligence Scale, calls the average score 100.

Inferential Statistics

7-2 How do we know whether an observed difference can be generalized to other populations?

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

inferential statistics numerical data that allow one to generalize—to infer from sample data the probability of something being true of a population.

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.** As noted in Module 5, 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 instructors at each institution) tells friends about the "great instructors" at the first school, and the "bores" at the second. Again, we know it but we ignore it: *Averages based on many cases are more reliable* (less variable) than averages based on only a few cases.

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

When Is a Difference Significant?

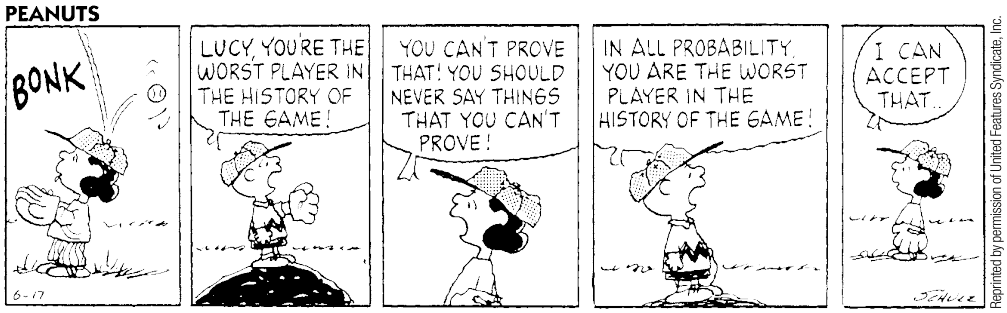
Perhaps you've compared men's and women's scores on a laboratory test of aggression, and found a gender difference. But individuals differ. How likely is it that the gender difference you found was just a fluke? Statistical testing can estimate the probability of the result occurring by chance.

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

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

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

In judging statistical significance, psychologists are conservative. They are like juries who must presume innocence until guilt is proven. For most psychologists, proof beyond a



reasonable doubt means not making much of a finding unless the odds of its occurring by chance, if no real effect exists, are less than 5 percent.

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

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

AP® Exam Tip

Sometimes a phrase that is frequently used in the media has a more specific meaning when used in psychology. That's the case with the phrase "statistically significant." Make sure you know the precise meaning.

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

Can you solve this puzzle?

The registrar's office at the University of Michigan has found that usually about 100 students in Arts and Sciences have perfect 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 (Jepson et al., 1983)?

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

Module 7 Review

7-1 How do we describe data using three measures of central tendency, and what is the relative usefulness of the two measures of variation?

- A measure of central tendency is a single score that represents a whole set of scores. Three such measures are the *mode* (the most frequently occurring score), the *mean* (the arithmetic average), and the *median* (the middle score in a group of data).
- Measures of variation tell us how diverse data are. Two measures of variation are the *range* (which describes the gap between the highest and lowest scores) and the *standard deviation* (which states how much scores vary around the mean, or average, score).
- Scores often form a *normal* (or bell-shaped) *curve.*

7-2 How do we know whether an observed difference can be generalized to other populations?

- To feel confident about generalizing an observed difference to other populations, we would want to know that
  - the sample studied was representative of the larger population being studied;
  - the observations, on average, had low variability;
  - the sample consisted of more than a few cases; and
  - the observed difference was *statistically significant.*

Multiple-Choice Questions

1. Which of the following is a measure of variation?

a. Range
b. Mean
c. Mode
d. Frequency
e. Median

2. Which statistical measure of central tendency is most affected by extreme scores?

a. Mean
b. Median
c. Mode
d. Skew
e. Correlation

3. A researcher calculates statistical significance for her study and finds a 5 percent chance that results are due to chance. Which of the following is an accurate interpretation of this finding?

a. This is well beyond the range of statistical significance.
b. This is the minimum result typically considered statistically significant.
c. This is not statistically significant.
d. There is no way to determine statistical significance without replication of the study.
e. Chance or coincidence is unrelated to statistical significance.

4. Descriptive statistics _____, while inferential statistics _____.

a. indicate the significance of the data; summarize the data
b. describe data from experiments; describe data from surveys and case studies
c. are measures of central tendency; are measures of variance
d. determine if data can be generalized to other populations; summarize data
e. summarize data; determine if data can be generalized to other populations

5. In a normal distribution, what percentage of the scores in the distribution falls within one standard deviation on either side of the mean?

a. 34 percent
b. 40 percent
c. 50 percent
d. 68 percent
e. 95 percent

Practice FRQs

1. Explain the difference between descriptive and inferential statistics in research.

Answer (2 points)

1 point: Descriptive statistics organize and summarize the data collected during research.

1 point: Inferential statistics are used to help determine whether results can be generalized to a larger population through the calculation of statistical significance.

2. The following data set includes information from survey research in a psychology course regarding how many hours each individual in the class spent preparing for the exam.

Student	Amount of hours reported studying
1	2
2	3
3	6
4	8
5	9
6	9
7	21

Examine the data and respond to the following:

- What is the middle score in this distribution? What term is used to describe the middle score?
- What would be the most useful statistic for measuring the variation of the hours spent studying? Why is this statistic a better measure of variation than the range?

(3 points)

Module 8

Frequently Asked Questions About Psychology


wdstock/E+/Getty Images

Module Learning Objectives

- **8-1** Explain the value of simplified laboratory conditions in illuminating everyday life.
- **8-2** Discuss whether psychological research can be generalized across cultures and genders.
- **8-3** Explain why psychologists study animals, and describe the ethical guidelines that safeguard animal research participants.
- **8-4** Describe the ethical guidelines that safeguard human research participants.
- **8-5** Examine whether psychology is free of value judgments.

We have reflected on how a scientific approach can restrain biases. We have seen how case studies, naturalistic observations, and surveys help us describe behavior. We have also noted that correlational studies assess the association between two variables, 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.

Psychology Applied

8-1 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? If, after playing violent video games in the lab, teens become more willing to push buttons that they think electrically shock someone, does this indicate that playing shooter games makes someone more likely to commit violence in everyday life?

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

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

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

The point to remember: Psychological science focuses less on particular behaviors than on seeking general principles that help explain many behaviors. And remember: Although psychological principles may help predict behaviors for groups of people, they minimally predict behavior for any individual. Knowing students' grade level may clue us to their average vocabulary level, but individual students' word power will vary.

8-2 Does behavior depend on one's culture and gender?

What can psychological studies done in one time and place—often with people from what researchers call the WEIRD (*W*estern, *E*ducated, *I*ndustrialized, *R*ich, and *D*emocratic) cultures (Henrich et al., 2010) 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. *Collectivist* cultures, for example, emphasize group goals, while *individualist* cultures put a priority on individual goals. 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.

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

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 specific learning disorder (formerly called dyslexia) 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).

Alistair Berg/Alamy

Soccer shoes? Because culture shapes social behavior, actions that seem ordinary to some may seem odd to others. Yet underlying these differences are powerful similarities. Children everywhere love to play sports such as soccer. But many American children would only play with athletic shoes on a field, not barefoot in the street, as do these Burkina Faso boys.

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 use disorder, depression, and eating disorders. Gender differences fascinate us, and studying them is potentially beneficial. For example, many researchers believe that women carry on conversations more readily to build relationships, while men talk more to give information and advice (Tannen, 2001). Knowing this difference can help us prevent conflicts and misunderstandings in everyday relationships.

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

"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.

Ethics in Research

8-3 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. We humans are not *like* animals, we *are* animals, sharing a common biology. Animal experiments have therefore led to treatments for human diseases—insulin for diabetes, vaccines to prevent polio and rabies, transplants to replace defective organs.

"Rats are very similar to humans except that they are not stupid enough to purchase lottery tickets." -Dave Barry, July 2, 2002

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

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.

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 Module 79 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, what safeguards should protect the well-being of animals in research? One survey of animal researchers gave an answer. Some 98 percent supported government regulations protecting primates, dogs, and cats, and 74 percent supported regulations providing for the humane care of rats and mice (Plous & Herzog, 2000). Many professional associations and funding agencies already have such guidelines. British Psychological Society guidelines call for housing animals under reasonably natural living conditions, with companions for social animals (Lea, 2000). American Psychological Association (APA) guidelines state that researchers must ensure the "comfort, health, and humane treatment" of animals and minimize "infection, illness, and pain" (APA, 2002). The European Parliament now mandates standards for animal care and housing (Vogel, 2010).

"The greatness of a nation can be judged by the way its animals are treated." -Mahatma Gandhi, 1869–1948

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

AP Photo/Mary Altaffer

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.

8-4 What ethical guidelines safeguard human participants?

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

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

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

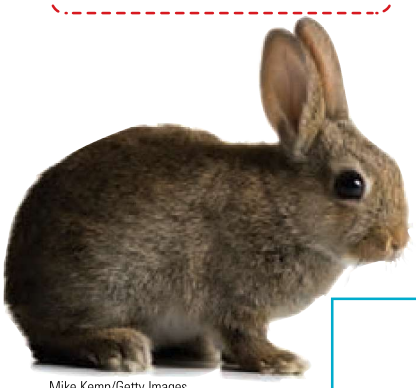
debriefing the postexperimental explanation of a study, including its purpose and any deceptions, to its participants.

Ethical principles developed by the American Psychological Association (2010), by the British Psychological Society (2009), and by psychologists internationally (Pettifor, 2004), urge researchers to (1) obtain potential participants' **informed consent,** (2) protect them from physical or emotional harm and discomfort, (3) keep information about individual participants confidential, and (4) fully **debrief** people (explain the research afterward). Moreover, most universities (where a great deal of research is conducted) now have an ethics committee—an Institutional Review Board (IRB)—that screens research proposals and safeguards participants' well-being.

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.

8-5 Is psychology free of value judgments?

"It is doubtless impossible to approach any human problem with a mind free from bias."
-Simone de Beauvoir, *The Second Sex*, 1953



Mike Kemp/Getty Images

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 8.1**).

Even the words we use to describe something can reflect our values. In psychology and in everyday speech, labels describe and labels evaluate: One person's *rigidity* is another's *consistency*. One person's *faith* is another's *fanaticism*. One country's *enhanced interrogation techniques,* such as cold-water immersion, become *torture* when practiced by its enemies. Our labeling someone as *firm* or *stubborn, careful* or *picky, discreet* or *secretive* reveals our own attitudes.

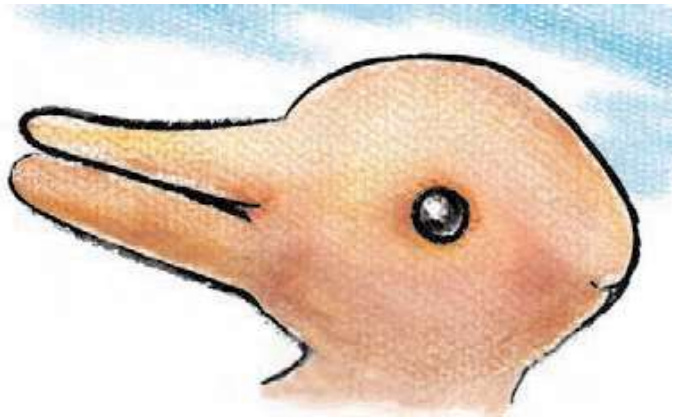


Figure 8.1
What do you see? Our expectations influence what we perceive. Did you see a duck or a rabbit? Show some friends this image with the rabbit photo above covered up and see if they are more likely to perceive a duck head instead. (From Shepard, 1990.)

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

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

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

Library of Congress

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.

Before You Move On

▶ ASK YOURSELF

Were any of this module's Frequently Asked Questions your questions? Do you have other questions or concerns about psychology?

▶ TEST YOURSELF

How are human and animal research participants protected?

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

Module 8 Review

8-1 Can laboratory experiments illuminate everyday life?

- Researchers intentionally create a controlled, artificial environment in the laboratory in order to test general theoretical principles. These general principles help explain everyday behaviors.

8-2 Does behavior depend on one's culture and gender?

- Attitudes and behaviors may vary somewhat by gender or across *cultures,* but because of our shared human kinship, the underlying processes and principles are more similar than different.

8-3 Why do psychologists study animals, and is it ethical to experiment on animals?

- Some psychologists are primarily interested in animal behavior; others want to better understand the physiological and psychological processes shared by humans and other species.
- Government agencies have established standards for animal care and housing. Professional associations and funding agencies also establish guidelines for protecting animals' well-being.

8-4 What ethical guidelines safeguard human participants?

- The APA ethics code outlines standards for safeguarding human participants' well-being, including obtaining their *informed consent* and *debriefing* them later.

8-5 Is psychology free of value judgments?

- Psychologists' values influence their choice of research topics, their theories and observations, their labels for behavior, and their professional advice.
- Applications of psychology's principles have been used mainly in the service of humanity.

Multiple-Choice Questions

1. Which of the following is more likely to be emphasized in individualist cultures than in collectivist cultures?

a. Gender differences
b. Shared goals
c. Personal achievement
d. Cooperation with the group
e. Preservation of tradition

2. What must a researcher do to fulfill the ethical principle of informed consent?

a. Keep information about participants confidential.
b. Allow participants to choose whether to take part.
c. Protect participants from potential harm.
d. Provide participants with a pre-experimental explanation of the study.
e. Provide participants with a post-experimental explanation of the study.

3. Which ethical principle requires that at the end of the study participants be told about the true purpose of the research?

a. Institutional Review Board approval
b. Informed consent
c. Confidentiality
d. Debriefing
e. Protection from physical harm

4. Which of the following beliefs would most likely be held by an individual in a collectivist culture?

a. Children should be encouraged to focus on personal goals and aspirations.
b. Children should be encouraged to develop harmonious relationships.
c. It is important to be competitive and assertive in order to get ahead in life.
d. If you want something done well, you should do it yourself.
e. It is important to satisfy personal needs before those of the larger community.

Practice FRQs

1. Provide three reasons why nonhuman animals are sometimes used in psychological research.

Answer

1 point: Some researchers use nonhuman animals because they are interested in understanding the animals themselves, including their thinking and behaviors.

1 point: Others use nonhuman animals to reduce the complexity that is part of human research. They hope to understand principles that may be similar to those that govern human psychological phenomena.

1 point: Researchers also study nonhuman animals in order to apply the findings in ways that will help both humans and the other animals themselves.

2. Researchers interested in studying stress gave 150 high school seniors a very difficult math exam. After the test, the researchers measured stress by examining physiological changes with extensive medical testing that included drawing blood samples.

- What ethical principle governs what students must be told before the research takes place? What should the potential participants be told?
- What ethical principle governs the appropriate use of the results of the medical testing? What would that principle say about the use of these results?

(8 points)

Unit II Review

Key Terms and Concepts to Remember

hindsight bias, p. 31
critical thinking, p. 35
theory, p. 38
hypothesis, p. 38
operational definition, p. 39
replication, p. 39
case study, p. 40
naturalistic observation, p. 40
survey, p. 42
sampling bias, p. 43
population, p. 43
random sample, p. 43
correlation, p. 46
correlation coefficient, p. 46
scatterplot, p. 46
illusory correlation, p. 50
experiment, p. 51
experimental group, p. 51
control group, p. 51
random assignment, p. 51
double-blind procedure, p. 51
placebo [pluh-SEE-bo] effect, p. 52
independent variable, p. 52
confounding variable, p. 52
dependent variable, p. 52
validity, p. 53
descriptive statistics, p. 57
mode, p. 57
mean, p. 57
median, p. 57
skewed distribution, p. 58
range, p. 58
standard deviation, p. 58
normal curve, p. 59
inferential statistics, p. 60
statistical significance, p. 60
culture, p. 65
informed consent, p. 68
debriefing, p. 68

AP® Exam Practice Questions

Multiple-Choice Questions

1. Which descriptive statistic would a researcher use to describe how close a student's SAT score is to a school's average SAT score?

a. Correlation coefficient
b. Mean
c. Median
d. Standard deviation
e. Range

2. Which method should a psychology researcher use if she is interested in testing whether a specific reward in a classroom situation causes students to behave better?

a. Case study
b. Experiment
c. Survey
d. Naturalistic observation
e. Correlation

3. When a distribution of scores is skewed, which of the following is the most representative measure of central tendency?

a. Inference
b. Standard deviation
c. Mean
d. Median
e. Correlation coefficient

4. A researcher wants to conduct an experiment to determine if eating a cookie before class each day improves student grades. He uses two psychology classes for the experiment, providing daily cookies to one and nothing to the other. At the end of the semester, the researcher compares the final grades of students in the two classes. What is the independent variable for this experiment?

a. The students in the class that received cookies
b. The presence or absence of cookies
c. The students in the class that didn't receive cookies
d. The period of the day that the two classes met
e. Semester grades

5. Which of the following represents naturalistic observation?

a. Researchers watch and record how elementary school children interact on the playground.
b. Researchers bring participants into a laboratory to see how they respond to a puzzle with no solution.
c. A principal looks at the relationship between the number of student absences and their grades.
d. A social worker visits a family home and gives feedback on family interactions.
e. Two grandparents sit in the front row to watch their grandson's first piano recital.

6. "Monday morning quarterbacks" rarely act surprised about the outcome of weekend football games. This tendency to believe they knew how the game would turn out is best explained by which psychological principle?

a. Overconfidence
b. Hindsight bias
c. Intuition
d. Illusory correlation
e. Random sampling

7. Researchers studying gender have found that

a. there are more similarities than differences between the genders.
b. there are no significant cognitive differences between the genders.
c. there are no significant emotional differences between the genders.
d. research tools are not capable of determining if there are true differences or not.
e. differences between the genders are becoming more pronounced over time.

8. A journalism student is writing an article about her school's new cell-phone policy, and she'd like to interview a random sample of students. Which of the following is the best example of a random sample?

a. The writer arrives at school early and interviews the first five students who come through the main entrance.
b. The writer pulls the names of five students from a hat that contains all students' names. She interviews the five selected students.
c. The writer asks her teacher if she can distribute a brief survey to the students in her AP® Psychology class.
d. The writer passes out brief surveys to 50 students in the hall and uses the 18 surveys returned to her as the basis of her article.
e. The writer asks the principal for the names of 10 students who have had their cell phones confiscated for a day for violating the policy. She interviews these 10 students.

9. Which of the following is a positive correlation?

a. As study time decreases, students achieve lower grades.
b. As levels of self-esteem decline, levels of depression increase.
c. People who exercise regularly are less likely to be obese.
d. Gas mileage decreases as vehicle weight increases.
e. Repeatedly shooting free throws in basketball is associated with a smaller percentage of missed free throws.

10. Why is random assignment of participants to groups an important aspect of a properly designed experiment?

a. If the participants are randomly assigned, the researcher can assume that the people in each of the groups are pretty similar.
b. By randomly assigning participants, the researcher knows that whatever is learned from the experiment will also be true for the population from which the participants were selected.
c. Random assignment keeps expectations from influencing the results of the experiment.
d. If participants are not randomly assigned, it is impossible to replicate the experiment.
e. Statistical analysis cannot be performed on an experiment if random assignment is not used.

11. Which of the following demonstrates the need for psychological science?

a. Psychology's methods are unlike those of any other science.
b. Psychological experiments are less valuable without psychological science.
c. Our intuitions about human thinking and behavior are not always accurate.
d. Intuition does not provide correct answers unless it is applied through the scientific method.
e. Psychological science research is superior to that of other sciences like biology and physics.

12. Which of the following is a potential problem with case studies?

a. They provide too much detail and the researcher is likely to lose track of the most important facts.
b. They are generally too expensive to be economical.
c. They may be misleading because they don't fairly represent other cases.
d. They are technically difficult and most researchers don't have the skills to do them properly.
e. The dependent variable is difficult to operationally define in a case study.

13. Which of the following is *not* an ethical principle regarding research on humans?

a. Researchers must protect participants from needless harm and discomfort.
b. Participants must take part in the study on a voluntary basis.
c. Personal information about individual participants must be kept confidential.
d. Research studies must be fully explained to participants when the study is completed.
e. Participants should always be informed of the hypothesis of the study before they agree to participate.

14. There is a negative correlation between TV watching and grades. What can we conclude from this research finding?

a. We can conclude that a student who watches a lot of TV is likely to have lower grades.
b. We can conclude that TV watching leads to lower grades.
c. We can conclude that TV watching leads to higher grades.
d. We can conclude that the grades students get impact their TV watching habits.
e. We can conclude that this is an illusory correlation.

15. A scientist's willingness to admit that she is wrong is an example of

a. curiosity.
b. intelligence.
c. humility.
d. skepticism.
e. cynicism.

Free-Response Questions

1. Sam Greene noticed an ad for an Internet dating service that claimed more people who used its service are in long-term relationships than people who didn't. Sam, a good critical thinker, knows this isn't enough to claim that the service causes people to find long-term love and wants to create an experiment to investigate. Use the following terms to describe an experiment that would support or dispute the ad's claim.

- Hypothesis
- Random sample
- Random assignment
- Operational definitions
- Independent variable
- Dependent variable
- Inferential statistics

Rubric for Free-Response Question 1

1 point: The hypothesis in this context is that the Internet dating service causes (or leads to) long-term relationships. Page 38

1 point: Since the population of interest for this study should be people who are looking for long-term relationships, selecting a random sample of adults seeking relationships would help assure that the conclusions could be fairly generalized to the dating public. Page 43

1 point: In this case, participants should be randomly assigned to use of the Internet service (the experimental group) or not (the control group). Page 51

1 point: Sam would need to operationally define what is meant by use of the Internet service, possibly including a precise number of visits to the website or time spent on the website. The phrase *long-term relationship* would also need an operational definition, possibly by the number of months together or a formal commitment (like engagement or marriage). Page 39

1 point: In Sam's study, the use of the online dating service is the independent variable. Page 52

1 point: The number of long-term relationships is the dependent variable. Page 52

1 point: Sam will need to calculate statistical significance for the experimental findings. In order to claim support for the hypothesis, the results need to show that there is no more than a 5 percent chance the findings are due to chance. Page 60

2. Dr. Tabor wanted to investigate the relationship between sleep and levels of alertness during a class for American university students. She gave surveys to 150 college freshmen in her introduction to psychology course, asking them to report how many hours they slept each night during a two-week period. Dr. Tabor also had the participants rate their level of alertness on a scale of 1 to 10, with 10 being the most alert each day at the end of class. Dr. Tabor compared the average amount of sleep reported by each participant along with their average score on the alertness scale on a graph to examine the data. The resulting correlation coefficient for Dr. Tabor's data was +0.89. Define each of the following terms and explain how each concept might apply to Dr. Tabor's research.

- Random sample
- Scatterplot
- Wording effects
- Positive correlation
- Operational definition

(10 points)

3. Find at least five problems in the research study described below. Identify the problem and explain how it is a violation of accepted research principles.

Dr. Pauling wanted to study whether vitamin C affects self-esteem. She recruited 200 respondents who arrived at her lab. Participants were told that they were about to participate in a harmless research study, and they needed to sign a release form in case there were harmful side effects from the vitamin C pills. The 100 participants on the right side of the room received a pill with vitamin C and the others on the left received a pill with caffeine. She then gave each group a list of questions to answer in essay form about their self-esteem. When they were finished, she thanked the participants and sent them on their way. After compiling her findings, Dr. Pauling printed the names of the students and their results in the campus newspaper so they would know what the results of the test were. Dr. Pauling concluded that vitamin C had a positive affect on self-esteem.

(5 points)

Multiple-choice self-tests and more may be found at www.worthpublishers.com/MyersAP2e