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

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

that women on replacement hormones had lower rates of heart disease, stroke, and colon cancer. But these women may also have been more likely to receive medical care, to exercise, and to eat well. So, did the hormones make women healthy or did healthy women take the hormones? In 2002, the National Institutes of Health announced the surprising results of a massive experiment that randomly assigned 16,608 healthy women to either replacement hormones or a placebo. The shocking result: Compared with women in the control condition, women receiving the hormones had *more* health problems (Love, 2002).

Independent and Dependent Variables

OBJECTIVE 14 | Explain the difference between an independent and a dependent variable.

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

This simple experiment manipulated just one drug factor. We call this experimental factor the **independent variable** because we can vary it independently of other factors, such as the men's age, weight, and personality (which random assignment should control). 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 1.8** for another experiment's design.)

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 these 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).

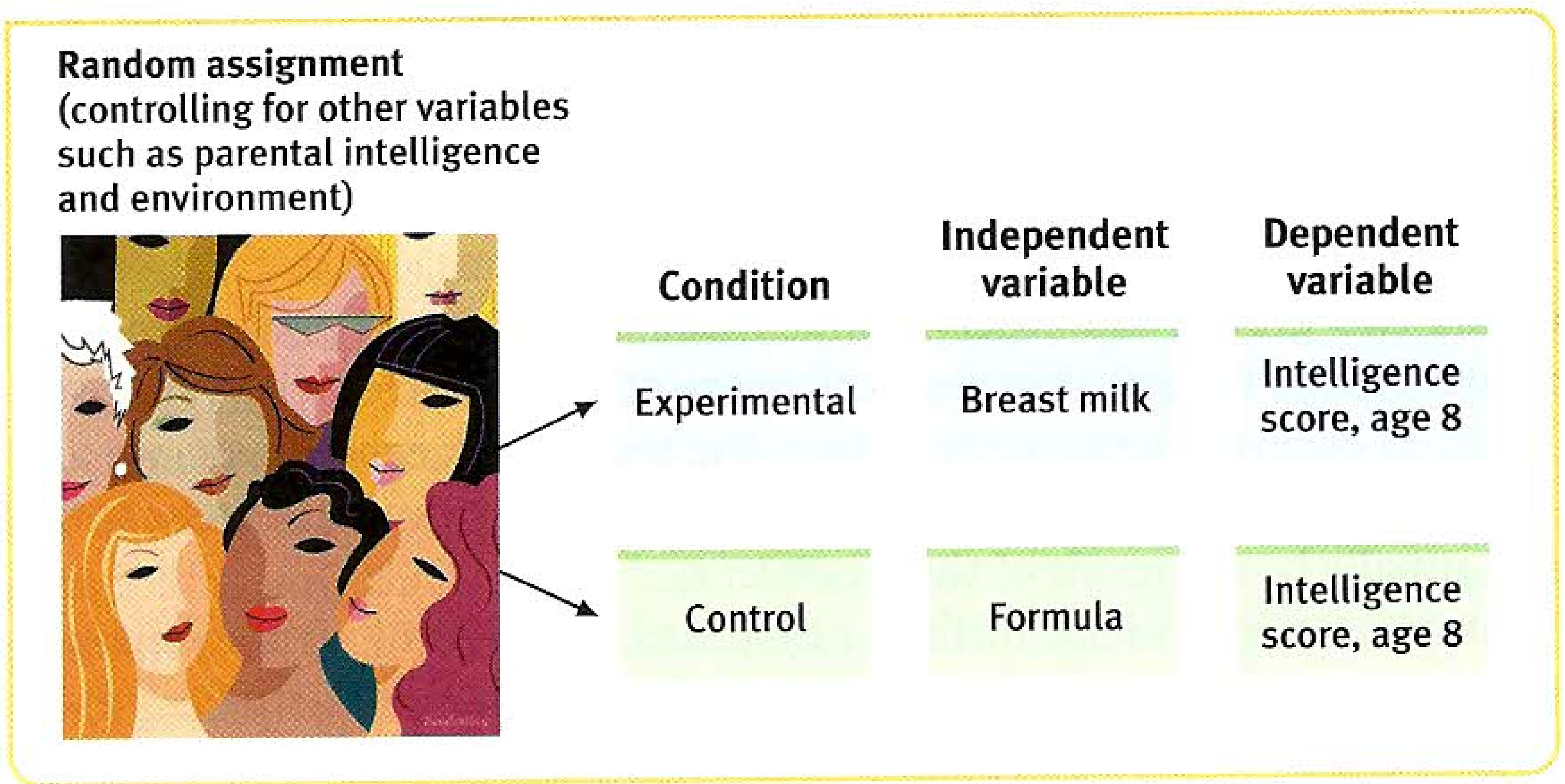


FIGURE 1.8
Experimentation To discern causation, psychologists may randomly assign some participants to an experimental treatment, others to a control condition. Measuring the dependent variable (intelligence score) will determine the effect of the independent variable (type of milk).

TABLE 1.3

COMPARING RESEARCH METHODS

Research Method	Basic Purpose	How Conducted	What Is Manipulated	Weaknesses
Descriptive	To observe and record behavior	Do case studies, surveys, or naturalistic observations	Nothing	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	Does not specify cause and effect
Experimental	To explore cause and effect	Manipulate one or more factors; use random assignment	The independent variable(s)	Sometimes not feasible; results may not generalize to other contexts; not ethical to manipulate certain variables

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* all other variables. An experiment has at least two different conditions: an experimental condition and a comparison or control condition. Random assignment works to equate the conditions 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 1.3** compares the features of psychology's research methods.

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

>> LEARNING OUTCOMES

Experimentation

OBJECTIVE 12 | Explain how experiments help researchers isolate cause and effect.
To discover cause-effect relationships, psychologists conduct *experiments*. By manipulating one or more factors of interest, and controlling other factors, experimenters can determine the effect on some behavior or mental process.

OBJECTIVE 13 | Explain why the double-blind procedure and random assignment build confidence in research findings.
In a double-blind procedure, neither the researchers nor the participants know whether participants are receiving the treatment or a placebo. This counteracts the possibility that a placebo effect or researchers' expectations will unintentionally influence the study's results. Random assignment minimizes preexisting differences between the groups by selecting people by chance for the experimental condition (the group exposed to the treatment) or the control condition (a group that experiences no treatment or a different version of the treatment).

OBJECTIVE 14 | Explain the difference between an independent and a dependent variable.
The independent variable is the factor you manipulate to study its effect. The dependent variable is the factor you measure to discover any changes that occur in response to these manipulations.

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

Statistical Reasoning

OBJECTIVE 15 | Explain the importance of statistical principles, and give an example of their use in everyday life.

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

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

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

- *One percent of Americans (2.7 million) are homeless.* Or is it 300,000, an earlier estimate by the federal government? Or 600,000, an estimate by the Urban Institute (Crossen, 1994)?
- *Ten percent of people are lesbians or gay men.* Or is it 2 to 3 percent, as suggested by various national surveys (Chapter 12)?
- *We ordinarily use but 10 percent of our brain.* Or is it closer to 100 percent? (Chapter 2; which 90 percent, or even 10 percent, would you be willing to sacrifice?)
- *We remember 10 percent of what we read, 20 percent of what we hear, 30 percent of what we see, and 80 percent of what we say.* So reported the British Audio Visual Society (Genovese, 2004). Or is it, as a book on accelerated learning declares, 20 percent of what we read, 30 percent of what we hear, 40 percent of what we see, and 50 percent of what we say?

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

Describing Data

OBJECTIVE 16 | Explain how bar graphs can misrepresent data.

Once researchers have gathered their data, their first task is to *organize* them. One way is to use a simple *bar graph,* as in **FIGURE 1.9**, which displays a distribution of trucks of different brands still on the road after a decade. When reading statistical graphs such as this, take care. As you can see, people can design a graph to make a difference look small or big, depending on what they want to emphasize.

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

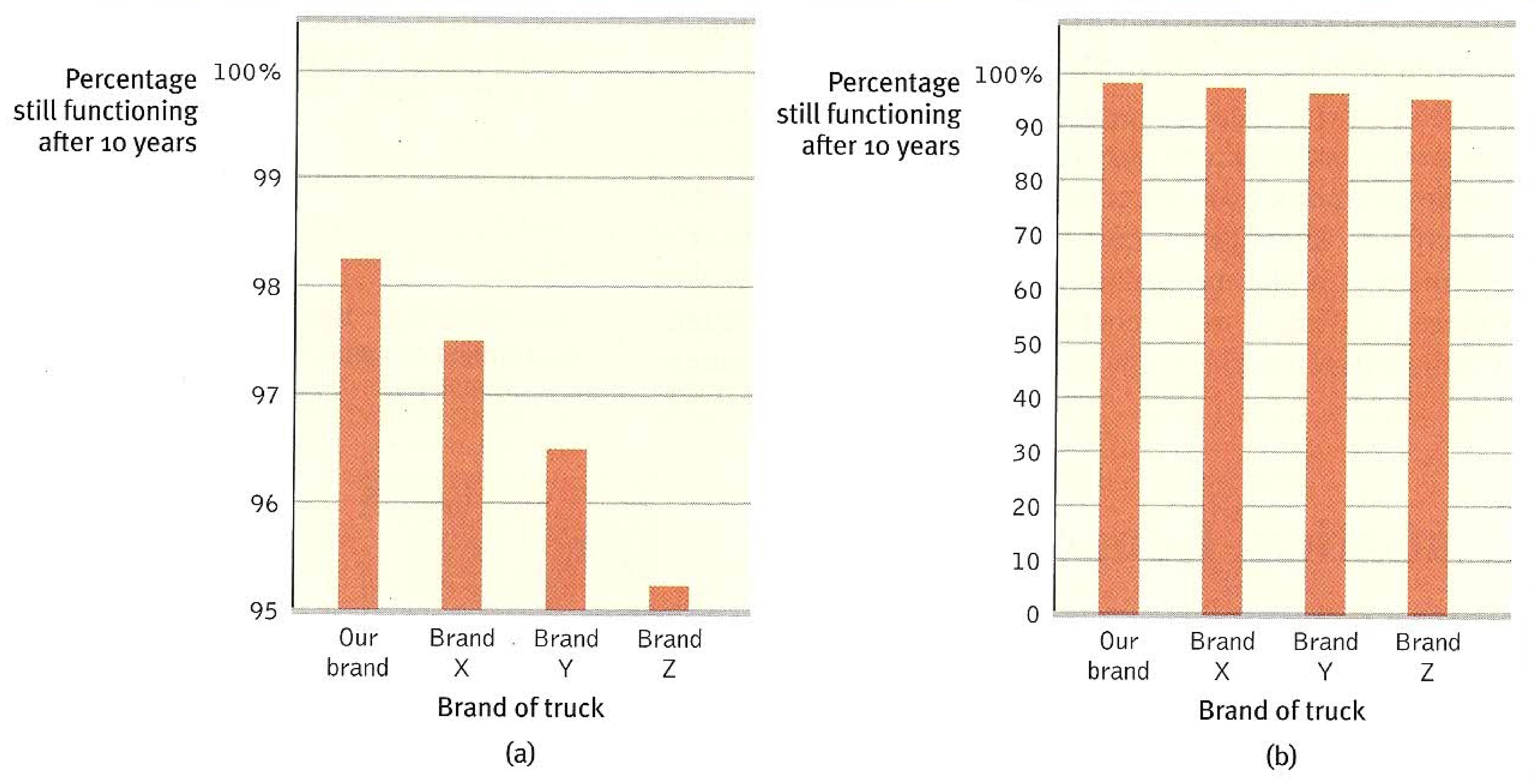


FIGURE 1.9 **Read the scale labels** An American truck manufacturer offered a graph (a)—with actual brand names included—to suggest the much greater durability of its trucks. Note, however, how the apparent difference shrinks as the vertical scale changes (graph b).

Measures of Central Tendency

OBJECTIVE 17 | Describe the three measures of central tendency, and tell which is most affected by extreme scores.

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

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

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

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

Measures of central tendency neatly summarize data. But consider what happens to the mean when a distribution is lopsided or *skewed.* With income data, for example, the mode, median, and mean often tell very different stories (**FIGURE 1.10**). This happens because the mean is biased by a few extreme scores. When Microsoft founder Bill Gates sits down in an intimate cafe, its average (mean) patron instantly becomes a billionaire. Understanding this, you can see how a British newspaper could accurately run the headline "Income for 62% Is Below Average" (Waterhouse, 1993). Because the bottom *half* of British income earners receive only a *quarter* of the national income cake, most British people, like most people everywhere, make less than the mean.

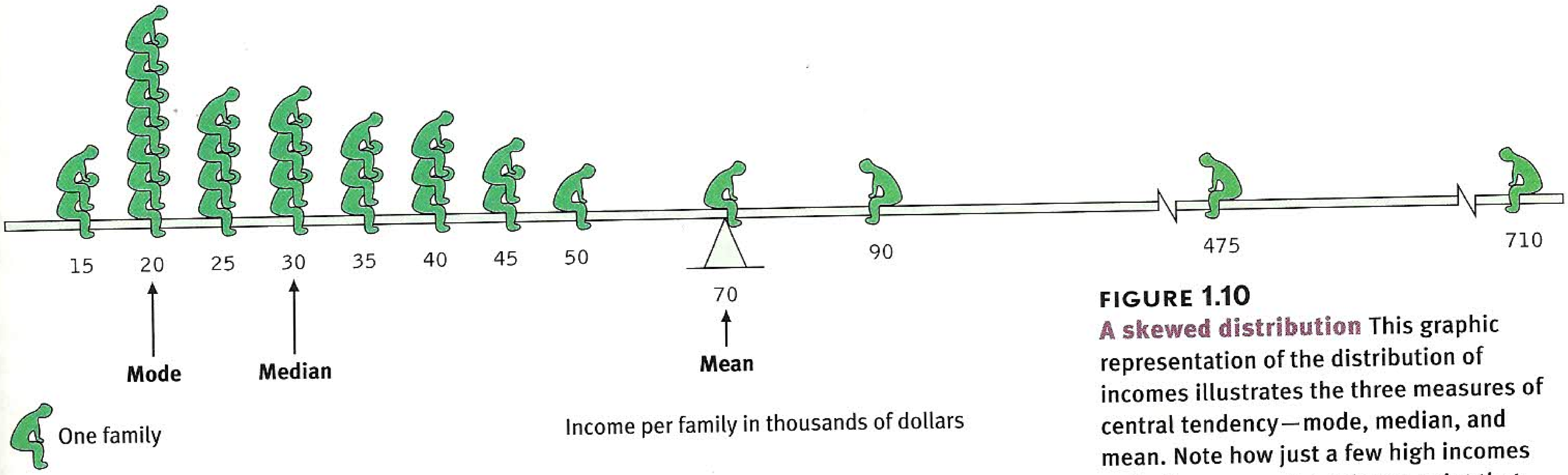


FIGURE 1.10
A skewed distribution This graphic representation of the distribution of incomes illustrates the three measures of central tendency—mode, median, and mean. Note how just a few high incomes make the mean—the fulcrum point that balances the incomes above and below—deceptively high.

In the United States, advocates and critics described the 2003 tax cut with different statistics, both true. The White House explained that "92 million Americans will receive an average tax cut of $1083." Critics agreed, but also noted that 50 million taxpayers got no cut, and half of the 92 million who did benefit received less than $100 (Krugman, 2003). Mean and median tell different true stories.

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.

The average person has one ovary and one testicle.

Measures of Variation

OBJECTIVE 18 | Describe two measures of variation.

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

TABLE 1.4

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	+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		

Class A: **Standard deviation =**

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

Class B: **Standard deviation =**

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

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

The **range** of scores—the gap between the lowest and highest scores—provides only a crude estimate of variation because a couple of extreme scores in an otherwise uniform group, such as the $475,000 and $710,000 incomes in Figure 1.10, 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 1.4**). (The computation assembles information about how much individual scores differ from the mean.) If your college or university attracts students of a certain ability level, their intelligence scores will have a smaller standard deviation than the one found in the more diverse community population outside your school.

Making Inferences

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

When Is an Observed Difference Reliable?

OBJECTIVE 19 | Identify three principles for making generalizations from samples.

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

1. ***Representative samples are better than biased samples.*** The best basis for generalizing is not from the exceptional and memorable cases one finds at the extremes (remember Bill Gates' income?) but from a representative sample of cases. No research involves a representative sample of the whole human population. Thus, it pays to keep in mind what population a study has sampled.

2. ***Less-variable observations are more reliable than those that are more variable.*** As we noted in the example of the basketball player whose points scored were consistent, an average is more reliable when it comes from scores with low variability.
3. ***More cases are better than fewer.*** An eager prospective university student visits two college campuses, each for a day. At the first, the student randomly attends two classes and discovers both instructors to be witty and engaging. At the next campus, the two sampled instructors seem dull and uninspiring. Returning home, the student (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.

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.

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

When Is a Difference Significant?

OBJECTIVE 20 | Explain how psychologists decide whether differences are meaningful.

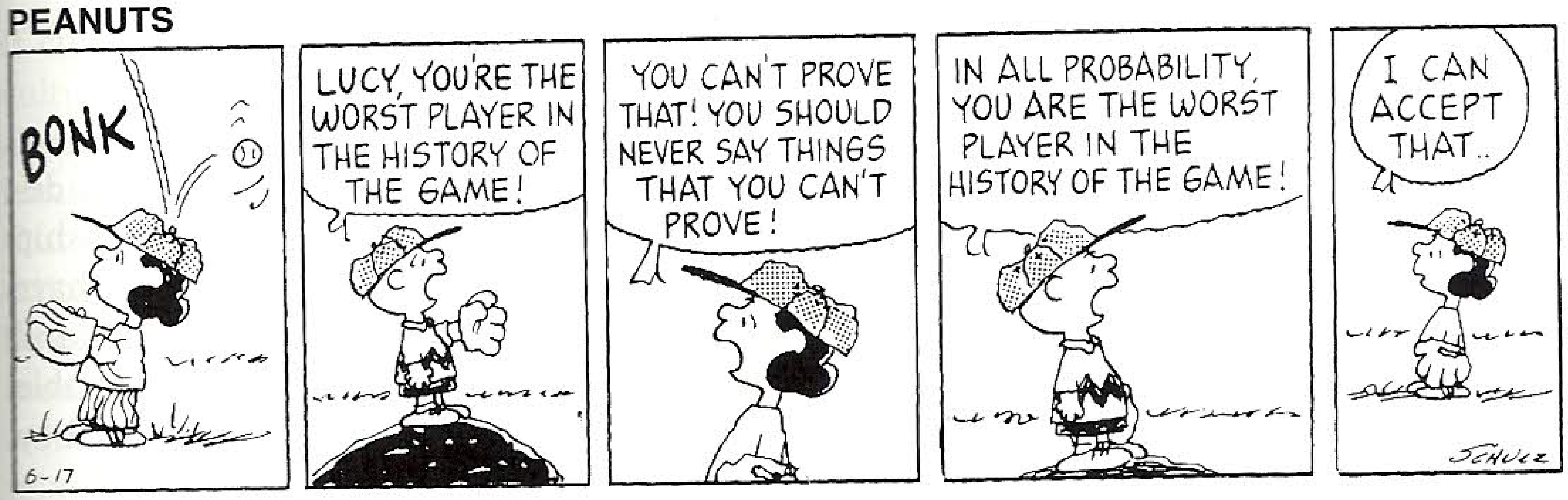
Statistical tests also help us determine whether differences are meaningful. Here is the underlying logic: When *averages* from two samples are each *reliable* measures of their respective populations (as when each is based on many observations that have small variability), then their difference (sometimes even a very small difference) is likely to be reliable as well. (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.) But when the *difference* between the sample averages is *large,* we have even more confidence that the difference between them reflects a real difference in their populations.

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

In judging statistical significance, psychologists are conservative. They are like juries who must presume innocence until guilt is proven. For most psychologists, proof beyond a reasonable doubt means not making much of a finding unless the odds of its occurring by chance are less than 5 percent (an arbitrary criterion).

When reading about research, you should remember that, given large enough or homogeneous enough samples, a difference between them may be "statistically significant" yet have little practical significance. For example, comparisons of intelligence test scores among 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 (Zajonc & Markus, 1975). But because the scores differ by only one or two points, the difference has little practical importance. Such findings have caused some psychologists to advocate alternatives to significance testing (Hunter, 1997). Better, they say, to use other ways to express a finding's "effect size"—its magnitude and reliability.

PEANUTS



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

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

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

>> LEARNING OUTCOMES

Statistical Reasoning

OBJECTIVE 15 | Explain the importance of statistical principles, and give an example of their use in everyday life.
Statistics help us to organize, summarize, and make inferences from data. We need not remember complicated formulas to think more clearly and critically about the data we encounter in everyday life. For example, understanding statistical concepts teaches us the importance of doubting big, round, undocumented numbers.

OBJECTIVE 16 | Explain how bar graphs can misrepresent data.
Scale labels and ranges used in bar graphs can be designed to minimize or maximize differences. When looking at statistical graphs in books and magazines and on TV and the Internet, think critically.

OBJECTIVE 17 | Describe the three measures of central tendency, and tell which is most affected by extreme scores.
The *median* is the middle score in a group of data. The *mode* is the most frequently occurring score. The *mean*, the arithmetic average, is most easily distorted by a few very high or very low scores.

OBJECTIVE 18 | Describe two measures of variation.
Measures of variation tell us how similar or diverse data are. A *range* describes the gap between the highest and lowest scores. The more useful measure, the *standard deviation*, states how much scores vary around the mean, or average, score.

OBJECTIVE 19 | Identify three principles for making generalizations from samples.
1. Representative samples are better than biased samples.
2. Less-variable observations are more reliable than those that are more variable.
3. More cases are better than fewer.

OBJECTIVE 20 | Explain how psychologists decide whether differences are meaningful.
When averages from two samples are each reliable measures of their own populations, and the difference between them is relatively large, we can assume the difference is significant—that the result did not occur by chance alone. Statistical significance indicates the likelihood of a result's occurring, not the importance of the result.

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

Frequently Asked Questions About Psychology

We have seen how case studies, surveys, and naturalistic observations help us describe behavior. We have also noted that correlational studies assess the relationship between two factors, which indicates how well 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. We have reflected on how a scientific approach, aided by statistics, can restrain biases.