Psychological Research Methods and Statistics

PSYCHOLOGY JOURN

For the next seven days, observe how statistics are used in the media. In your journal, describe the examples you find. ■

104,957

05,322
2,040
3,983
750
697
038
961
480

Online

Chapter Overview
Visit the *Understanding Psychology*
Web site at psychology.glencoe.com
and click on **Chapter 2—Chapter
Overviews** to preview the chapter.

What Is Research?

Reader's Guide

■ Main Idea

Psychologists must first decide how to approach the research issue. Then psychologists conduct the research in one of a variety of ways to test a hypothesis, solve a problem, or confirm previous findings.

■ Vocabulary
- sample
- naturalistic observation
- case study
- survey
- longitudinal study
- cross-sectional study
- correlation
- hypothesis
- variable
- experimental group
- control group

■ Objectives
- Describe the process of psychological research and the scientific method.
- Name the different types of psychological research.

EXPLORING PSYCHOLOGY

Do You Act This Way?

There are some chimps who, far more than others, constantly seem to try to ingratiate themselves with [win over] their superiors. Melissa, for one, particularly when she was young, used to hurry toward and lay her hand on the back or head of an adult male almost every time one passed anywhere near her. If he turned toward her, she often drew her lips back into a submissive grin as well. Presumably Melissa, like the other chimps who constantly attempt to ingratiate themselves in this way, is simply ill at ease in the presence of a social superior, so that she constantly seeks reassurance through physical contact. . . . There is much controversy as to how the human smile has evolved. It seems fairly certain, though, that we have two rather different kinds of smiles, . . . We smile when we are amused and we smile when we are slightly nervous, on edge, apprehensive. . . .

—from *In the Shadow of Man* by Jane Goodall, 1988

Jane Goodall observed the behavior of chimpanzees in Tanzania, Africa, to obtain data. She observed the behavior of chimps over a period of 30 years and provided much information about the animals' lives. Whereas Goodall used the research method of naturalistic observation, other scientists conduct experiments and surveys. All of these researchers, however, follow scientific methods.

Psychologists collect information somewhat like most people do in everyday life—only more carefully and more systematically. When you turn on the television and the picture is out of focus, you experiment with different knobs and dials until you find the one that works. When you ask a number of friends about a movie you are thinking of seeing, you are conducting an informal survey. Of course, there is more to doing scientific research than turning dials or asking friends what they think. Over the years psychologists, like other scientists, have transformed these everyday techniques for gathering and analyzing information into more precise tools.

PRE-RESEARCH DECISIONS

sample: the small group of participants, out of the total number available, that a researcher studies

Researchers must begin by asking a specific question about a limited topic or hypothesis. The next step is to look for evidence. The method a researcher uses to collect information partly depends on the research topic. For example, a social psychologist who is studying the effects of group pressure is likely to conduct an experiment. A psychologist who is interested in personality might begin with intensive case studies. Whatever approach to gathering data a psychologist selects, however, he or she must make certain basic decisions in advance.

Profiles In Psychology

Jane Goodall

1934–

"Every individual matters. Every individual has a role to play. Every individual makes a difference."

Jane Goodall, a British zoologist, became known for her work with chimpanzees in the wild. In 1960 she began her research at what is now Gombe Stream National Park in Tanzania. By living among the chimpanzees, she won their trust, observing their daily activities and writing detailed reports. She wrote, "The most wonderful thing about fieldwork, whether with chimps, baboons or any other wildlife, is waking up and asking yourself, 'What am I going to see today?'"

Goodall discovered while doing 30 years of research that chimps hunt and eat larger animals and make and use tools more than any other species except humans. Goodall also witnessed the first known instance in which one group of chimps systematically killed another group, even though the first group's survival was not threatened. This discovery surprised naturalists and suggested that behaviors like hunting, using tools, and warfare are not uniquely human.

Samples

Suppose a psychologist wants to know how the desire to get into college affects the attitudes of high school juniors and seniors. It would be impossible to study every junior and senior in the country. Instead, the researcher would select a **sample**, a relatively small group out of the total *population* under study—in this case, all high school juniors and seniors.

A sample must be *representative* of the population a researcher is studying. For example, if you wanted to know how tall American men were, you would want to make certain that your sample did not include a disproportionately large number of professional basketball players. Such a sample would be *nonrepresentative*; it would probably not represent American men in general.

There are two ways to avoid a nonrepresentative sample. One is to take a purely *random sample* so that each individual has an *equal chance* of being represented. For example, a psychologist might choose every twentieth name on school enrollment lists for a study of schoolchildren in a particular town. Random sampling is like drawing names or numbers out of a hat while blindfolded.

The second way to avoid a nonrepresentative sample is to deliberately pick individuals who represent the various subgroups in the population being studied. For example, the psychologist doing research on schoolchildren might select students of both sexes, of varying ages, of all social classes, and from all neighborhoods. This is called a stratified sample. In a *stratified sample*, subgroups in the population are represented proportionately in the sample. For example, if about 30 percent of schoolchildren in the United States are ages 5–8, then in a stratified sample of schoolchildren in the United States, 30 percent of those studied will be ages 5–8.

METHODS OF RESEARCH

The goals of research are to describe behavior, to explain its causes, to predict the circumstances under which certain behaviors may occur again, and to control certain behaviors. Psychologists use various methods of research to accomplish each of these goals.

Naturalistic Observation

Researchers need to know how people and animals behave naturally, when they are not conscious of being observed during an experiment. To obtain such information, a psychologist uses **naturalistic observation.** The cardinal rule of naturalistic observation is to avoid disturbing the people or animals you are studying by concealing yourself or by acting as unobtrusively as possible. Otherwise you may observe a performance produced because of the researcher's presence.

naturalistic observation: research method in which the psychologist observes the subject in a natural setting without interfering

Case Studies

A **case study** is an intensive study of a person or group. Most case studies combine long-term observations with diaries, tests, and interviews. Case studies can be a powerful research tool. Sigmund Freud's theory of personality development, discussed in Chapter 14, was based on case studies of his patients. Jean Piaget's theory of intellectual development, described in Chapter 3, was based in part on case studies of

case study: research method that involves an intensive investigation of one or more participants

survey: research method in which information is obtained by asking many individuals a fixed set of questions

longitudinal study: research method in which data is collected about a group of participants over a number of years to assess how certain characteristics change or remain the same during development

cross-sectional study: research method in which data is collected from groups of participants of different ages and compared so that conclusions can be drawn about differences due to age

his own children. By itself, however, a case study does not prove or disprove anything. The results cannot be generalized to anyone else. The researcher's conclusions may not be correct. Case studies, though, provide a wealth of descriptive material that may generate new hypotheses that researchers can then test under controlled conditions with comparison groups.

Surveys

One of the most practical ways to gather data on the attitudes, beliefs, and experiences of large numbers of people is through **surveys.** A survey may consist of interviews, questionnaires, or a combination of the two. Interviews allow a researcher to observe the participant and modify questions if the participant seems confused by them. On the other hand, questionnaires take less time to administer and the results are more uniform because everyone answers the same questions. Questionnaires also reduce the possibility that the researcher will influence the participant by unconsciously frowning at an answer he or she does not like. In interviews, there is always a danger that participants will give misleading answers in order to help themselves gain approval.

Longitudinal Studies

When conducting **longitudinal studies,** a psychologist studies the same group of people at regular intervals over a period of years to determine whether their behavior and/or feelings have changed and if so, how. Longitudinal studies are time-consuming and precarious; participants may disappear in midstudy. Longitudinal studies, however, are an ideal way to examine consistencies and inconsistencies in behavior over time. A good example was the New York Longitudinal Study begun in 1956. Psychologists followed 133 infants as they grew into adulthood, discovering that children are born with different temperaments (Thomas, Chess, & Birch, 1968).

Cross-Sectional Studies

An alternative approach to gathering data is cross-sectional studies. In a **cross-sectional study,** psychologists organize individuals into groups on the basis of age. Then, these groups are randomly sampled, and the members of each group are surveyed, tested, or observed simultaneously. Cross-sectional studies are less expensive than longitudinal studies and reduce the amount of time necessary for the studies.

In 1995 researchers conducted a cross-sectional study in which they showed three-, four-, six-, and seven-year-olds a picture of a serious-looking woman. The psychologists then asked the participants what they thought the woman was thinking about. The psychologists found that the older children seemed to have a clearer picture of mental processes. From

this discovery, the psychologists proposed that as children mature, their understanding of mental processes improves (Flavell, Green, & Flavell, 1995).

Correlations and Explanations

A researcher may simply want to observe people or animals and record these observations in a descriptive study. More often, however, researchers want to examine the *relationship* between two sets of observations—say, between students' grades and the number of hours they sleep.

Scientists use the word **correlation** to describe how two sets of data relate to each other. For example, there is a *positive correlation* between IQ scores and academic success. High IQ scores tend to go with high grades; low IQ scores tend to go with low grades. On the other hand, there is a *negative correlation* between the number of hours you spend practicing your tennis serve and the number of double faults you serve. As the hours of practice increase, errors decrease. In this case, a high rank on one measure tends to go with a low rank on the other (see Figure 2.1).

It is important to keep in mind that a correlation describes a relationship between two things. It does not mean, though, that one thing causes the other. In some cases, a third factor exists that may account for the positive correlation. Correlations do not identify what causes what. For example, although you might detect a positive correlation between sunny days and your cheerful moods, this does not mean that sunny days cause good moods.

correlation: the measure of a relationship between two variables or sets of data

Experiments

Why would a researcher choose experimentation over other research methods? Experimentation enables the investigator to *control* the situation and to decrease the possibility that unnoticed, outside variables will influence the results.

Figure 2.1   A Correlational Study
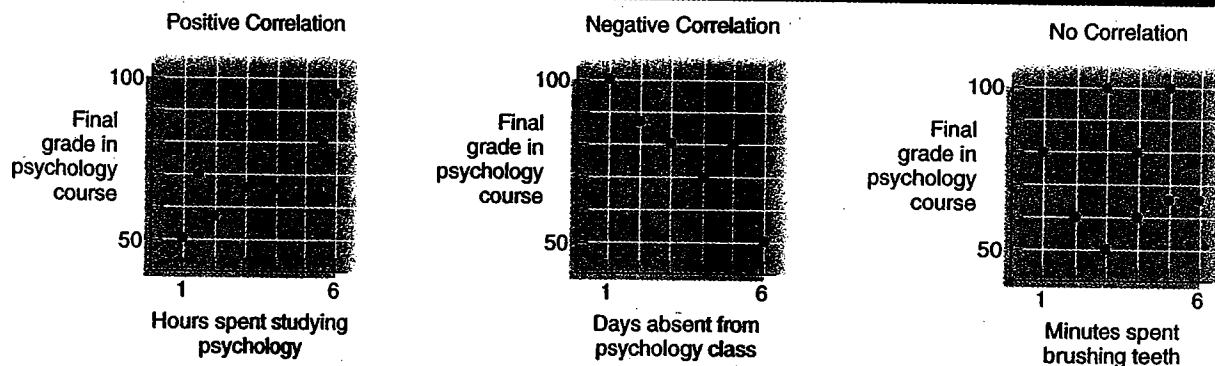
These charts display possible correlations between different variables. *How does time spent studying psychology correlate to the final grade in a psychology course?*

---

hypothesis: an educated guess about the relationship between two variables

variable: any factor that is capable of change

experimental group: the group to which an independent variable is applied

control group: the group that is treated in the same way as the experimental group except that the experimental treatment (the independent variable) is not applied

Every experiment has a **hypothesis,** or an educated guess, about the expected outcome; the researcher has some evidence for suspecting a specific answer. In a hypothesis, a psychologist will state what he or she expects to find. The hypothesis also specifies the important variables of the study.

In designing and reporting experiments, psychologists think in terms of **variables,** conditions and behaviors that are subject to change. There are two types of variables: independent and dependent. The *independent variable* is the one experimenters change or alter so they can observe its effects. If an effect is found, the *dependent variable* is the one that changes in relation to the independent variable. For example, the number of hours you study (the independent variable) affects your performance on an exam (the dependent variable).

Participants who are exposed to the independent variable are in the **experimental group.** Participants who are treated the same way as the experimental group, except that they are not exposed to the independent variable, make up the **control group** (see Figure 2.2). A control group is necessary in all experiments. Without it, a researcher cannot be sure the experimental group is reacting to what he or she thinks it is reacting to—a change in the independent variable. By comparing the way control and experimental groups behaved in an experiment (statistically), the researchers can determine whether the independent variable influences behavior and how it does so.

The results of any experiment do not constitute the final word on the subject, however. Psychologists do not fully accept the results of their own or other people's studies until the results have been *replicated*—that is, duplicated by at least one other psychologist with different participants. Why? Because there is always a chance that the studies may have hidden flaws.

Figure 2.2   Experimental Research

Psychology is an experimental science. Psychologists follow the same general procedures when conducting experimental research. *What are the dependent and independent variables of this experiment?*
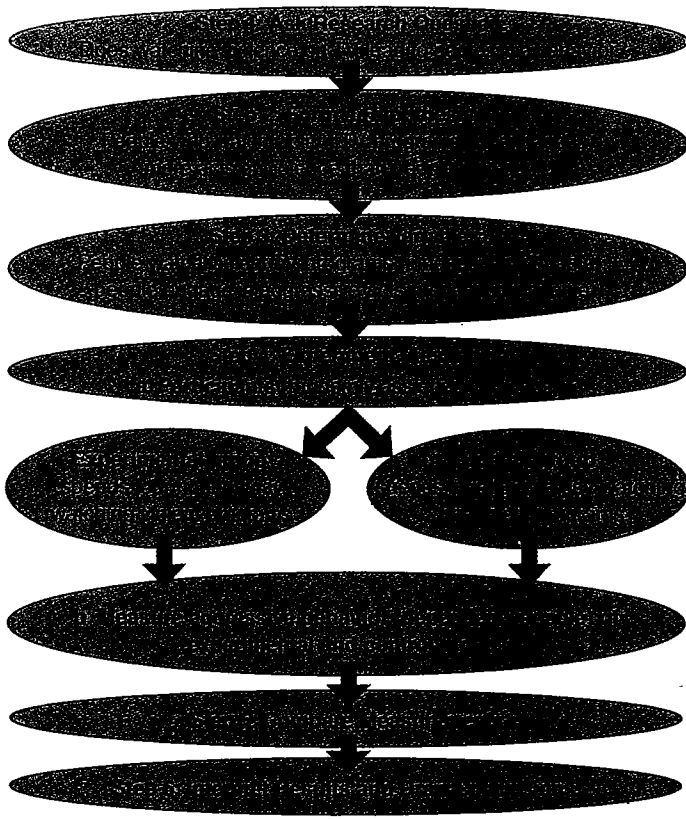

Ethical Issues

Ethics are the methods of conduct, or standards, for proper and responsible behavior. In 1992 the American Psychological Association published a set of ethical principles regarding the collection, storage, and use of psychological data. These principles are as follows:

• Using recognized standards of competence and ethics, psychologists plan research so as to minimize the possibility of misleading results. Any ethical

problems are resolved before research is started. The welfare and confidentiality of all participants are to be protected.

- Psychologists are responsible for the dignity and welfare of participants. Psychologists are also responsible for all research they perform or is performed by others under their supervision.

- Psychologists obey all state and federal laws and regulations as well as professional standards governing research.

- Except for anonymous surveys, naturalistic observations, and similar research, psychologists reach an agreement regarding the rights and responsibilities of both participants and researcher(s) before research is started.

- When consent is required, psychologists obtain a signed, informed consent before starting any research with a participant.

- Deception is used only if no better alternative is available. Under no condition is there deception about (negative) aspects that might influence a participant's willingness to participate.

- Other issues covered include sharing and utilizing data, offering inducements, minimizing evasiveness, and providing participants with information about the study.

Recently the use of animals in research has caused much concern and debate. Researchers have attempted to balance the rights of animals with the need for advancing the health of humans through research. While some people oppose subjecting animals to pain for research purposes, others point to the enormous gains in knowledge and reduction in human suffering that have resulted from such research.

Assessment

1. **Review the Vocabulary** Explain how a psychologist might select a sample for a survey.

2. **Visualize the Main Idea** In a chart similar to the one below, list and describe the advantages and disadvantages associated with each method of research.

3. **Recall Information** What pre-research decisions must a psychologist make?

4. **Think Critically** Why should psychologists question the results of an experiment that they have conducted for the first time?

5. **Application Activity** Suppose you wanted to find out whether there was a correlation between hours spent watching television and test grades in psychology class. Design a plan using one or more of the methods of research to help you study this correlation.

Problems and Solutions in Research

■ Main Idea
The investigation of psychological issues is a painstaking process. Psychologists must recognize and resolve errors while doing research.

■ Vocabulary
- self-fulfilling prophecy
- single-blind experiment
- double-blind experiment
- placebo effect

■ Objectives
- Summarize the methodological hazards of doing research.
- Examine experimental procedures psychologists use to avoid bias.

EXPLORING PSYCHOLOGY

Was She Doomed?

One young woman died of fear in a most peculiar way: When she was born, on Friday the 13th, the midwife who delivered her and two other babies that day announced that all three were hexed and would die before their 23rd birthday. The other two did die young. As the third woman approached her 23rd birthday, she checked into a hospital and informed the staff of her fears. The staff noted that she dealt with her anxiety by extreme hyperventilation (deep breathing). Shortly before her birthday, she hyperventilated to death.

—from *Introduction to Psychology* by James W. Kalat, 1999

Once an expectation is set, we tend to act in ways that are consistent with that expectation. How did the woman in the excerpt above die? Technically, when people do not breathe voluntarily, they breathe reflexively—the amount of carbon dioxide in the blood activates breathing. By breathing so deeply for so long (hyperventilating), the woman exhaled so much carbon dioxide that she did not have enough left in her bloodstream to trigger the breathing reflex. When she stopped breathing voluntarily, she stopped breathing altogether and died. In effect, the woman believed in the Friday the 13th hex and unintentionally fulfilled its prediction.

This is what we mean by a self-fulfilling prophecy. A **self-fulfilling prophecy** involves having expectations about a behavior and then acting in some way, usually unknowingly, to carry out that behavior.

In everyday life, we consciously or unconsciously tip off people as to what our expectations of them are. We give them cues, such as nodding

self-fulfilling prophecy: a situation in which a researcher's expectations influence that person's own behavior, and thereby influence the participant's behavior

and raising our eyebrows. People pick up on those cues and act as expected. Psychologists must be aware of such cues when conducting experiments. They must not allow their expectations to influence the results. The results must be unbiased. Science is a painstaking, exacting process. Every researcher must be wary of numerous pitfalls that can trap him or her into mistakes. In this section, we will look at some of the most common problems psychological researchers confront and how they cope with them.

AVOIDING A SELF-FULFILLING PROPHECY

Sometimes an experimenter's behavior may unwittingly influence the results. The experimenter may unintentionally raise an eyebrow or nod when posing a question, thus influencing the person being studied. One way to avoid this self-fulfilling prophecy is to use a double-blind technique. Suppose a psychologist wants to study the effects of a particular tranquilizer. She might give the drug to an experimental group and a placebo (a substitute for the drug that has no medical benefits) to a control group. The next step would be to compare their performances on a series of tests. This is a **single-blind experiment.** The participants are "blind" in the sense that they do not know whether they have received the tranquilizer or the placebo. What does it mean, then, if the participants taking the placebo drug report that they feel the effects of the tranquilizer? It means that their expectations have played a role—that they felt the effects because they believed they were taking a tranquilizing drug, not because of the drug itself.

The researcher will not know who takes the drug or the placebo. She may, for example, ask the pharmacist to number rather than label the pills. *After* she scores the tests, she goes back to the pharmacist to learn which participants took the tranquilizer and which took the placebo. This is a **double-blind experiment.** Neither the participants nor the experimenter knows which participants received the tranquilizer. This eliminates the possibility that the researcher will unconsciously find what she expects to find about the effects of the drug. The researcher remains unbiased.

THE MILGRAM EXPERIMENT

In the 1960s Stanley Milgram wanted to determine whether participants would administer painful shocks to others merely because an authority figure had instructed them to do so. Milgram collected nearly 1,000 male

---

More About...

The Hawthorne Study

In 1939 a group of industrial psychologists set out to determine how to increase workers' productivity at a General Electric plant in Hawthorne, Illinois (Roethlisberger & Dickson, 1939). The participants were eight assembly line workers. In the first experiment, the psychologists gradually increased the lighting in the room (the independent variable) and observed the effect on productivity (the dependent variable). Productivity improved as the lighting was increased. In a second experiment, the participants were permitted to take rest breaks. This also increased the productivity of the workers. Next, the psychologists reduced the lighting levels, and again productivity increased. The psychologists found that no matter what they did, productivity increased. Why? The psychologists soon recognized that the participants realized they were receiving special attention. This motivated the workers to work harder, thus increasing their productivity.

The results of the experiment in Hawthorne generated studies in human relations and management that apply to work situations today.

single-blind experiment: an experiment in which the participants are unaware of which participants received the treatment

double-blind experiment: an experiment in which neither the experimenter nor the participants know which participants received which treatment

Figure 2.3   Single-Blind and Double-Blind Experiments

Researchers must take measures during experimentation to guard against seeing only what they expect to see. *Why would a researcher conduct a double-blind experiment?*

	Experimenter	Participants	Organizer of Experiment
Single-Blind Experiment	aware	unaware	aware
Double-Blind Experiment	unaware	unaware	aware

participants, including college students and adults in different occupations. Milgram told the group of paid volunteers that he was studying the effects of punishment on learning. Milgram introduced each volunteer to a "learner"—actually someone posing as a learner. The volunteer watched the learner attempt to recite a list of paired words that he supposedly had memorized earlier. Each time the learner made a mistake, the volunteer, or "teacher," was ordered to push a button to deliver an electric shock to the learner. The shocks, mild at first, would increase with each mistake to a painful and dangerous level of 450 volts.

The volunteers at this point did not realize that the shocks were false because the learners displayed distress and pain, screaming and begging for the electric shocks to stop. Although the task did not seem easy for them, most of the volunteers delivered a full range of the electric shocks to the learners. (Sixty-five percent of the volunteers pushed the shock button until they reached maximum severity.)

Reading Check
Why can the Milgram experiment be classified as a single-blind experiment?

The results implied that ordinary individuals could easily inflict pain on others if such orders were issued by a respected authority. Later, Milgram informed the volunteers that they had been deceived and that no shocks had actually been administered. This was a good example of a single-blind experiment because the participants were unaware that they were not administering a shock. Critics raised the following questions, though. How would you feel if you had been one of Milgram's participants? Did Milgram violate ethical principles when he placed participants in a position to exhibit harmful behavior? Was the deception Milgram used appropriate? Did the information gained outweigh the deception? Before the start of any experiment today, the experimenter is required to submit a plan to a Human Subjects Committee that can either approve or reject the ethics of the experiment.

Milgram's hypothesis and experiment has been applied in similar studies. In Milgram's original study, more than half of the participants (26 of 40, or 65 percent) administered the highest level of shock. Researchers at Swarthmore College hypothesized that Milgram's findings were due, in part, to the fact that his participants were mostly

middle-aged, working-class men. Most had probably served in the military during World War II and thus had experience taking orders and obeying authority. Young, liberal, highly educated Swarthmore students would obey less. Yet, surprisingly, 88 percent of the Swarthmore undergraduates administered the highest level of shock!

THE PLACEBO EFFECT

When researchers evaluate the effects of drugs, they must always take into account a possible placebo effect. The **placebo effect** is a change in a patient's illness or physical state that results solely from the patient's knowledge and perceptions of the treatment. The placebo is some sort of treatment, such as a drug or injection, that resembles medical therapy yet has no medical effects.

In one study (Loranger, Prout, & White, 1961), researchers divided hospitalized psychiatric patients into two experimental groups and a control group. They gave the experimental groups either a "new tranquilizer" or a "new energizer" drug. The control group received no drugs at all. After a six-week period, the researchers evaluated the experimental groups. Fifty-three to eighty percent of the experimental groups reported that they had indeed benefited from the drugs. Yet all the drugs administered during the experiment were placebos. The participants had reacted to their own expectations of how the drug given to them would affect them. Neither the researchers nor the patients were aware that the drugs were placebos until after the experiment.

People spend millions of dollars a year on herbal remedies such as these, which have not been proven to cure their ills.

placebo effect: a change in a participant's illness or behavior that results from a belief that the treatment will have an effect, rather than the actual treatment

Assessment

1. **Review the Vocabulary** Explain how psychologists try to avoid the self-fulfilling prophecy.

2. **Visualize the Main Idea** Use a diagram similar to the one below to outline an experiment discussed in this section.

   Hypothesis:
   ▼
   Independent   Dependent
   Variables: Variables:
   ▼
   Results:
   ▼
   Conclusions:

3. **Recall Information** What questions about the Milgram experiment did critics raise? How are today's experiments restricted in regards to ethics?

4. **Think Critically** How can the expectations of the participants bias the results of an experiment? How can the expectations of the experimenter bias the results of an experiment?
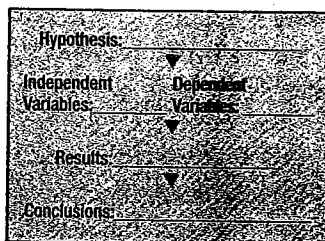
5. **Application Activity** Describe a single-blind experiment you might set up. Explain your hypothesis and the participants' tasks.

The Case of
Clever Hans

Period of Study: 1911

Introduction: A horse, Clever Hans, grew famous throughout Europe for his startling ability to answer questions. Taught by his owner, Mr. von Osten, Hans seemed to be able to add, subtract, multiply, divide, spell, and solve problems, even when his owner was not around. Oskar Pfungst decided to investigate the humanlike intelligence of the horse.

Hypothesis: Two different hypotheses are involved in this case. First, Mr. von Osten, believing that horses could be as intelligent as humans, hypothesized that he could teach Hans some problem-solving abilities. Pfungst, on the other hand, believed that horses could not learn such things and, while investigating this theory, developed a hypothesis that Hans, the horse, was reacting to visual cues to answer questions.

Method: Mr. von Osten, a German mathematics teacher, started by showing Hans an object while saying "One" and at the same time lifting Hans's foot once. Von Osten would lift Hans's foot twice for two objects, and so on. Eventually Hans learned to tap his hoof the correct number of times when von Osten called out a number. For four years, von Osten worked with Hans on more and more complex problems, until Hans was able to answer any question given him.

Upon hearing of the amazing horse, Pfungst grew skeptical and investigated. Pfungst soon discovered that Hans responded correctly to questions only when the questioner had calculated the answer first. Then Pfungst realized that Hans's answers proved wrong when the horse could not see the questioner. To test his hypothesis, Pfungst fitted the horse with blinders. The horse failed to answer the questions. Eventually Pfungst realized that the questioner would unknowingly give Hans clues as to the right answer. For example, after asking a question, the questioner would lean forward to watch Hans's foot. This was a cue for Hans to start tapping. Pfungst observed that "as the experimenter straightened up, Hans would stop tapping, he found that even the raising of his eyebrows was sufficient. Even the dilation of the questioner's nostrils was a cue for Hans to stop tapping." (Pfungst, 1911) Questioners involuntarily performed these actions, and Hans responded to the visual signals.

Results: Von Osten believed that he had been teaching the horse how to solve problems and answer questions, when in fact he had been teaching Hans to make simple responses to simple signals. Pfungst had uncovered errors in von Osten's experiments. Von Osten had practiced a self-fulfilling prophecy—he had unintentionally communicated to Hans how he expected the horse to behave. Pfungst had learned the truth by isolating the conditions under which Hans correctly and incorrectly answered questions. He had carefully observed the participant's reactions under controlled conditions.

Analyzing the Case Study

1. How did Mr. von Osten test his hypothesis?

2. What errors did von Osten make while testing his hypothesis?

3. **Critical Thinking** If Pfungst had not come along and found the truth, how could we discover today how Hans answered the questions?

Statistical Evaluation

■ Main Idea
Psychologists must collect and evaluate evidence to support their hypotheses.

■ Vocabulary
* statistics
* descriptive statistics
* frequency distribution
* normal curve
* central tendency
* variance
* standard deviation
* correlation coefficient
* inferential statistics

■ Objectives
* Recognize types of descriptive statistics.
* Describe inferential statistics.

EXPLORING PSYCHOLOGY

When Statistics Lie

Long ago, when Johns Hopkins University had just begun to admit women students, someone not particularly enamored of [happy with] coeducation reported a real shocker: Thirty-three and one-third percent of the women at Hopkins had married faculty members! The raw figures gave a clearer picture. There were three women enrolled at the time, and one of them had married a faculty man.

—from *How to Lie With Statistics* by Darrell Huff, 1954

Although people may use statistics to distort the truth (such as in the example above), people may also use statistics honestly to support their hypotheses. In order to allow statistics to validly support a hypothesis, psychologists must collect meaningful data and evaluate it correctly.

How many times have you been told that in order to get good grades, you have to study? A psychology student named Kate has always restricted the amount of TV she watches during the week, particularly before a test. She has a friend, though, who does not watch TV before a test but who still does not get good grades. This fact challenges Kate's belief. Although Kate hypothesizes that among her classmates, those who watch less TV get better grades, she decides to conduct a survey to test the accuracy of her hypothesis. Kate asks 15 students in her class to write down how many hours of TV they watched the night before a psychology quiz and how many hours they watched on the night after the quiz. Kate collects additional data. She has her participants check off familiar products on a

Baseball Statistics

Let's look at how statistics are used in one of our most popular sports, baseball. A batting average is the number of hits per official "at bats" (walks do not count). If a player has a batting average of .250, it means that on average he or she gets a hit every fourth time at the plate.

The earned run average represents the number of runs a pitcher allows per 9 innings of play. Consider the pitcher who pitches 180 innings in a season and allows 60 runs. On the average, this pitcher allows one run every 3 innings (180 innings divided by 60 runs). One run every 3 innings equals 3 runs every 9 innings, so the earned run average is 3. The next time you watch your favorite sport, think about the part that statistics plays in it.

list of 20 brand-name items that were advertised on TV the night before the quiz. Kate also asks her participants to give their height.

When the data are turned in, Kate finds herself overwhelmed with the amount of information she has collected. Her data are presented in Figure 2.4. How can she organize it all so that it makes sense? How can she analyze it to see whether it supports or contradicts her hypothesis? The answers to these questions are found in **statistics,** a branch of mathematics that enables researchers to organize and evaluate the data they collect. We will explore the statistical procedures that help psychologists make sense out of the masses of data they collect.

statistics: the branch of mathematics concerned with summarizing and making meaningful inferences from collections of data

descriptive statistics: the listing and summarizing of data in a practical, efficient way

DESCRIPTIVE STATISTICS

When a study such as Kate's is completed, the first task is to organize the data in as brief and clear a manner as possible. For Kate, this means that she must put her responses together in a logical format. When she does this, she is using **descriptive statistics,** the listing and summarizing of data in a practical, efficient way, such as through graphs and averages.

Figure 2.4	Kate's Data			
Before	**After**	**Grade***	**Products**	**Height**
0.0	1.5	5	2	71
0.5	2.5	10	4	64
0.5	2.5	9	6	69
1.0	2.0	10	14	60
1.0	2.5	8	10	71
1.0	1.5	7	9	63
1.5	3.0	9	7	70
1.5	2.5	8	12	59
1.5	2.5	8	9	75
1.5	3.0	6	14	60
2.0	3.0	5	13	68
2.5	2.5	3	17	65
2.5	3.5	4	10	72
3.0	3.0	0	18	62
4.0	4.0	4	20	67

Kate's data show the number of hours of television watched before and after the quiz, the grade on the quiz, the number of products recognized, and participants' height in inches. *How much television did the two students with the best grades watch the night before the quiz?*

* Highest grade possible is 10.

Figure 2.5   A Frequency Distribution

A frequency distribution shows how often a particular observation occurs. *How many students watched three or more hours of television the night before the quiz?*

Hours	Frequency Before*	Frequency After*
0.0	1	0
0.5	2	0
1.0	3	0
1.5	4	2
2.0	1	1
2.5	2	6
3.0	1	4
3.5	0	1
4.0	1	1
Total	15	15

*Number of students

Distributions of Data

One of the first steps that researchers take to organize their data is to create frequency tables and graphs. Tables and graphs provide a rough picture of the data. Are the scores bunched up or spread out? What score occurs most often? Frequency distributions and graphs provide researchers with their initial look at the data.

Kate is interested in how many hours of TV her participants watched the night before and the night after the quiz. She uses the numbers of hours of TV viewing as categories, and then she counts how many participants reported each category of hours before and after the quiz. She has created a table called a frequency distribution (see Figure 2.5). A **frequency distribution** is a way of arranging data so that we know how often a particular score or observation occurs.

What can Kate do with this information? A commonly used technique is to figure out percentages. This is done simply by dividing the frequency of participants within a category by the total number of participants and multiplying by 100. Before the quiz, about 13 percent of her participants (2 divided by 15) watched TV for 2.5 hours. On the night after the quiz, 40 percent of her participants watched 2.5 hours of TV (6 divided by 15). If you are familiar with the use of percentages, you know that test grades are often expressed as percentages (the number of correct points divided by the total number of questions times 100). Sometimes frequency distributions include a column giving the percentage of each occurrence.

frequency distribution: an arrangement of data that indicates how often a particular score or observation occurs

Figure 2.6   A Frequency Polygon

This graph shows the number of hours of TV watched the night before the quiz and the night after the quiz. *How do the two lines compare?*

— Hours of TV watched before quiz

— Hours of TV watched after quiz
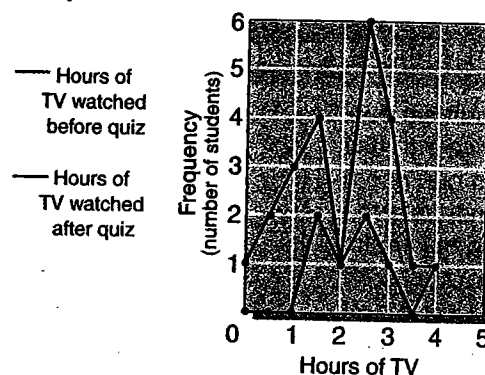
Figure 2.7   A Normal Curve

The maximum frequency lies in the center of a range of scores in a perfect normal curve. The frequency tapers off as you reach the edges of the two sides. **Where is the mean located in a normal curve?**
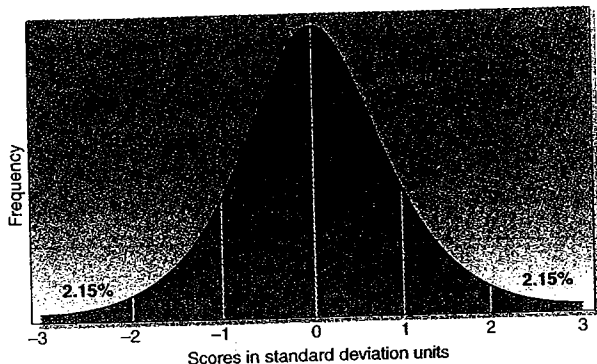

It is often easier to visualize frequency information in the form of a graph. Since Kate is most interested in how much TV her classmates watched, she decides to graph the results. Kate constructs a histogram. *Histograms* are very similar to bar graphs except that histograms show frequency distribution by means of rectangles whose widths represent class intervals and whose areas are proportionate to the corresponding frequencies.

Another kind of graph is the frequency polygon or frequency curve. Figure 2.6 is a frequency polygon. It shows the same information presented in a different way. Instead of boxes, a point is placed on the graph where the midpoint of the top of each histogram bar would be. Then the points are connected with straight lines.

Frequency polygons are useful because they provide a clear picture of the shape of the data distribution. Another important feature is that more than one set of data can be graphed at the same time. For example, Kate might be interested in comparing how much TV was watched the night before the quiz with the amount watched the evening after the quiz. She can graph the "after quiz" data using a different kind of line. The comparison is obvious; in general, her participants watched more TV on the night after the quiz than on the night before the quiz.

Figure 2.8   **Measures of Central Tendency**
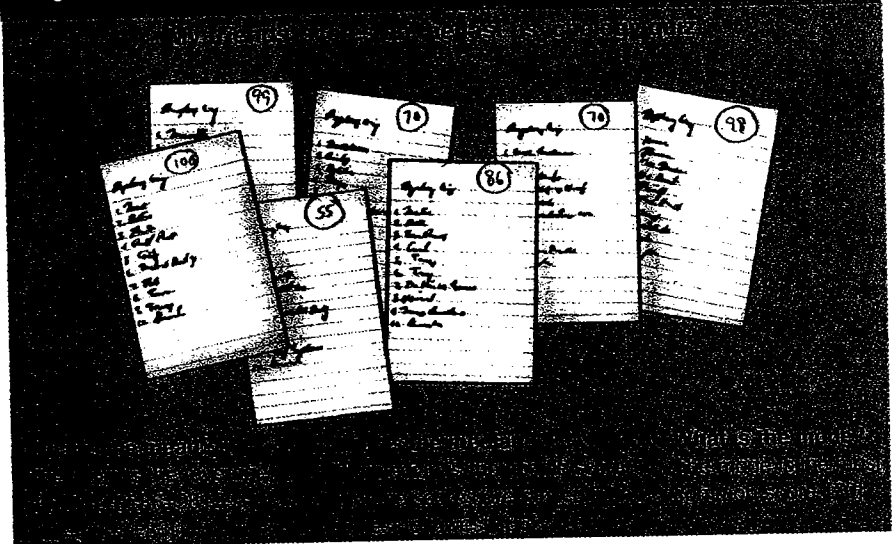
It is often useful to summarize a set of scores by identifying a number that represents the center, average, or most frequently occurring number of the distribution. **If your score matched the median on the last psychology quiz, how did you do in comparison to your classmates?**

Imagine that Kate could measure how much TV everyone in Chicago watched one night. If she could graph that much information, her graph would probably look something like Figure 2.7. A few people would watch little or no TV, a few would have the TV on all day, while most would watch a moderate amount of TV. Therefore, the graph would be highest in the middle and taper off toward the tails, or ends, of the distribution, giving it the shape of a bell.

This curve is called the **normal curve** (or bell-shaped curve). Many variables, such as height, weight, and IQ, fall into such a curve if enough people are measured. The normal curve is symmetrical. This means that if a line is drawn down the middle of the curve, one side of the curve is a mirror image of the other side. It is an important distribution because of certain mathematical characteristics. We can divide the curve into sections and predict how much of the curve, or what percentage of cases, falls within each section.

normal curve: a graph of frequency distribution shaped like a symmetrical, bell-shaped curve; a graph of normal distribution

Measures of Central Tendency

Most of the time, researchers want to do more than organize their data. They want to be able to summarize information about the distribution into statistics. For example, researchers might want to discuss the average height of women or the most common IQ test score. One of the most common ways of summarizing is to use a measure of **central tendency**— a number that describes something about the "average" score. We shall use Kate's quiz grades (refer back to Figure 2.4) in the examples that follow.

central tendency: a number that describes something about the "average" score of a distribution

The *mode* is the most frequent score. In a graphed frequency distribution, the mode is the peak of the graph. The most frequently occurring quiz grade is 8; that is, more students received an 8 than any other score. Distributions can have more than one mode. The data for height presented in Figure 2.4 have two modes: 60 and 71. Distributions with two modes are called *bimodal*.

✔️ **Reading Check**
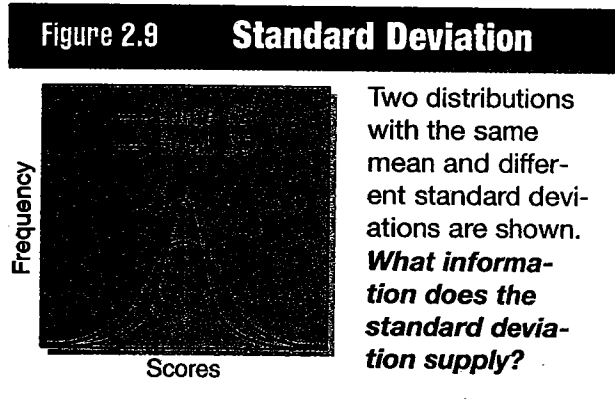What is the difference between the mean and the mode?

When scores are put in order from least to most, the *median* is the middle score. Since the median is the midpoint of a set of values, it divides the frequency distribution into two halves. Therefore, 50 percent of the scores fall below the median, and 50 percent fall above the median. For an odd number of observations, the median is the exact middle value.

The *mean* is what most people think of as an average and is the most commonly used measure of central tendency. To find the mean (or $\overline{X}$), add up all the scores and then divide by the number of scores added. The mean equals the sum of the scores on variable X divided by the total number of observations. For the quiz grades, the sum of the scores is 96, and the number of scores is 15. The mean equals 96 divided by 15, giving us a mean quiz grade of 6.4.

The mean can be considered the balance point of the distribution, like the middle of a seesaw, since it does reflect all the scores in a set of data. If the highest score in a data set is shifted higher, the mean

Figure 2.9   Standard Deviation

Two distributions with the same mean and different standard deviations are shown. *What information does the standard deviation supply?*

More About...

Transforming Scores

Suppose you take the ACT and score a 26. Then you take the SAT and get a 620. The college you want to go to will accept either test score. Which score should you send? (Which score is better?) To make a comparison between two scores that have different distributions, different means, and different variabilities, you must transform the scores.

ACT

18 = Mean
6 = Standard deviation

SAT

500 = Mean
100 = Standard deviation
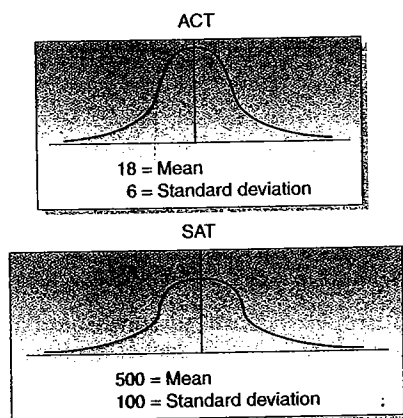
If you look at the distributions of the ACT and SAT, you will find that the ACT has a mean of 18 and a standard deviation of 6. So you take your score on the ACT (26) and subtract the mean from it (26−18) to get 8; 8 is 1.33 standard deviations above the mean (8/6). Do the same for your SAT score [620−500 = 120; so 620 is 1.2 standard deviations above the mean (120/100)]. So which score would you submit to the college of your choice? (The correct answer is your ACT score because 1.33 is greater than 1.2.)

What we just did was to make a standard score. A standard score is a transformed score that provides information about its location in a distribution.

variance: a measure of difference, or spread

standard deviation: a measure of variability that describes an average distance of every score from the mean

will shift upward also. If we change the highest quiz grade from 10 to 20, the mean changes from 6.4 to 7.1.

Measures of Variance

Distributions differ not only in their average score but also in terms of how spread out, or how variable, the scores are. Figure 2.9 shows two distributions drawn on the same axis. Each is symmetrical, and each has the same mean. However, the distributions differ in terms of their variance. Measures of **variance** provide an index of how spread out the scores of a distribution are.

Two commonly used measures of variance are the range and the standard deviation. To compute the *range,* subtract the lowest score in a data set from the highest score and add 1. The highest quiz grade is 10 and the lowest is 0, so the range is 11, representing 11 possible scores 0–10. The range uses only a small amount of information, and it is used only as a crude measure of variance.

The **standard deviation** is a better measure of variance because, like the mean, it uses all the data points in its calculation. It is the most widely used measure of variance. The standard deviation is a measure of distance. It is like (but not exactly like) an average distance of every score to the mean of the scores. This distance is called a *deviation* and is written: $X - \bar{X}$. Scores above the mean will have a positive deviation, and scores below the mean will have a negative deviation. The size of the typical deviation depends on how variable, or spread out, the distribution is. If the distribution is very spread out, deviations tend to be large. If the distribution is bunched up, deviations tend to be small. The larger the standard deviation, the more spread out the scores are (see Figure 2.9).

Correlation Coefficients

A **correlation coefficient** describes the direction and strength of the relationship between two sets of observations (recall the discussion of correlations in Section 1). The most commonly used measure is the Pearson correlation coefficient (r). A coefficient with a plus (+) sign indicates a *positive correlation.* This means that as one variable *increases,* the second variable also *increases.* For example, the more you jog, the better your cardiovascular system works. A coefficient with a minus (−) sign indicates a *negative correlation;* as one variable *increases,* the second variable *decreases.* For example, the more hours a person spends watching TV, the fewer hours are available for studying. Correlations can take any value between +1 and −1 including 0. An r near +1 or −1 indicates a strong relationship (either positive or negative), while an r near 0 indicates a weak relationship.

Generally, an r from ±0.60 to ±1.0 indicates a strong correlation, from ±0.30 to ±0.60 a moderate correlation, and from 0 to ±0.30 a weak correlation. A correlation of ±1.0 indicates a perfect relationship between two variables and is very rare.

To get an idea of how her data look, Kate draws some scatterplots. A *scatterplot* is a graph of participants' scores on the two variables, and it demonstrates the direction of the relationship between them. Figure 2.10 illustrates one of Kate's correlations. Note that each point represents one person's score on two variables.

correlation coefficient: describes the direction and strength of the relationship between two sets of variables

INFERENTIAL STATISTICS

The purpose of descriptive statistics is to describe the characteristics of a sample. Psychologists, however, are not only interested in the information they collect from their participants, but they also want to make generalizations about the population from which the participants come. To make such generalizations, they need the tools of inferential statistics. Using **inferential statistics,** researchers can determine whether the data they collect support their hypotheses, or whether their results are merely due to chance outcomes.

inferential statistics: numerical methods used to determine whether research data support a hypothesis or whether results were due to chance

Probability and Chance

If you toss a coin in the air, what is the probability that it will land with heads facing up? Since there are only two possible outcomes, the probability of heads is 0.50. If you toss a coin 100 times, you would expect 50 heads and 50 tails. If the results were 55 heads and 45 tails, would you think the coin is fair? What if it were 100 heads and zero tails?

When a researcher completes an experiment, he or she is left with lots of data to analyze. The researcher must determine whether the findings from the experiment support the hypothesis (for example, the coin is fair) or whether the results are due to chance. To do this, the researcher must perform a variety of statistical tests, called measures of statistical significance. When researchers conclude that their findings are statistically significant, they are stating, at a high level of confidence, that their results are not due to chance.

Statistical Significance

For many traits in a large population, the frequency distribution follows a characteristic pattern, called the normal curve (see Figure 2.7). For example, if you measured the heights of 500 students chosen at random from your high school, you would find very few extremely tall people and very few extremely short people. The majority of students

Figure 2.10   A Scatterplot

When there is little or no relationship between two variables, the points in the scatterplot do not seem to fall into any pattern. *What conclusions can you draw from this scatterplot?*
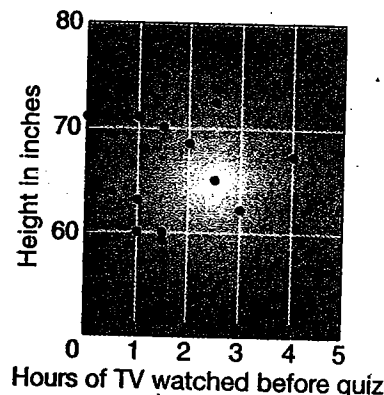
Quick Lab

Do some people really have psychic powers?

A well-known psychic sometimes explains his performance by saying the following: Think of a number between 1 and 50. Both digits must be odd numbers, but they must not be the same. For example, it could be 15, but it could not be 11. Please choose a number and I will tell you what number you are thinking of.

Procedure

1. Develop a hypothesis that explains how the psychic is performing this feat. (Hint: The psychic uses statistics, not magic.)

2. Try out the psychic's act on several of your classmates and record their responses.

Analysis

1. Based on the psychic's directions, decide which numbers can be used and which numbers will most likely be used.

2. How do your observations support or contradict your hypothesis?

See the Skills Handbook, page 622, for an explanation of designing an experiment.

---

would fall somewhere in the middle. Suppose Kate wants to know if her classmates watch more TV than the "average American." Since daily TV viewing is probably normally distributed, she can compare her results to the normal distribution if she knows the population's mean number of TV viewing hours.

When psychologists evaluate the results of their studies, they ask: Could the results be due to chance? What researchers really want to know is whether the results are so extreme, or so far from the mean of the distribution, that they are more likely due to chance.

The problem is that this question cannot be answered with a yes or no. This is why researchers use some guidelines to evaluate probabilities. Many researchers say that if the probability that their results were due to chance is less than 5 percent (0.05), then they are confident that the results are not due to chance. Some researchers want to be even more certain, and so they use 1 percent (0.01) as their level of confidence. When the probability of a result is 0.05 or 0.01 (or whatever level the researcher sets), we say that the result is *statistically significant*. It is important to remember that probability tells us how likely it is that an event or outcome is due to chance, but not whether the event is *actually* due to chance.

When does a statistically significant result *not* represent an important finding? Many statistical tests are affected by sample size. A small difference between groups may be magnified by a large sample and may result in a statistically significant finding. The difference, however, may be so small that it is not a meaningful difference.

Assessment

1. **Review the Vocabulary** What is the difference between a frequency distribution and a histogram? Between a normal curve and a scatterplot?

2. **Visualize the Main Idea** Using an organizer similar to the one at right, list and describe the measures of central tendency.

   ① _____
   ② _____
   ③ _____
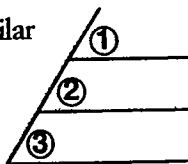
3. **Recall Information** What is the importance of the normal curve?

4. **Think Critically** What does correlation tell you about the relationship between two variables?

5. **Application Activity** Conduct a class or family survey on an issue, then display your findings in a frequency distribution, frequency polygon, or scatterplot. Apply evaluation rules. What conclusions can you reach from your results?

Summary and Vocabulary

Psychologists learn about what they do not know by carefully and systematically collecting information. They then must describe and analyze their research findings through various statistical measurements and interpret their results.

Section 1 What Is Research?

Main Idea: Psychologists must first decide how to approach the research issue. Then psychologists conduct the research in one of a variety of ways to test a hypothesis, solve a problem, or confirm previous findings.

- Researchers begin their research by asking a specific question about a limited topic; determining the validity of a claim, hypothesis, or theory; and choosing an unbiased sample.
- Psychologists use several methods of research to accomplish their research goals. These methods include naturalistic observation, case studies, surveys, and experiments.
- Psychologists follow a set of ethical principles that govern their research.

Section 2 Problems and Solutions in Research

Main Idea: The investigation of psychological issues is a painstaking process. Psychologists must recognize and resolve errors while doing research.

- In a self-fulfilling prophecy, an experimenter has expectations about a participant's behavior and then acts in some way, usually unknowingly, to influence that behavior.
- In single-blind experiments, the participants do not know which participants have received the treatment.
- Researchers can avoid a self-fulfilling prophecy by using the double-blind technique in their experiments.
- When researchers evaluate the effects of drugs, they must always take into account a possible placebo effect.

Section 3 Statistical Evaluation

Main Idea: Psychologists must collect and evaluate evidence to support their hypotheses.

- Researchers use descriptive statistics to organize data in a practical, efficient way.
- Descriptive statistics include distributions of data, measures of central tendency, measures of variance, and correlation coefficients.
- Researchers use inferential statistics to make generalizations about the population from which the participants come.
- Researchers perform a variety of statistical tests, called measures of statistical significance, to determine whether findings from their experiment support the hypothesis or whether the results are due to chance.

Chapter Vocabulary

sample (p. 36)
naturalistic observation (p. 37)
case study (p. 37)
survey (p. 38)
longitudinal study (p. 38)
cross-sectional study (p. 38)
correlation (p. 39)
hypothesis (p. 40)
variable (p. 40)
experimental group (p. 40)
control group (p. 40)
self-fulfilling prophecy (p. 42)
single-blind experiment (p. 43)
double-blind experiment (p. 43)
placebo effect (p. 45)
statistics (p. 48)
descriptive statistics (p. 48)
frequency distribution (p. 49)
normal curve (p. 51)
central tendency (p. 51)
variance (p. 52)
standard deviation (p. 52)
correlation coefficient (p. 52)
inferential statistics (p. 53)

Reviewing Vocabulary

Choose the letter of the correct term or concept
below to complete the sentence.

a. variance
b. sample
c. longitudinal study
d. control group
e. single-blind experiment

f. double-blind experiment
g. placebo effect
h. statistics
i. normal
j. frequency distribution

1. _____ is a branch of mathematics that helps researchers organize and evaluate data.

2. In a(n) _____, only the participants of the experiment do not know whether they are in the experimental group or the control group.

3. Measures of _____ indicate how spread out the scores of a distribution are.

4. A bell-shaped curve is a(n) _____ curve.

5. In an experiment, the _____ includes the participants who are not exposed to experimental variables.

6. The _____ is a change in a patient's physical state that results from the patient's perceptions of the treatment.

7. Researchers use a(n) _____ to arrange data so that they know how often a particular observation occurs.

8. Researchers generally select a(n) _____, which is a relatively small group of the total population that is being studied.

9. In a(n) _____, neither the participants nor the experimenter knows whether the participants are in the experimental group or the control group.

10. In a(n) _____, a researcher studies a group of people over a period of years.

Recalling Facts

1. What are two ways that a researcher can avoid a biased sample?

2. When do researchers use naturalistic observation?

3. How does a self-fulfilling prophecy present a problem for researchers?

4. Using a graphic organizer similar to the one below, identify and explain the kinds of descriptive statistics.

5. Why do researchers use inferential statistics? How do inferential statistics describe data differently than descriptive statistics?

Critical Thinking

1. Synthesizing Information How could you attempt to disprove the following hypothesis? You can raise blood pressure by making a participant anxious.

2. Analyzing Statements Explain the following statement: "Correlation does not imply causation."

3. Making Inferences What correlation would you expect between students' grades and class attendance?

4. Applying Concepts How are statistics used within your classroom? Within your school?

5. Analyzing Information Various kinds of statistics are used in sports. Provide examples of statistics from various sports.

Assessment

Psychology Projects

1. What is Research? Choose a traffic intersection near your home or school that has a stop sign. Design a study to assess whether or not motorists stop at the posted sign. Consider the research questions you need to answer, such as how to determine whether motorists comply with the sign, the number of vehicles, and the time of day. Conduct your study and record your observations.

2. Statistical Evaluation Collect heights from 20 women and 20 men. Create a frequency distribution for each group, and divide them into 5-inch intervals before counting. Graph your data for men and women separately as frequency polygons on the same axis. Compute means, medians, modes, ranges, and standard deviations for women and men separately. How are the two distributions alike and different?

Technology Activity

Does smoking cause lung cancer? Some scientists cite animal studies as proving that it does. Representatives of the tobacco industry state that animal studies cannot be generalized to humans. Search the Internet to find arguments and data from each side of this debate. Use that information to support both viewpoints in an essay.

Psychology Journal

For each of the examples of statistics you listed in your journal (at the beginning of the chapter), indicate whether you feel that enough information was provided to evaluate the validity of any reported claims. What other information should have been provided? How might additional information change the reported conclusions?

Building Skills

Interpreting Graphs Review the graphs, then answer the questions that follow.

1. What does each of the graphs illustrate?
2. Which age group owns or uses the most cellular phones?
3. How has the number of cellular subscribers changed since 1988?
4. Do most cell-phone owners talk on the phone while driving? What percentage never use their phones while driving?
5. What arguments might these statistics be used to support or refute?

Practice and **assess** key social studies skills with **Glencoe Skillbuilder Interactive Workbook CD-ROM, Level 2.**

See the Skills Handbook, page 628, for an explanation of interpreting graphs.