UNIT 0

An Introduction to Psychological Science Practices: Research Methods and Data Interpretation

As you will learn through this course, psychology is the *science* of behavior and mental processes. What do psychologists do? You might think that they offer counseling, analyze personality, give out child-raising advice, examine crime scenes, and testify in court. Do they? *Yes*—and much more. Consider some of psychology's questions that you may wonder about:

- Have you ever awakened from a nightmare and wondered why you had such a crazy dream? ***Why do we dream?***
- Have you ever played peekaboo with a 6-month-old and wondered why the baby finds your disappearing/reappearing act so delightful? ***What do babies perceive and think?***
- Have you ever wondered what fosters school and work success? ***Does inborn intelligence explain why some people get richer, think more creatively, or relate more sensitively? Or does gritty effort and a belief in the power of persistence matter more?***
- Have you ever become depressed or anxious and wondered whether you'll ever feel "normal"? ***What triggers our bad moods—and our good ones? What's the line between a routine mood swing and a psychological disorder?***

Psychology seeks to answer such questions about us all—how and why we think, feel, and act as we do. But as a *science*, psychology does more than speculate: It uses research and interpretation of the resulting data to separate uninformed opinions from examined conclusions.

MODULES

UNIT 0 Overview Video

Oli Scarff/Getty Images

◀ Leo, aged 9 months, wears a Geodesic Sensor Net as he prepares to take part in an experiment at the Birkbeck Babylab Centre for Brain and Cognitive Development in London, England. The experiment uses an electroencephalogram (EEG) to study brain activity while the baby examines different objects of varying complexity. The scientists use various experiments to test the babies' physical or cognitive responses with sensors including: eye-tracking, brain activation, and motion capture.

Module 0.1 The Scientific Attitude, Critical Thinking, and Developing Arguments

AP® Exam Tip

There are four science practices that all students should develop throughout the AP® Psychology course: Concept Application, Research Methods & Design, Data Interpretation, and Argumentation. Look for the helpful AP® Science Practice features that are incorporated throughout the text. Transferring these skills to the AP® Psychology Exam will be key to your success.

LEARNING TARGETS

0.1-1 Explain how psychology is a science.

0.1-2 Describe the three key elements of the scientific attitude and how they support scientific inquiry.

0.1-3 Explain how critical thinking feeds a scientific attitude, and smarter thinking for everyday life.

AP® Science Practice

Research

Because the scientific approach is so foundational to psychology, we will provide research tips (such as this one) throughout the modules in this text. This will allow you to learn about and understand research terminology in the context of psychological theories and concepts. Research methods and design will be an important part of the AP® exam!

Psychology Is a Science

0.1-1 How is psychology a science?

Underlying all science is, first, a passion for exploring and understanding without misleading or being misled. Some questions (*Is there life after death?*) are beyond science. Answering them in any way requires a leap of faith. With many other ideas (*Can some people demonstrate extrasensory perception [ESP]?*), the proof is in the pudding. Let the facts speak for themselves. In other words, look for *scientifically derived* evidence.

Magician James Randi (1928–2020) used an evidence-based approach that drew on observation and experimentation when testing those claiming to see glowing auras around people's bodies:

Randi: *Do you see an aura around my head?*

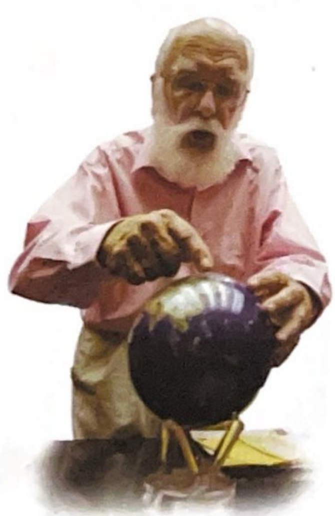
Aura seer: *Yes, indeed.*

Randi: *Can you still see the aura if I put this magazine in front of my face?*

Aura seer: *Of course.*

Randi: *Then if I were to step behind a wall barely taller than I am, you could determine my location from the aura visible above my head, right?*

Randi once told me [DM] that no aura seer agreed to take this simple test.


Alan Diaz/AP Photo

The Amazing Randi The late magician James Randi was an enthusiastic skeptic. During his life, he tested and debunked supposed psychic phenomena.

Key Elements of the Scientific Attitude

0.1-2 What are the three key elements of the scientific attitude and how do they support scientific inquiry?

No matter how sensible-seeming or how wild an idea is, the smart thinker asks: *Does it work?* When put to the test, do the data support its predictions? When subjected to scrutiny, crazy-sounding ideas sometimes find support. More often, science is self-cleansing. Right ideas stick around. Wrong ideas head to the waste heap, where they're discarded atop previous claims of miracle cancer cures and out-of-body travels into centuries past. To sift reality from fantasy and fact from fiction requires a scientific attitude: being skeptical but not cynical, open-minded but not gullible.

Putting a scientific attitude into practice requires not only curiosity and skepticism but also humility—awareness of our vulnerability to error and openness to surprises and new perspectives. What matters is not our opinion or yours, but rather the truths revealed by our questioning and testing. If people or other animals don't behave as our ideas predict, then so much the worse for our inaccurate ideas—and so much the better for scientific progress. One of psychology's early mottos expressed this humble attitude: "The rat is always right." See Developing Arguments: The Scientific Attitude.

AP® Science Practice

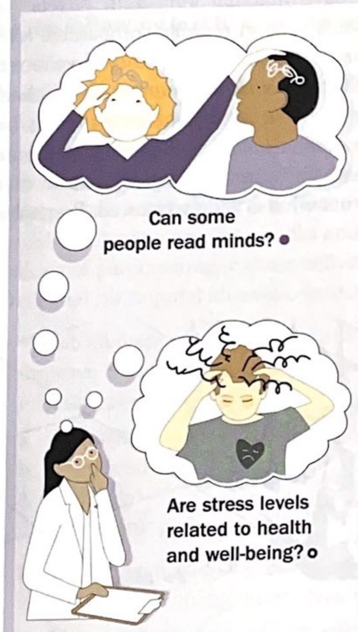
Developing Arguments

The Scientific Attitude

1 CURIOSITY:

Does it work?

When put to the test, can its predictions be confirmed?



• *No one has yet been able to demonstrate extrasensory mind-reading.*

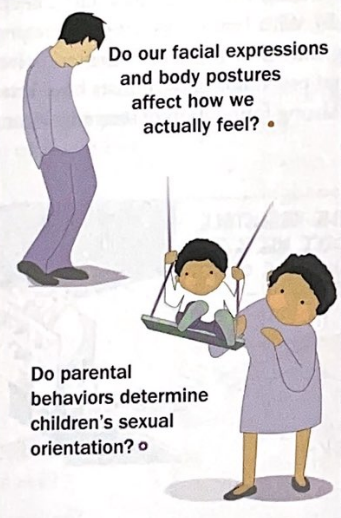
o *Many studies have found that higher stress relates to poorer health.*

2 SKEPTICISM:

What do you mean?

How do you know?

Sifting reality from fantasy requires a healthy skepticism — an attitude that is not cynical (doubting everything), but also not gullible (believing everything).



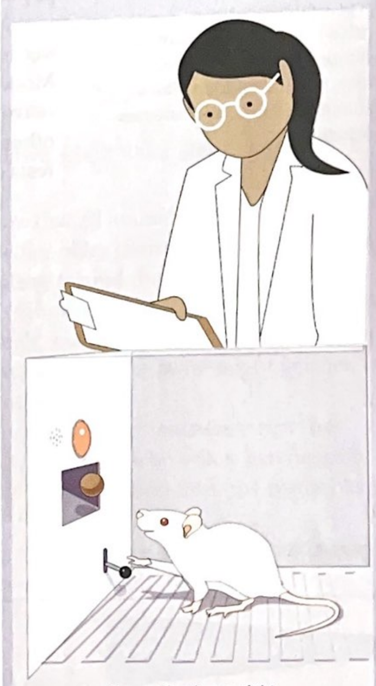
• *Our facial expressions and body postures can affect how we feel.*

o *Module 3.3 explains that there is not a relationship between parental behaviors and children's sexual orientation.*

3 HUMILITY:

That was unexpected! Let's explore further.

Researchers must be willing to be surprised and follow new ideas. People and other animals don't always behave as our ideas and beliefs would predict.



The rat is always right.

Developing Arguments Questions

By addressing these critical thinking questions, you will enhance your ability to develop arguments based on scientifically derived evidence. This is one of the skills that will be emphasized on the AP® exam.

1. Identify the reasoning against extrasensory mind-reading.
2. Using scientifically derived evidence, explain why skepticism is important in science.

Critical Thinking

0.1-3 How does critical thinking feed a scientific attitude, and smarter thinking for everyday life?

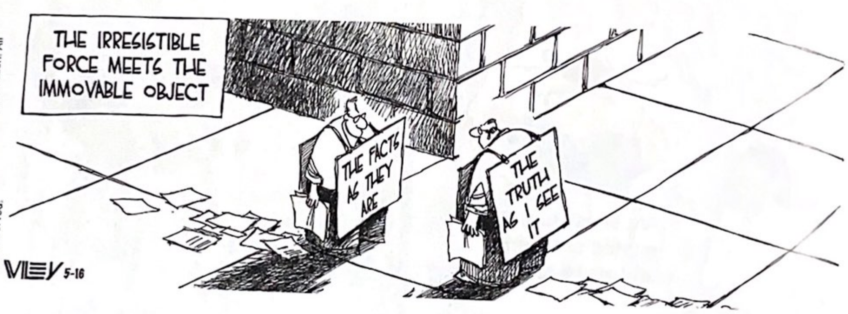
The scientific attitude—curiosity + skepticism + humility—prepares us to think harder and smarter. This smart thinking, called **critical thinking**, examines assumptions, appraises the source, discerns hidden biases, evaluates evidence, and assesses conclusions. When reading a research report, an online opinion, or a news story, critical thinkers ask questions: *How do they know that? What is this person's agenda? Is the conclusion based on an anecdote, or scientifically derived evidence? Does the evidence justify a cause-effect conclusion? What alternative explanations are possible?*

From a tongue-in-cheek Twitter feed:
"The problem with quotes on the internet is that you never know if they're true."
Abraham Lincoln

Critical thinkers wince when people make factual claims based on their gut: "I *feel like* climate change is [or isn't] happening." "I *feel like* self-driving cars are more [or less] dangerous." Such beliefs (commonly mislabeled as feelings) may or may not be true. Critical thinkers realize that they might be wrong. Sometimes, the best evidence confirms our beliefs. At other times, it beckons us to a different way of thinking. Cynics sometimes seem smart, yet most demonstrate less cognitive ability and academic competence than average (Stavrova & Ehlebracht, 2019).

critical thinking thinking that does not automatically accept arguments and conclusions. Rather, it examines assumptions, appraises the source, discerns hidden biases, evaluates evidence, and assesses conclusions.

Critical inquiry can surprise us. Here are some examples from psychological science: Massive losses of brain tissue early in life may have minimal long-term effects (see Module 1.4). People of differing age, gender, and wealth report roughly comparable levels of personal happiness (see Module 5.2). Depression touches many people, but most recover (see Module 5.4). Critical inquiry also sometimes debunks popular presumptions, by checking intuitive fiction with scientific fact: Sleepwalkers are *not* acting out their dreams (see Module 1.5). With brain stimulation or hypnosis, someone *cannot* immediately replay and relive long-buried or repressed memories (see Module 2.7). In these instances, and many others, what psychological scientists have learned is not what is widely believed. Psychology rests on a strong foundation of scientific inquiry.



AP® Science Practice

Check Your Understanding

Examine the Concept

- Explain what's involved in critical thinking.

Apply the Concept

- Were you surprised to learn that psychology is a science? How would you defend the point that psychology is a science if someone asked you about this?

Answers to the Examine the Concept questions can be found in Appendix C at the end of the book.

Module 0.1 REVIEW

0.1-1 How is psychology a science?

- Psychology's findings are the result of a scientific approach—based on careful observation and testing. Sifting reality from fantasy requires a scientific attitude.

0.1-2 What are the three key elements of the scientific attitude, and how do they support scientific inquiry?

- The scientific attitude equips us to be curious, skeptical, and humble in scrutinizing competing ideas or our own observations.

0.1-3 How does critical thinking feed a scientific attitude, and smarter thinking for everyday life?

- Critical thinking puts ideas to the test by examining assumptions, appraising the source, discerning hidden biases, evaluating evidence, and assessing conclusions.

AP® Practice Multiple Choice Questions

Use the following text to answer questions 1 and 2:

Harrison observed that people who read fiction novels were also artists. Whenever Harrison shared this observation with his friends, they said that teenagers only do those activities to meet school requirements. Harrison disagreed, instead believing that personality explained an association between reading novels and being artistic. Even when Harrison learned that having access to books and art was the biggest predictor of participating in these activities, he continued to believe that his original idea was correct.

1. Which element of the scientific attitude should Harrison improve if he wants to engage in scientific inquiry more effectively?

a. Curiosity
b. Skepticism
c. Humility
d. Questioning

2. If Harrison wanted to increase his use of critical thinking, which of the following questions might he ask?

a. How can I share my results with more people?
b. How can I sell my ideas to make a profit?
c. How can I interpret this result?
d. How can I minimize the impact of this result?

3. What is the name of the approach that uses observation and testing to draw conclusions?

a. Qualitative
b. Quantitative
c. Scientific
d. Critical

4. Which of the following best illustrates a skeptical attitude?

a. Evelyn believes that all animals turn into balloons at night because her older brother told her that he saw their dog floating around the house at night.
b. After Matsuo's mother told him that chewing on his fingernails might make him sick, Matsuo asked her to provide an explanation of how he might get sick from chewing his nails.
c. Hayden argues with all her teachers when they lecture because she doubts how they can possibly remember all the information from the textbooks they have read.
d. After reading a blog post online about the dangers of walking near busy streets, Beck decided to avoid walking near busy streets for the next five years.

5. Which of the elements of a scientific attitude is most associated with the question, "Does it work?"

a. Humility
b. Skepticism
c. Argumentation
d. Curiosity

6. Why is psychology considered a science?

a. Psychological researchers rely on observation and experimentation when testing claims.
b. Psychological researchers ask questions about behavior and mental processes.
c. Psychological researchers speculate about how and why people act as they do.
d. Psychological researchers utilize critical thinking.

Module 0.2 The Need for Psychological Science

LEARNING TARGETS

0.2-1 Explain how cognitive biases, such as hindsight bias, overconfidence, and the tendency to perceive order in random events illustrate why science-based answers are more valid than those based on common sense.

0.2-1 How do cognitive biases, such as hindsight bias, overconfidence, and the tendency to perceive order in random events illustrate why science-based answers are more valid than those based on common sense?

Some people suppose that psychology is mere common sense—documenting and dressing in jargon what people already know: "You get paid for using fancy methods to prove what everyone knows?" Indeed, our intuition is often right. As the baseball great Yogi Berra (1925–2015) once said, "You can observe a lot by watching." (We have Berra to thank for other gems, such as "Nobody ever goes there any more—it's too crowded," and "If the people don't want to come out to the ballpark, nobody's gonna stop 'em.") Because we're all behavior watchers, it would be surprising if many of psychology's findings had *not* been foreseen. Many people believe that love breeds happiness, for example, and they are right (we have what Module 4.7 calls a deep "need to belong").

But sometimes what seems like common sense, informed by countless casual observations, is wrong. In many other modules in this text, we will see how research has overturned popular ideas—that familiarity breeds contempt, that dreams predict the future, and that most of us use only 10 percent of our brain. We will also see how research has surprised us with discoveries about how the brain's chemical messengers control our moods and memories, about other animals' abilities, and about the relationship between social media use and depression.

Other things seem like commonsense truth only because we so often hear them repeated. Mere repetition of statements—whether true or false—makes them easier to

MYTH: GOING OUTSIDE IN COLD WEATHER WILL MAKE YOU SICK

THE FACTS:

- Colds and the flu are caused by viruses.
- Viruses spread more in the winter because people are indoors more, humidity is lower, and having a lower body temperature weakens the immune system.
- You can get hypothermia or frostbite from the cold.

THE BOTTOM-LINE:

Cold weather can make you more susceptible to viruses, but going outside while it's cold will not make you sick.

Critical thinking beats common sense Psychological scientists use critical thinking to determine whether scientifically derived evidence supports their assumptions. Critical thinking helps us discard myths and seek the truth.

hindsight bias the tendency to believe, after learning an outcome, that one would have foreseen it. (Also known as the *I-knew-it-all-along phenomenon.*)

process and remember, and thus more true-seeming (Dechêne at al., 2010; Fazio et al., 2015). Easy-to-remember misconceptions ("Bundle up before you go outside, or you will catch a cold!") can therefore overwhelm hard truths.

Three roadblocks to critical thinking—*hindsight bias, overconfidence,* and *perceiving patterns in random events*—help illustrate why we cannot rely solely on common sense.

Did We Know It All Along? Hindsight Bias

Consider how easy it is to draw the bull's-eye *after* the arrow strikes. After a couple breaks up, their friends say, "They weren't a good match." After the game, we credit the coach if a "gutsy play" wins the game, and criticize the same "stupid play" if it doesn't. After a war or an election, its outcome usually seems obvious. Although history may therefore seem like a series of inevitable events, the actual future is seldom foreseen. No one's diary recorded, "Today the Hundred Years War began."

This **hindsight bias** (also known as the *I-knew-it-all-along phenomenon*) is easy to demonstrate by giving half the members of a group some purported psychological finding and giving the other half an opposite result. Tell the first group, for example, "Psychologists have found that separation weakens romantic attraction. As the saying goes, 'Out of sight, out of mind.'" Ask them to imagine why this might be true. Most people can, and after hearing an explanation, nearly all will then view this true finding as unsurprising.

Tell the second group the opposite: "Psychologists have found that separation strengthens romantic attraction. As the saying goes, 'Absence makes the heart grow fonder.'" People given this untrue result can also easily imagine it, and most will also see it as unsurprising. When opposite findings both seem like common sense, there is a problem.

Such errors in people's recollections and explanations show why we need psychological research. It's not that common sense is usually wrong. Rather, common sense describes, after the fact, what *has* happened better than it predicts what *will* happen.

AP® Science Practice

Developing Arguments

Identify the reasoning, using the scientifically derived evidence presented here, that supports the claim that hindsight bias is a roadblock to critical thinking.

Everett Collection/Newscom

Hindsight bias When drilling its Deepwater Horizon oil well in 2010, BP employees took shortcuts and ignored warning signs, without intending to harm the environment, their employees, or their company's reputation. After the resulting oil spill in the Gulf of Mexico, with the benefit of 20/20 hindsight, the foolishness of those judgments became obvious.

More than 800 scholarly papers have shown hindsight bias in people young and old from around the world (Roese & Vohs, 2012). As physicist Niels Bohr reportedly jested, "Prediction is very difficult, especially if it's about the future."

Overconfidence

We humans tend to think we know more than we do. Asked how sure we are of our answers to factual questions (*Is Boston north or south of Paris?*), we tend to be more confident than correct.[1] And our confidence often drives us to quick—rather than correct—thinking (Rahnev et al., 2020). Consider these three anagrams, shown beside their solutions (from Goranson, 1978):

WREAT → WATER
ETRYN → ENTRY
GRABE → BARGE

How many seconds do you think it would have taken you to unscramble each of these? Knowing the answer tends to make us overconfident. (Surely, the solution would take only 10 seconds or so?) In reality, the average problem solver spends 3 minutes, as you also might, given a similar anagram without the solution: OCHSA.[2]

Are we any better at predicting social behavior? Psychologist Philip Tetlock (1998, 2005) collected more than 27,000 expert predictions of world events, such as whether Quebec would separate from Canada. His repeated finding: These predictions, which experts made with 80 percent confidence on average, were right less than 40 percent of the time. It turns out that only about 2 percent of people excel at predicting social behavior. Tetlock (with Gardner, 2016) calls them "superforecasters." Superforecasters avoid overconfidence. Faced with a difficult prediction, a superforecaster "gathers facts, balances clashing arguments, and settles on an answer."

Perceiving Order in Random Events

We're born with an eagerness to make sense of our world. People see a face on the Moon, or believe their watch shows some number groups more than others. Even in random

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

1. Boston is south of Paris.
2. The anagram solution: CHAOS.

> "The really unusual day would be one where nothing unusual happens."
>
> Statistician Persi Diaconis (2002)

AP® Exam Tip

Multiple-choice questions on the AP® exam sometimes test your knowledge of "media myths." Pay particular attention when psychological findings run counter to "common sense." You will also need to explain *why* claims are or are not supported, using scientifically derived evidence.

data, we often find patterns, because—here's a curious fact of life—*random sequences often don't look random* (Falk et al., 2009; Nickerson, 2002, 2005). Flip a coin 50 times and you may be surprised at the streaks of heads and tails—much like supposed "hot" and "cold" streaks in basketball shooting and baseball hitting. In actual random sequences, patterns and streaks (such as repeating digits) occur more often than people expect (Oskarsson et al., 2009). That also makes it hard for people to generate random-like sequences. When embezzlers try to simulate random digits when specifying how much to steal, their nonrandom patterns can alert fraud experts (Poundstone, 2014).

Why are we so prone to pattern-seeking? For most people, a random, unpredictable world is unsettling (Tullett et al., 2015). Making sense of our world relieves stress and helps us get on with daily living (Ma et al., 2017).

The point to remember: Our commonsense thinking is flawed due to three powerful tendencies: hindsight bias, overconfidence, and our tendency to perceive patterns in random events. But scientific inquiry can help us sift reality from illusion.

AP® Science Practice

Check Your Understanding

Examine the Concept

▶ Explain the difference between hindsight bias and overconfidence.

Apply the Concept

▶ Do you have a hard time believing you may be overconfident? Could overconfidence be at work in that self-assessment? How might reading this section about overconfidence help reduce your tendency to be overconfident?

▶ Explain why, after friends start dating, we often feel that we *knew* they were meant to be together.

Answers to the Examine the Concept questions can be found in Appendix C at the end of the book.

Module 0.2 REVIEW

0.2-1 How do cognitive biases, such as hindsight bias, overconfidence, and the tendency to perceive order in random events illustrate why science-based answers are more valid than those based on common sense?

- *Hindsight bias* (also called the "I-knew-it-all-along phenomenon") is the tendency to believe, after learning an outcome, that we would have foreseen it.
- Overconfidence in our judgments results partly from our bias to seek information that confirms them.
- These tendencies, along with our eagerness to perceive patterns in random events, lead us to overestimate the importance of commonsense thinking. Although limited by the testable questions it can address, scientific inquiry can help us overcome such biases and shortcomings.

AP® Practice Multiple Choice Questions

1. The tendency to exaggerate the correctness or accuracy of our beliefs and predictions is called

a. hindsight bias.
b. overconfidence.
c. critical thinking.
d. skepticism.

2. While sitting at a stoplight, Nancy believes that the next car she sees will be blue because the previous three cars have been blue. Which psychological concept best explains her belief?

a. Hindsight bias
b. Critical thinking
c. Perceiving order in random events
d. Overconfidence

3. After the student council election, a friend tells you he could have guessed who would be elected president. Which psychological phenomenon might this scenario best illustrate?

a. Common sense
b. Hindsight bias
c. Overconfidence
d. Perceiving order in random events

4. While taking a standardized test with randomly scrambled answers, you notice that your last four answers have been "c." Which of the following is true concerning the probability of the next answer being "c"?

a. It is higher because previous answers have been "c." Once a streak begins, it is likely to last for a while.
b. It is lower since prior answers have been "c." Since answers are distributed randomly, "c" answers become less common.
c. It is unaffected by previous answers. It is as likely to be "c" as any other answer.
d. It is higher based on previous answers being "c." Test constructors' default answer choice is "c."

5. Which of the following is an example of hindsight bias?

a. Armend is certain that electric cars will represent 80 percent of vehicles in 20 years and only reads research studies that support his hypothesis.
b. Liza underestimates how much time it will take her to finish writing her college application essays and fails to meet an important deadline.
c. Alliyah, after reading a definition on one of her flashcards, turns the card over to see the term and then tells herself she knew what the answer was all along.
d. Dr. Grace overestimates the effectiveness of her new treatment method because she fails to seek out any evidence refuting her theory.

Module 0.3 The Scientific Method

LEARNING TARGETS

0.3-1 Describe how theories advance psychological science.

0.3-2 Explain how psychologists use case studies, naturalistic observations, and surveys to observe and describe behavior, and why random sampling is important.

The Scientific Method

AP® Exam Tip

As you read this module, keep in mind that the scientific method is a set of principles and procedures, not a list of facts. You will be expected to understand how the science of psychology is done, not just what it has discovered.

Psychological scientists use the *scientific method*—a self-correcting process for evaluating ideas with observation and analysis. Psychological science welcomes hunches and plausible-sounding theories. And it puts them to the test. If a theory works—if the data support its predictions—so much the better for that theory. If the predictions fail, the theory gets revised or rejected. When researchers submit their work to a scientific journal, **peer reviewers**—other scientists who are experts—evaluate a study's theory, originality, and accuracy. The journal editor then uses the peer reviews to decide whether the research deserves publication.

Constructing Theories

0.3-1 How do theories advance psychological science?

In everyday conversation, we often use *theory* to mean "mere hunch." In science, a **theory** *explains* behaviors or events by offering ideas that *organize* observations. By using deeper principles to organize isolated facts, a theory summarizes and simplifies. As we connect the observed dots, a coherent picture emerges. A theory of how sleep affects memory, for example, helps us organize countless sleep-related observations into a short list of principles. Imagine that we observe over and over that people with good sleep habits tend to answer questions correctly in class and do well at test time. We might therefore theorize that sleep improves memory. So far, so good: Our principle neatly summarizes a list of observations about the effects of a good night's sleep.

Yet no matter how reasonable a theory may sound—and it does seem reasonable to suggest that sleep boosts memory—we must put it to the test. A good theory produces testable *predictions*, called **hypotheses**. Such predictions specify which results would support the theory and which results would disconfirm it. The **falsifiability** of a hypothesis is a mark of its scientific strength. (Can it be proven false?) To test our theory about sleep effects on memory, we might hypothesize that when sleep deprived, people will remember less from the day before. To test that hypothesis, we might assess how well people remember class materials they studied either before a good night's sleep or before

peer reviewers scientific experts who evaluate a research article's theory, originality, and accuracy.

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

hypothesis a testable prediction, often implied by a theory.

falsifiable the possibility that an idea, hypothesis, or theory can be disproven by observation or experiment.

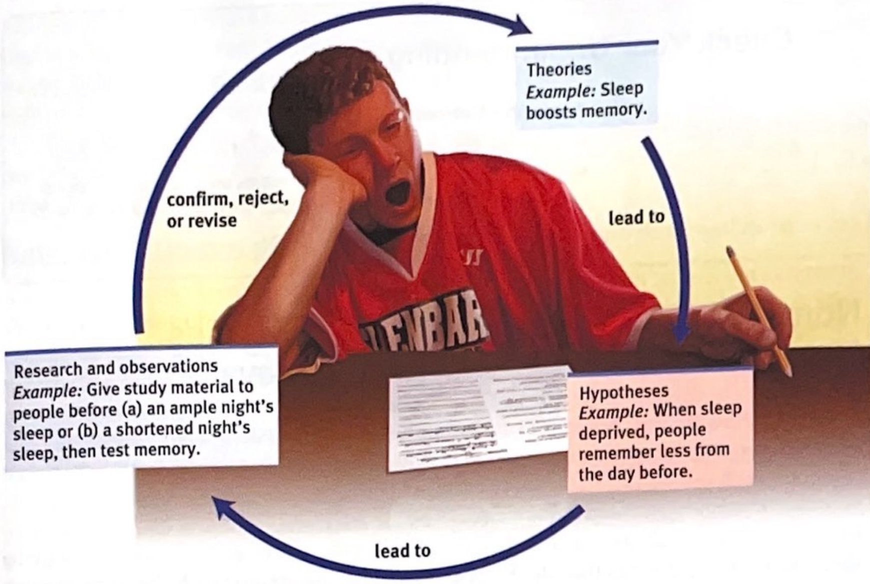


Figure 0.3-1
The scientific method
This self-correcting process asks questions and observes nature's answers.

a shortened night's sleep (**Figure 0.3-1**). The results will either support our theory or lead us to revise or reject it.

Our theories can bias our observations. Having theorized that better memory springs from more sleep, we may see what we expect: Sleep-deprived people's answers are less accurate. The urge to see what we expect is strong, both inside and outside the laboratory, as when people's views of climate change influence their interpretation of local weather events.

In the end, our theory will be useful if it (1) *organizes* observations and (2) implies *predictions* that anyone can use to check the theory or to derive practical applications. (Does people's sleep predict their retention?) Eventually, our research may (3) stimulate further research that leads to a revised theory that better organizes and predicts.

As a check on their own biases, psychologists report their research with precise, measurable **operational definitions** of research procedures and concepts. *Sleep deprived*, for example, may be defined as "at least 2 hours less than a person's natural sleep"; a study of *aggression* may observe how many pins a person stabs into a doll representing a lab partner; and a study of *helping* may record the number of dollars a person donates. By using carefully worded statements, others can **replicate** (repeat) the original observations with different participants, materials, and circumstances. If they get similar results, confidence in the finding's reliability grows. The first study of sleep deprivation aroused psychologists' curiosity. Now, after many successful replications with different people and questions, we feel sure of the phenomenon's power. Replication is confirmation.

We can test our hypotheses and refine our theories using *non-experimental* methods or *experimental* methods. Some non-experimental methods describe behaviors via *case studies, surveys,* or *naturalistic observations;* others compute *correlations* that assess associations among different factors. Experimental methods manipulate variables to see their effects. (As we will see in Module 0.6, *meta-analyses* may be used to analyze the results of multiple studies to reach an overall conclusion.)

operational definition a carefully worded statement of the exact procedures (operations) used in a research study. For example, *human intelligence* may be operationally defined as what an intelligence test measures. (Also known as *operationalization*.)

replication repeating the essence of a research study, usually with different participants in different situations, to see whether the basic finding can be reproduced.

AP® Science Practice

Check Your Understanding

Examine the Concept

- ► Explain the role of peer review in the scientific process.
- ► Explain why replication is important.

Apply the Concept

- ► What are two operational definitions of academic success?
- ► Will what you've learned about theories and replication change the way you read about research results, such as in your news feed?

Answers to Examine the Concept questions can be found in Appendix C at the end of the book.

Non-Experimental Methods: Case Studies, Naturalistic Observations, and Surveys

0.3-2 How do psychologists use case studies, naturalistic observations, and surveys to observe and describe behavior, and why is random sampling important?

In everyday life, we all observe and describe people, often drawing conclusions about why they think, feel, and act as they do. Professional psychologists do much the same, though more objectively and systematically, using *non-experimental methods*, such as:

- *case studies* (in-depth analyses of individuals or groups),
- *naturalistic observations* (recording the natural behavior of many individuals), and
- *surveys* and interviews (asking people questions).

The Case Study

Among the oldest research methods, the **case study** examines one individual or group in depth in the hope of revealing things true of us all. Some examples:

- *Brain damage.* Much of our early knowledge about the brain came from case studies of individuals who suffered a particular impairment after damage to a certain brain region.
- *Children's minds.* Developmental psychologist Jean Piaget taught us about children's thinking after carefully observing and questioning only a few children.
- *Animal intelligence.* Studies of various animals, including only a few chimpanzees, have revealed their capacity for understanding and language.

Skye Hohmann/Alamy Stock Photo

Freud and Little Hans Sigmund Freud's case study of 5-year-old Hans' extreme fear of horses led Freud to his theory of childhood sexuality. He conjectured that Hans felt unconscious desire for his mother, feared castration by his rival father, and then transferred this fear into his phobia about being bitten by a horse. As Module 4.5 will explain, today's psychological science discounts Freud's theory of childhood sexuality but does agree that much of the human mind operates outside our conscious awareness.

Intensive case studies are sometimes very revealing, and they often suggest directions for further study. But atypical individual cases may mislead us. Both in our everyday lives and in science, unrepresentative information can lead to mistaken conclusions. Indeed, anytime a researcher mentions a finding (*Smokers die younger: 95 percent of men over 85 are nonsmokers*) someone is sure to offer a contradictory anecdote (*Well, I have an uncle who smoked two packs a day and lived to be 89!*).

case study a non-experimental technique in which one individual or group is studied in depth in the hope of revealing universal principles.

Dramatic stories and personal experiences (even psychological case examples) command our attention and are easily remembered. Journalists understand this point, so they often begin their articles with compelling stories. Stories move us, but stories can also mislead. Which of the following do you find more memorable? (1) "In one study of 1300 dream reports concerning a kidnapped child, only 5 percent correctly envisioned the child as dead" (Murray & Wheeler,

1937) or (2) "I know a man who dreamed his sister was in a car accident, and two days later she died in a head-on collision!" Numbers can be numbing, but *the plural of anecdote is not evidence.* As psychologist Gordon Allport (1954, p. 9) said, "Given a thimbleful of [dramatic] facts we rush to make generalizations as large as a tub."

The point to remember: Individual cases can suggest fruitful ideas. What's true of all of us can be glimpsed in any one of us. But to find those general truths, we must employ other research methods.

> **AP® Science Practice**
>
> **Research**
>
> You will encounter many case studies as you work your way through the modules in this textbook. Be on the lookout for them.

Naturalistic Observation

A second non-experimental method involves recording responses in natural environments. These **naturalistic observations** have traditionally ranged from watching chimpanzee societies in the jungle, to videotaping and analyzing parent–child interactions in different cultures, to recording racial differences in students' self-seating patterns in a school lunchroom. Until recently, such naturalistic observation was mostly "small science"—possible to do with pen and paper rather than fancy equipment and a big budget (Provine, 2012). But today's digital technologies—thanks to "big data" harvested from phone apps, social media, online searches, and more—have transformed naturalistic observations into big science. Anonymously tapping into 15 million cell phones' GPS (global positioning system) data allowed scientists to track how often people in different geological regions obeyed stay-at-home orders and social distancing recommendations during the Covid pandemic (Glanz et al., 2020). New technologies—wearable cameras and fitness sensors, and internet-connected smart-home sensors—offer increasing possibilities for people to allow accurate recording of their activity, relationships, sleep, and stress (Nelson & Allen, 2018; Yokum et al., 2019).

The billions of people entering personal information online have also enabled big-data observations (without disclosing anyone's identity). One research team studied the ups and downs of human moods by counting positive and negative words in 504 million tweets from 84 countries (Golder & Macy, 2011). As **Figure 0.3-2** shows, people seemed happier on weekends, shortly after waking, and in the evenings. (Are late Saturday

> **AP® Science Practice**
>
> **Data Interpretation**
>
> Graphs, such as the one shown in Figure 0.3-2, provide a lot of information. The ability to interpret data from a graph is an important skill in psychology and one you will be asked to demonstrate on the AP® exam. Graphs reveal the variables in the study (here, the time of day, the day of the week, and the number of positive words in tweets) as well as the relationship between the variables. We will be providing opportunities for you to work with graphs throughout the modules in this text.

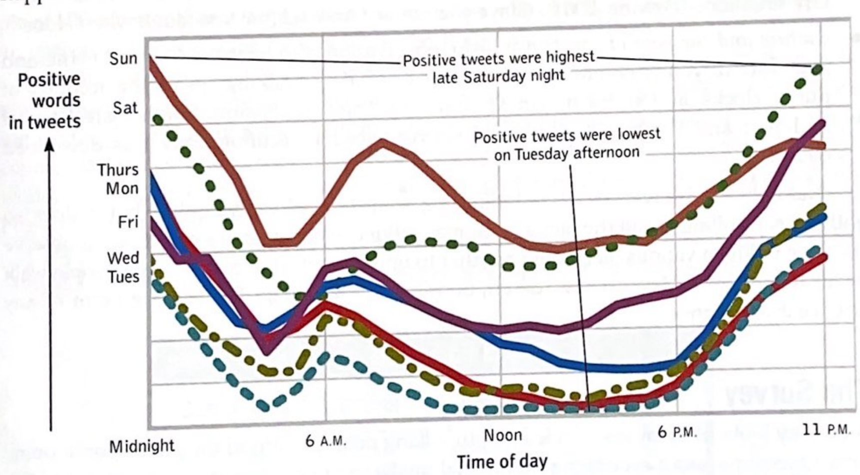


Figure 0.3-2
Twitter message moods, by time and by day
This graph illustrates how, without knowing anyone's identity, researchers can use big data to study human behavior on a massive scale. It is now possible to associate people's moods with, for example, their locations or the weather, and to study the spread of ideas through social networks. (Data from Golder & Macy, 2011.)

naturalistic observation a non-experimental technique of observing and recording behavior in naturally occurring situations without trying to manipulate and control the situation.

But asking questions is tricky. People may shade their answers in a socially desirable direction, such as by underreporting their cigarette consumption or overreporting their voting. And the answers often depend on how questions are worded and how respondents are chosen.

Wording Effects

Even small changes in the order or wording of questions can make a big difference (Table 0.3-1). Researchers attempt to phrase questions in a way that reduces **social desirability bias** (people answering in a way they think will please the researcher). To counter **self-report bias** (when people don't accurately report or remember their behaviors), researchers may pair surveys with other means of measuring behaviors.

Random Sampling

In everyday thinking, we tend to generalize from samples we observe, especially vivid cases. Given (1) a statistical summary of auto owners' evaluations of their car model and (2) the vivid comments of two frustrated owners, our impression may be influenced as much by the two unhappy owners as by the many more summarized evaluations. The temptation to succumb to **sampling bias**—to generalize from a few vivid but unrepresentative cases—is nearly irresistible. *Convenience sampling* is also tempting—collecting research from a group that is readily available, such as your friends at school, rather than a sample that would represent *all* the students at your school.

So how do you obtain a *representative sample?* Say you want to learn how students at your high school feel about online instruction. How could you choose a group that would represent the total student body? Typically, you would seek a **random sample**, in which every person in the entire **population** has an equal chance of being included in the sample group. You might number the names in the school directory and use a random-number generator to pick your survey participants. (Sending each student a questionnaire wouldn't work, because the conscientious people who returned it would not be a random sample.) Large representative samples are better than small ones, but a smaller representative sample of 100 is better than a larger unrepresentative sample of 500. You cannot compensate for an unrepresentative sample by simply adding more people.

Political pollsters sample voters in national election surveys just this way. Without random sampling, large samples—such as from "opt-in" website polls—often give misleading results. But by using some 1500 randomly sampled people, drawn from all areas of a country, they can provide a reasonably accurate snapshot of the nation's opinions. In today's world, however, with so many people not answering phones, door knocks, and emails, getting a random sample is a challenge.

The point to remember: Before accepting survey findings, think critically. Consider the sample. The best basis for generalizing is from a representative, random sample.

TABLE 0.3-1 Survey Wording Effects

Garners More Approval	Garners Less Approval
"aid to those in need"	"welfare"
"undocumented workers"	"illegal aliens"
"gun safety laws"	"gun control laws"
"revenue enhancers"	"taxes"
"enhanced interrogation"	"torture"
"pre-owned"	"used"

social desirability bias bias from people's responding in ways they presume a researcher expects or wishes.

self-report bias bias when people report their behavior inaccurately.

sampling bias a flawed sampling process that produces an unrepresentative sample.

random sample a sample that fairly represents a population because each member has an equal chance of inclusion.

population all those in a group being studied, from which random samples may be drawn. (*Note:* Except for national studies, this does not refer to a country's whole population.)

AP® Science Practice

Check Your Understanding

Examine the Concept

► Explain why we cannot assume that case studies always reveal general principles that apply to all of us.

► Explain how the wording can change the results of a survey.

► Explain the differences among case studies, naturalistic observation, and surveys.

Apply the Concept

► From your observations of people, can you think of a "case study" that has taught you something about people in general?

► Can you recall a misleading survey you have experienced or read about?

Answers to the Examine the Concept questions can be found in Appendix C at the end of the book.

Module 0.3 REVIEW

0.3-1 How do theories advance psychological science?

- Psychological *theories* are explanations that apply an integrated set of principles to organize observations and generate *hypotheses*—predictions that are falsifiable and can be used to check the theory or produce practical applications of it. By testing their hypotheses, researchers can confirm, reject, or revise their theories.
- To enable other researchers to *replicate* their studies, researchers report them using precise *operational definitions* of their procedures and concepts. If others achieve similar results, confidence in the conclusion will be greater.

0.3-2 How do psychologists use case studies, naturalistic observations, and surveys to observe and describe behavior, and why is random sampling important?

- Non-experimental methods, which include *case studies*, *naturalistic observations*, and *surveys*, show us what can happen, and they may offer ideas for further study.
- The best basis for generalizing about a *population* is a representative sample; in a *random sample*, every person in the entire population being studied has an equal chance of participating.
- Non-experimental methods describe but do not *explain* behavior, because these methods do not control for the many variables that can affect behavior.

AP® Practice Multiple Choice Questions

1. Why is an operational definition necessary when reporting research findings?

a. An operational definition allows others to replicate the study.
b. An operational definition provides many examples of the concept.
c. An operational definition uses more scientific language than a dictionary definition.
d. An operational definition considers contextual elements that may affect the study's results.

2. Which of the following questions is best investigated by means of a survey?

a. Are people more likely to vote Republican or Democrat in the next U.S. election?
b. Does extra sleep improve memory?
c. What is the most effective study technique for AP® exams?
d. What role does exercise play in heart health?

3. A testable prediction that drives research is known as a(n)

a. theory.
b. hypothesis.
c. operational definition.
d. random sample.

4. Researchers are interested in finding out if voters are more likely to vote for candidates who have more pleasant facial expressions. The researchers contact every hundredth person on the voter list to ask about candidate facial expressions. Which method are the researchers using in choosing the people they will call?

a. Random sample
b. Biased sample
c. Survey
d. Population

5. An individual with an exceptional memory is identified. For any given date, she is capable of recalling major events, the weather, and what she did that day. Which research method is being used if a psychologist conducts an in-depth investigation of this individual using questionnaires, brain scans, and memory tests?

a. Naturalistic observation
b. Survey
c. Interview
d. Case study

6. Dr. Tazurphase asked 100 people who were in line to ride the largest roller coaster in the world if they would be willing to purchase and drive a flying car. He also gave each person a survey asking questions about their personalities. Based on his results, he claimed that people who are efficient and open to new things were the most willing to purchase and drive a flying car. Which of the following describes one issue with Dr. Tazurphase's study?

a. Dr. Tazurphase did not choose a representative sample, so he is unable to make claims about the larger population's willingness to purchase a flying car.
b. Dr. Tazurphase did not choose a large enough sample, so he is unable to make claims about the larger population's willingness to buy a flying car.
c. Dr. Tazurphase did not include a control group to compare his results against, so he cannot determine the cause of someone's willingness to purchase a flying car.
d. Dr. Tazurphase did not randomly assign participants to groups, so he cannot determine the cause of someone's willingness to purchase a flying car.

7. Town City is considering installing new street lamps. Which question might Town City officials ask to ensure that they do not bias respondents?

a. "How much do you want to replace the broken, old lamps with working, new lamps?"
b. "Are you in favor of the inflated taxes associated with the installation of the existing street lamps?"
c. "To what extent are you in favor of installing street lamps throughout Town City?"
d. "When should we replace the street lamps?"

8. Dr. Buzz wanted to understand the impact of stressful life events on irritability. He asked college students to reflect on three major stressors in their lives, and then he asked the students how many times in a week they yelled at other people. Which of the following captures how Dr. Buzz operationally defined irritability?

a. Reflecting on three major stressors
b. Irritability
c. The number of times the students yelled in a week
d. Stressful life events

Module 0.4 Correlation and Experimentation

LEARNING TARGETS

0.4-1 Explain what it means when we say two things are correlated, and describe positive and negative correlations.

0.4-2 Explain illusory correlations and regression toward the mean.

0.4-3 Describe the characteristics of experimentation that make it possible to isolate cause and effect.

Psychologists use different methods to describe, predict, and explain how we think, feel, and act. Correlational research (a *non-experimental* method) describes the relationship between two or more variables. *Experiments* attempt to establish a cause-and-effect connection.

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

correlation coefficient a statistical index of the relationship between two variables (from –1.00 to +1.00).

variable anything that can vary and is feasible and ethical to measure.

scatterplot a graphed cluster of dots, each of which represents the values of two variables. The slope of the points suggests the direction of the relationship between the two variables. The amount of scatter suggests the strength of the correlation (little scatter indicates high correlation).

Correlation

0.4-1 What does it mean when we say two things are correlated, and what are positive and negative correlations?

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

Throughout this book, we often ask how strongly two **variables** are related: How closely related are the personality test scores for identical twins? How well do intelligence test scores predict career achievement? In such cases, **scatterplots** can be very revealing.

Each dot in a scatterplot represents the values of two variables. The three scatterplots in **Figure 0.4-1** illustrate the range of possible correlations from a perfect positive to a

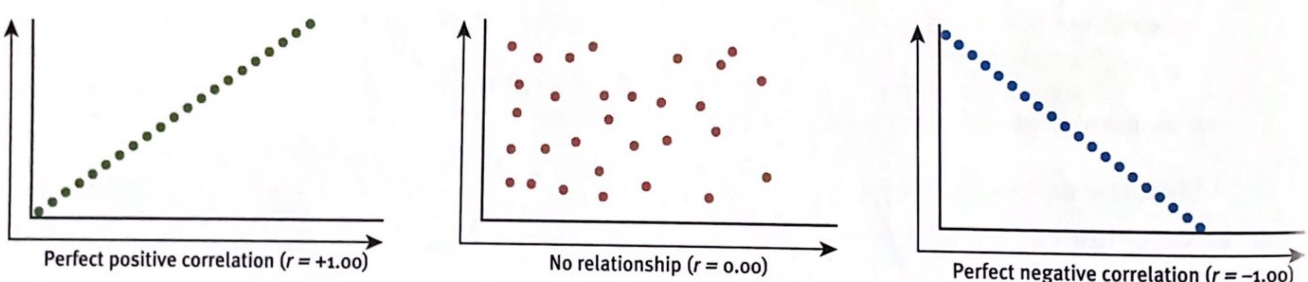


Figure 0.4-1
Scatterplots, showing patterns of correlation

Correlations — abbreviated r — can range from +1.00 (scores for one variable increase in direct proportion to scores for another variable) to 0.00 (no relationship) to –1.00 (scores for one variable decrease precisely as scores for the other variable rise).

perfect negative. (Perfect correlations rarely occur in the real world.) A correlation is positive if two sets of scores, such as for height and weight, tend to rise or fall together.

Saying that a correlation is "negative" says nothing about its strength. A negative correlation isn't "bad." It simply means two sets of scores relate inversely, one set going up as the other goes down. The correlation between people's height and the distance from their head to the ceiling is strongly (perfectly, in fact) negative.

Statistics can reveal what we might miss with casual observation. To demonstrate, consider the responses of 2291 Czech and Slovakian volunteers who were asked to rate, on a 1 to 7 scale, their *fear* and *disgust* related to each of 24 animals (Polák et al., 2020). With all the relevant data right in front of you (**Table 0.4-1**), can you tell whether the correlation between participants' fear and their disgust is positive, negative, or close to zero?

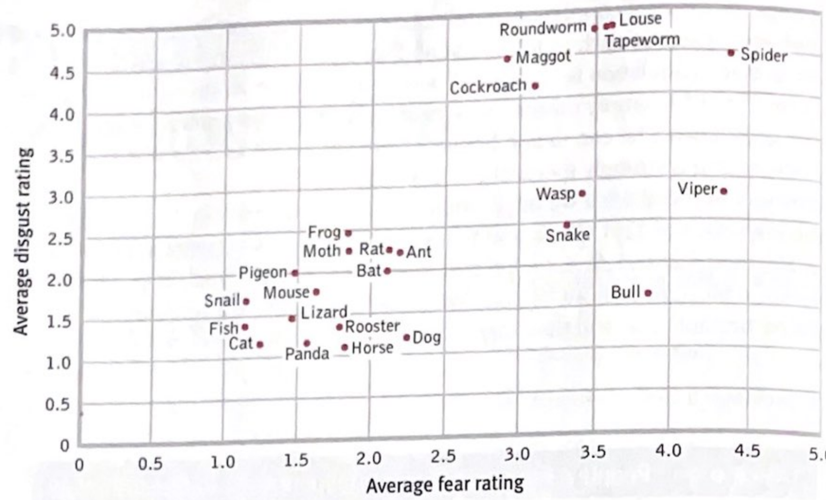
TABLE 0.4-1 People's Fear and Disgust Responses to Various Animals

Animal	Average Fear	Average Disgust
Ant	2.12	2.26
Bat	2.11	2.01
Bull	3.84	1.62
Cat	1.24	1.17
Cockroach	3.10	4.16
Dog	2.25	1.20
Fish	1.15	1.38
Frog	1.84	2.48
Grass snake	3.32	2.47
Horse	1.82	1.11
Lizard	1.46	1.46
Louse	3.58	4.83
Maggot	2.90	4.49
Mouse	1.62	1.78
Panda	1.57	1.17
Pigeon	1.48	2.01
Rat	2.11	2.25
Rooster	1.78	1.34
Roundworm	3.49	4.79
Snail	1.15	1.69
Spider	4.39	4.47
Tapeworm	3.60	4.83
Viper	4.34	2.83
Wasp	3.42	2.84

Figure 0.4-2

Scatterplot for fear and disgust felt toward 24 animals

This display of average self-reported fear and disgust (each represented by a data point) reveals an upward slope, indicating a positive correlation. The considerable scatter of the data indicates the correlation is much lower than +1.00.



AP® Science Practice

Data

You can tell from the scatterplot in Figure 0.4-2 that fear and disgust are positively correlated because as one increases, so does the other. One variable (fear) increases in direct proportion to scores for another (disgust).

AP® Exam Tip

This is the first of several times in your psychology course that you will see something labeled as being positive or negative. We often think that if something is positive it is good and if it is negative it is bad. That is not always the case in psychology. Here, positive and negative refer only to the direction of the correlation. They say nothing about whether the relationship is desirable or not.

AP® Exam Tip

Take note of how much emphasis is put on this idea: Correlation does not imply a cause-effect relationship. This is an important point to understand when it comes to thinking critically about conclusions from research.

When comparing the columns in Table 0.4-1, you might not detect much of a relationship between fear and disgust. In fact, the correlation in this example is positive ($r = +.72$), as we can see if we display the data as a scatterplot (**Figure 0.4-2**).

If we fail to see a relationship when data are presented as systematically as in Table 0.4-1, how much less likely are we to notice them in everyday life? To see what is right in front of us, we sometimes need statistical illumination. We can easily see evidence of gender discrimination when given statistically summarized information about job level, seniority, performance, gender, and salary. But we often see no discrimination when the same information dribbles in, case by case (Twiss et al., 1989). Single events or individuals catch our attention, especially if we want to see (or deny) bias. In contrast, statistics calculate patterns by counting every case equally. See **Table 0.4-2** to test your understanding further.

Correlations can help us see the world more clearly by revealing the extent to which two things relate. However, correlational research has a *directionality problem*—it cannot tell us which variable is the cause, and which one is the effect. If teen social media use correlates with (predicts) teen risk of depression, that may—or may not—indicate that social media use causes an increased risk of depression. The research may also have a *third variable problem* (see Developing Arguments: Correlation and Causation).

TABLE 0.4-2 Test Your Understanding of Correlation

Which of the following news reports are examples of a *positive* correlation, and which are examples of a *negative* correlation? (Check your answers below.)

1. The more college students sleep, the better their academic performance (Okano et al., 2019). ___________
2. The more time teen girls spend absorbed with online social media, the more at risk they are for depression and suicidal thoughts (Kelly et al., 2019; Twenge & Campbell, 2019). ___________
3. The longer children were breast-fed, the greater their later academic achievement (Horwood & Fergusson, 1998). ___________
4. The more leafy vegetables older adults eat, the less their mental decline over the ensuing 5 years (Morris et al., 2018). ___________

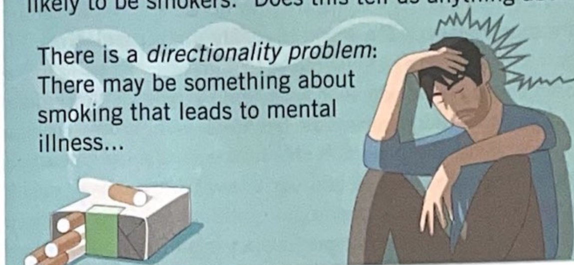
ANSWERS: 1. positive; 2. positive; 3. positive; 4. negative.

AP® Science Practice — Developing Arguments

Correlation and Causation

Mental illness *correlates* with smoking—meaning that those who experience mental illness are also more likely to be smokers.[1] Does this tell us anything about what *causes* mental illness or smoking? **NO.**

There is a *directionality problem*: There may be something about smoking that leads to mental illness...



Or those with mental illness may be more likely to smoke.

OR

There may be some *third variable*, such as a stressful home life, for example, that triggers *both* smoking and mental illness.

So, then, how would you interpret these recent findings:
a) sexual hook-ups correlate with college women's experiencing depression, and
b) *delaying* sexual intimacy correlates with positive outcomes such as greater relationship satisfaction and stability?[2]

Possible explanations:

1. Sexual restraint	→	Better mental health and stronger relationships
2. Depression	→	People being more likely to hook up
3. Some third variable, such as lower impulsivity	→	Sexual restraint, psychological well-being, and better relationships

Correlations do help us predict.
Consider: Self-esteem correlates negatively with (and therefore predicts) depression. The lower people's self-esteem, the greater their risk for depression.

Possible interpretations:

1. Low self-esteem	→	Depression	Directionality problem
2. Depression	→	Low self-esteem	Directionality problem
3. Some third variable, such as distressing events or biological predispostion	→	Both low self-esteem and depression	

You try it!
A survey of over 12,000 adolescents found that the more teens feel loved by their parents, the less likely they are to behave in unhealthy ways—having early sex, smoking, abusing alcohol and drugs, exhibiting violence.[3] What are three possible ways we could interpret that finding?

***The point to remember:* Correlation does not equal causation.**
Correlation suggests a possible cause-effect relationship but does not prove it. Remember this principle and you will be wiser as you read and hear news of scientific studies.

Developing Arguments Questions

1. What were your responses to the "You Try It" question? Were you able to identify three possible ways to interpret that finding?

2. Using scientifically derived evidence presented above, explain why correlation does not equal causation.

1. Belluck, 2013. 2. Fielder et al., 2013; Willoughby et al., 2014. 3. Resnick et al., 1997.

Nancy Brown/The Image Bank/Getty Images

Correlation need not mean causation Length of marriage positively correlates with hair loss in men. Does this mean that marriage causes men to lose their hair (or that balding men make better husbands)?[3]

"Once you become sensitized to it, you see regression everywhere."
Psychologist Daniel Kahneman (1985)

Illusory Correlations and Regression Toward the Mean

0.4-2 What are illusory correlations, and what is regression toward the mean?

Correlations make clear the relationships we might otherwise miss; they also keep us from falsely assuming a relationship exists where there really is none. When we believe there is a relationship between two things, we are likely to notice and recall instances that confirm our belief. If we believe that dreams forecast actual events, we may notice and recall confirming instances more than disconfirming instances. The result is an **illusory correlation**.

Illusory correlations can feed an illusion of control—that we can personally influence chance events. Gamblers, remembering past lucky rolls, may come to believe they influenced the roll of the dice by throwing gently for low numbers and hard for high numbers. The illusion that uncontrollable events correlate with our actions is also fed by a statistical phenomenon called **regression toward the mean**. Extreme results, such as a lower-than-expected test score, are caused by unfortunate combinations—test topic, question difficulty, our sleep (or lack thereof), the weather. The same combination may not happen again, so our next test score should be higher. Simply said, extraordinary happenings tend to be followed by more ordinary ones. Outlier grades will usually regress toward students' average grades. And a team's unusually poor performance in one game will usually improve the next.

Failure to recognize regression can cause superstitious thinking. After berating a team for poorer-than-usual performance, a coach may—when the team regresses to normal—think the scolding actually worked. After lavishing praise for an exceptionally fine performance, the coach may be disappointed when a team's behavior migrates back toward its average. In an unexpected twist, then, regression toward the average can mislead us into feeling rewarded after criticizing others ("That criticism really made them work harder!") and feeling punished after praising them ("All those compliments made them slack off!") (Tversky & Kahneman, 1974).

The point to remember: When a fluctuating behavior returns to normal, fancy explanations for why it does so are often wrong. Regression toward the mean is probably at work.

AP® Science Practice

Check Your Understanding

Examine the Concept

- How would you interpret a *correlation coefficient* of –0.87?
- Describe a scatterplot.

Apply the Concept

- Can you think of a popular media report you've read that confused correlation with causation?
- You hear the school basketball coach telling her friend that she rescued her team's winning streak by yelling at the players after an unusually bad first half. What is another explanation of why the team's performance improved?

Answers to the Examine the Concept questions can be found in Appendix C at the end of the book.

Experimentation

0.4-3 What are the characteristics of experimentation that make it possible to isolate cause and effect?

Happy are they, remarked the Roman poet Virgil, "who have been able to perceive the causes of things." How might psychologists sleuth out the causes in correlational studies, such as the correlation between teen girls' social media use and their risk of depression and self-harm (Odgers & Jensen, 2020)? To establish cause and effect, psychologists use experimentation.

illusory correlation perceiving a relationship where none exists, or perceiving a stronger-than-actual relationship.

regression toward the mean the tendency for extreme or unusual scores or events to fall back (regress) toward the average.

3. In this case, as in many others, a third variable can explain the correlation: Golden anniversaries and baldness both accompany aging.

Experimental Manipulation

Our sleuthing starts with two plain facts:

1. Beginning in 2010, worldwide smartphone and social media use swelled.
2. Simultaneously, Canadian, American, and British teen girls' rates of depression, anxiety, self-harm, and suicide also mushroomed (Mercado et al., 2017; Morgan et al., 2017; Statistics Canada, 2016).

What do such findings mean? Is there a cause-effect connection, perhaps above a certain amount of screen time? Should parents limit their children's screen time? Even big correlational data from a million teens couldn't tell us. Moving beyond the simple correlation, in seven of nine *longitudinal* (over time) studies, teens' current social media use predicted future mental health issues (Haidt & Twenge, 2019; Zhou et al., 2020). Even so, to identify cause and effect, researchers must **experiment**. Experiments enable researchers to isolate the effects of one or more factors by (1) *manipulating the factors of interest* and (2) *holding constant ("controlling") other factors*. To do so, they often create an **experimental group**, in which people receive the treatment (such as reduced screen time), and a contrasting **control group**, in which they do not.

To minimize any preexisting differences between the two groups, experimenters **randomly assign** people to each condition. Random assignment—whether with a random numbers table or the flip of a coin—effectively equalizes the two groups. If one-third of the volunteers for an experiment can wiggle their ears, then about one-third of the people in each group will be ear wigglers. So, too, with age, attitudes, and other characteristics, which will be similar in the experimental and control groups. Thus, if the groups differ at the experiment's end, we can surmise that the treatment had an effect. (Note the difference between random *sampling*, which creates a representative survey sample, and random *assignment*, which equalizes the experimental and control groups.)

So, what do *experiments* reveal about the relationship between girls' social media use and their risk of depression and self-harm? One experiment identified nearly 1700 people who agreed to deactivate their Facebook account for 4 weeks (Allcott et al., 2020). Compared with people in the control group, those randomly assigned to the deactivation group spent more time watching TV and socializing with friends and family—and they reported lower depression, and greater happiness and satisfaction with their lives (and less post-experiment Facebook use). Less Facebook time meant a happier life.

The debate over the effects of prolonged social media use is ongoing. For now, most researchers agree that unlimited teen social media use poses a modest mental health risk. With more large correlational and longitudinal studies, and more experiments, researchers will refine this tentative conclusion.

The point to remember: Correlational studies, which uncover naturally occurring relationships, are complemented by experiments, which manipulate a factor to determine its effect.

experiment a research method in which an investigator manipulates one or more factors (independent variables) to observe the effect on some behavior or mental process (the dependent variable). By *random assignment* of participants, the experimenter aims to control other relevant factors.

experimental group in an experiment, the group exposed to the treatment—that is, to one version of the independent variable.

control group in an experiment, the group *not* exposed to the treatment; contrasts with the experimental group and serves as a comparison for evaluating the effect of the treatment.

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

Procedures and the Placebo Effect

Consider, then, how we might assess therapeutic interventions. Our tendency to seek new remedies when we are ill or emotionally down can produce misleading testimonies. If three days into a cold we start taking zinc tablets and find our cold symptoms lessening, we may credit the pills rather than the cold naturally subsiding. In the 1700s, bloodletting *seemed* effective. When the patient actually survived, this "treatment" was credited for the recovery. When patients didn't survive, the practitioner inferred the disease was too advanced to be reversed. So, whether or not a remedy is truly effective, enthusiastic users will probably endorse it. To determine a treatment's effect, we must control for other factors.

And that is precisely how new drugs and new methods of psychological therapy are evaluated (Module 5.5). Investigators randomly assign participants in these

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

studies to research groups. One group receives a pseudotreatment—an inert *placebo* (perhaps a pill with no drug in it). The other group receives a treatment, such as an antidepressant medication. The participants are often *blind* (uninformed) about what treatment, if any, they are receiving, which is considered a **single-blind procedure**. If the study is using a **double-blind procedure**, neither the participants nor those who administer the drug and collect the data will know which group is receiving the treatment.

In double-blind studies, researchers check a treatment's actual effects apart from the participants' and the staff's belief in its healing powers. Just *thinking* you are getting a treatment can boost your spirits, relax your body, and relieve your symptoms. This **placebo effect** is well documented in reducing pain, depression, anxiety, and even auditory hallucinations in schizophrenia (Dollfus et al., 2016; Kirsch, 2010). Athletes have run faster when given a supposed performance-enhancing drug (McClung & Collins, 2007). Decaf-coffee drinkers have reported increased vigor and alertness when they thought their brew had caffeine in it (Dawkins et al., 2011). And the more expensive the placebo, the more "real" it seems to us—a fake pill that costs $2.50 works better than one costing 10 cents (Waber et al., 2008). To know how effective a therapy really is, researchers must control for a possible placebo effect.

single-blind procedure an experimental procedure in which the research participants are ignorant (blind) about whether they have received the treatment or a placebo.

double-blind procedure an experimental procedure in which both the research participants and the research staff are ignorant (blind) about whether the research participants have received the treatment or a placebo. Commonly used in drug-evaluation studies.

placebo [pluh-SEE-bo; Latin for "I shall please"] **effect** experimental results caused by expectations alone; any effect on behavior caused by the administration of an inert substance or condition, which the recipient assumes is an active agent.

independent variable in an experiment, the factor that is manipulated; the variable whose effect is being studied.

confounding variable in an experiment, a factor other than the factor being studied that might influence a study's results.

experimenter bias bias caused when researchers may unintentionally influence results to confirm their own beliefs.

dependent variable in an experiment, the outcome that is measured; the variable that may change when the independent variable is manipulated.

Independent and Dependent Variables

Here is a practical experiment: Victor Benassi and his colleagues gave college psychology students frequent in-class quizzes. Some items served merely as review—students were given questions with answers. Other self-testing items required students to actively produce the answers. When tested weeks later on a final exam, students did far better on material on which they had been tested (75 percent correct) rather than merely reviewed (51 percent correct). By a wide margin, testing beat restudy.

This simple experiment manipulated just one factor: the study procedure (reading answers versus self-testing). We call this experimental factor the **independent variable** because we can vary it *independently* of other factors, such as the students' memories, intelligence, and age. Other factors that can potentially influence a study's results are called **confounding variables**. Single-blind procedures help control for the *social desirability bias* (participants affecting results by trying to please the researchers). Double-blind procedures reduce **experimenter bias** (when researchers may unintentionally influence results to confirm their own beliefs). In experiments, random assignment ensures that confounding variables have an equal chance of appearing in the experimental and control conditions. Therefore, random assignment controls for possible confounding variables.

Experiments examine the effect of one or more independent variables on some measurable behavior, called the **dependent variable** because it can vary *depending* on what takes place during the experiment. Both variables are given precise *operational definitions*, which specify the procedures that manipulate the independent variable (the review versus self-testing study method in this experiment) and measure the dependent variable (final exam performance). These definitions offer a level of precision that enables others to replicate the study. (**Figure 0.4-3** depicts the previously mentioned Facebook experiment design.)

Figure 0.4-3
Experimentation
To establish causation, psychologists control for confounding variables by randomly assigning some participants to an experimental group and others to a control group. Measuring the dependent variable (depression score) will determine the effect of the independent variable (social media exposure).

Random assignment (controlling for other confounding variables, such as temperament and environment)


Lucy Lambriex/Getty Images

Group	Independent variable	Dependent variable
Experimental	Deactivated Facebook account	Depression test score 4 weeks later
Control	Did not deactivate Facebook account	Depression test score 4 weeks later

Let's pause to check your understanding using a simple psychology experiment: To test the effect of perceived ethnicity on the availability of rental housing, researchers sent identically worded email inquiries to 1115 Los Angeles–area landlords (Carpusor & Loges, 2006). The researchers varied the ethnic connotation of the sender's name and tracked the percentage of positive replies (invitations to view the apartment in person). "Patrick McDougall," "Said Al-Rahman," and "Tyrell Jackson" received, respectively, 89 percent, 66 percent, and 56 percent invitations. In this experiment, what was the independent variable? What was the dependent variable?[4]

A key goal of experimental design is **validity**, which means the experiment tests what it is supposed to test. In the rental housing experiment, we might ask, "Did the email inquiries test the effect of perceived ethnicity? Did the landlords' responses actually vary with the ethnicity of the name?"

Let's recap. A *variable* is anything that can vary (social media exposure, test performance, landlord responses—anything within the bounds of what is feasible and ethical). Experiments aim to *manipulate* an *independent* variable, *measure* a *dependent* variable, and *control confounding* variables. An experiment has at least two different conditions: an *experimental condition* and a *comparison* or *control condition*. *Random assignment* works to minimize preexisting differences between the groups before any treatment effects occur. 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).

AP® Science Practice

Research

Independent and dependent variables are only relevant to experimental methods in which an investigator manipulates one or more factors (the independent variable) to observe the effect on some behavior or mental process (the dependent variable).

AP® Science Practice

Research

Researchers strive for both validity (the extent to which a test or experiment measures what it is supposed to measure) *and* reliability (the extent to which findings can be replicated, as described in Module 0.3).

AP® Science Practice

Check Your Understanding

Examine the Concept

▶ By using *random assignment*, researchers are able to control for ____________ ____________, which are other factors besides the independent variable(s) that may influence research results.

▶ Match the term on the left (i through iii) with the description on the right (a through c).

i. Double-blind procedure
ii. Dependent variable
iii. Random assignment

a. in an experiment, the outcome that is measured; the variable that may change when the independent variable is manipulated
b. helps minimize preexisting differences between experimental and control groups
c. controls for the placebo effect; neither researchers nor participants know who receives the real treatment

▶ Explain the difference between random assignment and random sampling.

Apply the Concept

▶ Explain why, when testing a new drug to control blood pressure, we would learn more about its effectiveness by giving it to half the participants in a group of 1000 rather than to all 1000 participants.

▶ If you became a research psychologist, which questions would you like to explore with experiments?

▶ Can you think of a time when you may have been fooled by the placebo effect?

Answers to the Examine the Concept questions can be found in Appendix C at the end of the book.

> *[We must guard] against not just racial slurs, but . . . against the subtle impulse to call Johnny back for a job interview, but not Jamal.*
>
> *U.S. President Barack Obama, Eulogy for Clementa Pinckney, June 26, 2015*

AP® Exam Tip

The identification of independent and dependent variables will likely be tested on the AP® exam. Experiments are critical to psychology, and independent and dependent variables are critical to experiments.

validity the extent to which a test or experiment measures or predicts what it is supposed to.

4. The independent variable, which the researchers manipulated, was the implied ethnicity of the applicants' names. The dependent variable, which the researchers measured, was the rate of positive responses from the landlords.

AP® Science Practice

Exploring Research Methods & Design

This feature, which shows up throughout the modules in this text, allows you to apply your knowledge of research methods in the context of psychological content.

Imagine you are a researcher interested in stress and the immune system. You want to answer this question: Does stress cause a decrease in immune functioning?

- Explain why you would need to use the experimental method to address this question.
- What would be your independent variable? What would be your dependent variable?

Module 0.4 REVIEW

0.4-1 What does it mean when we say two things are correlated, and what are positive and negative correlations?

- *Correlation* is the degree to which two variables are related, and how well one predicts the other.
- In a positive correlation, two factors increase or decrease together. In a negative correlation, one variable increases as the other decreases.
- A *correlation coefficient* describes the strength and direction of a relationship between two variables, from +1.00 (a perfect positive correlation) through zero (no correlation at all) to −1.00 (a perfect negative correlation).
- Data on a relationship may be displayed in a *scatterplot*, in which each dot represents a value for the two variables.
- Correlational research is a non-experimental method. Correlations enable prediction because they show how two factors are related—either positively or negatively. A correlation can indicate the possibility of a cause-effect relationship, but it does not prove the direction of the influence, or whether an underlying third variable may explain the correlation.

0.4-2 What are illusory correlations, and what is regression toward the mean?

- *Illusory correlations* are random events that we notice and falsely assume are related.
- *Regression toward the mean* is the tendency for extreme or unusual scores to fall back toward their average.

0.4-3 What are the characteristics of experimentation that make it possible to isolate cause and effect?

- To discover cause-effect relationships, psychologists conduct *experiments*, manipulating one or more variables of interest and controlling other variables.
- Using *random assignment*, they can minimize *confounding variables*, such as preexisting differences between the *experimental group* (exposed to the treatment) and the *control group* (given a placebo or different version of the treatment).
- The *independent variable* is the factor that the experimenter manipulates to study its effect; the *dependent variable* is the factor that the experimenter measures to discover any changes occurring in response to the manipulation of the independent variable.
- Studies may use a *single-blind procedure* to control for the social desirability bias, and they may use a *double-blind procedure* to avoid the *placebo effect* and experimenter bias.
- An experiment has *validity* if it tests what it is supposed to test.

AP® Practice Multiple Choice Questions

1. Which of the following is an example of negative correlation?
   a. People who spend more time exercising tend to experience less depression.
   b. Students with lower IQ scores tend to have lower grades.
   c. As hours studying for a test decrease, so do grades on that test.
   d. Students' shoe sizes are not related to their grades.

2. In an experiment to test the effects of room temperature on test performance, the independent variable is
   a. the scores on the test before the experiment begins.
   b. the scores on the test at the end of the experiment.
   c. the role of the teacher.
   d. the temperature of the room.

3. Researchers have discovered that individuals with lower income levels report having fewer hours of total sleep. Therefore,
   a. income and sleep levels are positively correlated.
   b. income and sleep levels are negatively correlated.
   c. income and sleep levels are not correlated.
   d. lower income levels cause individuals to have fewer hours of sleep.

4. Which of the following correlation coefficients represents the strongest relationship between two variables?
   a. +0.75
   b. +1.3
   c. −0.85
   d. −0.05

5. The purpose of random assignment is to
   a. give every member of the population an equal chance to participate in the research.
   b. eliminate the placebo effect.
   c. reduce potential confounding variables.
   d. generate operational definitions for the independent and dependent variables.

6. In a drug study, neither the participants nor the person distributing the pills knows who is receiving the new drug and who is receiving the placebo. This type of research design is said to be a(n) ___________ study.
   a. correlational
   b. confounding
   c. double-blind
   d. single-blind

7. Which of the following best describes the purpose of a control group in an experimental design?
   a. Having a control group allows researchers to reduce the effect of confounding variables on the dynamic between the independent variable and the dependent variable.
   b. Having a control group allows researchers to determine a cause-and-effect relation between the independent variable and dependent variable.
   c. Having a control group allows researchers to better generalize their results to the population of interest.
   d. Having a control group allows researchers to replicate the results from past research.

Module 0.5 Research Design and Ethics in Psychology

LEARNING TARGETS

0.5-1 Explain the process of determining which research design to use.

0.5-2 Explain the value of simplified laboratory conditions in illuminating everyday life.

0.5-3 Explain why psychologists study animals, and explain the ethical research guidelines that safeguard human and animal welfare.

0.5-4 Describe how psychologists' values influence what they study and how they apply their results.

Psychologists use research to ask questions about how we think, feel, and act. Before doing their research, psychologists need to consider which research design to use. They need to reflect on whether their laboratory findings may generalize across different populations, time periods, and cultures. And they need to carefully think through how their research will follow ethical research guidelines.

Research Design

AP® Exam Tip

Table 0.5-1 summarizes over 13 pages of coverage. Spend some time with it, as it is information you will likely encounter on the AP® exam.

0.5-1 How would you know which research design to use?

Throughout this book, you will read about amazing psychological science discoveries. How do psychological scientists choose research methods and design their studies in ways that provide meaningful results? Understanding how research is done—how testable questions are developed and studied—is key to appreciating all of psychology. **Table 0.5-1**

TABLE 0.5-1 Comparing Research Methods

Research Method	Basic Purpose	How Conducted	What Is Manipulated	Weaknesses
Non-experimental: Case Studies, Naturalistic Observations, Surveys	To observe and record behavior	Do case studies, naturalistic observations, or surveys	Nothing	No control of variables; single cases may be misleading
Non-experimental: Correlational Studies	To detect naturally occurring relationships; to assess how well one variable predicts another	Collect data on two or more variables; no manipulation	Nothing	Cannot 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

compares the features of psychology's main research methods. In later modules, you will read about other research designs, including *twin studies* (Module 1.1) and *cross-sectional and longitudinal research* (Module 2.8).

In psychological research, no questions are off limits, except untestable (or unethical) ones: Does free will exist? Are people born evil? Is there an afterlife? Psychologists can't test those questions. But they *can* test whether free-will beliefs, aggressive personalities, and a belief in life after death influence how people think, feel, and act (Dechesne et al., 2003; Shariff et al., 2014; Webster et al., 2014).

Having chosen their question, psychologists then select the most appropriate research design—*experimental, correlational, case study, naturalistic observation, twin study, longitudinal,* or *cross-sectional*—and determine how to set it up most effectively. They consider the amount of money and time available, ethical issues, and other limitations. For example, it wouldn't be ethical for a researcher studying child development to use the experimental method and randomly assign children to loving versus punishing homes.

Next, psychological scientists decide how to measure the behavior or mental process being studied, using either quantitative or qualitative methods. **Quantitative research** methods use numerical data to represent degrees of a variable, for example using a *Likert scale*, where questionnaire responses fall on a continuum (such as from "strongly disagree" to "strongly agree"). **Qualitative research** methods rely on in-depth, narrative data. For example, psychologists may conduct *structured interviews* to understand the causes and consequences of individuals' aggression. Both quantitative and qualitative methods provide valuable information about human behavior and often complement each other.

Regardless of the methods and measures they choose, researchers want to have confidence in their findings. Therefore, they carefully consider confounding variables—factors other than those being studied that may affect their interpretation of results.

Finally, researchers should strive to plan and conduct their studies with diversity, equity, and inclusion in mind—considering potential biases and working to represent marginalized groups.

AP® Science Practice

Research

How researchers choose to measure their variables reflects their operational definitions. Recall from Module 0.3 that an operational definition is a carefully worded statement of the exact procedures (operations) used in a research study. Researchers create operational definitions in a way that ensures broad groups are represented, and equitable procedures are followed.

Predicting Everyday Behavior

0.5-2 How can simplified laboratory conditions 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 say anything useful about flying a plane at night? Or imagine that, after playing violent video games in the lab, teens become more willing to push buttons that they think blast someone with unpleasant noise. Would this behavior indicate that playing shooter games makes someone more likely to commit violence in everyday life?

Before you answer, consider this: 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 rather to test *theoretical principles* (Mook, 1983). In aggression studies, deciding whether to push a button that delivers a noise blast 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.*

quantitative research a research method that relies on quantifiable, numerical data.

qualitative research a research method that relies on in-depth, narrative data that are not translated into numbers.

When psychologists apply laboratory research on aggression to actual violence, they are applying theoretical principles of aggressive behavior, principles 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 typically generalize to the everyday world (Mitchell, 2012).

The point to remember: Psychological science focuses less on specific behaviors than on revealing general principles that help explain many behaviors.

Psychology's Research Ethics

0.5-3 Why do psychologists study animals, and what ethical research guidelines safeguard human and animal welfare?

We have reflected on how using 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 factors, showing how well one predicts another. We have examined the logic that underlies experiments, which use control conditions and random assignment of participants to isolate the causal effects of an independent variable on a dependent variable.

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 common questions about psychology's ethics and values.

Protecting Research Participants

Studying and Protecting Animals

Many psychologists study nonhuman 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, 2002

Humans are complex. But some of the same processes by which we learn are present in other animals, even sea slugs and honeybees. The simplicity of the sea slug's nervous system is precisely what makes it so revealing of the neural mechanisms of learning. Ditto for the honeybee, which resembles us humans in how it learns to cope with stress (Dinges et al., 2017).

Sharing such similarities, should we not respect our animal relatives? The animal protection movement protests the use of animals in psychological, biological, and medical research. "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).

"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 other animals. In experiments on stress and cancer, is it right that mice get tumors in the hope that people might not? Was it right that researchers exposed monkeys to a coronavirus in the search for a Covid vaccine (Shandrashekar et al., 2020)? Humans raise and slaughter 56 billion animals each year (Thornton, 2019). Is our use and consumption of other animals as natural as the behavior of carnivorous hawks, cats, and whales?

For those who give human life top priority, a second question emerges: 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 also supported regulations providing humane care for rats and mice (Plous & Herzog, 2000). Many professional associations and funding agencies already have such guidelines. British Psychological Society (BPS) 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 provide "humane care and healthful conditions" and that testing should "minimize discomfort" (APA, 2012). The European Parliament also mandates standards for animal care and housing (Vogel, 2010). Most universities screen research proposals, often through an animal care ethics committee or *Institutional Review Board* (more on this below), and laboratories are regulated and inspected.

Animals have themselves benefited from animal research. One 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® Science Practice

Research

You will come across examples of animal research throughout the modules in this text. When you do, remember that contemporary psychologists conduct this research according to ethical guidelines.

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

Mahatma Gandhi, 1869–1948

MARY ALTAFFER/AP Photo

Animal research benefiting animals Psychologists have helped zoos enrich animal environments — for example, by giving animals more choices to reduce the learned helplessness of captivity (Kurtycz, 2015; Weir, 2013). 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.

Studying and Protecting Humans

What about human participants? Does the image of white-coated scientists seeming to deliver electric shocks trouble you? Actually, most psychological studies are free of such stress. Blinking lights, flashing words, and pleasant social interactions are more common.

Occasionally, though, researchers do temporarily stress or deceive people (sometimes with the help of *confederates*, who pretend to be fellow participants but are actually part of the experiment), but only when they believe it is essential to a justifiable end, such as understanding and controlling violent behavior or studying mood swings. Many experiments won't work if participants know everything beforehand. (Wanting to be helpful, the participants might try to confirm the researcher's predictions, thus causing the *social desirability bias*.)

Some of psychology's famous early experiments used stressful and deceptive methods that are considered unacceptable today. These psychologists deprived baby monkeys of their mothers, conditioned human babies to burst into tears, and semi-starved men who refused to perform military service during World War II. More to come on each of these in later modules.

Today's ethics codes, from the APA and Britain's BPS, urge researchers to (1) obtain potential participants' **informed consent** (called *informed assent* in the case of minors) to take part, (2) protect participants from greater-than-usual harm and discomfort, (3) keep information about individual participants confidential, and (4) fully **debrief** people (explain the research afterward, including any temporary deception). To enforce these ethical standards, universities and research organizations have established *Institutional Review Boards* (IRBs) comprised of at least five people, which must include one scientist, one non-scientist, and one community representative. IRBs screen research proposals and safeguard "the rights, welfare, and well-being of human research participants" (NIEHS, 2019).

informed consent giving potential participants enough information about a study 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.

Ensuring Scientific Integrity

AP® Exam Tip

The ability to explain the differences among similar concepts is an important skill that will be tested on the AP® exam. Can you explain the difference between informed consent and debriefing? Informed consent happens *before* the study and allows participants to make an informed choice about whether to participate. Debriefing occurs *after* the study and aims to educate participants about the true nature of the study.

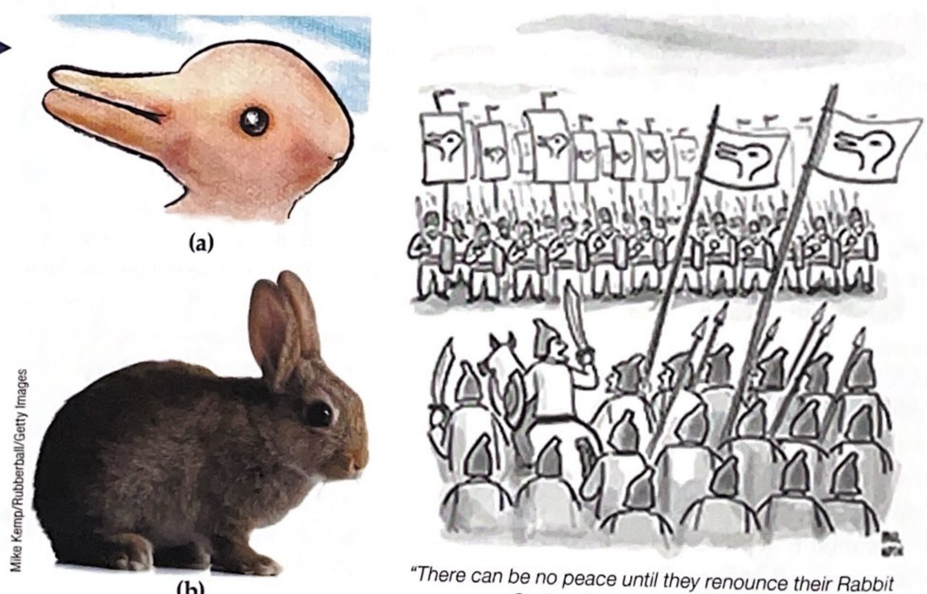
In science, as in everyday life, mistakes happen. When data get accidentally miscomputed or misreported, that's forgivable and correctable. What's not acceptable—and will get a scientist banished from the profession—is fraud. Leading scientists cite honesty as the most important scientific value, followed by curiosity and perseverance (*Nature*, 2016). Community members rate scientists as the most trusted professionals, followed by doctors, judges, and members of the armed forces (Ipsos, 2019). To seek career advancement by plagiarizing another's words or ideas, or to make up data, is to risk a swift end to one's career.

Fake science also has the potential to cause great harm. This happened in 1998 when a now-disbarred British physician published an article in the prestigious journal *The Lancet*, reporting a dozen cases in which British children given the measles, mumps, and rubella (MMR) vaccine supposedly developed autism afterward. Other studies failed to reproduce the finding (replication matters!) (Hviid et al., 2019). An investigation revealed a fraud—with falsified data—and the journal retracted the report (Godlee, 2011). Alas, by then the widely publicized finding—"the most damaging medical hoax of the last 100 years" (Flaherty, 2011)—had produced declining vaccination rates. Instead of following the typical path toward disease elimination, U.S. measles rates in 2019 rose to their highest levels in 25 years (CDC, 2019; Graham et al., 2019). Though the science was self-correcting, the damage lingers on. Nevertheless, the good news is that scientific scrutiny, complete with replication, can inform and protect us.

Values in Psychology

0.5-4 How do psychologists' values influence what they study and how they apply their results?

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? Cultural differences or social injustice? Conformity or independence? Values can also color "the facts"—our observations and interpretations. Sometimes we see what we want or expect to see (**Figure 0.5-1**).



"There can be no peace until they renounce their Rabbit God and accept our Duck God."

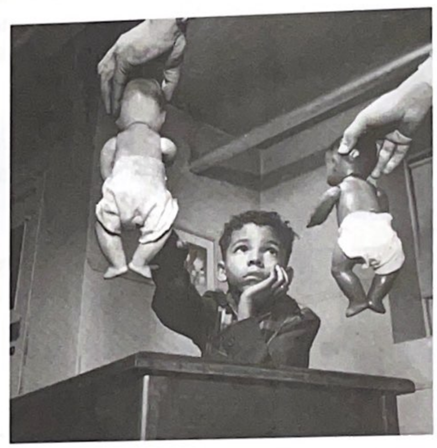
Figure 0.5-1
What do you see?
Our expectations influence what we perceive in (a). Did you see a duck or a rabbit? Show some friends this image with the rabbit photo (b) covered up and see if they are more likely to perceive a duck. (Inspired by Shepard, 1990.)

Even the words we use to describe traits and tendencies 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*. Our labeling someone as *firm* or *stubborn*, *careful* or *picky*, *discreet* or *secretive* reveals our own attitudes.

So, values inform psychological science—and psychological science has the power to persuade. This may lead some to feel distrustful: Is psychology dangerously powerful? Might it 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 have the power to deceive, its purpose is to enlighten. Every day, psychologists explore ways to enhance learning, creativity, and compassion. Psychology speaks to many of our world's great problems—extremist terrorism, political corruption, economic inequality, climate change, prejudice, refugee crises—all of which involve attitudes and behaviors. Psychology also speaks to our deepest longings—for love, for happiness, for meaning. Psychology cannot address all of life's great questions, but it speaks to some mighty important ones.

Macmillan Learning

Macmillan Learning



Gordon Parks Foundation

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 indicated that they had likely absorbed and internalized anti-Black prejudice.

AP® Science Practice

Check Your Understanding

Examine the Concept

▶ Explain the difference between quantitative and qualitative research methods.

▶ Describe informed consent and debriefing, and explain their importance to research.

Apply the Concept

▶ In what ways do values affect researchers? Does this surprise you?

▶ What other questions do you have about psychological research?

Answers to the Examine the Concept questions can be found in Appendix C at the end of the book.

Module 0.5 REVIEW

0.5-1 How would you know which research design to use?

- Psychological scientists design studies and choose research methods that will best provide meaningful results.
- Researchers generate *testable questions*, and then carefully consider the best *design* to use in studying those questions (experimental, correlational, case study, naturalistic observation, twin study, longitudinal, or cross-sectional).
- Researchers next *measure* the variables they are studying, and finally they *interpret* their results, keeping possible confounding variables in mind.

0.5-2 How can simplified laboratory conditions illuminate everyday life?

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

0.5-3 Why do psychologists study animals, and what ethical research guidelines safeguard human and animal welfare?

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

0.5-4 How do psychologists' values influence what they study and how they apply their results?

- 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.

AP® Practice Multiple Choice Questions

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

a. Keep information about participants confidential.
b. Protect participants from potential harm and compensate them for participation in a study.
c. Provide participants with enough information about a study to enable them to make a rational decision about whether to participate.
d. Explain the purpose of study and any deception in the study to participants after they participate.

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

a. Informed consent
b. Informed assent
c. Debriefing
d. Protection from physical harm

3. The laboratory environment is designed to
   a. exactly re-create the events of everyday life.
   b. re-create psychological forces under controlled conditions.
   c. re-create psychological forces under random conditions.
   d. provide the opportunity to do case study research.

4. Which of the following animal studies is most likely to meet ethical principles and receive Institutional Review Board approval?
   a. Do monkeys that smoke get cancer?
   b. Will rats deprived of food for one week survive?
   c. What are the effects of raising kittens in isolation?
   d. Can dolphins learn simple language?

5. Which of the following accurately illustrates the correct order of the scientific process in psychological research?
   a. First, researchers interpret their results, and then they measure their variables, after which they identify hypotheses that align with their research design.
   b. First, researchers choose the best design, and then they interpret their results, after which they measure their variables.
   c. First, researchers measure their variables, and then they identify their hypotheses, after which they choose the best design for their study.
   d. First, researchers create hypotheses, and then they design their study to measure their variables, after which they interpret their results.

Module 0.6 Statistical Reasoning in Everyday Life

LEARNING TARGETS

0.6-1 Describe descriptive statistics.

0.6-2 Explain how we describe data using three measures of central tendency, and percentile rank.

0.6-3 Explain the relative usefulness of the two measures of variation.

0.6-4 Describe inferential statistics.

0.6-5 Explain how we determine whether an observed difference can be generalized to other populations.

Statistics are important tools for psychological scientists. But statistics also benefit us all, by helping us see what the unaided eye might miss. To be an educated person today is to be able to apply simple statistical principles to everyday reasoning. We needn't memorize complicated formulas to think more clearly and critically about data.

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

- ***We ordinarily use only 10 percent of our brain.*** Or is it closer to 100 percent (Module 1.4)?
- ***To be healthy, walk 10,000 steps a day.*** Or will 8500 or 13,000 steps do the trick, or how about swimming or jogging (Mull, 2019)?

If you see an attention-grabbing headline presented without scientifically derived evidence—that nationally there are 1 million teen pregnancies, 2 million homeless seniors, or 3 million alcohol-related motor vehicle accidents each year—you can be pretty sure that someone is estimating. If they want to emphasize the problem, they will be motivated to estimate high. If they want to minimize the problem, they will estimate low. *The point to remember:* Use critical thinking when presented with big, round, undocumented numbers.

Statistical illiteracy also feeds needless health scares (Gigerenzer, 2010). In the 1990s, the British press reported a study showing that women taking a particular contraceptive pill had a 100 percent increased risk of blood clots that could produce strokes. The story went viral, causing thousands of women to stop taking the pill. What happened as a result? A wave of unwanted pregnancies and an estimated 13,000 additional abortions (which, like other medical procedures, also are associated with increased blood clot risk). Distracted by big, round numbers, few people focused on the study's actual findings: A 100 percent increased risk, indeed—but only from 1 in 7000 to 2 in 7000. Such false alarms underscore the need to think critically, to learn statistical reasoning, and to present statistical information more transparently.

More recently, statistical confusion about health information also infected people's understanding of Covid vaccine effectiveness. If a vaccine is "95 percent effective," does that mean a recipient has a 5 percent chance of contracting the virus? As a *New York Times* story explained, Pfizer/BioNTech's clinical trial enrolled nearly 44,000 people, half of whom received its vaccine and half of whom received a placebo (Thomas, 2020). "Out of 170 cases of Covid, 162 were in the placebo group, and eight were in the vaccine group." So, there was a 162 to 8 (95 percent to 5 percent) ratio—which defined the vaccine as 95 percent effective. Of those vaccinated, only 8 of nearly 22,000 people—less than 1/10th of 1 percent (not 5 percent)—contracted the virus during the study period. And of the 32,000 people who received either the Moderna or Pfizer vaccine, how many during the study contracted a *severe* case of Covid? The grand total, noted a follow-up *New York Times* report, was *one* (Leonhardt, 2021).

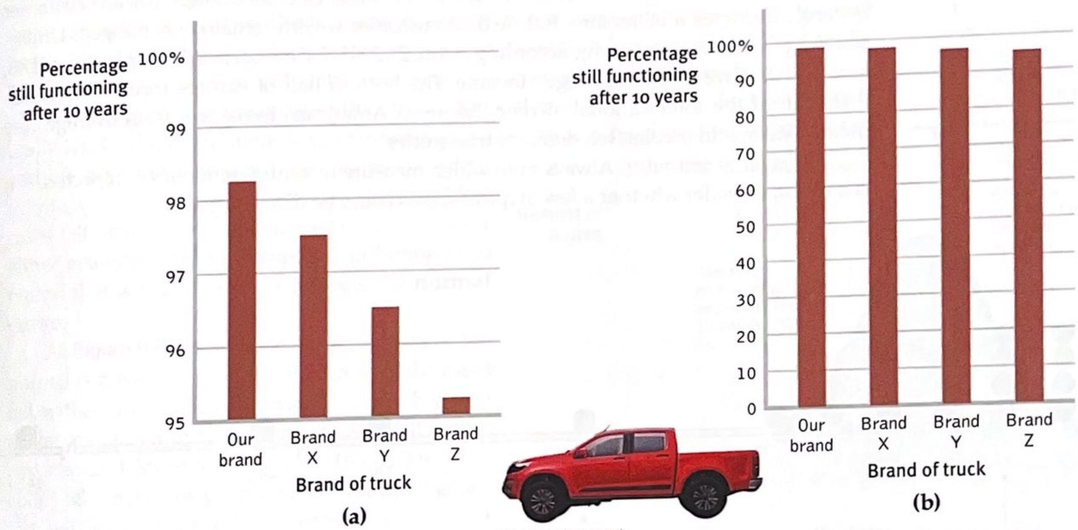
Descriptive Statistics

0.6-1 What are descriptive statistics?

Once researchers have gathered their data, they may use **descriptive statistics** to measure and describe characteristics of the group under study—similar to the way teachers use descriptive statistics to assess how their students have performed. One way to do this is to show the data in a simple *bar graph*, called a **histogram**. **Figure 0.6-1** is a histogram that displays a distribution of different brands of trucks still on the road after a decade. When reading statistical graphs such as this one, take care. It's easy to design a graph to make a difference look big (Figure 0.6-1a) or small (Figure 0.6-1b). The secret lies in how you label the vertical scale (the *y-axis*).

The point to remember: Think smart. When interpreting graphs, consider the scale labels and note their *range*.

Figure 0.6-1
Read the scale labels
A truck manufacturer offered graph (a) — with actual brand names included — to suggest the much greater durability of its trucks. Note, however, how the *y*-axis of each graph is labeled. The range for the *y*-axis label in graph (a) is only from 95 to 100. The range for graph (b) is from 0 to 100. All the trucks rank as 95 percent and up, so almost all are still functioning after 10 years, which graph (b) makes clear.



Whitevector/Shutterstock

AP® Science Practice

Data Interpretation

The ability to evaluate graphical representations of data is important in psychology, and it will show up in the AP® exam. Study the histograms in Figure 0.6-1 and answer the following questions.

- Identify the two variables represented in these graphs.
- Are these data quantitative or qualitative?
- What conclusions can you draw from the data depicted in these graphs?

descriptive statistics numerical data used to measure and describe characteristics of groups; include 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.

percentile rank the percentage of scores that are lower than a given score.

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

Measures of Central Tendency

0.6-2 How do we describe data using three measures of central tendency, and percentile rank?

After organizing and describing their data, researchers' 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. (A *bimodal distribution* occurs when there are two frequently occurring scores.) The most familiar measure of central tendency is the **mean**, or arithmetic average—the total sum of all the scores divided by the number of scores. The midpoint of a data distribution—the 50th percentile—is the **median**. 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. **Percentile rank** is the percentage of scores that are less than a given score. So, if you are in the 79th percentile in a math competition in your state, your score is higher than 79 percent of your peers.

Measures of central tendency neatly summarize data. But consider what happens to the mean when a distribution is lopsided, when it's **skewed** by a few way-out scores. With income data, for example, the mode, median, and mean often tell very different stories (Figure 0.6-2). This happens because the mean is biased by a few extreme incomes. When SpaceX and Tesla CEO Elon Musk sits down in a small café, its average (mean) customer instantly becomes a billionaire. But median customer wealth remains unchanged. Understanding this, you can see why, according to the 2020 U.S. Census, nearly 60 percent of U.S. households have "below average" income. The bottom half of earners receive much less than half of the total national income. So, most Americans make less than average (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.

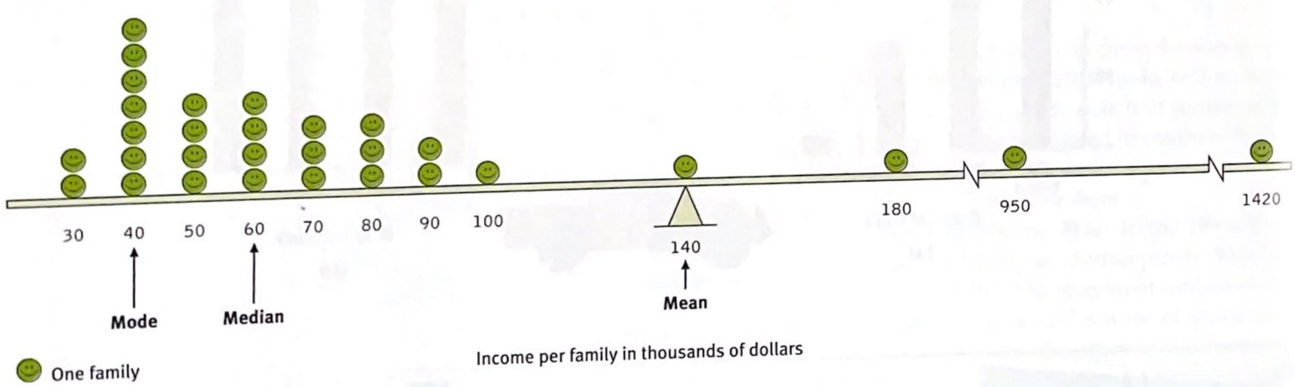


Figure 0.6-2
A skewed distribution
This graphic representation of the distribution of a village's incomes illustrates the three measures of central tendency: mode, median, and mean. Note how just a few high incomes make the mean — the fulcrum point that balances the incomes above and below — deceptively high.

AP® Science Practice

Data Interpretation

Consider the following data set representing scores on an exam:

24 87 27 89 80 92 85 94 87 99

- Calculate the mean, median, and mode for this set of data.
- Explain why the mean might be misleading in interpreting this data set.

Measures of Variation

0.6-3 What is the relative usefulness of the two 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 the season's first 10 games. 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—provides only a crude estimate of variation. In an otherwise similar group, a couple of extreme scores, such as the $950,000 and $1,420,000 incomes in Figure 0.6-2, will create a deceptively large range.

A more useful standard for measuring how much scores deviate (differ) from one another is the **standard deviation**. It better gauges whether scores are packed together or dispersed, because it incorporates information from each score. The computation[5] assembles information about how much individual scores differ from the mean, which can be very telling. Let's say test scores from Class A and Class B both have the same mean (75 percent correct), but very different standard deviations (5.0 for Class A and 15.0 for Class B). Have you ever had test experiences like that—where two-thirds of your classmates in one class score in the 70 to 80 percent range, but scores in another class are more spread out (two-thirds between 60 and 90 percent)? The standard deviation, as well as the mean score, tell us about how each class is faring.

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

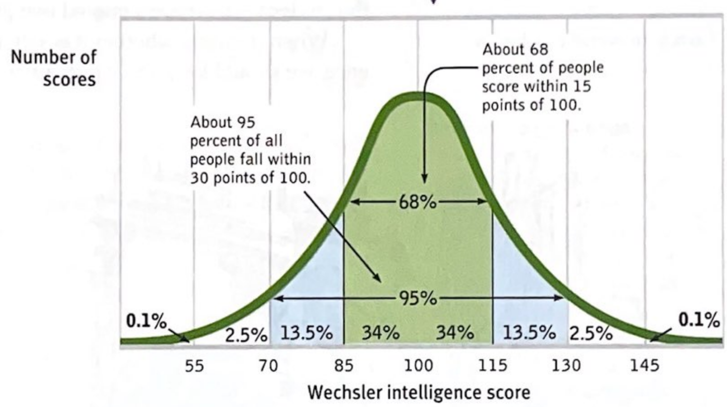
As **Figure 0.6-3** shows, a useful property of the normal curve is that roughly 68 percent of the cases fall within one standard deviation on either side of the mean. About 95 percent of cases fall within two standard deviations. Thus, as Module 2.8 notes, about 68 percent of people taking an intelligence test will score within ±15 points of 100. About 95 percent will score within ±30 points.

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.

normal curve 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 scores lie near the extremes. (Also called a *normal distribution*.)

Figure 0.6-3
The normal curve
Scores on aptitude tests tend to form a normal, or bell-shaped, curve. The most commonly used intelligence test, the Wechsler Adult Intelligence Scale, calls the average score 100.



AP® Science Practice

Check Your Understanding

Examine the Concept

▶ Explain what is meant by mean, mode, median, and percentile rank.

▶ We determine how much scores vary around the average in a way that includes information about the ________ of scores (difference between highest and lowest) by using the ________ ________ formula.

Apply the Concept

▶ Find a graph in an online or print magazine, newspaper, or advertisement. How does the advertiser use (or misuse) statistics to make a point?

Answers to the Examine the Concept questions can be found in Appendix C at the end of the book.

5. The actual standard deviation formula is: $\sqrt{\dfrac{\text{Sum of (deviations)}^2}{\text{Number of scores} - 1\ (n-1)}}$

Inferential Statistics

0.6-4 What are inferential statistics?

> **AP® Science Practice**
>
> **Data**
>
> Inferential statistics help researchers draw conclusions about the population based on the sample in their study. They are quite different from descriptive statistics, which are used to simply describe a sample's characteristics.

Data are noisy. The average score from an experimental group (such as those who deactivated their Facebook account, in the experiment we mentioned in Module 0.4) could conceivably differ from the average score from the control group (those who didn't) 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 from the research sample? For guidance, we can ask how reliable and statistically significant the differences are. These **inferential statistics** help us determine if results can be generalized to a larger population (all those in a group being studied).

When Is an Observed Difference Reliable?

0.6-5 How do we determine whether an observed difference can be generalized to other populations?

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

meta-analysis a statistical procedure for analyzing the results of multiple studies to reach an overall conclusion.

Imagine an eager high school senior who visits two university campuses, each for a day. At the first school, the student randomly samples two classes and finds that both instructors are witty and engaging. At the second school, the two sampled instructors seem dull and uninspiring. Should the student conclude that the first school's teachers are "great" and the second school's teachers are "bores"?

You might respond that the student should sample more classes—and you'd be right. It's possible that the populations of teachers at the two universities are equal. Just by chance, the student could have sampled two great (and two boring) teachers.

When deciding whether it is safe to infer a population difference from a sample difference, we should keep three principles in mind:

1. ***Representative samples are better than biased (unrepresentative) samples.*** The best basis for generalizing is from a representative sample of cases, not from the exceptional and memorable cases one finds at the extremes. Research never randomly samples the whole human population. Thus, it pays to remember which population a study has sampled.
2. ***Bigger samples are better than smaller ones.*** We know it but we ignore it: Averages based on many cases are more precise than averages based on a few. More (randomly sampled) cases make the sample's estimate more precise. Larger samples also make for a more *replicable* study—one that will find a similar estimate the next time.
3. ***More estimates are better than fewer estimates.*** A study gives one brief peek at what's going on in the population. But the best thing to do is conduct multiple studies and combine all the estimates, using **meta-analysis**. Better to consider an entire forest of findings rather than focusing on a single study.

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

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

When Is an Observed Difference Significant?

Suppose you sampled men's and women's scores on a laboratory test of aggression and found a gender difference. But samples can vary. So, how likely is it that your observed gender difference was just a fluke?

Researchers use statistical testing to estimate the probability of the result occurring by chance. They begin with the assumption that no difference exists between groups, an assumption called the *null hypothesis*. Then, using statistics, they evaluate whether the observed gender difference is so rare that it's unlikely to fit the null hypothesis. If so, they reject the null hypothesis of no differences, and they say that the result is **statistically significant**. Such a large difference would support an *alternative hypothesis*—that the populations of men and women really do differ in aggression.

What factors determine statistical significance? When averages from two samples are precise estimates of their respective populations (as when each is based on many observations that have low variability), then any *difference* between the two samples is more likely to be statistically significant. (For our example: The less the variability in women's and in men's aggression scores, and the more scores we observe, the more precisely we will estimate that the observed gender difference is real.) When the difference we estimate is *large, it's also more likely to reflect* a real difference in the population.

In short, when estimates are precise and when the difference between them is relatively large, we're more likely to find that the difference is statistically significant. This means that the observed difference in the sample is probably more than just chance variation, so we reject the original null hypothesis of no existing differences.

In judging statistical significance, psychologists are conservative. They are like juries who must presume innocence until guilt is proven. Many psychological tests provide *p*-values, which indicate the probability of the result, given the null hypothesis. For most psychologists, strong evidence that we can reject the null (no-difference) hypothesis occurs when the probability (*p*-value) of that result is very low. "Very low" is usually set at less than 5 percent ($p < .05$). When a sample's result would occur less than 5 percent of the time assuming the null hypothesis, we say it is significant.

When learning about research, you should remember that a "statistically significant" result may have little *practical significance*. Especially when a sample is very large, a result might be statistically significant but have a tiny **effect size**. One large study tested the intelligence of first-born and later-born individuals. Researchers revealed a statistically significant tendency for first-born individuals to have higher average scores than their later-born siblings (Rohrer et al., 2015; Zajonc & Markus, 1975). But the difference was only about 1.5 IQ points, so the vast majority of IQ is determined by factors other than birth order. There were 20,000 people in the study, so this difference was "significant," but it had little practical importance.

To interpret results, researchers also use the *confidence interval*—a range of values that likely includes the population's true mean value. If one high school class has a mean score of 80 on an academic achievement test, we would not assume that 80 is the true mean value of all

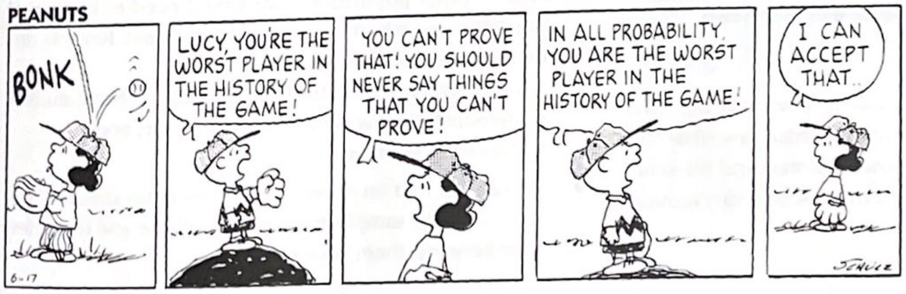
AP® Exam Tip

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 of this. It has been on the AP® exam in the past.

statistical significance a statistical statement of how likely it is that a result (such as a difference between samples) occurred by chance, assuming there is no difference between the populations being studied.

effect size the strength of the relationship between two variables. The larger the effect size, the more one variable can be explained by the other.

PEANUTS



high school test-takers (the population). But statistical analyses can provide a range of values—for example, between 70 and 90—that gives researchers a degree of "confidence" that the population's true mean falls in this interval. Thus, a confidence interval helps researchers estimate whether a sample's range of scores likely includes the population's true mean value.

The point to remember: Statistical significance indicates the likelihood that the result would have happened by chance if the null hypothesis (of no difference) were true. But statistically significant is not the same as *important* or *strong*.

AP® Science Practice

Check Your Understanding

Examine the Concept

▶ ___________ statistics summarize data, while ___________ statistics determine whether data can be generalized to other populations.

▶ Explain the three principles we should keep in mind when deciding if it is safe to infer a population difference from a sample difference.

Apply the Concept

▶ Can you think of a situation where you were fooled by writers or speakers attempting to persuade you with statistics? What have you learned in this module that will be most helpful in the future to avoid being misled? Explain what is meant by a skewed distribution.

Answers to the Examine the Concept questions can be found in Appendix C at the end of the book.

Module 0.6 REVIEW

0.6-1 What are descriptive statistics?

- Researchers use descriptive statistics to measure and describe characteristics of groups under study, often using a histogram to display their data.
- Descriptive statistics include measures of central tendency, percentile rank, and measures of variation.

0.6-2 How do we describe data using three measures of central tendency, and percentile rank?

- 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). *Percentile rank* indicates what percentage of scores falls beneath a given score.

0.6-3 What is the relative usefulness of the two measures of variation?

- 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).
- The range offers only a crude measure of how much the data vary; the standard deviation is far better at giving researchers a clear understanding of variation.
- Scores often form a *normal* (or bell-shaped) *curve*.

0.6-4 What are inferential statistics?

- Researchers use *inferential statistics* to determine the probability of their findings being also true of the larger population.
- Inferential statistics include ways of determining the reliability and significance of an observed difference between the results for different groups.

0.6-5 How do we determine whether an observed difference can be generalized to other populations?

- To feel confident about generalizing an observed difference to other populations, we would need to know that the difference is both reliable and significant. Reliable differences are based on samples that:
  - are representative of the larger population being studied,
  - demonstrate low variability, on average; and
  - consist of many cases.
- We can say that an observed difference has *statistical significance* if the sample averages are reliable and the difference between them is large.

AP® Practice Multiple Choice Questions

For questions 1–3, use the following data: 33, 40, 12, 25, 80.

1. What is the mean of the data?

a. 68
b. 98
c. 33
d. 38

2. What is the median of the data?

a. 33
b. 68
c. 38
d. 80

3. What is the mode of the data?

a. 33
b. 25
c. 12
d. There is no mode.

4. Which measure of central tendency is most influenced by skewed data or extreme scores in a distribution?

a. Mean
b. Median
c. Mode
d. Percentile rank

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

a. This result is highly statistically significant.
b. This result reflects the minimum standard typically considered statistically significant.
c. This result is not statistically significant.
d. This result cannot be evaluated on statistical significance without replication of the study.

6. Descriptive statistics __________, while inferential statistics __________.

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

7. 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. 50 percent
c. 68 percent
d. 95 percent

UNIT 0 Review

KEY TERMS AND CONTRIBUTORS TO REMEMBER

critical thinking, p. 0-6
hindsight bias, p. 0-10
peer reviewers, p. 0-14
theory, p. 0-14
hypothesis, p. 0-14
falsifiable, p. 0-14
operational definition, p. 0-15
replication, p. 0-15
case study, p. 0-16
naturalistic observation, p. 0-17
survey, p. 0-18
social desirability bias, p. 0-19
self-report bias, p. 0-19
sampling bias, p. 0-19
random sample, p. 0-19
population, p. 0-19
correlation, p. 0-22
correlation coefficient, p. 0-22
variable, p. 0-22
scatterplot, p. 0-22
illusory correlation, p. 0-26
regression toward the mean, p. 0-26
experiment, p. 0-27
experimental group, p. 0-27
control group, p. 0-27
random assignment, p. 0-27
single-blind procedure, p. 0-28
double-blind procedure, p. 0-28
placebo effect, p. 0-28
independent variable, p. 0-28
confounding variable, p. 0-28
experimenter bias, p. 0-28
dependent variable, p. 0-28
validity, p. 0-29
quantitative research, p. 0-33
qualitative research, p. 0-33
informed consent, p. 0-35
debriefing, p. 0-35
descriptive statistics, p. 0-41
histogram, p. 0-41
mode, p. 0-42
mean, p. 0-42
median, p. 0-42
percentile rank, p. 0-42
skewed distribution, p. 0-42
range, p. 0-43
standard deviation, p. 0-43
normal curve, p. 0-43
inferential statistics, p. 0-44
meta-analysis, p. 0-44
statistical significance, p. 0-45
effect size, p. 0-45

Unit 0 AP® Practice Multiple Choice Questions

1. A student is interested in knowing how widely the academic aptitude of college-bound students varies at her school. Which of the following statistical methods should she use to determine how much students' SAT scores vary from the school's average SAT score?
   a. Correlation coefficient
   b. Mean
   c. Percentile rank
   d. Standard deviation

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. Correlation

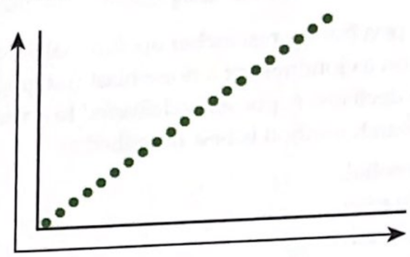
3. In a perfectly normal distribution of scores, which of the following statements is true?
   a. The mean, median, and mode are all the same number.
   b. The mode is equal to the standard deviation.
   c. The scores are positively correlated.
   d. There is a positive skew to the distribution of data.

4. Which of the following describes the placebo effect?
   a. Students in art class are not told that their work will be evaluated for a scholarship so they do not submit their best work.
   b. Participants in an experiment do not know if they are in the experimental or control group so their attitudes about the study are unaffected.
   c. Participants in a drug study are given an inert pill instead of the drug and behave as though they were given the drug.
   d. Only women are chosen for a study, even though the population included men.

5. Which of the following represents naturalistic observation?
   a. Researchers go to the mall and distribute surveys about the stores in the mall.
   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. Researchers observe students' seating patterns in the cafeteria.

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. Illusory correlation
   d. Random sampling

7. Which of the following statements best describes the graph below?



   a. This is a scatterplot of a perfect positive correlation.
   b. This is a scatterplot of a weak negative correlation.
   c. This is a histogram of a weak positive correlation.
   d. This is a histogram of a perfect positive correlation.

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.

9. Which of the following is a positive correlation?
   a. As study time increases, students achieve lower grades.
   b. As levels of self-esteem decline, levels of depression increase.
   c. The more people exercise, the better they sleep.
   d. Gas mileage decreases as vehicle weight increases.

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 each group is similar to each other at the beginning of the study.
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. If participants are not randomly assigned, it is impossible to replicate the experiment.
d. Statistical analysis cannot be performed on an experiment if random assignment is not used.

11. A social psychology researcher operationally defines aggression as loudness of a noise blast (ranging from 0 to 105 decibels) supposedly delivered to a stranger. This research method is best described as

a. inferential.
b. qualitative.
c. quantitative.
d. replication.

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 feasible.
c. The information learned may not apply to the wider population.
d. The dependent variable is difficult to operationally define in a case study.

13. Which of the following is an ethical principle regarding experimental research on humans?

a. Researchers must protect participants from needless harm and discomfort.
b. Regardless of the research design, all participants sign an informed consent form.
c. Personal information about individual participants can only be revealed in peer-reviewed journals.
d. 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 this is an illusory correlation.
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 a student who watches a lot of TV is likely to have lower grades.

15. Which of the following groups of scores would have the smallest standard deviation?

a. 20, 40, 60, 80, 100
b. 5, 15, 25, 35, 45
c. 2, 4, 6, 8, 10
d. 100, 200, 300, 400, 500

Use this scenario to answer questions 16–20:

Researchers wanted to find out if eating sugary foods would increase a person's ability to remember the names of U.S. presidents. The experiment involved 30 female and 30 male participants. A third of the participants (Group A) were given cookies while studying the names. Another third (Group B) were given nothing while studying the names. The final third (Group C) were given mint-flavored candy while studying the names. They were tested on the names a day later. The researchers found that Group A did substantially better than Group B, but about the same as Group C.

16. The dependent variable in this study is

a. the mint candy.
b. the test scores.
c. the cookies.
d. the list of presidents.

17. The independent variable in this study is

a. the list of presidents.
b. the test scores.
c. food given.
d. gender.

18. Which of the following is/are the experimental group(s) of this study?

a. Group B
b. Group C
c. Groups A & C
d. Groups B & C

19. Which of the following is/are the control group(s) of this study?

a. Group A
b. Group B
c. Groups A & B
d. Groups B & C

20. Which of the following is the best conclusion for this study?

a. Only eating cookies tends to improve memory recall.
b. Eating cookies or mint candy tends to improve memory recall.
c. Eating nothing tends to improve memory recall.
d. Only eating mint-flavored candy tends to improve memory recall.

21. Kai scored in the 90th percentile in a math competition in her state. Which of the following statements is true of Kai's score?

a. Her score is higher than 10 percent of others in the competition.
b. Her score is lower than 89 percent of others in the competition.
c. Her score is higher than 90 percent of others in the competition.
d. Her score is an outlier in this competition.

22. Paulette is taking a survey. Instead of being honest, she is answering it in a way that she thinks will please the researchers. Paulette is showing a(n)

a. hindsight bias.
b. sampling bias.
c. experimenter bias.
d. social desirability bias.