Unit XI

Testing and Individual Differences

Modules

Three huge controversies have sparked recent debate in and beyond psychology. First is the "memory war," over whether traumatic experiences are repressed and can later be recovered, with therapeutic benefit. The second great controversy is the "gender war," over the extent to which nature and nurture shape our behaviors as men and women. In this unit, we meet the "intelligence war": Does each of us have an inborn general mental capacity (intelligence), and can we quantify this capacity as a meaningful number?

School boards, courts, and scientists debate the use and fairness of tests that assess people's mental abilities and assign them a score. Is intelligence testing a constructive way to guide people toward suitable opportunities? Or is it a potent, discriminatory weapon camouflaged as science? First, some basic questions:

- What is intelligence?
- How can we best assess it?
- To what extent does it result from heredity and from environmental influence?

- What do test score differences among individuals and groups really mean? Should we use such differences to track the abilities of public school students? To admit them to colleges or universities? To hire them?

This unit offers answers. It identifies a variety of mental gifts. And it concludes that the recipe for high achievement blends talent and grit.

Module 60


Jetta Productions/Getty Images

Introduction to Intelligence

Module Learning Objectives

60-1 Discuss the difficulty of defining *intelligence.*

60-2 Present arguments for and against considering intelligence as one general mental ability.

60-3 Compare Gardner's and Sternberg's theories of intelligence.

60-4 Describe the four components of emotional intelligence.

60-5 Describe the relationship between intelligence and brain anatomy.

60-6 Describe the relationship between intelligence and neural processing speed.

60-1 How is *intelligence* defined?

Psychologists debate: Should we consider intelligence as one aptitude or many? As linked to cognitive speed? As neurologically measurable? On this much, intelligence experts agree: Intelligence is a concept and not a "thing."

In many research studies, *intelligence* has been operationally defined as whatever intelligence tests measure, which has tended to be school smarts. But intelligence is not a quality like height or weight, which has the same meaning to everyone around the globe. People assign the term *intelligence* to the qualities that enable success in their own time and in their own culture (Sternberg & Kaufman, 1998). In the Amazon rain forest, *intelligence* may be understanding the medicinal qualities of local plants. In a North American high school, it may be mastering difficult concepts in tough courses. In both locations, **intelligence** is the ability to learn from experience, solve problems, and use knowledge to adapt to new situations. An **intelligence test** assesses people's mental abilities and compares them with others, using numerical scores.

intelligence mental quality consisting of the ability to learn from experience, solve problems, and use knowledge to adapt to new situations.

intelligence test a method for assessing an individual's mental aptitudes and comparing them with those of others, using numerical scores.


Hemis/Alamy

Hands-on healing The socially constructed concept of intelligence varies from culture to culture. This natural healer in India displays his intelligence in his knowledge about his medicinal plants and understanding of the needs of the people he is helping.

Is Intelligence One General Ability or Several Specific Abilities?

60-2 What are the arguments for and against considering intelligence as one general mental ability?

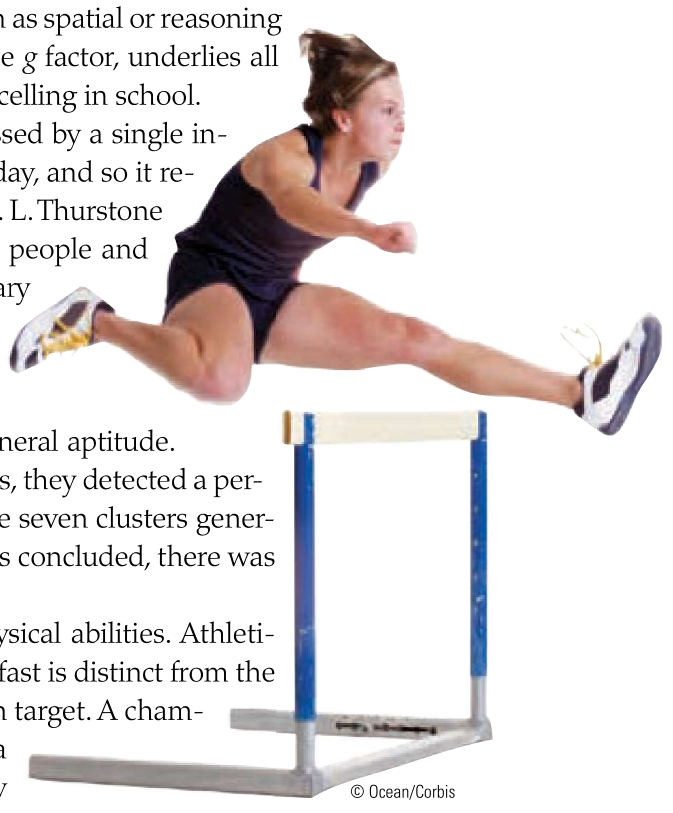
You probably know some people with talents in science, others who excel in social studies, and still others gifted in athletics, art, music, or dance. You may also know a talented artist who is stumped by the simplest math problem, or a brilliant math student with little aptitude for literary discussion. Are all these people intelligent? Could you rate their intelligence on a single scale? Or would you need several different scales?

Charles Spearman (1863–1945) believed we have one **general intelligence** (often shortened to *g*). He granted that people often have special abilities that stand out and he helped develop **factor analysis,** a statistical procedure that identifies clusters of related items. But Spearman also found that those who score high in one area, such as verbal intelligence, typically score higher than average in other areas, such as spatial or reasoning ability. Spearman believed a common skill set, the *g* factor, underlies all intelligent behavior, from navigating the sea to excelling in school.

AP® Exam Tip

David Myers identified three "huge controversies" in the unit opener. All three are covered extensively in this book, and all three will probably show up on the AP® exam.

This idea of a general mental capacity expressed by a single intelligence score was controversial in Spearman's day, and so it remains. One of Spearman's early opponents was L. L. Thurstone (1887–1955). Thurstone gave 56 different tests to people and mathematically identified seven clusters of primary mental abilities (word fluency, verbal comprehension, spatial ability, perceptual speed, numerical ability, inductive reasoning, and memory). Thurstone did not rank people on a single scale of general aptitude. But when other investigators studied these profiles, they detected a persistent tendency: Those who excelled in one of the seven clusters generally scored well on the others. So, the investigators concluded, there was still some evidence of a *g* factor.


© Ocean/Corbis

We might, then, liken mental abilities to physical abilities. Athleticism is not one thing but many. The ability to run fast is distinct from the eye-hand coordination required to throw a ball on target. A champion weightlifter rarely has the potential to be a skilled ice skater. Yet there remains some tendency

general intelligence *(g)* a general intelligence factor that, according to Spearman and others, underlies specific mental abilities and is therefore measured by every task on an intelligence test.

factor analysis a statistical procedure that identifies clusters of related items (called *factors*) on a test; used to identify different dimensions of performance that underlie a person's total score.

for good things to come packaged together—for running speed and throwing accuracy to correlate, thanks to general athletic ability. So, too, with intelligence. Several distinct abilities tend to cluster together and to correlate enough to define a general intelligence factor.

Satoshi Kanazawa (2004, 2010) argues that general intelligence evolved as a form of intelligence that helps people solve *novel* problems—how to stop a fire from spreading, how to find food during a drought, how to reunite with one's tribe on the other side of a flooded river. More common problems—such as how to mate or how to read a stranger's face or how to find your way back to camp—require a different sort of intelligence. Kanazawa asserts that general intelligence scores *do* correlate with the ability to solve various novel problems (like those found in academic and many vocational situations) but do *not* much correlate with individuals' skills in *evolutionarily familiar* situations—such as marrying and parenting, forming close friendships, and navigating without maps. No wonder academic and social skills may come in different bodies.

"***g*** is one of the most reliable and valid measures in the behavioral domain . . . and it predicts important social outcomes such as educational and occupational levels far better than any other trait." -Behavior geneticist Robert Plomin (1999)

Theories of Multiple Intelligences

60-3 How do Gardner's and Sternberg's theories of multiple intelligences differ?

Since the mid-1980s, some psychologists have sought to extend the definition of *intelligence* beyond Spearman's and Thurstone's academic smarts.

GARDNER'S EIGHT INTELLIGENCES

Howard Gardner (1983, 2006) views intelligence as multiple abilities that come in different packages. Brain damage, for example, may destroy one ability but leave others intact. And consider people with **savant syndrome,** who often score low on intelligence tests but have an island of brilliance (Treffert & Wallace, 2002). Some have virtually no language ability, yet are able to compute numbers as quickly and accurately as an electronic calculator, or identify the day of the week corresponding to any given historical date, or render incredible works of art or musical performance (Miller, 1999). About 4 in 5 people with savant syndrome are males, and many also have autism spectrum disorder (ASD; see Module 47).

savant syndrome a condition in which a person otherwise limited in mental ability has an exceptional specific skill, such as in computation or drawing.

The late memory whiz Kim Peek, a savant who did not have ASD, was the inspiration for the movie *Rain Man.* In 8 to 10 seconds, he could read and remember a page. During his lifetime, he memorized 9000 books, including Shakespeare and the Bible. He learned maps from the front of phone books and could provide GPS-like travel directions within any major U.S. city. Yet he could not button his clothes. And he had little capacity for abstract concepts. Asked by his father at a restaurant to "lower your voice," he slid lower in his chair to lower his voice box. Asked for Lincoln's Gettysburg Address, he responded, "227 North West Front Street. But he only stayed there one night—he gave the speech the next day" (Treffert & Christensen, 2005).

Islands of genius: Savant syndrome Matt Savage, an award-winning jazz musician, is a Berklee College of Music graduate who has released many albums. His success has been hard-won given his early childhood diagnosis of what is now called autism spectrum disorder, which came with struggles to communicate and an initial inability to tolerate sounds of any kind.

Joanne Rathe/The Boston Globe via Getty Images

Using such evidence, Gardner argues that we do not have *an* intelligence, but rather *multiple intelligences* (**FIGURE 60.1** on the next page), including the verbal and mathematical aptitudes assessed by standard tests. Thus, the computer programmer, the poet, the street-smart adolescent who becomes a crafty executive, and the basketball team's point guard exhibit different kinds of intelligence (Gardner, 1998a).

Wouldn't it be nice if the world were so just that being weak in one area would be compensated by genius in another? Alas, say Gardner's critics, the world is not just (Ferguson, 2009; Scarr, 1989). Recent research, using factor analysis, has confirmed that there *is* a general intelligence factor (Johnson et al., 2008): *g* matters. It predicts performance on various complex tasks and in various jobs (Gottfredson,

Figure 60.1
Gardner's eight intelligences

grit in psychology, grit is passion and perseverance in the pursuit of long-term goals.

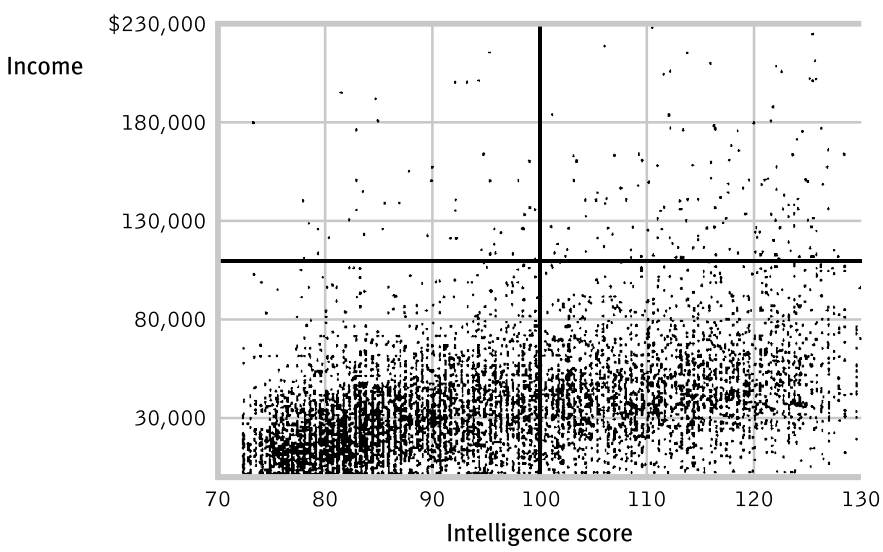
Try This

For more on how self-disciplined grit feeds achievement, see Module 82.

2002a,b, 2003a,b; see also **FIGURE 60.2**). Much as jumping ability is not a predictor of jumping performance when the bar is set a foot off the ground—but becomes a predictor when the bar is set higher—so extremely high cognitive ability scores predict exceptional attainments, such as doctoral degrees and publications (Kuncel & Hezlett, 2010).

Even so, "success" is not a one-ingredient recipe. High intelligence may help you get into a good college and ultimately a desired profession, but it won't make you successful once there. The recipe for success combines talent with **grit:** Those who become highly successful tend also to be conscientious, well-connected, and doggedly energetic.

Figure 60.2
Smart and rich? Jay Zagorsky (2007) tracked 7403 participants in the U.S. National Longitudinal Survey of Youth across 25 years. As shown in this scatterplot, their intelligence scores shared a small but significant correlation (+.30) with their later income. Each dot indicates a given youth's intelligence score and later adult income.



K. Anders Ericsson (2002, 2007; Ericsson et al., 2007) reports a *10-year rule:* A common ingredient of expert performance in chess, dancing, sports, computer programming, music, and medicine is "about 10 years of intense, daily practice." Various animal species, including bees, birds, and chimps, likewise require time and experience to acquire peak expertise in skills such as foraging (Helton, 2008). As with humans, animal performance therefore tends to peak near midlife.


Courtesy of Cameras on Wheels

Spatial intelligence genius In 1998, World Checkers Champion Ron "Suki" King of Barbados set a new record by simultaneously playing 385 players in 3 hours and 44 minutes. Thus, while his opponents often had hours to plot their game moves, King could only devote about 35 seconds to each game. Yet he still managed to win all 385 games!

"You have to be careful, if you're good at something, to make sure you don't think you're good at other things that you aren't necessarily so good at. . . . Because I've been very successful at [software development] people come in and expect that I have wisdom about topics that I don't."
-Bill Gates (1998)

STERNBERG'S THREE INTELLIGENCES

Robert Sternberg (1985, 1999, 2003) agrees that there is more to success than traditional intelligence and also agrees with Gardner's idea of multiple intelligences. But he proposes a *triarchic theory* of three, not eight, intelligences:

- ***Analytical (academic problem-solving) intelligence*** is assessed by traditional intelligence tests, which present well-defined problems having a single right answer. Such tests predict school grades reasonably well and vocational success more modestly.
- ***Creative intelligence*** is demonstrated in reacting adaptively to novel situations and generating novel ideas. Many inventions result from such creative problem solving.
- ***Practical intelligence*** is required for everyday tasks, which may be ill-defined, with multiple solutions. Managerial success, for example, depends less on academic problem-solving skills than on a shrewd ability to manage oneself, one's tasks, and other people. Sternberg and Richard Wagner (1993, 1995) offer a test of practical managerial intelligence that measures skill at writing effective memos, motivating people, delegating tasks and responsibilities, reading people, and promoting one's own career. Business executives who score relatively high on this test tend to earn high salaries and receive high performance ratings.


© The New Yorker Collection, 1988, Donald Reilly from cartoonbank.com. All Rights Reserved.

"You're wise, but you lack tree smarts."

With support from the U.S. College Board® (which administers the Advanced Placement® Program as well as the widely used SAT Reasoning Test™ to U.S. college and university applicants), Sternberg (2006, 2007, 2010) and a team of collaborators have developed new measures of creativity (such as thinking up a caption for an untitled cartoon) and practical thinking (such as figuring out how to move a large bed up a winding staircase). Their initial data indicate that these more comprehensive assessments improve prediction of American students' first-year college grades, and they do so with reduced ethnic-group differences.

Although Gardner and Sternberg differ on specific points, they agree that multiple abilities can contribute to life success. They also agree that the differing varieties of giftedness add spice to life and challenges for education. Under their influence, many teachers have been trained to appreciate such variety and to apply multiple intelligence theory in their classrooms.


David R. Frazier Photolibrary, Inc./Alamy

Street smarts This child selling candy on the streets of Manaus, Brazil, is developing practical intelligence at a very young age.

Emotional Intelligence

60-4 What are the four components of emotional intelligence?

Also distinct from academic intelligence is *social intelligence*—the know-how involved in successfully comprehending social situations. People with high social intelligence can read social situations the way a skilled football player reads the defense or a seafarer reads the weather. The concept was first proposed in 1920 by psychologist Edward Thorndike, who noted, "The best mechanic in a factory may fail as a foreman for lack of social intelligence" (Goleman, 2006, p. 83). Later psychologists have marveled that high-aptitude people are "not, by a wide margin, more effective . . . in achieving better marriages, in successfully raising their children, and in achieving better mental and physical well-being" (Epstein & Meier, 1989). Others have explored the difficulty that some smart people have processing and managing social information (Cantor & Kihlstrom, 1987; Weis & Süß, 2007). This idea is especially significant for an aspect of social intelligence that John Mayer, Peter Salovey, and David Caruso (2002, 2008) have called **emotional intelligence.** They have developed a test that assesses four emotional intelligence components:

emotional intelligence the ability to perceive, understand, manage, and use emotions.

- *Perceiving* emotions (to recognize them in faces, music, and stories)
- *Understanding* emotions (to predict them and how they change and blend)
- *Managing* emotions (to know how to express them in varied situations)
- *Using* emotions to enable adaptive or creative thinking

Mayer, Salovey, and Caruso caution against stretching "emotional intelligence" to include varied traits such as self-esteem and optimism. Rather, emotionally intelligent people are both socially and self-aware. And in both the United States and Germany, those scoring high on managing emotions enjoy higher-quality interactions with friends (Lopes et al., 2004). They avoid being hijacked by overwhelming depression, anxiety, or anger. Being sensitive to emotional cues, they know what to say to soothe a grieving friend, encourage a colleague, and manage a conflict.

Emotional intelligence is less a matter of conscious effort than of one's unconscious processing of emotional information (Fiori, 2009). Yet the outgrowths of this automatic processing become visible. Across dozens of studies in many countries, those scoring high in emotional intelligence exhibit somewhat better job performance (Joseph & Newman, 2010; Van Rooy & Viswesvaran, 2004; Zeidner et al., 2008). They also can delay gratification in pursuit of long-range rewards, rather than being overtaken by immediate impulses. They are emotionally in tune with others, and thus often succeed in career, marriage, and parenting situations where academically smarter (but emotionally less intelligent) people fail (Cherniss, 2010a,b; Ciarrochi et al., 2006).

Brain damage reports have provided extreme examples of the results of diminished emotional intelligence in people with high general intelligence. Neuroscientist Antonio Damasio (1994) tells of Elliot, who had a brain tumor removed: "I never saw a tinge of emotion in my many hours of conversation with him, no sadness, no impatience, no frustration." Shown disturbing pictures of injured people, destroyed communities, and natural disasters, Elliot showed—and realized he felt—no emotion. He knew but he could not feel. Unable to intuitively adjust his behavior in response to others' feelings, Elliot lost his job. He went bankrupt. His marriage collapsed. He remarried and divorced again. At last report, he was dependent on a disability check and custodial care from a sibling.

"I worry about [intelligence] definitions that collapse assessments of our cognitive powers with statements about the kind of human beings we favor."
-Howard Gardner, "Rethinking the Concept of Intelligence," 2000

Some scholars, however, are concerned that emotional intelligence stretches the concept of intelligence too far. Multiple-intelligence man Howard Gardner (1999b) welcomes our stretching the concept into such realms as music and information about ourselves and others. But let us also, he says, respect emotional sensitivity, creativity, and motivation as important but different. Stretch "intelligence" to include everything we prize and it will lose its meaning.

Is Intelligence Neurologically Measurable?

You know it: You are smarter than some people and not as smart as others. Question: What in that heart of smarts—your brain—creates this difference? Is it your brain's relative *size?* The amount of certain brain *tissue?* Your brain networks' *efficiency?*

Brain Size and Complexity

60-5 To what extent is intelligence related to brain anatomy?

After the brilliant English poet Lord Byron died in 1824, doctors discovered that his brain was a massive 5 pounds, not the normal 3 pounds. Three years later, Beethoven died and his brain was found to have exceptionally numerous and deep convolutions. Such observations set brain scientists off studying the brains of other geniuses (Burrell, 2005). Do people with big brains have big smarts?

Alas, some geniuses had small brains, and some dim-witted criminals had brains like Byron's. More recent studies that directly measure brain volume using MRI scans do reveal correlations of about +.33 between brain size (adjusted for body size) and intelligence score (Carey, 2007; McDaniel, 2005). Bigger is better.

One review of 37 brain-imaging studies revealed associations between intelligence and brain size and activity in specific areas, especially within the frontal and parietal lobes (Jung & Haier, 2007; Tang et al., 2010). Intelligence is having ample gray matter (mostly neural cell bodies) plus ample white matter (axons) that make for efficient communication between brain centers (Deary et al., 2009; Haier et al., 2009).

Sandra Witelson would not have been surprised. With the brains of 91 Canadians as a comparison base, Witelson and her colleagues (1999) seized an opportunity to study Einstein's brain. Although not notably heavier or larger in total size than the typical Canadian's brain, Einstein's brain was 15 percent larger in the parietal lobe's lower region—which just happens to be a center for processing mathematical and spatial information.

AP® Exam Tip

Do not continue on if you can't remember what terms like *MRI*, *parietal lobe,* and *axon* mean. Now is the time to head back to Unit III for a review. If you do this sort of review frequently, you'll have much better command of the material on the day of the AP® exam.

Brain Function

60-6 To what extent is intelligence related to neural processing speed?

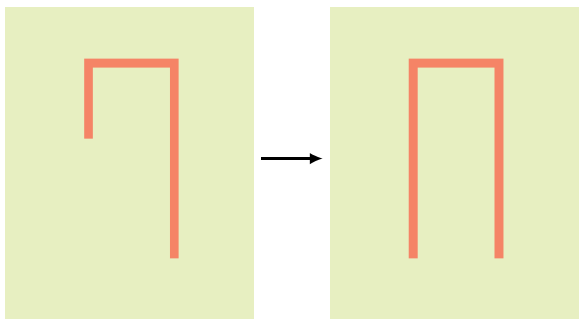
The correlations between brain anatomy and intelligence only begin to explain intelligence differences. Searching for other explanations, neuroscientists are studying the brain's functioning.

As people contemplate a variety of questions like those found on intelligence tests, a frontal lobe area just above the outer edge of the eyebrows becomes especially active—in the left brain for verbal questions, and on both sides for spatial questions (Duncan et al., 2000). Information from various brain areas seems to converge here, suggesting to researcher John Duncan (2000) that it may be a "global workspace for organizing and coordinating information" and that some people may be "blessed with a workspace that functions very, very well."

Functioning well means functioning efficiently. Brain scans reveal that smart people use less energy to solve problems (Haier, 2009). They are like skilled athletes, for whom agile moves can seem effortless. Agile minds come with agile brains.

So, are more intelligent people literally more quick-witted, much as today's speedier computer chips enable ever more powerful computing? On some tasks they seem to be. Verbal intelligence scores are predictable from the speed with which people retrieve information from memory (Hunt, 1983). Those who recognize quickly that *sink* and *wink* are different words, or that *A* and *a* share the same name, tend to score high in verbal ability. Extremely precocious 12- to 14-year-old college students are especially quick in responding to such tasks (Jensen, 1989). To try to define *quick-wittedness,* researchers are taking a close look at speed of perception and speed of neural processing.

Figure 60.3
An inspection time task
A stimulus is flashed before being overridden by a masking image. How long would you need to glimpse the stimulus at the left to answer the question? People who can perceive the stimulus very quickly tend to score somewhat higher on intelligence tests. (Adapted from Deary & Stough, 1996.)



Across many studies, the correlation between intelligence score and the speed of taking in perceptual information tends to be about +.3 to +.5 (Deary & Der, 2005; Sheppard & Vernon, 2008). A typical experiment flashes an incomplete stimulus, as in **FIGURE 60.3,** then a *masking image*—another image that overrides the lingering afterimage of the incomplete stimulus. The researcher then asks participants whether the long side appeared on the right or left. Those whose brains require the least inspection time to register a simple stimulus tend to score somewhat higher on intelligence tests (Caryl, 1994; Deary & Caryl, 1993; Reed & Jensen, 1992).

Perhaps people who process more quickly accumulate more information. Or perhaps, as one Australian-Dutch research team has found, processing speed and intelligence correlate not because one causes the other but because they share an underlying genetic influence (Luciano et al., 2005).

* * *

For a summary of Spearman's, Thurstone's, Gardner's, and Sternberg's theories, see **TABLE 60.1.**

Table 60.1 Comparing Theories of Intelligence

Theory	Summary	Strengths	Other Considerations
Spearman's general intelligence (g)	A basic intelligence predicts our abilities in varied academic areas.	Different abilities, such as verbal and spatial, do have some tendency to correlate.	Human abilities are too diverse to be encapsulated by a single general intelligence factor.
Thurstone's primary mental abilities	Our intelligence may be broken down into seven factors: word fluency, verbal comprehension, spatial ability, perceptual speed, numerical ability, inductive reasoning, and memory.	A single *g* score is not as informative as scores for seven primary mental abilities.	Even Thurstone's seven mental abilities show a tendency to cluster, suggesting an underlying *g* factor.
Gardner's multiple intelligences	Our abilities are best classified into eight independent intelligences, which include a broad range of skills beyond traditional school smarts.	Intelligence is more than just verbal and mathematical skills. Other abilities are equally important to our human adaptability.	Should all of our abilities be considered *intelligences*? Shouldn't some be called talents?
Sternberg's triarchic theory	Our intelligence is best classified into three areas that predict real-world success: analytical, creative, and practical.	These three facets can be reliably measured.	1. These three facets may be less independent than Sternberg thought and may actually share an underlying *g* factor. 2. Additional testing is needed to determine whether these facets can reliably predict success.

Before You Move On

▶ ASK YOURSELF

The modern concept of multiple intelligences (as proposed by Gardner and Sternberg) assumes that the analytical school smarts measured by traditional intelligence tests are important abilities but that other abilities are also important. Different people have different gifts. What are yours?

▶ TEST YOURSELF

Joseph, a Harvard Law School student, has a straight-A average, writes for the *Harvard Law Review,* and will clerk for a Supreme Court justice next year. His grandmother, Judith, is very proud of him, saying he is way more intelligent than she ever was. But Joseph is also very proud of Judith: As a young woman, she was imprisoned by the Nazis. When the war ended, she walked out of Germany, contacted an agency helping refugees, and began a new life in the United States as an assistant chef in her cousin's restaurant. According to the definition of intelligence in this unit, is Joseph the only intelligent person in this story? Why or why not?

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

Module 60 Review

60-1 How is *intelligence* defined?

- *Intelligence* is a mental quality consisting of the ability to learn from experience, solve problems, and use knowledge to adapt to new situations.
- An *intelligence test* aims to assess these qualities and compare them with those of others, using a numerical score.

60-2 What are the arguments for and against considering intelligence as one general mental ability?

- Charles Spearman proposed that we have one *general intelligence* (*g*). He helped develop *factor analysis,* a statistical procedure that identifies clusters of related mental abilities.
- L. L. Thurstone disagreed and identified seven different clusters of mental abilities. Yet a tendency remained for high scorers in one cluster to score high in other clusters.
- Studies indicate that *g* scores are most predictive in novel situations and do not much correlate with skills in evolutionarily familiar situations.

60-3 How do Gardner's and Sternberg's theories of multiple intelligences differ?

- *Savant syndrome* seems to support Howard Gardner's view that we have multiple intelligences. He proposed eight independent intelligences: linguistic, logical-mathematical, musical, spatial, bodily-kinesthetic, intrapersonal, interpersonal, and naturalist.
- Robert Sternberg's triarchic theory proposes three intelligence areas that predict real-world skills: analytical (academic problem solving), creative, and practical.

60-4 What are the four components of emotional intelligence?

- *Emotional intelligence,* which is an aspect of social intelligence, is the ability to perceive, understand, manage, and use emotions.
- Emotionally intelligent people achieve greater personal and professional success.
- Some critics question whether calling these abilities "intelligence" stretches that concept too far.

60-5 To what extent is intelligence related to brain anatomy?

- Some studies have found a positive correlation between intelligence score and brain size and activity, especially in the frontal and parietal lobes.
- Ample gray matter and white matter enable efficient communication between brain circuits.

60-6 To what extent is intelligence related to neural processing speed?

- People who score high on intelligence tests tend also to have agile brains and score high in speed of perception and speed of neural processing.
- The direction of correlation has not been determined, and some third factor may influence both intelligence and processing speed.

Multiple-Choice Questions

1. According to Robert Sternberg, what kind of intelligence is assessed by traditional intelligence tests?

a. Linguistic
b. Practical
c. Creative
d. Spatial
e. Analytical

2. According to Charles Spearman and others, which of the following underlies specific mental abilities and is measured by every task on an intelligence test?

a. Savant syndrome
b. General intelligence (*g*)
c. Factor analysis
d. Intelligence
e. Emotional intelligence

3. Of the following, which term best describes the condition in which a person with limited mental ability excels at a specific skill such as computation?

a. Savant syndrome
b. *g* factor
c. Creative intelligence
d. Emotional intelligence
e. Street smarts

4. Which of the following is not a component of emotional intelligence?

a. Understanding emotions
b. Perceiving emotions
c. Using emotions
d. Managing emotions
e. Inventing emotions

Practice FRQs

1. Give a summary, a strength, and a weakness of Charles Spearman's idea of general intelligence.

Answer

1 point: General intelligence is basic intelligence that predicts our abilities in varied academic areas.

1 point: A strength of this idea is that different abilities, such as verbal and spatial, tend to correlate.

1 point: A weakness of this idea is that human abilities are too diverse to be explained by a single general intelligence factor.

2. Name and describe Robert Sternberg's three intelligences.

(3 points)

Module 61

Assessing Intelligence


Marmaduke St. John/Alamy

Module Learning Objectives

- **61-1** Discuss the history of intelligence testing.
- **61-2** Distinguish between aptitude and achievement tests.
- **61-3** Explain the meaning of standardization, and describe the normal curve.
- **61-4** Explain the meanings of reliability and validity.

How do we assess intelligence? And what makes a test credible? Answering these questions begins with a look at why psychologists created tests of mental abilities and how they have used those tests.

The Origins of Intelligence Testing

61-1 When and why were intelligence tests created?

Some societies concern themselves with promoting the collective welfare of the family, community, and society. Other societies emphasize individual opportunity. Plato, a pioneer of the individualist tradition, wrote more than 2000 years ago in *The Republic* that "no two persons are born exactly alike; but each differs from the other in natural endowments, one being suited for one occupation and the other for another." As heirs to Plato's individualism, people in Western societies have pondered how and why individuals differ in mental ability.

Western attempts to assess such differences began in earnest over a century ago. The English scientist Francis Galton (1822–1911) had a fascination with measuring human traits. When his cousin Charles Darwin proposed that nature selects successful traits through the survival of the fittest, Galton wondered if it might be possible to measure "natural ability" and to encourage those of high ability to mate with one another. At the 1884 London Exposition, more than 10,000 visitors received his assessment of their "intellectual strengths" based on such things as reaction time, sensory acuity, muscular power, and body proportions. But alas, on these measures, well-regarded adults and students did not outscore others. Nor did the measures correlate with one another.

Although Galton's quest for a simple intelligence measure failed, he gave us some statistical techniques that we still use (as well as the phrase "nature and nurture"). And his persistent belief in the inheritance of genius—reflected in his book, *Hereditary Genius*—illustrates an important lesson from both the history of intelligence research and the history of science: Although science itself strives for objectivity, individual scientists are affected by their own assumptions and attitudes.

Alfred Binet "Some recent philosophers have given their moral approval to the deplorable verdict that an individual's intelligence is a fixed quantity, one which cannot be augmented. We must protest and act against this brutal pessimism" (Binet, 1909, p. 141).

"The IQ test was invented to predict academic performance, nothing else. If we wanted something that would predict life success, we'd have to invent another test completely." -Social psychologist Robert Zajonc (1984b)

mental age a measure of intelligence test performance devised by Binet; the chronological age that most typically corresponds to a given level of performance. Thus, a child who does as well as the average 8-year-old is said to have a mental age of 8.

Stanford-Binet the widely used American revision (by Terman at Stanford University) of Binet's original intelligence test.

intelligence quotient (IQ) defined originally as the ratio of mental age *(ma)* to chronological age *(ca)* multiplied by 100 (thus, $IQ = ma/ca \times 100$). On contemporary intelligence tests, the average performance for a given age is assigned a score of 100, with scores assigned to relative performance above or below average.

Alfred Binet: Predicting School Achievement

The modern intelligence-testing movement began at the turn of the twentieth century, when France passed a law requiring that all children attend school. Some children, including many newcomers to Paris, seemed incapable of benefiting from the regular school curriculum and in need of special classes. But how could the schools objectively identify children with special needs?

The French government hesitated to trust teachers' subjective judgments of children's learning potential. Academic slowness might merely reflect inadequate prior education. Also, teachers might prejudge children on the basis of their social backgrounds. To minimize bias, France's minister of public education in 1904 commissioned Alfred Binet (1857–1911) and others to study the problem.

Binet and his collaborator, Théodore Simon, began by assuming that all children follow the same course of intellectual development but that some develop more rapidly. On tests, therefore, a "dull" child should perform as does a typical younger child, and a "bright" child as does a typical older child. Thus, their goal became measuring each child's **mental age,** the level of performance typically associated with a certain chronological age. The average 9-year-old, then, has a mental age of 9. Children with below-average mental ages, such as 9-year-olds who perform at the level of typical 7-year-olds, would struggle with age-appropriate schoolwork.

To measure mental age, Binet and Simon theorized that mental aptitude, like athletic aptitude, is a general capacity that shows up in various ways. After testing a variety of reasoning and problem-solving questions on Binet's two daughters, and then on "bright" and "backward" Parisian schoolchildren, Binet and Simon identified items that would predict how well French children would handle their schoolwork.

Note that Binet and Simon made no assumptions concerning *why* a particular child was slow, average, or precocious. Binet personally leaned toward an environmental explanation. To raise the capacities of low-scoring children, he recommended "mental orthopedics" that would help develop their attention span and self-discipline. He believed his intelligence test did not measure inborn intelligence as a meter stick measures height. Rather, it had a single practical purpose: to identify French schoolchildren needing special attention. Binet hoped his test would be used to improve children's education, but he also feared it would be used to label children and limit their opportunities (Gould, 1981).

Lewis Terman: The Innate IQ

Binet's fears were realized soon after his death in 1911, when others adapted his tests for use as a numerical measure of inherited intelligence. This began when Stanford University professor Lewis Terman (1877–1956) found that the Paris-developed questions and age norms worked poorly with California schoolchildren. Adapting some of Binet's original items, adding others, and establishing new age norms, Terman extended the upper end of the test's range from teenagers to "superior adults." He also gave his revision the name it retains today—the **Stanford-Binet.** For Terman, intelligence tests revealed the intelligence with which a person was born.

From such tests, German psychologist William Stern derived the famous **intelligence quotient,** or **IQ.** The IQ is simply a person's mental age divided by chronological age and multiplied by 100 to get rid of the decimal point:

$$\text{IQ} = \frac{\text{mental age}}{\text{chronological age}} \times 100$$

Thus, an average child, whose mental and chronological ages are the same, has an IQ of 100. But an 8-year-old who answers questions as would a typical 10-year-old has an IQ of 125.

The original IQ formula worked fairly well for children but not for adults. (Should a 40-year-old who does as well on the test as an average 20-year-old be assigned an IQ of only 50?) Most current intelligence tests, including the Stanford-Binet, no longer compute an IQ in this manner (though the term *IQ* still lingers as a shorthand expression for "intelligence test score"). Instead, they represent the test-taker's performance *relative to the average performance of others the same age.* This average performance is arbitrarily assigned a score of 100, and about two-thirds of all test-takers fall between 85 and 115.

© Jason Love

Mrs. Randolph takes mother's pride too far.

Terman promoted the widespread use of intelligence testing. His motive was to "take account of the inequalities of children in original endowment" by assessing their "vocational fitness." In sympathy with Francis Galton's *eugenics*—a much-criticized nineteenth-century movement that proposed measuring human traits and using the results to encourage only smart and fit people to reproduce—Terman (1916, pp. 91–92) envisioned that the use of intelligence tests would "ultimately result in curtailing the reproduction of feeble-mindedness and in the elimination of an enormous amount of crime, pauperism, and industrial inefficiency" (p. 7).

With Terman's help, the U.S. government developed new tests to evaluate both newly arriving immigrants and World War I army recruits—the world's first mass administration of an intelligence test. To some psychologists, the results indicated the inferiority of people not sharing their Anglo-Saxon heritage. Such findings were part of the cultural climate that led to a 1924 immigration law that reduced Southern and Eastern European immigration quotas to less than one-fifth of those for Northern and Western Europe.

Binet probably would have been horrified that his test had been adapted and used to draw such conclusions. Indeed, such sweeping judgments became an embarrassment to most of those who championed testing. Even Terman came to appreciate that test scores reflected not only people's innate mental abilities but also their education, native language, and familiarity with the culture assumed by the test. Abuses of the early intelligence tests serve to remind us that science can be value-laden. Behind a screen of scientific objectivity, ideology sometimes lurks.

AP® Exam Tip

David Myers did not use the terms *nature* or *nurture* as he described the contributions of Binet and Terman, but he dropped many hints that should allow you to figure out who leaned toward the nature position and who leaned toward the nurture position. Can you do it?

Modern Tests of Mental Abilities

61-2 What's the difference between achievement and aptitude tests?

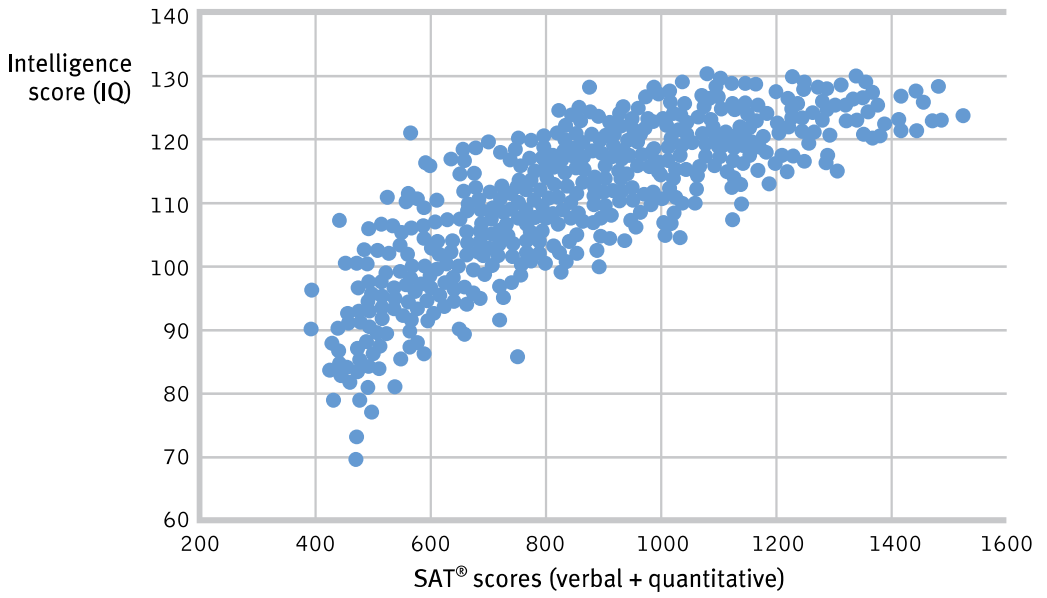
By this point in your life, you've faced dozens of ability tests: school tests of basic reading and math skills, course exams, intelligence tests, and driver's license exams, to name just a few. Psychologists classify such tests as either **achievement tests,** intended to *measure* what you have learned, or **aptitude tests,** intended to *predict* your ability to learn a new skill. Exams covering what you have learned in this course (like the AP® Exam) are achievement tests. A college entrance exam, which seeks to predict your ability to do college work, is an aptitude test—a "thinly disguised intelligence test," says Howard Gardner (1999a). Indeed, total scores on the U.S. SAT® correlated +.82 with general intelligence scores in a national sample of 14- to 21-year-olds (Frey & Detterman, 2004; **FIGURE 61.1** on the next page).

achievement test a test designed to assess what a person has learned.

aptitude test a test designed to predict a person's future performance; aptitude is the capacity to learn.

Figure 61.1

Close cousins: Aptitude and intelligence scores A scatterplot shows the close correlation between intelligence scores and verbal and quantitative SAT® scores. (From Frey and Detterman, 2004.)



Wechsler Adult Intelligence Scale (WAIS) the WAIS is the most widely used intelligence test; contains verbal and performance (nonverbal) subtests.

Richard T. Nowitz/Corbis

Matching patterns Block design puzzles test visual abstract processing ability. Wechsler's individually administered intelligence test comes in forms suited for adults and children.

Psychologist David Wechsler created what is now the most widely used individual intelligence test, the **Wechsler Adult Intelligence Scale (WAIS),** with a version for school-age children (the *Wechsler Intelligence Scale for Children [WISC]*), and another for preschool children. The latest (2008) edition of the WAIS consists of 15 subtests, including these:

- *Similarities*—Reasoning the commonality of two objects or concepts, such as "In what way are wool and cotton alike?"
- *Vocabulary*—Naming pictured objects, or defining words ("What is a guitar?")
- *Block design*—Visual abstract processing, such as "Using the four blocks, make one just like this."

- *Letter-number sequencing*—On hearing a series of numbers and letters, repeat the numbers in ascending order, and then the letters in alphabetical order: "R-2-C-1-M-3."

It yields not only an overall intelligence score, as does the Stanford-Binet, but also separate scores for verbal comprehension, perceptual organization, working memory, and processing speed. Striking differences among these scores can provide clues to cognitive strengths or weaknesses that teachers or therapists can build upon. For example, a low verbal comprehension score combined with high scores on other subtests could indicate a reading or language disability. Other comparisons can help a psychologist or psychiatrist establish a rehabilitation plan for a stroke patient. Such uses are possible, of course, only when we can trust the test results.

Principles of Test Construction

61-3 What are standardization and the normal curve?

To be widely accepted, psychological tests must meet three criteria: They must be *standardized, reliable,* and *valid.* The Stanford-Binet and Wechsler tests meet these requirements.

Standardization

The number of questions you answer correctly on an intelligence test would tell us almost nothing. To evaluate your performance, we need a basis for comparing it with others' performance. To enable meaningful comparisons, test-makers first give the test to a representative sample of people. When you later take the test following the same procedures, your score

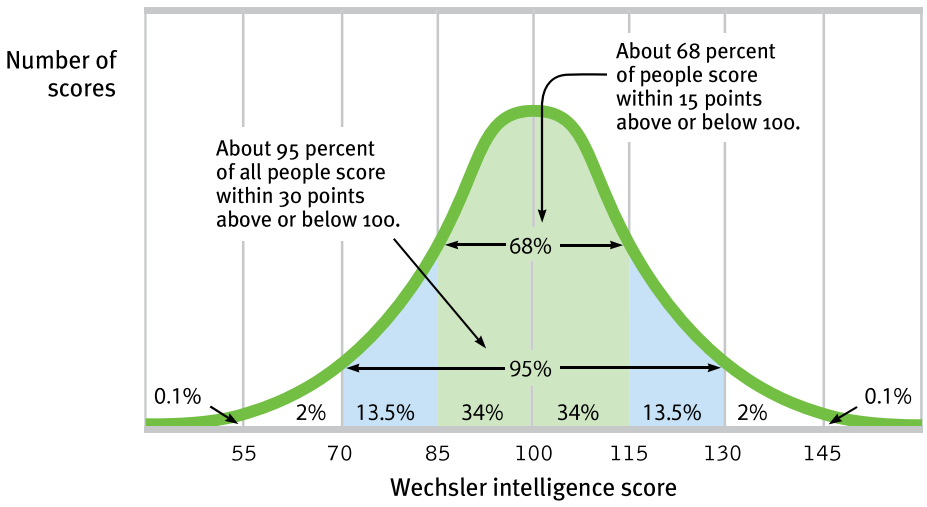


Figure 61.2

The normal curve Scores on aptitude tests tend to form a normal, or bell-shaped, curve around an average score. For the Wechsler scale, for example, the average score is 100.

can be compared with the sample's scores to determine your position relative to others. This process of defining meaningful scores relative to a pretested group is called **standardization.**

Group members' scores typically are distributed in a bell-shaped pattern that forms the **normal curve** shown in **FIGURE 61.2**. No matter what we measure—height, weight, or mental aptitude—people's scores tend to form this roughly symmetrical shape. On an intelligence test, we call the midpoint, the average score, 100. Moving out from the average toward either extreme, we find fewer and fewer people. For both the Stanford-Binet and Wechsler tests, a person's score indicates whether that person's performance fell above or below the average. As Figure 61.2 shows, a performance higher than all but 2 percent of all scores earns an intelligence score of 130. A performance lower than 98 percent of all scores earns an intelligence score of 70.

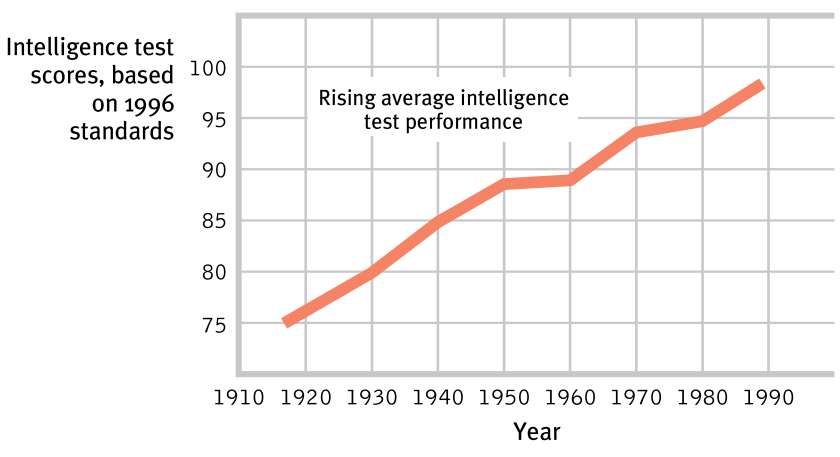
standardization defining uniform testing procedures and meaningful scores by comparison with the performance of a pretested group.

normal curve the symmetrical, bell-shaped curve that describes the distribution of many physical and psychological attributes. Most scores fall near the average, and fewer and fewer scores lie near the extremes.

To keep the average score near 100, the Stanford-Binet and Wechsler scales are periodically restandardized. If you took the WAIS Fourth Edition recently, your performance was compared with a standardization sample who took the test during 2007, not to David Wechsler's initial 1930s sample. If you compared the performance of the most recent standardization sample with that of the 1930s sample, do you suppose you would find rising or declining test performance? Amazingly—given that college entrance aptitude scores were dropping during the 1960s and 1970s—intelligence test performance was improving. This worldwide phenomenon is called the *Flynn effect,* in honor of New Zealand researcher James Flynn (1987, 2009b, 2010), who first calculated its magnitude. As **FIGURE 61.3** indicates, the average person's intelligence test score in 1920 was—by today's standard—only a 76! Such rising performance has been observed in 29 countries, from Canada to rural Australia (Ceci & Kanaya, 2010). Although the gains have recently reversed in Scandinavia, the historic increase is now widely accepted as an important phenomenon (Lynn, 2009; Teasdale & Owen, 2005, 2008).

AP® Exam Tip

Can you remember why the intelligence test scores in Figure 61.2 are marked off in 15-point intervals? Do the 68 percent and 95 percent areas seem familiar? They should—you've seen this graph before. It's Figure 7.3 from the module on statistical reasoning. Intelligence tests are being used to illustrate that 68 percent of a population will be within one standard deviation of the mean for normally distributed data. Ninety-five percent will be within two standard deviations.



Lewis J. Merrim/Science Source

Figure 61.3

Getting smarter? In every country studied, intelligence test performance rose during the twentieth century, as shown here with U.S. Wechsler and Stanford-Binet test performance between 1918 and 1989. In Britain, test scores have risen 27 points since 1942. (From Hogan, 1995.) Very recent data indicate this trend may have leveled off or may even be reversing.

The Flynn effect's cause has been a mystery. Did it result from greater test sophistication? (But the gains began before testing was widespread and have even been observed among preschoolers.) Better nutrition? As the nutrition explanation would predict, people have gotten not only smarter but taller. But in post-war Britain, notes Flynn (2009a), the lower-class children gained the most from improved nutrition but the intelligence performance gains were greater among upper-class children. Or did the Flynn effect stem from more education? More stimulating environments? Less childhood disease? Smaller families and more parental investment (Sundet et al., 2008)?

Regardless of what combination of factors explains the rise in intelligence test scores, the phenomenon counters one concern of some hereditarians—that the higher twentieth-century birthrates among those with lower scores would shove human intelligence scores downward (Lynn & Harvey, 2008). Seeking to explain the rising scores, and mindful of global mixing, one scholar has even speculated about the influence of a genetic phenomenon comparable with "hybrid vigor," which occurs in agriculture when cross-breeding produces corn or livestock superior to the parent plants or animals (Mingroni, 2004, 2007).

reliability the extent to which a test yields consistent results, as assessed by the consistency of scores on two halves of the test, on alternate forms of the test, or on retesting.

validity the extent to which a test measures or predicts what it is supposed to. (See also *content validity* and *predictive validity*.)

content validity the extent to which a test samples the behavior that is of interest.

predictive validity the success with which a test predicts the behavior it is designed to predict; it is assessed by computing the correlation between test scores and the criterion behavior. (Also called *criterion-related validity*.)

Reliability

61-4 What are reliability and validity?

Knowing where you stand in comparison to a standardization group still won't tell us much about your intelligence unless the test has **reliability**—unless it yields dependably consistent scores. To check a test's reliability, researchers retest people. They may use the same test or they may split the test in half to see whether odd-question scores and even-question scores agree. If the two scores generally agree, or *correlate,* the test is reliable. The higher the correlation between the *test-retest* or the *split-half* scores, the higher the test's reliability. The tests we have considered so far—the Stanford-Binet, the WAIS, and the WISC—all have reliabilities of about +.9, which is very high. When retested, people's scores generally match their first score closely.

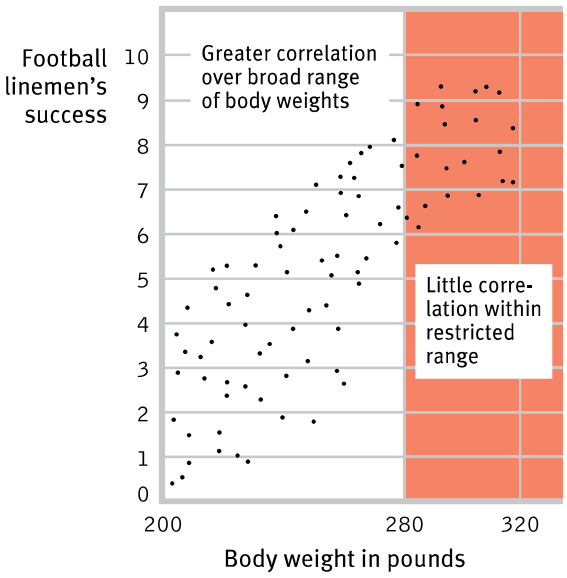
AP® Exam Tip

Be careful! The terms *reliability* and *validity* have more precise meanings to psychologists than they do to the general public.

Validity

High reliability does not ensure a test's **validity**—the extent to which the test actually measures or predicts what it promises. If you use an inaccurate tape measure to measure people's heights, your height report would have high reliability (consistency) but low validity. It is enough for some tests that they have **content validity,** meaning the test taps the pertinent behavior, or *criterion.* The road test for a driver's license has content validity because it samples the tasks a driver routinely faces. Course exams have content validity if they assess one's mastery of a representative sample of course material. But we expect intelligence tests to have **predictive validity:** They should predict the criterion of future performance, and to some extent they do.

Are general aptitude tests as predictive as they are reliable? As critics are fond of noting, the answer is plainly *No.* The predictive power of aptitude tests is fairly strong in the early school years, but later it weakens. Academic aptitude test scores are reasonably good predictors of achievement for children ages 6 to 12, where the correlation between intelligence score and school performance is about +.6 (Jensen, 1980). Intelligence scores correlate even more closely with scores on *achievement tests:* +.81 in one comparison of 70,000 English children's intelligence scores at age 11 with their academic achievement in national exams at age 16 (Deary et al., 2007, 2009). The SAT® exam, used in the United States as a college entrance exam, is less successful in predicting first-year college grades. (The correlation, which is less than +.5, is, however, a bit higher when adjusting for high scorers electing tougher courses [Berry & Sackett, 2009; Willingham et al., 1990].) By the time we get to the Graduate Record Examination® (GRE®; an aptitude test similar to the SAT® exam but for those applying to graduate school), the correlation with graduate school performance is an even more modest but still significant +.4 (Kuncel & Hezlett, 2007).



AP Photo/Bill Kostroun

Figure 61.4
Diminishing predictive power Let's imagine a correlation between football linemen's body weight and their success on the field. Note how insignificant the relationship becomes when we narrow the range of weight to 280 to 320 pounds. As the range of data under consideration narrows, its predictive power diminishes.

Why does the predictive power of aptitude scores diminish as students move up the educational ladder? Consider a parallel situation: Among all American and Canadian football linemen, body weight correlates with success. A 300-pound player tends to overwhelm a 200-pound opponent. But within the narrow 280- to 320-pound range typically found at the professional level, the correlation between weight and success becomes negligible (**FIGURE 61.4**). The narrower the *range* of weights, the lower the predictive power of body weight becomes. If an elite university takes only those students who have very high aptitude scores, those scores cannot possibly predict much. This will be true even if the test has excellent predictive validity with a more diverse sample of students. So, when we validate a test using a wide range of people but then use it with a restricted range of people, it loses much of its predictive validity.

Before You Move On

▶ ASK YOURSELF

Are you working to the potential reflected in your standardized test scores? What, other than your aptitude, is affecting your school performance?

▶ TEST YOURSELF

What was the purpose of Binet's pioneering intelligence test?

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

Module 61 Review

61-1 When and why were intelligence tests created?

- In the late 1800s, Francis Galton, who believed that genius was inherited, attempted but failed to construct a simple intelligence test.
- In France in 1904, Alfred Binet, who tended toward an environmental explanation of intelligence differences, started the modern intelligence-testing movement by developing questions to measure children's *mental age* and thus predict progress in the school system.
- During the early twentieth century, Lewis Terman of Stanford University revised Binet's work for use in the United States.
  - Terman believed intelligence is inherited, and he thought his *Stanford-Binet* could help guide people toward appropriate opportunities.
  - During this period, intelligence tests were sometimes used to "document" scientists' assumptions about the innate inferiority of certain ethnic and immigrant groups.

61-2 What's the difference between achievement and aptitude tests?

- *Achievement tests* are designed to assess what you have learned.
- *Aptitude tests* are designed to predict what you can learn.
- The *WAIS (Wechsler Adult Intelligence Scale)*, an aptitude test, is the most widely used intelligence test for adults.

61-3 What are standardization and the normal curve?

- *Standardization* establishes a basis for meaningful score comparisons by giving a test to a representative sample of future test-takers.
- The distribution of test scores often forms a *normal* (bell-shaped) *curve* around the central average score, with fewer and fewer scores at the extremes.

61-4 What are reliability and validity?

- *Reliability* is the extent to which a test yields consistent results (on two halves of the test, or when people are retested).
- *Validity* is the extent to which a test measures or predicts what it is supposed to.
  - A test has *content validity* if it samples the pertinent behavior (as a driving test measures driving ability).
  - It has *predictive validity* if it predicts a behavior it was designed to predict. (Aptitude tests have predictive ability if they can predict future achievements.)

Multiple-Choice Questions

1. A test-developer defines uniform testing procedures and meaningful scores by comparison with the performance of a pretested group. Which of the following best describes this process?

a. Reliability testing
b. Validation
c. Content validation
d. Standardization
e. Predictive validity

2. Which of the following best describes the extent to which a test yields consistent results upon retesting?

a. Content validity
b. Validity
c. Reliability
d. Predictive validity
e. Normal curve

3. Which of the following can be used to demonstrate that only about 2 percent of the population scores are at least two standard deviations above the mean on an intelligence test?

a. Reliability test
b. Aptitude test
c. Predictive validity test
d. Test-retest procedure
e. Normal curve

Practice FRQs

1. What are the fundamental differences between achievement and aptitude tests?

Answer

1 point: Achievement tests are designed to assess what a person has learned.

1 point: An aptitude test is designed to predict a person's future performance.

2. Name and briefly describe the three essential principals of test construction.

(3 points)