

CHAPTER 1

What Are Psychological Tests?

CHAPTER 1: WHAT ARE PSYCHOLOGICAL TESTS?

After completing your study of this chapter, you should be able to do the following:

- Define what a psychological test is, and understand that psychological tests extend beyond personality and intelligence tests.
- Trace the history of psychological testing from Alfred Binet and intelligence testing to the tests of today.
- Describe the ways in which psychological tests can be similar to and different from one another.
- Describe the three characteristics that are common to all psychological tests, and understand that psychological tests can demonstrate these characteristics to various degrees.
- Describe the assumptions that must be made when using psychological tests.
- Describe the different ways that psychological tests can be classified.
- Describe the differences among four commonly used terms that students often get confused: psychological assessment, psychological tests, psychological measurement, and surveys.
- Identify and locate print and online resources that are available for obtaining information about psychological tests.

“When I was in the second grade, my teacher recommended that I be placed in the school’s gifted program. As a result, the school psychologist interviewed me and had me take an intelligence test.”

“Last semester I took a class in abnormal psychology. The professor had all of us take several personality tests, including the MMPI [Minnesota Multiphasic Personality Inventory]. It was awesome! We learned about different types of psychological disorders that the MMPI can help diagnose.”

4 SECTION I: OVERVIEW OF PSYCHOLOGICAL TESTING

“This year I applied for a summer job with a local bank. As a part of the selection process, I had to participate in a structured interview and an assessment center.”

“Yesterday I took my driving test—both the written and the road test. I couldn’t believe everything they made me do. I had to parallel park, switch lanes, and make both right and left turns.”

If your instructor asked whether you have ever taken a psychological test, you would probably report the intelligence test you took as an elementary school student or the personality test you took in your abnormal psychology class. If your instructor asked what the purpose of psychological testing is, you would probably say it is to determine whether someone is gifted or has a psychological disorder. Intelligence tests and personality tests are indeed psychological tests—and they are indeed used to identify giftedness and diagnose psychological disorders. However, this is only a snapshot of what psychological testing is all about. There are many types of psychological tests, and they have many different purposes.

In this chapter, we introduce you to the concept of psychological testing. We discuss what a psychological test is and introduce some tests you might never have considered to be psychological tests. Then, after exploring the history of psychological testing, we discuss the three defining characteristics of psychological tests and the assumptions that must be made when using these tests. We then turn our attention to the many ways of classifying tests. We also distinguish four concepts that students often get confused: psychological assessment, psychological tests, psychological measurement, and surveys. We conclude this chapter by sharing with you some of the resources (print and online) that are available for locating information about psychological testing and specific psychological tests.

Why Should You Care About Psychological Testing?

Before discussing what a psychological test is, we would like to increase your understanding of just how important it is for you to understand the foundations of psychological testing. Psychological testing is not just another subject that you may study in college; rather, it is a topic that personally affects many individuals. Each day, different types of professionals administer psychological tests to many different individuals, and the results of these tests are used in ways that significantly affect you and those around you. For example, test scores are used to diagnose mental disorders, to determine whether medicines should be prescribed (and, if so, which ones), to treat mental and emotional illnesses, to select individuals for jobs, to select individuals for undergraduate and professional schools (for example, medical school, law school), and to determine grades. Good tests facilitate high-quality decisions, and bad tests facilitate low-quality decisions.

The consequences of bad decisions can be significant. For example, a poor hiring decision can dramatically affect both the person being hired and the hiring organization. From the organization’s perspective, a poor hiring decision can result in increased absenteeism, reduced morale of other staff, and lost productivity and revenue. From the employee’s perspective, a poor hiring decision may result in a loss of motivation, increased stress leading to depression and anxiety, and perhaps loss of opportunity to make progress in his or her career. Although you might never administer, score, or interpret a test, it is very likely that you or someone you know may have a life-altering decision made about him or her

based on test scores. Therefore, it is important that you understand the foundations of psychological testing, specifically how to tell whether a decision is a good or bad one. Being able to do this requires that you understand the foundations of psychological testing.

What Are Psychological Tests?

Each anecdote at the beginning of this chapter involves the use of a psychological test. Intelligence tests, personality tests, interest and vocational inventories, college entrance exams, classroom tests, structured interviews, assessment centers, and driving tests all are psychological tests. Even the self-scored tests that you find in magazines such as *Glamour* and *Seventeen* (tests that supposedly tell you how you feel about your friends, stress, love, and more) can be considered psychological tests. Although some are more typical, all meet the definition of a psychological test. Together, they convey the very different purposes of psychological tests. For a continuum of some of the most and least commonly recognized types of psychological tests, see Figure 1.1.

Similarities Among Psychological Tests

While different testing professionals might define psychological testing in various ways, psychological testing is historically and best defined as “the process of administering, scoring, and interpreting psychological tests” (Maloney & Ward, 1976, p. 9). But what exactly is a psychological test? We can easily answer this question by considering what all psychological tests do.

First, all psychological tests require a person to perform some **behavior**—an observable and measurable action. For example, when students take a multiple-choice midterm exam, they must read the various answers for each item and identify the best one. When individuals take an intelligence test, they may be asked to define words or solve math problems. When participating in a structured job interview, individuals must respond to questions from the interviewer—questions such as “Tell me about a time when you had to deal with an upset customer. What was the situation, what did you do, and what was the outcome?” In each of these cases, individuals are performing some observable and measurable behavior.

Second, the behavior an individual performs is used to measure some personal attribute, trait, or characteristic that is thought to be important in describing or understanding human behavior. For example, the questions on a multiple-choice exam might measure your knowledge of a particular subject

Figure 1.1 A Continuum of Psychological Tests

<i>More Typical</i>	←—————→		<i>Less Typical</i>
Personality tests Intelligence tests	Vocational tests Interest inventories Achievement tests Ability tests	Self-scored magazine tests Classroom quizzes and exams	Road portion of driving test Structured employment interviews Assessment centers

6 SECTION I: OVERVIEW OF PSYCHOLOGICAL TESTING

area such as psychological testing. The words you defined or the math problems you solved might measure your verbal ability or quantitative reasoning. It is also important to note that sometimes the behavior an individual performs is also used to make a prediction about some outcome. For example, the questions you answered during a structured job interview may be used to predict your success in a management position.

So, what is a psychological test? It is something that requires you to perform a behavior to measure some personal attribute, trait, or characteristic or to predict an outcome.

Differences Among Psychological Tests

Although all psychological tests require that you perform some behavior to measure personal attributes, traits, or characteristics or to predict outcomes, these tests can differ in various ways. For example, they can differ in terms of the behavior they require you to perform, what they measure, their content, how they are administered and formatted, how they are scored and interpreted, and their psychometric quality (**psychometrics** is the quantitative and technical aspect of mental measurement).

Behavior Performed

The behaviors a test taker must perform vary by test. For example, a popular intelligence test, the Wechsler Adult Intelligence Scale–Fourth Edition (WAIS-IV), a general test of adult intelligence requires test takers to (among other things) define words, repeat lists of digits, explain what is missing from pictures, and arrange blocks to duplicate geometric card designs (Pearson Education, 2012e). The Thematic Apperception Test (TAT), a widely used and researched projective personality test designed at Harvard University in the 1930s, requires test takers to look at ambiguous pictures showing a variety of social and interpersonal situations and to tell stories about each picture (Pearson Education, 2012d). The 2011 revised Graduate Record Examination (GRE) General Test, a graduate school admissions test that measures verbal reasoning, quantitative reasoning, and analytical writing skills, requires test takers to answer multiple-choice questions and respond to analytical writing tasks (Educational Testing Service, 2012b). The road portion of an auto driving test typically requires test takers to do things such as start a car, change lanes, make right and left turns, use turn signals properly, and parallel park. Assessment centers require job applicants to participate in simulated job-related activities (that mimic the activities they would perform in the job) such as engaging in confrontational meetings with disgruntled employees, processing e-mail and paperwork, and conducting manager briefings.

Attribute Measured and Outcome Predicted

What a test measures or predicts can vary. For example, the WAIS-IV asks individuals to explain what is missing from pictures to measure verbal intelligence. The TAT requires individuals to tell stories about pictures to identify conscious and unconscious drives, emotions, conflicts, and so on in order to ultimately measure personality. The road portion of a driving test requires individuals to perform various driving behaviors to measure driving ability. The GRE requires students to answer different types of questions to determine if they are ready for graduate-level work and predict success in graduate school.

Some of the characteristics, attributes, and traits commonly measured by psychological tests include personality, intelligence, motivation, mechanical ability, vocational preference, spatial ability, and anxiety. Some of the outcomes that tests typically predict include worker productivity, success in college, and who will benefit from specialized services such as clinical treatment programs.



More detail about the WAIS-IV can be found in Test Spotlight 1.1 in Appendix A.

Content

Two tests that measure the same characteristic, attribute, or trait can require individuals to perform significantly different behaviors or to answer significantly different questions. Sometimes how the test developers define the particular characteristic, attribute, or trait affects how the test is structured. For example, the questions on two intelligence tests may differ because one author may define intelligence as the ability to reason and another author may define it in terms of **emotional intelligence**—one's ability to understand one's own feelings and the feelings of others and to manage one's emotions (Gibbs, 1995).

The difference in content may also be due to the theoretical orientation of the test. (We talk more about theoretical orientation and its relation to test content in Chapter 9.)

Administration and Format

Psychological tests can differ in terms of how they are administered and their format. A test can be administered in paper-and-pencil format (individually or in a group setting), on a computer, or verbally. Similarly, a psychological test may consist of multiple-choice items, agree/disagree items, true/false items, open-ended questions, or some mix of these. There are also tests that ask respondents to perform some behavior such as sorting cards, playing a role, or writing an essay.

Scoring and Interpretation

Psychological tests can differ in terms of how they are scored and interpreted. Some tests require test takers to document answers on scannable sheets that are then computer scored. Some tests are hand-scored by the person administering the test. Other tests are scored by the test takers themselves. In terms of interpretation, some tests generate results that can be interpreted easily by the test taker, and others require a knowledgeable professional to explain the results to the test taker.

Psychometric Quality

Last, but extremely important, psychological tests can differ in terms of their psychometric quality. For now, let us just say that there are a lot of really good tests out there that measure what they say they measure and do so consistently, but there are also a lot of really poor tests out there that do not measure what they say they measure. Good tests measure what they claim to measure, and any conclusions that are drawn from the test scores about the person taking the test are appropriate (they are what we call *valid*). Good tests also measure whatever they measure consistently (they are what we

8 SECTION I: OVERVIEW OF PSYCHOLOGICAL TESTING

call *reliable*). The concepts of reliability and validity are central to determining whether a test is “good” or “bad” and are covered in detail later in this textbook. These concepts are so important that four chapters are devoted to them (Chapter 6 covers reliability, and Chapters 7–9 cover validity).

Because tests can differ in so many ways, to make informed decisions about tests, you must know how to properly critique a test. A critique of a test is an analysis of the test. A good critique answers many of the questions in Table 1.1. (These questions are also in Appendix B.) Your instructor may have additional ideas about what constitutes a good critique.

INTERIM SUMMARY 1.1 SIMILARITIES AND DIFFERENCES AMONG PSYCHOLOGICAL TESTS

Similarities

- All psychological tests require an individual to perform a behavior.
- The behavior performed is used to measure some personal attribute, trait, or characteristic.
- This personal attribute, trait, or characteristic is thought to be important in describing or understanding behavior.
- The behavior performed may also be used to predict outcomes.

Differences

Psychological tests can differ in terms of the following:

- The behavior they require the test taker to perform
- The attribute they measure
- Their content
- How they are administered and formatted
- How they are scored and interpreted
- Their psychometric quality

Table 1.1 Guidelines for Critiquing a Psychological Test

General descriptive information

- What is the title of the test?
- Who is the author of the test?
- Who publishes the test, and when was it published? (Include dates of manuals, norms, and supplementary materials.)
- How long does it take to administer the test?
- How much does it cost to purchase the test? (Include the cost of the test, answer sheets, manual, scoring services, and so on.)
- Is the test proprietary or nonproprietary?

Purpose and nature of the test

- What does the test measure? (Include scales.)
- What does the test predict?
- What behavior does the test require the test taker to perform?
- What population was the test designed for (for example, age, type of person)?
- What is the nature of the test (for example, maximal performance, behavior observation, self-report, standardized or nonstandardized, objective or subjective)?
- What is the format of the test (for example, paper-and-pencil or computer, multiple choice or true/false)?

Practical evaluation

- Is the test manual comprehensive? (Does it include information on how the test was constructed, its reliability and validity, composition of norm groups, whether it is easy to read?)
- Is the test easy or difficult to administer?
- How clear are the administration directions?
- How clear are the scoring procedures?
- What qualifications and training does a test administrator need to have?
- Does the test have face validity?

Technical evaluation

- Is there a norm group?
- Who comprises the norm group?
- What types of norms are there (for example, percentiles, standard scores)?
- How was the norm group selected?
- Are there subgroup norms (for example, by age, gender, region, occupation, and so on)?
- What is the estimate of the test's reliability?
- How was reliability determined?
- What is the evidence for the validity of the test?
- How was the evidence for validity gathered?
- What is the standard error of measurement?
- What are the confidence intervals?

Test reviews

- What do reviewers say are the strengths and weaknesses of the test?
- What studies that use the test as a measurement instrument have been published in peer-reviewed journals?
- How did the test perform when researchers or test users, other than the test developer or publisher, used it?

Summary

- Overall, what do you see as being the strengths and weaknesses of the test?

The History of Psychological Testing

Some scholars believe that the use of psychological tests can be traced to 2200 BCE in ancient China. For a summary of this history, see For Your Information Box 1.1. Most scholars agree that serious research efforts on the use and usefulness of psychological tests did not begin until the 20th century with the advent of intelligence testing.

Intelligence Tests

Alfred Binet and the Binet–Simon Scale

Late in the 19th century, Alfred Binet founded the first experimental psychology research laboratory in France. In his lab, Binet attempted to develop experimental techniques to measure intelligence and reasoning ability. He believed that intelligence was a complex characteristic that could be determined by evaluating a person's reasoning, judgment, and problem-solving abilities. Binet tried a variety of tasks to measure reasoning, judgment, and problem solving on his own children as well as on other children in the French school system.

Binet was successful in measuring intelligence, and in 1905 he and Théodore Simon published the first test of mental ability, the Binet–Simon Scale. Parisian school officials used this scale to decide which children, no matter how hard they tried, were unable to profit from regular school programs (Binet & Simon, 1905).

FYI

FOR YOUR INFORMATION BOX 1.1

Psychological Tests: From Ancient China to the 20th Century

2200 BCE: Xia Dynasty

The use of psychological tests may date back approximately 4,000 years to 2200 BCE, when the Chinese emperor Yushun examined officials every third year to determine whether they were suitable to continue in office (DuBois, 1970; W. A. P. Martin, 1870). However, modern ancient China scholars say little archaeological evidence exists to support these claims. Reliable writing systems were developed by the Chinese somewhere between 1766 and 1122 BCE (Shang dynasty; Bowman, 1989). Nowhere in the writings were there any hints suggesting that leaders were examined as just described. Even in 1115 BCE, with the advent of more elaborate writing systems, there were no inscriptions or writings to suggest the existence of such an examination process (W. A. P. Martin, 1870).

**200–100 BCE: Late Qin, Early Han Dynasty**

Most modern ancient China scholars agree that royal examinations began around 200 to 100 BCE, in the late Qin (Ch'in) or early Han dynasty (Eberhard, 1977; Franke, 1960; Hucker, 1978; Pirazzoli-t'Serstevens, 1982; Rodzinski, 1979).

618–907 CE: T'ang Dynasty

The examination systems seem to have been discontinued until the T'ang dynasty, when their use increased significantly (Bowman, 1989).

1368–1644: Ming Dynasty

During the Ming dynasty, the examinations became more formal, with different levels of examinations (municipal, county, provincial, and national). The examination results became associated with granting formal titles, similar to today's university degrees. Upon passing each level of examination, people received more titles and increasingly more power in the civil service (Bowman, 1989). The examinations were distressful, and this distress became a part of Chinese culture and also a part of folk stories and the literature (poems, comedies, and tragedies). Nonetheless, the examination system seemed to work well. Today, many scholars believe that the examination system kept talented men in the national government (Kracke, 1963) and kept members of the national government from becoming nobility because of their descent.

Seeing the value of the examinations for making important decisions, European governments, and eventually the governments of the United Kingdom, the United States, Canada, and other countries, adopted the use of such examination systems.

1791: France and Britain

France initially began using this kind of examination system in 1791. However, soon after, Napoleon temporarily abolished them. The system adopted by France served as a model for a British system started in 1833 to select trainees for the Indian civil service—the beginning of the British civil service.

1860s: United States

Due to the success of the British system, Senator Charles Sumner and Representative Thomas Jenckes proposed to Congress in 1860 that the United States use a similar system. Jenckes's report, *Civil Service in the United States*, described the British and Chinese systems in detail. This report laid the foundation for the establishment of the Civil Service Act Health and Psychosocial Instruments (HAPI), passed in January 1883.

20th Century: Western Europe and the United States

In 1879, Wilhelm Wundt introduced the first psychological laboratory, in Leipzig, Germany. At this time, psychology was the study of the similarities among people. For example, physiological psychologists studied how the brain and the nervous system function, and experimental psychologists conducted research to discover how people learn and remember. Strongly influenced by James McKeen Cattell, an American researcher in Wundt's laboratory, psychologists turned their attention to exploring individual differences. Cattell and others realized that learning about the differences among people was just as important as learning about the similarities among people. They believed that developing formal psychological tests to measure individual differences could help solve many social problems, such as who should be placed in remedial programs, who should be sent to battlefields, and who should be hired for particular jobs. At this time, scientists were particularly interested in finding a quantitative way of measuring general intelligence.

During the early 20th century, serious research efforts began on the use and usefulness of various testing procedures. Research conducted by scholars in the United States and Germany eventually led to Alfred Binet's research on intelligence in children.

Lewis Terman and the Stanford–Binet

Binet's work influenced psychologists across the globe. Psychological testing became a popular method of evaluation, and the Binet–Simon Scale was adapted for use in many countries. In 1916, Lewis Terman, an American psychologist, produced the Stanford–Binet Intelligence Scales, an adaptation of Binet's original test. This test, developed for use with Americans ages 3 years to adulthood, was used for many years. A revised edition of the Stanford–Binet remains one of the most widely used intelligence tests today.



More detail about the Stanford-Binet Intelligence Scales can be found in Test Spotlight 1.2 in Appendix A.

The Wechsler–Bellevue Intelligence Scale and the Wechsler Adult Intelligence Scale

By the 1930s, thousands of psychological tests were available, and psychologists and others were debating the nature of intelligence (what intelligence was all about). This dispute over defining intelligence prompted

the development in 1939 of the original Wechsler–Bellevue Intelligence Scale (WBIS) for adults, which provided an index of general mental ability (as did the Binet–Simon Scale) and revealed patterns of a person’s intellectual strengths and weaknesses. David Wechsler, the chief psychologist at Bellevue Hospital in New York City, constructed the WBIS believing that intelligence is demonstrated based on an individual’s ability to act purposefully, think logically, and interact/cope successfully with the environment (Hess, 2001; B. G. Rogers, 2001; Thorne & Henley, 2001). Wechsler published the second edition, the WBIS-II, in 1946.

In 1955, Wechsler revised the WBIS-II and renamed it the Wechsler Adult Intelligence Scale (WAIS). In 1981 and 1991 the WAIS was updated and published as the WAIS-R and WAIS-III, respectively. In a continuing effort to improve the measurement of intelligence, as well as the clinical utility and user-friendliness of the test, the fourth edition was published in 2008 (Pearson Education, 2012e).

Personality Tests

In addition to intelligence testing, the early 1900s brought about an interest in measuring personality.

The Personal Data Sheet

During World War I, the U.S. military wanted a test to help detect soldiers who would not be able to handle the stress associated with combat. To meet this need, the American Psychological Association (APA) commissioned an American psychologist, Robert Woodworth, to design such a test, which came to be known as the Personal Data Sheet (PDS). The PDS was a paper-and-pencil psychiatric interview that required military recruits to respond *yes* or *no* to a series of 200 questions (eventually reduced to 116 questions) that searched for mental disorders. The questions covered topics such as excessive anxiety, depression, abnormal fears, impulse problems, sleepwalking, nightmares, and memory problems (Segal & Coolidge, 2004). One question asked, “Are you troubled with the idea that people are watching you on the street?” (cited in R. J. Cohen, Swerdlik, & Phillips, 1996). During a pilot study of the test, new recruits on average showed 10 positive psychoneurotic symptoms; recruits who were deemed unfit for service generally showed 30 to 40 positive psychoneurotic symptoms (Segal & Coolidge, 2004). Unfortunately, because Woodworth did not complete the final design of this test until too late in the war, the PDS was never implemented or used to screen new recruits.

After World War I, Woodworth developed the Woodworth Psychoneurotic Inventory, a version of the PDS. Unlike the PDS, the Woodworth Psychoneurotic Inventory was designed for use with civilians and was the first self-report test. It was also the first widely used personality inventory.

The Rorschach Inkblot Test and the TAT

During the 1930s, interest also grew in measuring personality by exploring the unconscious. With this interest came the development of two important projective tests: the Rorschach Inkblot Test and the TAT. The Rorschach, a projective personality test (described further in Chapter 14), was developed by Swiss psychiatrist Hermann Rorschach. The TAT, also a projective personality test, was developed by two American psychologists, Henry A. Murray and C. D. Morgan. Both tests are based on the personality theories of Carl Jung and continue to be widely used today for personality assessment.

Vocational Tests

During the 1940s, a need developed for **vocational tests** to help predict how successful an applicant would be in specific occupations. The Public Employment Services needed such tests because thousands of people had lost their jobs due to the Great Depression and thousands more were coming out of school and seeking work. Because there were not enough jobs, people were forced to look for new lines of work. As a result, psychologists developed large-scale programs to design vocational aptitude tests that would predict how successful a person would be at an occupation before entering it. In 1947, the Department of Labor developed the General Aptitude Test Battery (GATB) to meet this need. The GATB was used for a variety of purposes, including vocational counseling and occupational selection.

By the mid-20th century, numerous tests were available and they were used by many to make important decisions about individuals. Because of the increased use of psychological tests, to help protect the rights of the test taker, the APA (1953) published *Ethical Standards of Psychologists*. (We discuss these ethical standards in more detail in Chapter 3.)

Testing Today

In the 21st century, psychological testing is a big business. There are thousands of commercially available, standardized psychological tests as well as thousands of unpublished tests. Tests are published by hundreds of test publishing companies that market their tests very proactively—on the web and in catalogs. While in 1995 test sales totaled approximately \$7 million, by 1997 sales had increased to \$263 million (Clarke, Madaus, Horn, & Ramos, 2002). By 2002, test sales were estimated to be somewhere between \$400 and \$700 million per year. In the early 1990s, 20 million Americans per year were taking psychological tests (Hunt, 1993). From 1991 to 2010, the number of individuals taking the SAT alone increased from approximately 1 million to 1.7 million (College Board, 2011).

For the names and web addresses of some of the most well-known test publishers, as well as some of the most popular tests they publish, see On the Web Box 1.1. Publishing and marketing companies are capitalizing on the testing trend, creating and marketing a bonanza of new products and study aids. To read about some of these products and study aids, see In the News Box 1.1.

Today, psychological testing is a part of the American culture. Psychological tests are in use everywhere. For example, let us take a look at Sylvan Learning, a provider of personal instructional services to children from kindergarten through 12th grade that has more than 1,100 centers worldwide. You might be familiar with Sylvan Learning because of the test preparation programs it offers (for example, preparation for the SAT). However, did you know that much of its business is focused on personalized programs to help children develop skills in areas such as reading, math, and writing? These personalized programs are created by administering and combining the results of standardized tests to capture a student's academic strengths and weaknesses and to identify skill gaps (Sylvan Learning, 2010). Sylvan Learning uses identified skill gaps, often the reason for underperformance in school, to create a blueprint for an individual child's unique tutoring program. The company also administers learning style inventories to help instructors understand how each child learns best. Trained and certified instructors integrate these learning styles into their tutoring sessions to promote individual student learning (Sylvan Learning, 2010).

Now let us take a look at the Society for Human Resources Management (SHRM). As the world's largest association devoted to human resources management, it provides human resources professionals with essential information and resources (SHRM, 2012a). One of these resources is an online testing center, which provides SHRM members who are qualified testing professionals with electronic access to

more than 400 tests, from over 50 test publishers, in areas such as personality and skills assessment, coaching and leadership, mechanical and technical skills, information technology skills, pre-employment screening, and career exploration (SHRM, 2012b). The testing center allows qualified testing professionals to purchase individual tests, administer the tests online, and receive electronic reports.

ON THE WEB BOX 1.1

Names and Web Addresses of Test Publishers



Open your web browser, go to your favorite search engine, and conduct a search for “test publishers” or “psychological test publishers.” You will find pages and pages of websites dedicated to psychological testing and publishing. You will also find the websites of hundreds of test publishers. Although there are many different publishers, some of the most well known, including some of the widely known tests they publish, are listed here:

<i>Publisher</i>	<i>Website</i>	<i>Popular Published Tests</i>
Educational Testing Service	www.ets.org	<ul style="list-style-type: none"> • Advanced Placement (AP) Program Tests • Graduate Management Admission Test (GMAT) • Graduate Record Examinations (GRE) • SAT • Test of English as a Foreign Language (TOEFL)
Pearson	www.pearsonassessments.com	<ul style="list-style-type: none"> • BarOn Emotional Quotient Inventory • Bayley Scales of Infant and Toddler Development—III • Bender Visual-Motor Gestalt Test—II • Watson–Glaser Critical Thinking Appraisal
Hogan Assessment Systems	www.hoganassessments.com	<ul style="list-style-type: none"> • Hogan Personality Inventory (HPI) • Hogan Development Survey (HDS) • Hogan Business Reasoning Inventory (HBRI) • Motives, Values, Preferences Inventory (MVPI)
IPAT	www.ipat.com	<ul style="list-style-type: none"> • 16 Personality Factors (16PF)
PAR	www4.parinc.com	<ul style="list-style-type: none"> • Self-Directed Search • NEO Personality Inventory • Personality Assessment Inventory • Slosson Intelligence Test—Revised for Children and Adults
Psytech International	www.psytech.co.uk	<ul style="list-style-type: none"> • Occupational Interest Profile • Clerical Test Battery • Values and Motives Inventory
PSI	www.psonline.com	<ul style="list-style-type: none"> • Customer Service Battery • Firefighter Selection Test • Police Selection Test
Hogrefe	www.testagency.com	<ul style="list-style-type: none"> • Rorschach Inkblot Test • Trauma Symptom Inventory (TSI) • WPQ Emotional Intelligence Questionnaire
University of Minnesota Press Test Division	www.upress.umn.edu/test-division	<ul style="list-style-type: none"> • Minnesota Multiphasic Personality Inventory (MMPI)
Wonderlic	www.wonderlic.com	<ul style="list-style-type: none"> • Wonderlic Personnel Test

IN THE NEWS

Box 1.1 SAT Prep Tools: From Cellphones to Handhelds to CDs

Early in 2005, the College Board introduced thousands of high school juniors to the new SAT. No longer containing the much-dreaded analogy questions, the new SAT is longer and more difficult and, for the first time, contains a writing section (College Board, 2012). The writing section contains multiple-choice questions that assess how well test takers use standard written English language and a handwritten essay to assess how well they can develop a point of view on a topic.



Not wanting to miss an opportunity, publishers capitalized on the updated SAT by creating and marketing a number of new and innovative products—products promising to appeal to today's technology-savvy, music-hungry, multitasking teens. In 2005, the *Wall Street Journal* published an article introducing some of these unique products. In 2009, the products were still being marketed to students preparing to take the SAT.



Princeton Review: This company offers private instruction and tutoring for standardized achievement tests. In partnership with Cocel, Princeton Review has developed a new software program called Prep for the SAT, which beams SAT practice questions, including reading passages, to cell phones so that students can prepare for the SAT at their convenience. Answers are quickly graded, and parents can even receive electronic reports. In 2005, Princeton Review also released Pocket Prep, an interactive, portable, handheld SAT prep device designed to help 21st-century high school students prepare for the SAT using a format and technology that suits their lifestyles and preferences. Pocket Prep features information about the new SAT; comprehensive verbal, math, and essay preparation; full-length timed practice exams; instant scoring; and personal diagnostic reports. It also includes practice drills, flash cards, and an extensive verbal and essay reference suite to help students maximize their grammar and essay scores.

Kaplan: Another test preparation company, Kaplan has designed software for cell phones and handheld devices and is publishing books, such as *Frankenstein* and *Wuthering Heights*, that contain SAT vocabulary words in bold print as well as their definitions. An example of a sentence containing an SAT vocabulary word (*desolation*) might be “Mr. Heathcliff and I are such a suitable pair to divide the desolation between us.”

SparkNotes: This Internet-based, youth-oriented education product (owned by Barnes & Noble) has published several Spuzzles books containing crossword puzzles in which the answers are commonly occurring SAT vocabulary words. For example, in U.S. History Spuzzle No. 56, the clue to 8 Across is “English Quaker who founded Pennsylvania in 1681.”

Wiley Publishing: A well-known publisher of print and electronic products, Wiley has published a teen novel, *The Marino Mission: One Girl, One Mission, One Thousand Words*, that contains 1,000 need-to-know SAT vocabulary words. Not only are vocabulary words defined at the bottom of each page, but there also are self-tests at the end of the novel to help readers retain what they have learned.

Defined Mind: These independent recording artists, along with Kaplan, have produced Vocabulary Accelerator, a 12-track CD full of rock, folk-funk, and techno beats. What is unique is that the lyrics are studded with SAT vocabulary.

SOURCE: Adapted from Kronhold (2005).

One of the most significant and controversial uses of psychological testing in the 21st century has been a result of the No Child Left Behind Act of 2001 (NCLB). NCLB, which President George W. Bush signed into law on January 8, 2002, was intended to improve the performance of America's primary and secondary schools. NCLB contains the following four basic strategies for improving the performance of schools—strategies that were intended to change the culture of America's schools by defining a school's success in terms of the achievement of its students (U.S. Department of Education, 2004):

1. Increase the accountability that states, school districts, and schools have for educating America's children by requiring that all states implement statewide systems that (a) set challenging standards for what children in Grades 3 to 8 should know and learn in reading and math, (b) test students in Grades 3 to 8 on a yearly basis to determine the extent to which they know and have learned what they should have according to state standards, and (c) include annual statewide progress objectives to ensure that all students are proficient by the 12th grade.
2. Ensure that all children have access to a quality education by allowing parents to send their children to better schools if their schools do not meet state standards.
3. Increase the amount of flexibility that high-performing states and school districts have for spending federal education dollars.
4. Place more emphasis on developing children's reading skills by making grants available to states to administer screening and diagnostic assessments to identify children who may be at risk for reading failure and by providing teachers with professional development and resources to help young children attain the knowledge and skills they need to be readers.

Ten years after implementing NCLB to improve the quality of education across the United States, there remains significant controversy regarding the “stringent, standardized testing that would effectively identify underachieving schools” (Chen, 2011, para. 1). While tests have always played a critical role in the assessment of student achievement, NCLB requires that students be tested more often and relies on test scores to make more important decisions than in the past. In Chapter 13, we talk more about how one state, Florida, has responded to NCLB, focusing primarily on the role that psychological tests have played in assessing the extent to which children and schools measure up to state standards.

The Defining Characteristics of Psychological Tests

As we have already discussed, a psychological test is anything that requires an individual to perform a behavior for the purpose of measuring some attribute, trait, or characteristic or to predict an outcome. All good psychological tests have three characteristics in common:

1. They representatively sample the behaviors thought to measure an attribute or thought to predict an outcome. For example, suppose we are interested in developing a test to measure your physical ability. One option would be to evaluate your performance in every sport you have ever played. Another option would be to have you run the 50-meter dash. Both of these options have drawbacks. The first option would be very precise, but not very practical. Can you imagine how much time and energy it would take to review how you performed in every sport you have ever played? The second option is too narrow and unrepresentative. How fast you run the 50-meter dash does not tell us much about your physical ability in general. A better method would be to take a representative sample of performance in sports. For example, we might require you to participate in

some individual sports (for example, running, tennis, gymnastics) and team sports (for example, soccer, basketball) that involve different types of physical abilities (for example, strength, endurance, precision). This option would include a more representative sample.

2. All good psychological tests include behavior samples that are obtained under standardized conditions. That is, a test must be administered the same way to all people. When you take a test, various factors can affect your score besides the characteristic, attribute, or trait that is being measured. Factors related to the environment (for example, room temperature, lighting), the examiner (for example, examiner attitude, how the instructions are read), the examinee (for example, illness, fatigue), and the test (for example, understandability of questions) all can affect your score. If everyone is tested under the same conditions (for example, the same environment), we can be more confident that these factors will affect all test takers similarly. If all of these factors affect test takers similarly, we can be more certain that a person's test score accurately reflects the attribute being measured. Although it is possible for test developers to standardize factors related to the environment, the examiner, and the test, it is difficult to standardize examinee factors. For example, test developers have little control over what test takers do the night before they take a test.
3. All good psychological tests have rules for scoring. These rules ensure that all examiners will score the same set of responses in the same way. For example, teachers might award 1 point for each multiple-choice question you answer correctly, and they might award or deduct points based on what you include in your response to an essay question. Teachers might then report your overall exam score either as the number correct or as a percentage of the number correct (the number of correct answers divided by the total number of questions on the test).

Although all psychological tests have these characteristics, not all exhibit these characteristics to the same degree. For example, some tests may include a more representative sample of behaviors than do others. Some tests, such as group-administered tests, may be more conducive to administration under standardized conditions than are individually administered tests. Some tests have well-defined rules for scoring, and other tests have general guidelines. Some tests have very explicit scoring rules, for example, "If Question 1 is marked true, then deduct 2 points." Other tests, such as those that include short answers, may have less explicit rules for scoring, for example, "Award 1 point for each concept noted and defined."

INTERIM SUMMARY 1.2 THE THREE DEFINING CHARACTERISTICS OF PSYCHOLOGICAL TESTS

All psychological tests have three common characteristics:

- First, a good test should representatively sample the behaviors thought to measure an attribute or predict an outcome. This ensures that the test measures what it says it measures.
- Second, the behavior samples should be obtained under standardized conditions.

That is, a test must be administered exactly the same way to all individuals so that we can be confident that a person's score accurately reflects the attribute being measured or the outcome being predicted.

- Third, there must be rules for scoring so that all examiners will score the test in the same way.

Assumptions of Psychological Tests

There are many assumptions that must be made when using psychological tests. The following are what we consider the most important assumptions:

1. *Psychological tests measure what they purport to measure or predict what they are intended to predict.* In addition, any conclusions or inferences that are drawn about the test takers based on their test scores must be appropriate. This is also called test validity. If a test is designed to measure mechanical ability, we must assume that it does indeed measure mechanical ability. If a test is designed to predict performance on the job, then we must assume that it does indeed predict performance. This assumption must come from a personal review of the test's validity data.
2. *An individual's behavior, and therefore test scores, will typically remain stable over time.* This is also called test–retest reliability. If a test is administered at a specific point in time and then we administer it again at a different point in time (for example, 2 weeks later), we must assume, depending on what we are measuring, that an individual will receive a similar score at both points in time. If we are measuring a relatively stable trait, we should be much more concerned about this assumption. However, there are some traits, such as mood, that are not expected to show high test–retest reliability.
3. *Individuals understand test items the same way.* For example, when asked to respond *true* or *false* to a test item such as “I am almost always healthy,” we must assume that all test takers interpret “almost always” similarly.
4. *Individuals will report accurately about themselves* (for example, about their personalities, about their likes and dislikes). When we ask people to remember something or to tell us how they feel about something, we must assume that they will remember accurately and that they have the ability to assess and report accurately on their thoughts and feelings. For example, if we ask you to tell us whether you agree or disagree with the statement “I have always liked cats,” you must remember not only how you feel about cats now but also how you felt about cats previously.
5. *Individuals will report honestly their thoughts and feelings.* Even if people are able to report correctly about themselves, they may choose not to do so. Sometimes people respond how they think the tester wants them to respond, or they lie so that the outcome benefits them. For example, if we ask test takers whether they have ever taken a vacation, they may tell us that they have even if they really have not. Why? Because we expect most individuals to occasionally take vacations, and therefore the test takers think we would expect most individuals to answer *yes* to this question. Criminals may respond to test questions in a way that makes them appear neurotic or psychotic so that they can claim they were insane when they committed crimes. When people report about themselves, we must assume that they will report their thoughts and feelings honestly, or we must build validity checks into the test.
6. *The test score an individual receives is equal to his or her true score plus some error, and this error may be attributable to the test itself, the examiner, the examinee, or the environment.* That is, a test taker's score may reflect not only the attribute being measured but also things such as awkward question wording, errors in administration of the test, examinee fatigue, and the temperature of the room in which the test was taken. When evaluating an individual's score, we must assume that it will include some error.

Although we must accept some of these assumptions at face value, we can increase our confidence in others by following certain steps during test development. For example, in Section III of this

textbook, which covers test construction, we talk about how to design test questions that are more likely to be understood universally. We also talk about the techniques that are available to promote honest answering. In Section II, which covers psychometric principles, we discuss how to measure a test's reliability and validity.

Test Classification Methods

As we have already discussed, there are tens of thousands of commercially available psychological tests, and professionals refer to these tests in various ways. Sometimes professionals refer to them as tests of maximal performance, behavior observation tests, or self-report tests. Sometimes professionals refer to tests as being standardized or nonstandardized, objective or projective. Other times professionals refer to tests based on what the tests measure. In this section, we discuss the most common ways that professionals classify and refer to psychological tests.

Maximal Performance, Behavior Observation, or Self-Report

Most psychological tests can be defined as being tests of maximal performance, behavioral observation tests, or self-report tests.

- **Tests of maximal performance** require test takers to perform a particular well-defined task such as making a right-hand turn, arranging blocks from smallest to largest, tracing a pattern, or completing mathematical problems. Test takers try to do their best because their scores are determined by their success in completing the task. Intelligence tests, tests of specific abilities (for example, mechanical ability), driving tests (road and written), and classroom tests all are good examples of tests of maximal performance.
- **Behavior observation tests** involve observing people's behavior and how people typically respond in a particular context. Unlike with tests of maximal performance, many times people do not know that their behavior is being observed and there is no single defined task for the individual to perform. Many restaurants use this technique to assess food servers' competence in dealing with customers. Sometimes managers hire trained observers to visit their restaurant disguised as a typical customer. In exchange for a free meal or some predetermined compensation, observers agree to record specific behaviors performed by a food server. For example, observers may document whether a food server greeted them in a friendly manner. Other examples of behavior observations include documenting job performance for performance appraisals or clinical interviews.
- **Self-report tests** require test takers to report or describe their feelings, beliefs, opinions, or mental states. Many personality inventories, such as the Hogan Personality Inventory (HPI), are self-report tests. The HPI, a test used primarily for personnel selection and individualized assessment, asks test takers to indicate whether each of more than 200 statements about themselves is true or false.

Most psychological tests fit one of the above categories, and some tests contain features of more than one category. For example, a structured job interview (which involves asking all job applicants a standard set of interview questions) could include both technical questions and questions about one's beliefs or opinions. Technical questions, which are well defined for the interviewee, qualify the interview as a test of maximal performance. Questions about beliefs and opinions qualify it as a self-report

test. The interviewer may also observe the interviewees' behaviors, such as their greetings, which would qualify the interview as a behavioral observation.

Standardized or Nonstandardized

Standardized tests are those that have been administered to a large group of individuals who are similar to the group for whom the test has been designed. For example, if a test is designed to measure the writing ability of high school students, the test would be administered to a large group of high school students. This group is called the **standardization sample**—people who are tested to obtain data to establish a frame of reference for interpreting individual test scores. These data, called **norms**, indicate the average performance of a group and the distribution of scores above and below this average.

For example, if you took the SAT, the interpretation of your score included comparing it with the SAT standardization sample to determine whether your score was high or low in comparison with others and whether you scored above average, average, or below average. In addition, standardized tests always have specific directions for administration and scoring.

Nonstandardized tests do not have standardization samples and are more common than standardized tests. Nonstandardized tests are usually constructed by a teacher or trainer in a less formal manner for a single administration. For example, in many cases, the exams you take in your college courses are nonstandardized tests.

Objective or Projective

Sometimes people make a distinction between objective and projective tests. **Objective tests** are structured and require test takers to respond to structured true/false questions, multiple-choice questions, or rating scales. What the test taker must do is clear, for example, answer *true* or *false*, circle the correct multiple-choice answer, or circle the correct item on the rating scale. The GRE, Stanford-Binet Intelligence Scales, General Aptitude Test Battery, and most classroom tests are examples of objective tests.

Another example of an objective test is the NEO Personality Inventory-3 (NEO-PI-3), an objective self-report instrument designed to identify what makes individuals unique in their thinking, feeling, and interaction with others. Although there are two forms of the inventory, both measure five broad personality dimensions: neuroticism, extroversion, openness, agreeableness, and conscientiousness (PAR, 2012b). Test takers are asked to indicate whether they strongly disagree, disagree, are neutral, agree, or strongly agree with each of 240 statements. These statements are about their thoughts, feelings, and goals. For sample questions from the NEO Personality Inventory, see For Your Information Box 1.2.



More detail about the NEO Personality Inventory can be found in Test Spotlight 1.3 in Appendix A.

On the other hand, **projective tests** are unstructured. They require test takers to respond to unstructured or ambiguous stimuli such as incomplete sentences, inkblots, and abstract pictures. The role of the test taker is less clear than with a standardized test. People who use projective tests believe that test takers project themselves into the task they are asked to perform and that their responses are based on what they believe the stimuli mean and on the feelings they experience while

responding. These tests tend to elicit highly personal concerns. They are often used to detect unconscious thoughts or personality characteristics, and they may be used to identify the need for psychological counseling. The TAT is an example of a projective test. (Chapter 14 contains more information on the TAT and other projective tests.)

FYI

FOR YOUR INFORMATION BOX 1.2

Sample Items From the NEO Personality Inventory

The NEO Personality Inventory is an objective self-report instrument designed to identify what makes individuals unique in their thinking, feeling, and interaction with others. The inventory measures five broad personality dimensions: neuroticism, extroversion, openness, agreeableness, and conscientiousness. Test takers are asked to indicate whether they strongly disagree (SD), disagree (D), are neutral (N), agree (A), or strongly agree (SA) with each of 240 statements. These statements are about their thoughts, feelings, and goals. In the following, we list a sample item from three of the five scales:

Neuroticism

Frightening thoughts sometimes come into my head.	SD	D	N	A	SA
---	----	---	---	---	----

Extroversion

I don't get much pleasure from chatting with people.	SD	D	N	A	SA
--	----	---	---	---	----

Openness

I have a very active imagination.	SD	D	N	A	SA
-----------------------------------	----	---	---	---	----

SOURCE: Reproduced by special permission of the publisher, Psychological Assessment Resources, Inc., 16204 North Florida Avenue, Lutz, Florida 33549, from the NEO Personality Inventory-Revised, by Paul T. Costa Jr., PhD, and Robert R. McCrae, PhD, Copyright 1978, 1985, 1989, 1991, 1992 by Psychological Assessment Resources (PAR). Further reproduction is prohibited without permission from PAR.

Dimension Measured

Psychological tests are often discussed in terms of the dimensions they measure. For example, sometimes we distinguish among achievement tests, aptitude tests, intelligence tests, personality tests, and interest inventories. We refer to these as dimensions because they are broader than a single attribute or trait level. Often these types of tests measure various personal attributes or traits.

Achievement Tests

Achievement tests measure a person's previous learning in a specific academic area (for example, computer programming, German, trigonometry, psychology). A test that requires you to list the three characteristics of psychological tests would be considered an achievement test. Achievement tests are also referred to as tests of knowledge.

Achievement tests are used primarily in educational settings to determine how much students have learned or what they can do at a particular point in time. Many elementary schools and high schools rely on achievement tests to compare what students know at the beginning of the year with what they know at the end of the year, to assign grades, to identify students with special educational needs, and to measure students' progress.

Aptitude Tests

Achievement tests measure a test taker's knowledge in a specific area at a specific point in time. **Aptitude tests** assess a test taker's potential for learning or ability to perform in a new job or situation.

Aptitude tests measure the product of cumulative life experiences—or what one has acquired over time. They help determine what “maximum” can be expected from a person.

Schools, businesses, and government agencies often use aptitude tests to predict how well someone will perform or to estimate the extent to which an individual will profit from a specified course of training. Vocational guidance counseling may involve aptitude testing to help clarify the test taker’s career goals. If a person’s score is similar to scores of others already working in a given occupation, the test will predict success in that field.

Intelligence Tests

Intelligence tests, like aptitude tests, assess the test taker’s ability to cope with the environment, but at a broader level. Intelligence tests are often used to screen individuals for specific programs (for example, gifted programs, honors programs) or programs for the mentally challenged. Intelligence tests are typically used in educational and clinical settings.

Interest Inventories

Interest inventories assess a person’s interests in educational programs for job settings and provide information for making career decisions. Because these tests are often used to predict satisfaction in a particular academic area or employment setting, they are administered primarily to students by counselors in high schools and colleges. Interest inventories are not intended to predict success; rather, they are intended only to offer a framework for narrowing career possibilities.

Personality Tests

Personality tests measure human character or disposition. The first personality tests were designed to assess and predict clinical disorders. These tests remain useful today for determining who needs counseling and who will benefit from treatment programs. Newer personality tests measure “normal” personality traits. For example, the Myers–Briggs Type Indicator (MBTI) is often used by industrial/organizational psychologists to increase employees’ understanding of individual differences and to promote better communication between members of work teams. Career counselors also use the MBTI to help students select majors and careers consistent with their personalities.



More detail about the MBTI can be found in Test Spotlight 1.4 in Appendix A.

Personality tests can be either objective or projective. The MBTI is an example of an objective personality test. Projective personality tests, such as the TAT, serve the same purpose as some objective personality tests, but they require test takers to respond to unstructured or ambiguous stimuli.

Subject Tests

Many popular psychological testing reference books also classify tests by subject. For example, the *Eighteenth Mental Measurements Yearbook* (Buros Institute, 2010) classifies thousands of tests into 19 major subject categories:

- Achievement
- Behavior assessment
- Developmental
- Education

- English
- Fine arts
- Foreign languages
- Intelligence
- Mathematics
- Miscellaneous (for example, courtship and marriage, driving and safety education, etiquette)
- Multiaptitude batteries
- Neuropsychological
- Personality
- Reading
- Science
- Sensorimotor
- Social studies
- Speech and hearing
- Vocations

Reference books such as the *Mental Measurements Yearbook* often indicate whether a test is (a) a test of maximal performance, a behavior observation test, or a self-report test; (b) standardized or nonstandardized; and (c) objective or projective. We discuss the *Mental Measurements Yearbook*, as well as other reference books, later in this chapter.

INTERIM SUMMARY 1.3 ASSUMPTIONS AND TEST CLASSIFICATION METHODS

When using psychological tests, the following assumptions must be made:

- Psychological tests measure what they say they measure, and any inferences that are drawn about test takers based on their test scores are appropriate.
- An individual's behavior, and therefore test scores, will remain unchanged over time.
- Individuals understand test items similarly.
- Individuals can report about themselves accurately.
- Individuals will report their thoughts and feelings honestly.

- The test score an individual receives is equal to his or her true ability plus some error.

Psychological tests can be classified in many different ways:

- As tests of maximal performance, behavior observation tests, or self-report tests
- As standardized or nonstandardized
- As objective or projective
- Based on the dimensions they measure
- Based on subject

Psychological Assessment, Psychological Tests, Measurements, and Surveys

Before discussing much more, we should spend some time discussing some terms that students often confuse—psychological assessment, psychological tests, measurement, and surveys. Students often think of psychological assessment and psychological testing as one and the same. Similarly, students often do

not understand the difference between psychological tests and surveys. This section is designed to help you distinguish among these terms that are commonly used in psychological testing.

Psychological Assessments and Psychological Tests

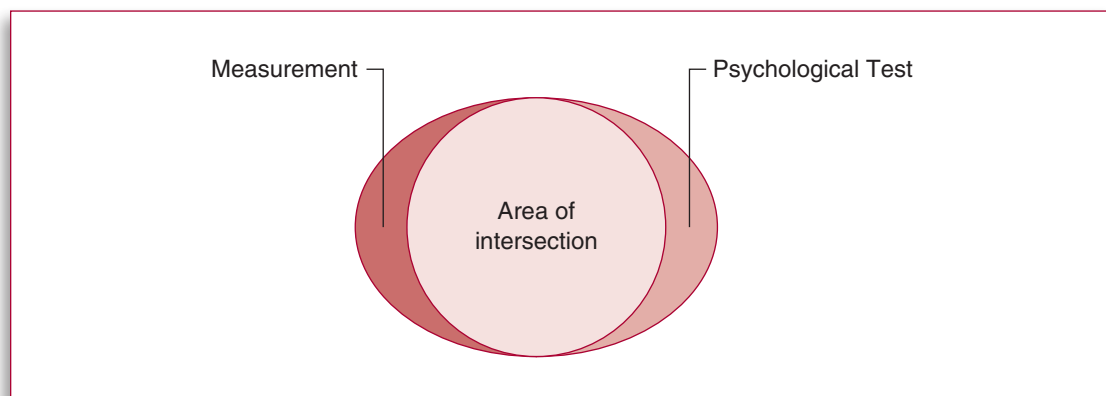
Psychological assessments and psychological tests both are methods of collecting important information about people, and both are also used to help understand and predict behavior (Kline, 2000; Maloney & Ward, 1976). Assessment, however, is a broader concept than psychological testing. **Psychological assessment** involves multiple methods, such as personal history interviews, behavioral observations, and psychological tests, for gathering information about an individual. Psychological assessment involves *both* an objective component and a subjective component (Matarazzo, 1990), and psychological tests are only one tool in the assessment process. For example, a clinical psychologist may conduct a psychological assessment of a patient and, as a part of this assessment, may administer a psychological test such as the Minnesota Multiphasic Personality Inventory (MMPI).

Psychological Tests and Measurements

Although the meanings overlap, *psychological test* and *measurement* are not synonyms. **Measurement**, broadly defined, is the assignment of numbers according to specific rules. The concept of measurement is represented by the darker circle in Figure 1.2.

Psychological tests require test takers to answer questions or perform tasks to measure personal attributes. The concept of a psychological test is represented by the lighter circle in the figure. With psychological tests, test takers' answers to questions or their performance on some task is not initially expressed in physical units of any kind; instead, scores are derived according to some predetermined method. In some cases, the end result of a psychological test is not a derived score at all, but rather a verbal description of an individual. For example, there are some personality tests that, although they

Figure 1.2 A Comparison of Measurement and Psychological Testing (the area of intersection represents samples of behavior expressed as numerical scores)



have rules for scoring or summarizing information, do not produce overall scores. Instead, these tests yield profiles. The MBTI is an example of such a test.

Psychological tests can be considered psychological measurements when a sample of behavior can be expressed as a numerical score. This is represented by the overlapping section of the two circles in Figure 1.2.

You will find that many people use the terms *psychological test* and *psychological measurement* interchangeably. Although most psychological tests are measurements, not all psychological tests, strictly defined, meet the definition of a measurement. Throughout the remainder of this text, we follow the common practice of referring to all psychological tests as measurements because most of them are, but keep in mind the distinctions we have drawn in this section.

Psychological Tests and Surveys

Surveys, like psychological tests (and psychological assessments), are used to collect important information from individuals. Surveys differ from psychological tests in two important ways. First, psychological tests focus on individual outcomes, and surveys focus on group outcomes. Psychological tests provide important information about individual differences and help individuals and institutions make important decisions about individuals. For example, a psychological test may suggest that a child is unusually intelligent and therefore should be placed in a gifted or honors program. Surveys, on the other hand, provide important information about groups and help us make important decisions about groups. For example, an organizational survey may suggest that employees are displeased with a company benefits program and that a new benefits program is needed.

Second, the results of a psychological test are often reported in terms of an overall derived score or scaled scores. Results of surveys, on the other hand, are often reported at the question level by providing the percentage of respondents who selected each answer alternative. Of course, in some cases, surveys focus on individual outcomes and are constructed using scales. In such cases, the survey approximates a psychological test. (Chapter 10 is devoted to an in-depth discussion of surveys.)

Locating Information About Tests

With so many psychological tests available, we are sure you can imagine that finding the most appropriate one for your specific purpose can be a difficult task. To choose an appropriate test for a particular circumstance, you must know the types of tests that are available and their merits and limitations. Prior to the 1950s, test users had few resources for obtaining such information. Today, however, numerous resources are available. Although all have the same general purpose—to help test users make informed decisions—the information such resources contain varies. Some resources provide only general descriptive information about psychological tests, such as the test's name, author, and publisher, and others contain detailed information, including test reviews and detailed bibliographies. Some resources focus on commercially available, standardized published tests, and others focus on unpublished tests. Some references include information about tests for particular groups (for example, children), and others include a broad range of tests for various populations.

Some of the most commonly used resource books, including a brief synopsis of the contents, are described in For Your Information Box 1.3. The first four resource books, *Tests in Print (TIP)*, the *Mental Measurements Yearbook (MMY)*, *Tests*, and *Test Critiques*, are often viewed as the most useful and popular (APA, 2012a).

FYI


FOR YOUR INFORMATION BOX 1.3

Commonly Used Resource Books

<i>Book Title</i>	<i>Contents</i>
<i>Tests in Print</i> (multiple volumes)	<i>Tests in Print (TIP)</i> is published in multiple volumes. Each volume contains descriptive listings of commercially published tests that are available for purchase. <i>TIP</i> also serves as a comprehensive index to the contents of previously published <i>Mental Measurements Yearbooks</i> . Each descriptive listing, or test entry, contains extensive information, including but not limited to the title of the test, the purpose of the test, the intended population, publication dates, the acronym used to identify the test, scores the test provides, whether the test is an individual test or group test, whether the test has a manual, the author(s), the publisher, the cost of the test, and available foreign adaptations. Each entry also contains brief comments about the test as well as cross-references to reviews in the <i>Mental Measurements Yearbooks</i> .
<i>Mental Measurements Yearbook</i> (multiple volumes)	The <i>Mental Measurements Yearbook (MMY)</i> is published in multiple volumes. Each volume contains descriptive information and test reviews of new English-language, commercially published tests and tests that have been revised since the publication of the previous <i>MMY</i> edition. The <i>MMY</i> is cumulative, meaning that later volumes build on earlier ones rather than replace them. Each descriptive listing, or test entry, contains extensive information about a particular test. If the test is a revision of a previous test, the entry also includes the volume of the <i>MMY</i> in which the test was originally described. Each entry also typically includes information about the test's reliability and validity, one or two professional reviews, and a list of references to pertinent literature. For a guide to descriptive entries in the <i>MMY</i> , see Figure 1.3. The <i>MMY</i> is very likely accessible electronically through your college's library system.
<i>Tests</i>	<i>Tests</i> contains descriptions of a broad range of tests for use by psychologists, educators, and human resource professionals. Each entry includes the test title, the author, the publisher, the intended population, the test purpose, major features, the administration time, the cost, and the availability.
<i>Test Critiques</i> (multiple volumes)	<i>Test Critiques</i> is published in multiple volumes. Each volume contains reviews of frequently used psychological, business, and educational tests. Each review includes descriptive information about the test (for example, author, attribute measured, norms) and information on practical applications and uses. <i>Test Critiques</i> also contains in-depth information on reliability, validity, and test construction.
<i>Personality Test and Reviews</i> (multiple volumes)	<i>Personality Test and Reviews</i> is published in volumes. Each volume contains a bibliography of personality tests that are contained in the <i>MMY</i> . Each entry contains descriptive information about the test as well as test reviews.
<i>Tests in Education</i>	<i>Tests in Education</i> contains descriptive and detailed information about educational tests for use by teachers, administrators, and educational advisers.
<i>Testing Children</i>	<i>Testing Children</i> contains descriptions of tests available for children. These descriptions include the knowledge, skills, and abilities measured by each test; the content and structure of the test; the time required to administer the test; the scores that are produced; the cost; and the publisher.
<i>Tests and Measurements in Child Development: A Handbook</i>	<i>Tests and Measurements in Child Development</i> contains a listing of unpublished measures for use with children as well as detailed information about each measure.
<i>Measures for Psychological Assessment: A Guide to 3,000 Original Sources and Their Applications</i>	<i>Measures for Psychological Assessment</i> is a guide that contains annotated references to thousands of less recognized assessment devices developed and described in journal articles.

Whether you are trying to locate tests that measure intelligence, self-esteem, or some other attribute, trait, or characteristic, we suggest that you begin your search with one of the first four resource books in For Your Information Box 1.3. *TIP* and the *MMY* are two of the most helpful references, and students often find it most helpful to begin with *TIP*. Figure 1.3 includes a descriptive guide of the type of information you will find in the *MMY*. Figure 1.4 includes a summary of how to use *TIP* to find tests. You can find more information on how to use both of these resources, as well as how to use the information contained in these resources to evaluate a test, on the Buros Institute of Mental Measurements homepage, discussed in On the Web Box 1.2.

Figure 1.3 A Guide to Descriptive Entries in the *Mental Measurements Yearbook*



**BUROS
INSTITUTE
OF MENTAL MEASUREMENTS**

Mental Measurements Yearbook and Tests in Print
A Guide to the Descriptive Entries

Entry Number: The number cited in all indexes when referring to this test.

Title: Test titles are printed in boldface type; secondary or series titles are set off from main titles by colon.

Population: A description of the groups for which the test is intended.

Administration: Individual or group administration is indicated.

Distribution: This is noted only for tests that are put on a special market by the publisher.

Special Editions: Various types of special editions are listed here.

Author: All test authors' names are reported, exactly as printed on the test materials.

Cross References: For tests that have been previously listed in a Buros publication, cross references to the reviews, excerpts, and references will be noted here. "9:1410," for example, refers to test 1410 in the *Ninth Mental Measurements Yearbook*; "T4:3010" refers to test 3010 in *Tests in Print IV*.

[420]
The Hypothetical Test: Reading.
Purpose: Designed to "measure achievement in reading."
Population: Grades 9-12.
Publication Dates: 1989—1994.
Acronym: HYPE.
Scores, 3: Vocabulary, Comprehension, and Total.
Administration: Individual or group.
Forms, 3: Survey, Abbreviated, Complete Battery.
Restricted Distribution: Distribution of Survey Form restricted to school principals.
Price Data, 1995: \$70 per complete kit including 100 tests, scoring key, and manual ('94, 120 pages); \$9 per scoring key; \$32 per manual.
Special Editions: Braille edition available.
Time: 50(60) minutes.
Comments: May be self-scored.
Author: Jane J. Doe.
Publisher: Hypothetical Tests, Inc.
Cross References: See T4:3010 (2 references); for reviews by John Roe and Robert Smith of an earlier edition, see 9:1410 (6 references).

Purpose: A brief, clear statement describing the purpose of the test; often these are quotations from the test manual.

Publication Date: The inclusive range of publication dates.

Acronym: Acronym by which the test may be commonly known.

Scores: The number of explicit scores is presented along with the descriptions of what they are intended to measure.

Forms: All available forms, parts, and levels are listed.

Price Data: Price information is reported for test packages, answer sheets, accessories, and specimen sets.

Time: This is the amount of time to take, and administer, the test. The first number is the actual working time examinees are allowed, and the second (parenthesized) number is the total time needed to administer the test.

Comments: Special notations and comments.

Publisher: The publisher's full address can be found in the Publishers Directory and Index.


Buros Institute of Mental Measurements
University of Nebraska-Lincoln
21 Teachers College Hall
Lincoln, NE 68588-0348

This is not copyrighted material. Reproduction and dissemination are encouraged.
See the Buros Institute website at www.unl.edu/buros

SOURCE: Buros Institute of Mental Measurements, University of Nebraska-Lincoln. www.unl.edu/buros.

Because there is a wealth of psychological tests available, there is a wealth of resources available for you to use in gathering information about psychological tests. You are not limited to print resources; advances in technology now allow you to access the Internet and gather information about psychological tests on demand. On the Web Box 1.2 discusses some websites you can access to locate information on psychological tests. For Your Information Box 1.4 discusses where you can locate unpublished psychological tests.

Figure 1.4 How to Use *Tests in Print*



How to Use *Tests in Print*

Tests in Print (TIP) consists of descriptive listings, without reviews, of commercially published tests in print. TIP is also a comprehensive index to the contents of previously published Mental Measurements Yearbooks.

1. If you know the TEST TITLE:

Use the “**Index of Titles.**” The index lists all tests in that volume plus all tests out of print since last being listed. “2458,” for example, refers to test 2458 in that volume; “9:1128” refers to now out-of-print test 1128 in the *Ninth Mental Measurements Yearbook*. *Citation numbers refer to entry numbers, not to page numbers.*

Example from “Index of Titles”:

Short Tests of Clerical Ability, 2458
Shortened Edinburgh Reading Test, 2459
Shortened Aptitude Test, T4:2195
Signals Learning Test, 2461
Silver Burdett Music Competency Tests, 9:1128
Silver Drawing Test of Cognitive Skills and Adjustment, 2462
Simile Interpretations, T4:2198
Similes, T4:2199

2. If you know the TYPE OF TEST:

Use the “**Classified Subject Index**” to locate various categories of tests, such as achievement, intelligence, personality, etc. This index organizes all tests into 18 major categories; tests appear alphabetically within each category. *Citation numbers refer to entry numbers, not to page numbers.*

Example from “Classified Subject Index, Education”:

Gifted Program Evaluation Survey, Gifted and talented programs, see 1040
Graduate Records Examinations Education Test, Graduate School candidates, see 1063
High School Characteristics Index, Grades 9-13, 4-13, see 1157
How a Child Learns, Classroom teachers, see 1175
Hudson Educational Skills Inventory, Grades K-12, see 1184

3. If you know the NAME OF THE TEST AUTHOR OR REVIEWER:

Use the “**Index of Names.**” This index includes test authors (for example, “test, 1460”), review authors (“rev, 2589”), and authors of referenced articles (“ref, 2222”). (Parenthesized numbers indicate the reference number.) *Citation numbers refer to entry numbers, not to page numbers.*

Example from “Index of Names”:

Caeglio, G.: test, 1460
Caffey, C. A.: ref, 2222(1)
Caggiola, A. A.: ref, 2563(449)
Cahalane, J.: ref, 268(395), 1043(39)
Cahen, L. S.: rev, 2589
Cahill, C.: ref, 1705(65), 2937(935)
Cahir, N.: ref, 1135(14), 2674(188)
Cahn, T. S.: ref, 268(90)
Cain, J.: ref, 93(84), 1690(84)
Cain, L. F.: test, 2844

Buros Institute of Mental Measurements
University of Nebraska-Lincoln
21 Teachers College Hall
Lincoln, NE 68588-0348

This is not copyrighted material. Reproduction and dissemination are encouraged.
See the Buros Institute website at www.unl.edu/buros

SOURCE: Buros Institute of Mental Measurements, University of Nebraska-Lincoln. www.unl.edu/buros.

ON THE WEB BOX 1.2

Locating Information About Tests on the Web



Computer technology lets us connect to the Internet and locate websites containing valuable information about psychological tests. These websites include information such as the following:

- Frequently asked questions about psychological testing
- How to find a particular type of psychological test
- How to locate reviews of psychological tests
- How to select an appropriate test
- What qualifications are necessary to purchase psychological tests
- How to contact test publishers
- How to obtain copies of specific psychological tests

Although there are many available websites, here are four that we have found to be extremely valuable:

Website	Description
American Psychological Association www.apa.org/science/programs/testing/find-tests.aspx#findinfo	Although the American Psychological Association (APA) does not sell or endorse specific testing instruments, it does provide guidance on testing resources and how to find psychological tests. This website contains answers to the most frequently asked questions about psychological testing. One section focuses on questions about published psychological tests (those that can be purchased from a test publisher); here you will find advice on how to find information about a particular test and about the proper use of tests, how to contact test publishers and purchase tests, and available software and scoring services. Another section focuses on unpublished psychological tests and measures (those that are not commercially available); here you will find advice on how to find unpublished tests in your area of interest and important information regarding your responsibilities as a user of unpublished tests.
Buros Institute of Mental Measurement www.unl.edu/buros/bimm/index.html	The Buros Institute of Mental Measurement promotes the appropriate use of tests and provides professional assistance, expertise, and information to those who use commercially published tests. This website contains a number of instructional resources, tools, and links. For example, it contains detailed instructions on what information can be found in two popular Buros publications that we have already discussed: the <i>Mental Measurements Yearbook</i> and <i>Tests in Print</i> . This site also contains some great “how to” resources such as how to use <i>Tests in Print</i> and the <i>Mental Measurements Yearbook</i> and how to use the information in these resources to evaluate a test. In addition, it contains a link to <i>Test Reviews Online</i> , a service that provides access to more than 2,000 test reviews, beginning with those that were published in the <i>Ninth Mental Measurements Yearbook</i> . Likewise, there are links to the <i>Code of Fair Testing Practices</i> (discussed further in Chapter 3) and the APA’s frequently asked questions website mentioned previously.
Test Collection at ETS www.ets.org/test_link/about	The Test Collection at ETS is the world’s largest database of tests and measurement instruments that have been available since the early 1900s. This online database contains descriptions of more than 20,000 tests (published and unpublished) and research instruments, collected from test publishers and test authors from around the world. Each description includes the title of the test/instrument, the author, the publication date, availability (how to obtain the test or measurement), the intended population, and specific uses of the test/instrument. In addition to providing information about specific tests, this database contains valuable information on how to order tests.
O*Net Resource Center www.onetcenter.org/guides.html	The Occupational Information Network (O*Net) is sponsored by the U.S. Department of Labor and is a primary source for occupational information. Consisting of a comprehensive database of worker attributes and job characteristics, O*Net also provides valuable resources on testing and assessment—resources intended to support public and private sector efforts to identify and develop the skills of the American workforce. This website provides access to three extremely valuable testing and assessment guides: <ul style="list-style-type: none"> • <i>Testing and Assessment: A Guide to Good Practices for Workforce Investment Professionals</i> includes information on how assessment instruments can be used to promote talent development in career counseling, training, and other talent development activities. It discusses how to evaluate and select assessment instruments, administer and score assessments to meet business and individual client needs, and accurately and effectively interpret assessment results. It also lists the professional and legal standards related to assessment use in talent development. • <i>Tests and Other Assessments: Helping You Make Better Career Decisions</i> includes an explanation of how assessment instruments are used in employment selection and career counseling and provides tips and strategies for taking tests and other assessments. • <i>Testing and Assessment: An Employer’s Guide to Good Practices</i> helps managers and workforce development professionals understand and use employment testing and assessment practices to meet their organizations’ human resources goals.

F Y I

FOR YOUR INFORMATION BOX 1.4

Locating Unpublished Psychological Tests

Although there are thousands of commercially available tests, there are just as many, if not more, unpublished tests designed and used by researchers. A number of print and nonprint resources are available for locating information on unpublished tests.

Two of the most popular print resources are the *Directory of Unpublished Experimental Mental Measures* and *Measures for Psychological Assessment: A Guide to 3,000 Original Sources and Their Applications*. Three of the most popular nonprint resources for locating information about unpublished or noncommercial tests are Tests in Microfiche, the PsycINFO database, and the Health and Psychosocial Instruments database.

Directory of Unpublished Experimental Mental Measures (Goldman & Mitchell, 2007)

This directory provides easy access to more than 5,000 experimental mental measures, tests, and surveys that have been used by other researchers but are not commercially available. Topics range from educational adjustment and motivation to personality and perception. The measures, tests, and surveys are arranged in a 24-category system and grouped according to function and content, noting purpose, format, psychometric information (where available), and related research. First published in 1974 and currently in its ninth edition, this resource is updated periodically by the publisher.

Measures for Psychological Assessment: A Guide to 3,000 Original Sources and Their Applications (Chun, Cobb, & French, 1975)

This guide includes annotated references to psychological measures that have appeared in journal articles and other publications. Although a bit outdated, this can be a useful resource. It has two sections: primary references and applications. The primary references section includes the name of each measure, the reference in which the measure originally appeared, and one or more other researchers who have used the measure in experimental research. The applications section includes other research studies that have used the original measures and references other experimental tests.

Tests in Microfiche

This resource can be accessed through the Test Collection at ETS. It contains a variety of educational and psychological instruments that are cited in the literature but are either out of date or unpublished. It contains more than 800 tests, and new tests are added each year. For more information, go to www.ets.org/test_link/about or check with your college's library.

PsycINFO Database

This bibliographic database indexes published studies in psychology. By using the Form/Content field "Tests & Measures" to search the PsycINFO database, you can find tests that have been used in research and written about in the literature. For more information, go to www.apa.org/pubs/databases/psycinfo/index.aspx.

Health and Psychosocial Instruments Database (HAPI)

This computerized database includes citations to unpublished health and psychosocial evaluation and measurement tools (for example, questionnaires, interviews, tests, checklists, rating scales) that have appeared in journals and technical reports since 1985. HAPI is updated quarterly and contains more than 15,000 measurement instruments. HAPI is provided online by Ovid Technologies, which typically must be accessed through BRS Information Technologies at your college's library. Some libraries maintain the database on CD-ROM. For more information, see www.ovid.com/site/catalog/DataBase/866.jsp.

Another valuable resource for locating psychological tests is APA's new PsycTESTS database. For more information, see In the News Box 1.2

IN THE NEWS

Box 1.2 PsycTESTS Database

In September 2011, the APA issued the following news release:

The American Psychological Association (APA) announces the release of its new database, PsycTESTS. This new resource joins APA's suite of renowned research databases, providing a simple yet comprehensive solution for libraries and their patrons looking for access to test content.



PsycTESTS will be updated monthly and includes over 2,200 indexed test records, with more than 1,500 ready-to-use instruments. PsycTESTS is a research database that provides access to an abundance of full-text psychological tests, measures, scales, and other assessments as well as descriptive and administrative information for each. PsycTESTS spans the full breadth of assessment in psychological research including neuropsychological tests, personality assessments, skills and proficiency tests, and IQ and aptitude scales and measurements. The majority of the test records found in PsycTESTS are research instruments which are available for immediate use. The database also covers select commercially published tests, and provides direct links to their respective publisher sites. All test records in PsycTESTS are created by skilled indexers with deep knowledge of the psychology literature.

"PsycTESTS is an ideal resource for students, librarians, researchers, and practitioners," remarks Gary R. VandenBos, PhD, APA Publisher. "It goes a long way towards facilitating research and practice in the field of psychology and behavioral science."

Available tests include those compiled from a systematic review of peer-reviewed journals published by APA, Hogrefe Publishing Group, and other large publishers; technical reports from PsycEXTRA; contributions from individual test authors; collections from the Archives of the History of American Psychology; and more. Tests are available as printable PDFs, or as images, audio, or software for easy download and display. Records include information on the scope of the test, test implementation, a high-level overview of the test's development, and reliability and validity data.

"PsycTESTS fills an immediate need for a variety of full-text psychological tests, measures and assessments to be available and easily accessible. It will make finding relevant tests easier and more efficient," stated Linda Beebe, the Senior Director of PsycINFO.

PsycTESTS will be available through APA PsycNET as well as through several of APA's current vendor partners.

You can find more information on the PsycTESTS database at www.apa.org/pubs/databases/psyctests/index.aspx.

SOURCE: American Psychological Association (2011).

Chapter Summary

By now, we hope you understand that psychological testing extends well beyond the use of intelligence and personality tests. Anything that requires a test taker to perform a behavior that is used to measure some personal attribute, trait, or characteristic or to predict an outcome can be considered a psychological test. The quizzes and exams you take in class are psychological tests. The written and

road portions of driving exams are psychological tests. Even the structured job interviews you have participated in, or will participate in as you conduct your job search, qualify as psychological tests.

Psychological tests have various similarities and many differences. All psychological tests require an individual to perform one or more behaviors, and these behaviors are used to measure some personal attribute, trait, or characteristic thought to be important in describing or understanding behavior or to predict an outcome. However, psychological tests can and do differ in terms of the behaviors they require individuals to perform, the attributes they measure, their content, how they are administered and formatted, how they are scored and interpreted, and their psychometric quality.

Although the use of psychological tests can be traced to ancient China, most scholars agree that the advent of formal psychological testing did not begin until Binet published the first test of intelligence in 1905. Today, psychological testing is a big business, with tens of thousands of commercially available, standardized psychological tests as well as thousands of unpublished tests.

All good tests have three defining characteristics in common. First, they include a representative sample of behaviors. Second, they collect the sample under standardized conditions. Third, they have rules for scoring. When using psychological tests, we must make some assumptions. We must assume that a test measures what it says it measures, that any inferences that are drawn about test takers from their scores on the test are appropriate, that an individual's behavior (and therefore test scores) will remain stable over time, that individuals understand test items similarly, that individuals can and will report accurately about their thoughts and feelings, and that the test score an individual receives is equal to his or her true behavior/ability in the real world plus some error.

Testing professionals refer to psychological tests in various ways. Sometimes they refer to them as tests of maximal performance, behavior observations, or self-reports. Sometimes they refer to them as standardized or nonstandardized. Other times they refer to them as objective or projective. Professionals also refer to tests based on the dimensions they measure.

It is important to remember the distinctions among four commonly misunderstood terms: *psychological assessment*, *psychological test*, *measurement*, and *survey*. First, although both psychological assessments and psychological tests are used to gather information, a psychological test is only one of many tools in the psychological assessment process. Second, a psychological test can be considered to be a measurement when the sampled behavior can be expressed in a derived score. Third, psychological tests are different from surveys in that psychological tests focus on individual differences and often report one overall derived score (or scaled scores), and surveys focus on group similarities and typically report results at the question or item level.

Last, but not least, a number of resources are available, in print and online, to locate information about published and unpublished psychological tests and measures. The *Mental Measurements Yearbook* and *Tests in Print* are two of the most popular references for learning more about available tests.

Engaging in the Learning Process

KEY CONCEPTS

After completing your study of this chapter, you should be able to define each of the following terms. These terms are bolded in the text of this chapter and defined in the Glossary.

- achievement tests
- aptitude tests
- behavior
- behavior observation tests

- emotional intelligence
- intelligence tests
- interest inventories
- measurement
- nonstandardized tests
- norms
- objective tests
- personality tests
- projective tests
- psychological assessments
- psychological tests
- psychometrics
- self-report tests
- standardization sample
- standardized tests
- surveys
- tests of maximal performance
- vocational tests

LEARNING ACTIVITIES

The following are some learning activities you can engage in to support the learning objectives for this chapter.

<i>Learning Objectives</i>	<i>Study Tips and Learning Activities</i>
<i>After completing your study of this chapter, you should be able to do the following:</i>	<i>The following study tips will help you meet these learning objectives:</i>
Define what a psychological test is, and understand that psychological tests extend beyond personality and intelligence tests.	<ul style="list-style-type: none"> • Write your definition of a psychological test. List examples of psychological tests, from what comes to your mind first to what comes to your mind last. Compare your list of examples with Figure 1.1. • Ask various professionals, in and outside of the psychology field, to define what a psychological test is. Compare and contrast their definitions. Compare these definitions with the definitions provided in this textbook. Discuss why definitions might vary.
Trace the history of psychological testing from Alfred Binet and intelligence testing to the tests of today.	<ul style="list-style-type: none"> • Reflect on the history of testing. Create a timeline showing significant events in testing, beginning with testing in ancient China and ending with testing today.
Describe the ways in which psychological tests can be similar to and different from one another.	<ul style="list-style-type: none"> • Think about two exams you recently took. Make two lists: one of how they were similar and another of how they were different. Compare your lists with Interim Summary 1.1.
Describe the three characteristics that are common to all psychological tests, and understand that psychological tests can demonstrate these characteristics to various degrees.	<ul style="list-style-type: none"> • Recall the three characteristics common to all psychological tests. Make three columns, and label them Representative Sample of Behaviors, Standardized Conditions, and Rules for Scoring. Select one or two psychological tests that you have taken. Write how the test(s) demonstrates each characteristic. • Construct an eight-question quiz, with one question for each learning objective. Give the quiz to your classmates (your professor will determine the logistics of this). As a class, discuss whether the quiz meets all of the characteristics of a psychological test. What were the strengths of your quiz? How could your quiz have been improved?

(Continued)

(Continued)

<i>Learning Objectives</i>	<i>Study Tips and Learning Activities</i>
Describe the assumptions that must be made when using psychological tests.	<ul style="list-style-type: none"> Describe the six assumptions we must make when using psychological tests. Without looking in your book, see how many assumptions you can write. Compare your written assumptions with the assumptions in the book. Explain why we must make these assumptions.
Describe the different ways that psychological tests can be classified.	<ul style="list-style-type: none"> Review the test classification methods in your book. Think about the road portion of the driving test, the SAT, a job interview, the NEO Personality Inventory, and a multiple-choice test you took recently. Classify each test using the different test classification methods.
Describe the differences among four commonly used terms that students often get confused: psychological assessment, psychological tests, psychological measurement, and surveys.	<ul style="list-style-type: none"> Draw a picture or diagram illustrating how these four commonly confused terms overlap.
Identify and locate printed and online resources that are available for locating information about psychological tests.	<ul style="list-style-type: none"> Go to your college library and find <i>Tests in Print</i> and the <i>Mental Measurements Yearbook</i>. Write the names of three tests and what they measure. Go to each of the websites referenced in your book. Compare and contrast the information found on these websites. Select a psychological test that is mentioned in Chapter 1 or 2 or that is suggested by your instructor. Using reference books available at your college library and online, collect as much of the information as possible about your test. Keep track of where you found the information.

PRACTICE QUESTIONS

The following are some practice questions to assess your understanding of the material presented in this chapter.

Multiple Choice

Choose the one best answer to each question.

- What do all psychological tests require that you do?
 - Answer questions
 - Fill out a form
 - Perform a behavior
 - Sign a consent form
 - Personality tests
 - Intelligence tests
 - Structured interviews
 - Classroom tests
- According to the textbook, which one of the following is least typical of psychological tests?
 - Lewis Binet
 - Alfred Simon
- Who published the first test of intelligence in 1905?
 - Lewis Binet
 - Alfred Simon

- c. Robert Woodworth
 - d. Alfred Binet
4. Who published the Stanford–Binet?
 - a. Henry Murray
 - b. Robert Woodworth
 - c. Lewis Terman
 - d. Alfred Binet
 5. What test did Robert Woodworth develop during World War I to help the U.S. military detect soldiers who would not be able to handle the stress associated with combat?
 - a. Thematic Apperception Test
 - b. Stanford–Binet
 - c. Personal Data Sheet
 - d. Rorschach Inkblot Test
 6. What was the first widely used personality inventory?
 - a. Woodworth Psychoneurotic Inventory
 - b. Personal Data Sheet
 - c. Rorschach Inkblot Test
 - d. Thematic Apperception Test
 7. A test that requires you to demonstrate your driving ability can best be classified as what type of test?
 - a. Test of maximal performance
 - b. Self-report test
 - c. Behavior observation test
 - d. Projective test
 8. A test that requires you to respond to test questions about your feelings and beliefs can best be described as what type of test?
 - a. Test of maximal performance
 - b. Self-report test
 - c. Behavior observation test
 - d. Projective test
 9. The role of the test taker is least clear in which type of test?
 - a. Objective test
 - b. Projective test
 - c. Standardized test
 - d. Self-report test
 10. What type of test is administered to a large group of individuals who are similar to the group for which the test has been designed?
 - a. Nonstandardized test
 - b. Standardized test
 - c. Objective test
 - d. Subjective test
 11. What type of test would a classroom teacher most likely administer?
 - a. Achievement test
 - b. Aptitude test
 - c. Intelligence test
 - d. Interest inventory
 12. If Jose took a test to identify his potential for learning or his ability to perform in an area in which he had not been specifically trained, what type of test would he be taking?
 - a. Achievement test
 - b. Intelligence test
 - c. Aptitude test
 - d. Vocational test
 13. Joe took three tests. One required him to respond to true/false questions, one to multiple-choice questions, and one to rating scales. What type of tests did he take?
 - a. Projective tests
 - b. Nonstandardized tests
 - c. Subjective tests
 - d. Objective tests
 14. What type of test would a career development counselor most likely administer?
 - a. Achievement test
 - b. Aptitude test
 - c. Intelligence test
 - d. Interest inventory
 15. If you wanted to locate a professional test review for a commercially available published test, which one of the following would be the best source?
 - a. Tests in Print
 - b. Tests in Microfiche
 - c. Mental Measurements Yearbook
 - d. Measures for Psychological Assessment

Short Answer/Essay

Read each of the following, and consider your response carefully based on the information presented in this chapter. Write your answer to each question in two or three paragraphs.

1. What is a psychological test?
2. Why is it important for you to understand foundations of psychological testing?
3. Identify three different psychological tests. Discuss how they are similar and how they are different.
5. When using a psychological test, what assumptions must be made? Why are these assumptions important?
6. What are the similarities and differences among intelligence tests, aptitude tests, and achievement tests? Provide an example of each.
7. Identify a psychological assessment, psychological test, and measurement. Discuss how they are similar and different.
8. How are psychological tests and surveys similar? How are psychological tests and surveys different?

ANSWER KEYS*Multiple Choice*

1. c	2. c	3. d	4. c	5. c
6. a	7. a	8. b	9. b	10. b
11. a	12. c	13. d	14. d	15. c

Short Answer/Essay

Refer to your textbook for answers. If you are unsure of an answer and cannot generate the answer after reviewing your book, ask your professor for clarification.

CHAPTER 2

Why Is Psychological Testing Important?

CHAPTER 2: WHY IS PSYCHOLOGICAL TESTING IMPORTANT?

After completing your study of this chapter, you should be able to do the following:

- Describe different types of decisions that are made using the results of psychological tests.
- Explain which professionals use psychological tests, in what settings, and for what reasons.
- Describe some concerns individuals have regarding the use of psychological tests as well as the social and legal implications of psychological testing, especially as they relate to intelligence, achievement, aptitude, and integrity testing.

“I have a 4.0 grade point average, but I didn’t do so well on the LSAT. I’m going to apply to some top-rated law schools and, to be safe, to some law schools where LSAT scores aren’t as important.”

“The college I will be attending decided not to give me the scholarship I wanted. Because all the applicants had similar grade point averages and great letters of recommendation, the scholarship review committee decided to give the scholarship to the applicant who had the highest SAT score.”

“I told my academic counselor that I did not know what I wanted to do with my life. She sent me to the college career services office. A career counselor talked with me for a long time, asking me about my likes and dislikes and my hobbies. She had me take several interest and vocational tests, and she used test scores to help me focus on what I should major in.”

You have probably had to make some important decisions in your life, such as where to apply to college and what to major in. Likewise, others have probably made important decisions about you. For example, colleges may have decided to admit you, scholarship committees may have decided not to offer you a scholarship, organizations may have decided to hire you, or school psychologists may have decided to place you in a gifted program. There is a good chance you or others used your score on one or more psychological tests to make many of these decisions.

In this chapter, we discuss why psychological testing is important. We look at who uses psychological tests and for what reasons. We also discuss some of the concerns society has about the use of psychological tests.

The Importance of Psychological Testing

As you learned in Chapter 1, psychological testing is important because people use test results to make important decisions such as those just mentioned. These decisions affect every one of us. Consider some of the decisions we (the textbook authors) have made in our roles as college faculty, industrial/organizational practitioners, and organizational leaders using test scores:

- What grade a student has earned
- Whether to consider a job candidate for employment
- Whether to hire a job candidate
- Whether an employee will receive a merit increase (and if so, how much)
- What coaching advice to offer a business leader

Likewise, consider some of the decisions that have been made by others, about us or our families, based on the results of psychological tests:

- Whether we, or our children, will be admitted to a specific college
- Whether our children will be invited to participate in the elementary school gifted program
- Whether our children will receive a Florida high school diploma or a certificate of completion
- Whether our children will receive a college scholarship

We and others make many types of decisions using the results of psychological tests. These decisions are often classified as individual versus institutional and comparative versus absolute.

Individual and Institutional Decisions

Both individuals and institutions use the results of psychological tests to make decisions. If test takers use their test scores to make decisions about themselves, these are referred to as **individual decisions**. For example, in the future you may take the Law School Admissions Test (LSAT), a half-day standardized test required for admission to most law schools. Knowing that some law schools are more competitive than others, the score that you receive on this test might influence the law schools to which

you apply. If you do very well on the test, you may apply to more competitive law schools. Or perhaps you are having a difficult time deciding what career you would like to pursue. You might seek assistance from a career counselor to explore and discuss various career options. As part of the process, the career counselor may ask you to complete an interest inventory. Based on the results of this inventory (as well as other information), you may decide to pursue a career in, for example, teaching or computer science. In this case, you (the individual who took the test) used the test results to make a decision about your career.

Institutional decisions, on the other hand, are those made by another entity (for example, a company, an organization, an institution) about an individual based on his or her test results. For example, let us say that because you did well on your LSAT, you decided to apply to a highly competitive law school. Administrators at the law school to which you apply will use your LSAT score, among other things, to help them make a decision about whether you will be offered admission to their law school. Likewise, let us say that an acquaintance of yours is attending counseling sessions with a mental health professional. As part of these counseling sessions, the mental health professional may administer a number of psychological tests and use the results to develop a treatment program for your acquaintance. In each of these cases, someone else—usually representing an institution—has used the results of a psychological test to make a decision about another individual.

Comparative and Absolute Decisions

When institutions use test scores to make decisions about those who took a test, they do so using either a comparative method or an absolute method. **Comparative decisions** are made by comparing the test scores of a number of people to see who has the best score. For example, imagine you applied to and were accepted into the law school of your choice. Now imagine that the law school is going to offer an academic scholarship to only one individual who was offered admission. Based on interviews and letters of recommendation, you advance as one of four finalists for the scholarship. Who will get the scholarship now depends on LSAT scores. Because you scored higher than the other three finalists, you receive the scholarship. This is a comparative decision because all of the finalists' LSAT scores were compared and the individual with the highest score was selected. Or perhaps you applied for a job at an organization where psychological tests were used as part of the selection process. If, after you took these tests, the organization decided to continue to consider your application because you scored better than 75% of the other applicants, the organization would be using the test results to make decisions using a comparative method.

Absolute decisions, on the other hand, are decisions made by others (institutions) by looking at who has the minimum score needed to qualify. For example, let us consider the same scholarship example, with you advancing as one of four finalists for the scholarship. However, this time the school offers the scholarship to any finalist who has a score of at least 160 (where the minimum score is 120 and the maximum score is 180). Or suppose the organization to which you applied for a job called and informed you that the managers would like you to come in for an interview because you scored at least a 50 on one of the tests. In each of these cases, the institution made a decision about you not by comparing your score with the scores of other test takers but rather by basing their decision on some minimum score.

Who Uses Psychological Tests and for What Reasons?

A variety of professionals use psychological tests for many different purposes in a number of different settings. Psychiatrists, psychologists, social workers, mental health counselors, career counselors, human resources directors, administrators, and many other professionals all use psychological tests. As you can see from Table 2.1, professionals use psychological tests in three primary settings: educational, clinical, and organizational. We have devoted Chapters 13 to 15 to discussing how tests are used in these settings, and here we provide an overview to help you understand the importance of testing.

Table 2.1 Who Uses Psychological Tests and for What Purposes

<i>Educational Settings</i>		
<i>Who</i>	<i>Where</i>	<i>Why</i>
Administrators	Primary schools	To select students into schools
Teachers	Secondary schools	To award scholarships
School psychologists	Colleges and universities	To place students in programs
Career counselors		To measure student learning
		To identify problems
		To identify career interests
<i>Clinical Settings</i>		
<i>Who</i>	<i>Where</i>	<i>Why</i>
Clinical psychologists	Mental health clinics	To diagnose disorders
Psychiatrists	Residential programs	To plan treatment programs
Social workers	Private practices	To assess treatment outcomes
Counseling psychologists		To counsel and advise
Licensed professional counselors		
Marriage and family counselors		
<i>Organizational Settings</i>		
<i>Who</i>	<i>Where</i>	<i>Why</i>
Human resources professionals	Organizations	To make hiring decisions
Industrial/organizational practitioners	Consulting companies	To determine training needs
	Consulting practices	To evaluate employee performance

Educational Settings

Administrators, teachers, school psychologists, and career counselors in primary schools (elementary school), secondary schools (middle school and high school), and colleges and universities all use psychological tests. For example, in colleges and universities, administrators (for example, admissions officers, deans of admissions) use the results of tests such as the SAT to help make admissions decisions

and award scholarships. Originally developed in 1926 by Carl Campbell Brigham for the College Examination Board, today the SAT is the most widely used standardized college admission test. It measures the critical thinking skills students are believed to need for academic success in college.

In primary and secondary schools, teachers administer tests to measure student learning and assign course grades. The same is true for faculty at colleges and universities. In primary and secondary schools, school psychologists use the results of tests to determine eligibility for gifted programs and to identify developmental, visual, and auditory problems for which children might need special assistance. In colleges and universities, career counselors use the results of tests to help students identify career interests and select major areas of concentration and careers that are consistent with students' skills and interests. These are only some of the ways in which professionals use tests in educational settings. (Chapter 13 is devoted to discussing how educators use psychological tests.)



More detail about the SAT can be found in Test Spotlight 2.1 in Appendix A.

Clinical Settings

Clinical psychologists, psychiatrists, social workers, counseling psychologists, and other health care professionals use psychological tests in clinical settings such as mental health clinics, residential programs, and private practices. For example, clinical psychologists and psychiatrists administer psychological tests to help diagnose psychological disorders, plan treatment programs, and determine the effectiveness of treatment programs. Counseling psychologists administer psychological tests to individuals and use the results to counsel and advise those seeking help in how to deal with problems of everyday living such as marital and relationship problems. Again, these are only some of the ways professionals use psychological tests in clinical settings. (Chapter 14 focuses on how clinicians use psychological tests.)

Organizational Settings

Human resources professionals and industrial/organizational practitioners use psychological tests in organizations. For example, they administer tests to job applicants to measure the applicants' knowledge, skills, abilities, and personalities—to predict the likelihood of their success on the job. They use the test results, along with other information, to make hiring decisions. The NFL even uses psychological tests to make football player selection decisions (see In the News Box 2.1 for more information). Human resources professionals and industrial/organizational practitioners use psychological tests to identify employees' strengths and opportunities for development and ultimately to determine employees' training needs (for an example of such a test, the Leadership Practices Inventory, see On the Web Box 2.1). These professionals also use tests to measure employee performance and determine employee performance ratings. (Chapter 15 discusses how managers and others use psychological tests in organizational settings.)

ON THE WEB BOX 2.1

The Leadership Practices Inventorywww.lpionline.com

Developed by Jim Kouzes and Barry Posner, the Leadership Practices Inventory (LPI) is a 360°, or multirater, leadership assessment tool that measures the extent to which individuals demonstrate the Five Practices of Exemplary Leadership (John Wiley & Sons, 2000–2012).

1. Model the way
2. Inspire a shared vision
3. Challenge the process
4. Enable others to act
5. Encourage the heart

In a 360° process, surveys are sent to those people an individual works closely with, including managers, peers, subordinates, and customers, and these people are asked to provide feedback on the extent to which the individual demonstrates key behaviors.

Believing that leadership is a measurable, learnable, and teachable set of behaviors, the LPI provides leaders with the opportunity to gather feedback from those they work with on a day-to-day basis on the extent to which they engage in 30 leadership behaviors associated with the Five Practices of Exemplary Leadership. Using a rating scale from 1 (*almost never*) to 10 (*almost always*), peers, direct reports (subordinates), others the individual works with, and the leader's manager rate the extent to which the individual leader demonstrates each of the 30 leadership behaviors. Individuals also rate themselves.

Available for electronic implementation through a web-based application, the LPI Online offers easy-to-use administrator and leader websites as well as a just-in-time, streamlined, and easy-to-interpret personalized report. This report presents 360° results in a number of ways, including the following:

- Self and individual observer ratings (for example, manager, colleague, direct report, others in the workplace), and overall average scores for each of the five practices as well as each of the behaviors within each practice, in both numeric and graphical form
- Ranking of all leadership behaviors from most frequently demonstrated to least frequently demonstrated
- A comparison of how self-scores and observers' scores compare with the scores of thousands of individuals who have taken the same version of the LPI
- All responses to any open-ended questions that were included in the survey

If you want to learn more about the LPI, go to www.lpionline.com/demo/demo.html and view the online demo.

IN THE NEWS

Box 2.1 The NFL and the Wonderlic

As you may or may not know, American football is a physically demanding game that requires players to coordinate closely to carry out strategies they have learned and practiced on the field. According to Amy Shipley (2011) of the *Washington Post*, National Football League (NFL) executives who select players for professional football teams highly regard one psychological test, the Wonderlic Basic Skills Test (WBST). For three decades, the WBST has been as important to the NFL's selection requirement as the bench press and the 40-yard dash. The WBST provides a short measure of adult language and math skills and measures the job readiness of teenagers and adults. As you can see by reading Test Spotlight 15.2 in Appendix A, the questions are multiple choice and they focus on language and math. The NFL is the only one of the major U.S. sports organizations that gives prospects an intelligence test.



You may be interested to know that Ryan Fitzpatrick, who signed with the Buffalo Bills in 2009, scored 48 on the WBST, which correlates to an IQ of about 150. We might expect that a smart quarterback would be successful, but what about the Miami Dolphins' Dan Marino, who reportedly had a WBST score in the teens? Marino was a Hall of Fame quarterback who holds many NFL and team passing records (Miami Dolphins, 1995–2011).

Gil Brandt, former vice president of player personnel for the Dallas Cowboys, had a bias against players who scored exceptionally well on the Wonderlic. "When you have a player who is really extraordinarily smart, he's probably not as aggressive as a player with a lesser intelligence quotient," Brandt said. "I'm not sure that's correct—but I think it is" (quoted in Shipley, 2011, p. D03).

Lyons, Hoffman, and Michel (2009) investigated this question and published their study of 762 players from three NFL draft classes. They found no correlation between intelligence, as measured by the WBST, and performance on the field. There was, however, a negative correlation between football performance and the WBST for tight ends and defensive players. Yes, you understood correctly—the lower their scores on the Wonderlic, the better performance they achieved on the field! Even at the quarterback position, where many think cognitive abilities are needed, there was no significant relationship between high scores and high performance.

INTERIM SUMMARY 2.1

WHO USES PSYCHOLOGICAL TESTS AND FOR WHAT REASONS?

- Psychological testing is important because people use tests to make important decisions.
- Individual decisions are those made by the person who takes a test, and institutional decisions are those made by others as a result of an individual's performance on a test.
- Comparative decisions involve comparing people's scores with one another to see who has the best score, and absolute decisions involve seeing who has the minimum score to qualify.
- Psychological tests are used by professionals for a variety of purposes in educational, clinical, and organizational settings.
- In educational settings, administrators, teachers, school psychologists, and career counselors use psychological tests to make a variety of educational decisions, including admission, grading, and career decisions.
- In clinical settings, clinical psychologists, psychiatrists, social workers, and other health care professionals use psychological tests to make diagnostic decisions, determine interventions, and assess the outcomes of treatment programs.
- In organizational settings, human resources professionals and industrial/organizational psychologists use psychological tests to make decisions such as whom to hire for a particular position, what training individuals need, and what performance rating an individual will receive.

The Social and Legal Implications of Psychological Testing

As we discussed in Chapter 1, psychological testing today is a big business and a part of the American culture. Thousands of psychological tests are available, and they are taken by many individuals and professionals to make informed decisions. In educational settings alone, approximately three million students took the SAT in 2010 and 2011 (College Board, 2011).

As you just learned, individuals and institutions use the results of psychological tests for many purposes, including making important decisions. These decisions, of course, are meant to benefit people. The widespread use of tests suggests that, in general, psychological tests do serve their purpose. Nonetheless, psychological testing has always been, and probably always will be, controversial. Although some of this controversy stems from the general public's misunderstandings about the nature and use of psychological tests and can be easily eliminated through education, some of the controversy is deeply rooted in ongoing debates and occurs among professionals themselves.

One of the largest and most deeply rooted controversies pertains to discrimination. For years, some members of society have been concerned that standardized psychological tests unfairly discriminate against certain racial and economic groups, resulting in qualified members of these groups being passed over for admission to educational programs or not being hired at the same rate as members of other groups. The American public really began to express its concern that psychological tests were discriminatory when psychological testing became widespread during the 20th century. Much of this concern was, and continues to be, targeted at standardized tests of intelligence, aptitude, and achievement. As shared by Lemann (1999) in his book *The Big Test*,

In the 1940s standardized educational tests created a ranking of Americans, one by one from top to bottom on a single measure. If one analyzed the ranking by social or ethnic group, then at the bottom, always, were Negroes. Living mostly in the benighted, educationally inferior South, consigned to separate schools that operated only sporadically with ill-trained teachers, historically denied by law even the chance to learn to read and write and figure, disproportionately poor, ill-nourished, and broken-familied, Negroes as a group were in a uniquely bad position to perform well on tests designed to measure such school-bred skills as reading and vocabulary and mathematics fluency. So whenever goods were distributed on the basis of test scores, Negroes got a disproportionately low share of them. (pp. 155–156)

Today, the discrimination controversy continues. One of the most recent and visible cases started in 2003, when the city of New Haven, Connecticut, administered a promotional examination for firefighters aspiring to achieve the ranks of lieutenant and captain. When the tests were scored, the city made the decision to scrap the results for fear that the city would be sued for discrimination if it used the results because no African Americans and only 2 Hispanics had scores high enough to be eligible for promotion. Subsequently, 1 Hispanic and 19 White firefighters sued the city, claiming that they were discriminated against when they were denied promotions as a result of the city's refusal to use the test results. The firefighters claimed city officials denied them promotions due to the fear of potential Civil Rights Act violations rather than how they performed on the promotional exams (*Supreme Court Rules for White Firefighters*, 2009). The case was elevated to the Supreme Court in June 2009. We talk more about this case and the Supreme Court ruling in Chapter 15.

The Controversy Over Intelligence Tests

Intelligence Testing in Education

Researchers have documented that middle- and upper-class White people, on average, score higher on intelligence tests than do other economic and racial groups (Lemann, 1999). Early in the 20th century, believing that this difference in intelligence was due to heredity, elementary schools began administering intelligence tests to students and using the results to place those with higher IQ scores in special academic programs and those with lower scores in more vocationally related programs (Hunt, 1993). Individuals who believed that intelligence was inherited had no problem with using psychological tests in this manner. They believed that people who do better on such tests naturally have superior intellects. In their view, if intelligence is indeed inherited, using psychological tests in this manner is fair and in the best interest of individuals and society.

However, what if intelligence is not inherited but rather is the result of the environment in which one is raised? If this were the case, all people would be born with the same potential and only those who grew up in favorable backgrounds would, in general, score higher on intelligence and academic ability tests. Those who had disadvantaged backgrounds would score lower. In this case, using intelligence test scores, which are thought to reflect innate abilities, to determine an individual's educational opportunities would be unfair. Hence, we are sure you can understand the debate and the public's concern over the use of intelligence tests.

Over the years, activists who believe that intelligence is determined primarily by environment have worked to eliminate what they consider to be the unfair use of such tests. During the 1960s, in the heat of the civil rights movement, activist groups demanded that schools abandon the use of intelligence tests. New York, Los Angeles, and Washington, D.C., did just that (Hunt, 1993). For Your Information Box 2.1 includes discussion of a court case in which, for exactly this reason, schools in California were ordered not to use intelligence tests for student placement. However, continued efforts to eliminate intelligence testing failed when it became apparent that the placement of children with learning or physical disabilities in the same classrooms as average and gifted children slowed learning (Hunt, 1993).

Intelligence Testing in the Army

In an effort to improve the credibility of psychological testing and establish psychology as a true scientific movement, during World War I Robert Yerkes came across an opportunity to promote the use of mental testing. The American military gave Yerkes permission to administer mental tests to more than 1.75 million U.S. Army recruits. As a result, believing that individuals might be intelligent but not literate or proficient in English, Yerkes (1921) designed the Army Alpha and Beta tests, the first mental tests designed for group testing. The Army Alpha test was developed for use with literate groups, and the Army Beta test for use with those who were unable to read, write, or speak English. Yerkes argued that both tests measured native intellectual abilities—abilities unaffected by culture or educational opportunities. By the end of the war, the Army Alpha and Beta tests were being used to screen army recruits for officer training.

During the 1920s, Walter Lippmann, a popular newspaper columnist, criticized the Army Alpha and Beta tests as having a great potential for abusing the psychological testing process—a process that could be of great benefit to the army (Lippmann, 1922a–e). Like others, Lippmann questioned whether intelligence tests such as the Army Alpha and Beta tests actually measured “intelligence” and whether intelligence was determined by heredity or through life experiences—a question that came to be known as the **nature-versus-nurture controversy**. Data collected using army recruits suggested that average intelligence scores of African American males were much lower than average scores of White males. In addition,

FYI

FOR YOUR INFORMATION BOX 2.1

Can IQ Tests Be Illegal?

Unless you live in California, you are probably not aware of the controversy surrounding the use of IQ scores as a method for placing children in Educable Mentally Retarded (EMR) classes. In 1979, in the case of *Larry P. v. Riles*, testimony suggested that IQ tests are biased against African American children. The plaintiff, the party bringing the suit, showed that six African American children who scored low on one intelligence test scored much higher on the same test when it was revised to reflect the African American children's cultural background. The African American children's first scores placed them in the range labeled "retarded"; however, the scores from the revised test labeled them as "normal." In addition, evidence was given that a higher proportion of African American children, compared with the rest of the student body, were in EMR classes. This information caused the judge to rule that schools in California may not use IQ test scores to place African American children in EMR classes or their "substantial equivalent."

California abolished EMR classes, and in 1986 the same judge modified his ruling, this time banning the use of IQ tests to evaluate African American children referred for any special assessment. This ruling did not please all parents. For instance, Wendy Strong, the mother of a 7-year-old, tried to get help for her daughter, Brianna, who had problems in learning. Because her race was shown as African American on school records, school psychologists were not able to administer an IQ test to Brianna. Brianna's mother threatened to have her daughter's racial category changed so that she could be tested. Such a change was possible because Brianna had one African American parent and one White parent.

Eventually, another suit was brought by African American parents who wished to have their children tested. In 1994, the appeals court ruled that parents such as the Strongs were not adequately represented in the 1986 proceedings. Therefore, the court canceled the 1986 ruling but upheld the original 1979 ruling.

SOURCE: Adapted from *Crawford v. Honig*, 37 F.3d 485 (9th Cir. 1994).

when scores of foreign-born recruits were sorted by their countries of origin, those from Turkey, Greece, Russia, Italy, and Poland produced large numbers of scores that indicated a mental age of younger than 11 years, which for an adult was an indication of low intelligence (Yerkes, 1921). Political groups who opposed the immigration of large numbers of families from Europe following World War I used these data to support their arguments that immigration was harmful to the United States.

Later in the 20th century, Gould (1982) also criticized such mass intelligence testing, claiming that the intelligence tests were culturally biased. For immigrants, the language and customs of the United States were unfamiliar and what appeared to be stupidity was just lack of cultural knowledge and experience. Consider the following three examples from the Army Alpha test (Gould, 1982, para. 21):

1. Crisco is a
 - a. patent medicine.
 - b. disinfectant.
 - c. toothpaste.
 - d. food product.

2. Washington is to Adams as first is to _____.
3. Christy Mathewson is famous as a
 - a. writer.
 - b. artist.
 - c. baseball player.
 - d. comedian.

How did you do on those questions? Did you know that Crisco was a popular vegetable shortening and butter substitute? Washington was the first president of the United States, and Adams was the second. At the time these questions were used, Christy Mathewson was a well-known baseball player; however, his name is no longer general knowledge.

Critics' concern was that immigrants would not have the cultural knowledge and experience to answer such questions correctly. In addition, the tests themselves and the instructions given when administering them were usually incomprehensible to uninformed test takers. The Army Alpha and Beta tests, for example, required test takers to follow directions and perform a series of ballet movements that were confusing and distracting.

The Army Alpha and Beta tests were discontinued following World War I, but the nature-versus-nurture debate continued. Its connection to psychological tests and intelligence raised public controversy again nearly 50 years later at a time when the civil rights movement was changing the American experience.

In 1969, Arthur Jensen published an article in the *Harvard Education Review* that again pointed out a difference in average intelligence scores between Blacks and Whites. Although there have been numerous explanations for these findings, Jensen caused an uproar by implying that this difference in intelligence was almost exclusively (80%) due to genetic factors. This time, Jensen used later and more sophisticated tests than those the army had used during World War I, but the basic "pro-heredity" argument was still the same.

The debate that followed Jensen's (1969) article led professionals and the public to question how psychologists and test developers define and measure intelligence. A number of psychologists (Eells, Davis, Havighurst, Herrick, & Tyler, 1951; Harrington, 1975, 1976) also have pointed out that the intelligence tests administered to Blacks were invalid for measuring the intelligence of Blacks because the tests had been developed for middle-class White children whose experiences are different from those of children from other ethnic groups and socioeconomic classes. Furthermore, Hilliard (1984) questioned Jensen's underlying assumptions regarding an operational definition of race.

The same debate arose again in 1994 when Richard Herrnstein and Charles Murray published their book *The Bell Curve: Intelligence and Class Structure in American Life*, which reiterated many of the conclusions that Jensen drew in 1969. Herrnstein and Murray argue that IQ is extremely important, that it is somewhere between 40% and 80% heritable, and that it is related not only to school performance but also to jobs, income, crime, and illegitimacy. Herrnstein and Murray used intelligence research to substantiate their claim that some of the difference in average IQ scores between Whites and Blacks is likely attributable to genetic factors, suggesting that Blacks are genetically inferior in intellectual abilities and capabilities.

In response to the publication of *The Bell Curve*, the American Psychological Association (APA) convened a task force of psychologists representing the prevalent attitudes, values, and practices of the psychology profession. Based on the work of this task force, the APA published the report *Intelligence: Knowns and Unknowns* (Neisser et al., 1996). The report did not disagree with the data presented in *The Bell Curve*; however, it interpreted the data differently and concluded that although no one knows why the difference exists, there is no support for the notion that the 15-point IQ difference between Black and White Americans is due to genetics (Neisser et al., 1996; Yam, 1998). Furthermore, in a review of *The Bell Curve* in *Scientific American*, Leon Kamin (1995) states, "The caliber of the data in *The Bell Curve* is, at

many critical points, pathetic. Further, the authors repeatedly fail to distinguish between correlation and causation and thus draw many inappropriate conclusions” (p. 99).

Nisbett et al. (2012) published a review that discussed the most current research findings on intelligence. While the review observed that the Neisser et al. (1996) article remains a good summary of the state of our knowledge in intelligence, there have been some important new findings since then. Improvements in brain imaging techniques have enabled us to learn much more about the relationship between brain physiology and intelligence. We also now know much more about the roles of both heredity and environment in intelligence. In spite of these advances, there is still much we do not know about the nature of intelligence and its correlates.

For an interesting news story about IQ scores and The Flynn Effect, see In the News Box 2.2.

IN THE NEWS

Box 2.2 The Flynn Effect

You may have read or heard of a new book, *The Better Angels of Our Nature: Why Violence Has Declined*, written by a well-known psychologist, Steven Pinker (2011). In 800 pages, Pinker presents evidence that the inhabitants of our world are not as violent as they once were. He discusses many types of sociological data (for example, calculations of soldiers killed in wars 2,000 ago compared to today) to make his case. We do not have space to discuss all his arguments here; however, we would like to discuss a little known testing phenomena called the *Flynn effect* that Pinker uses to make his point.



In 1984, James Flynn published his discovery that over time IQ scores become greater. According to Flynn, Charles Murray named the phenomenon, for any number of groups (students, military enlistees, racial groups, and so on) whose IQ scores became greater over time, the Flynn effect. Flynn’s original study compared the same group of participants who had taken either the Wechsler or Stanford-Binet IQ tests two or more times over the course of many years. He examined the data from 73 studies containing 7,500 participants, ages 2 to 48 years, and concluded that between 1932 and 1978, White Americans had gained 14 IQ points.

The Flynn effect holds true for people of various cultural and racial backgrounds and for IQ tests written in various languages, including Chinese. Some researchers suggested that the effect was created by culture bias in the IQ tests themselves. Subsequent research shows that the IQ gains were greatest when researchers administered culture-reduced tests, so we cannot attribute the increase in IQ scores to cultural or racial test bias. Nor can we attribute the increase to better education, because the tests on which the scores increased most do not require a rich vocabulary or even mathematical ability. They do, however, require advanced powers of abstract reasoning, and Flynn believes that the spread of scientific reasoning has caused the increase in IQ scores (Singer, 2011).

Now we can return to Steven Pinker and his optimistic forecasts for our world. Pinker argues that enhanced powers of reasoning—suggested by research on the Flynn effect—enable us to detach ourselves from our immediate experience or perspective and frame our ideas in more abstract, universal terms. This detachment leads to a more overall view of others’ interests as well as our own. Finally, this universal view leads to better moral commitments, including nonviolence. Pinker suggests that the 20th century has seen, and the 21st century will see, a “moral Flynn effect” in which an acceleration of reasoning ability moves the world away from violence, persecution, and exploitation.

The Controversy Over Aptitude and Integrity Tests

As with intelligence tests, the American public has expressed concern over the use of aptitude and integrity tests.

Aptitude Tests and the U.S. Employment Service

During the 1940s, before the Equal Employment Opportunity Act became law, the U.S. Employment Service (USES) developed the General Aptitude Test Battery (GATB) to assist with career counseling and job referral. An occupationally oriented, multiaptitude test, the GATB consists of 12 tests measuring nine cognitive and manual aptitudes: general learning ability, verbal aptitude, numerical aptitude, spatial aptitude, form perception, clerical perception, motor coordination, finger dexterity, and manual dexterity (Nelson Education, n.d.).

As with intelligence tests, research showed that average GATB scores of minority groups were well below those of the other groups. Because the USES and many of its state and local offices used GATB scores to make referrals to employers, more Whites were being referred for particular jobs than were African Americans or Hispanics (Hunt, 1993). The amended Civil Rights Act of 1991 made it illegal to use GATB scores in this way because national policy required giving the disadvantaged compensatory advantages (Wigdor, 1990). Rulings by the Equal Employment Opportunity Commission and several court decisions resulted in a solution called **within-group norming** or “race norming”—the practice of administering the same test to every test taker but scoring the test differently according to the race of the test taker. Using within-group norming, test users would not be able to refer test takers for jobs using their raw test scores—the scores calculated according to the test instructions—or based on how their scores compared with others in the overall norm group. Instead, the test users were required to compare each test taker’s score with the scores of other test takers only within the same racial or ethnic group. (We talk more about norms in Chapter 5.) Using race norming, a minority test taker who scored the same as a White test taker would in fact rank higher than the White test taker. Employment services in 38 states used this race norming.

Many psychologists were outraged about the use of race norming. They claimed it was a disgrace to the psychological testing industry, a distortion of a test’s measure of job fitness (L. S. Gottfredson, 1991), and an illegal quota system that unfairly discriminated against Whites. Nonetheless, in 1989 the National Research Council conducted a study that supported the use of race norming. However, the council recommended that referrals by employment services be based not only on an applicant’s GATB score but also on the applicant’s experience, skills, and education. Several years later, race norming was outlawed, but not because it was unfair. In a struggle to pass the Civil Rights Act of 1991, members of Congress who favored race norming needed to yield to those who did not. As passed, Section 106 of the Civil Rights Act of 1991 prohibited employers from adjusting scores on the basis of race, color, religion, sex, or national origin (Hunt, 1993) when the sole purpose was to refer or select people for jobs. Use of the GATB in the United States has declined considerably because it became evident that parts of the GATB discriminated against minorities. Canadians continue to use the GATB as a pre-employment test.

The U.S. Armed Forces now uses a similar instrument, the Armed Services Vocational Aptitude Battery (ASVAB), a series of tests used primarily by the military to help determine whether individuals qualify for service in certain military branches and, if so, what jobs they qualify for (Military Advantage, 2012). For more information on the ASVAB, read On the Web Box 2.2.

ON THE WEB BOX 2.2

The Armed Services Vocational Aptitude Battery (ASVAB)

www.military.com/join-armed-forces/asvab/



The United States Department of Defense developed and maintains the ASVAB, which is an aptitude test administered at over 14,000 schools and military processing stations across the United States. A person's ASVAB scores are used to determine qualification for certain military occupations and are also used to determine enlistment bonuses. Individuals seriously interested in joining the military often take the ASVAB practice tests to determine strengths and areas of improvement to better prepare for the ASVAB test.

To learn more about the ASVAB, take available practice tests, and receive some tips and strategies, go to www.military.com/join-armed-forces/asvab. To find out more about minimum scores for specific military jobs in the Air Force, Army, Marine Corps, and Navy, go to www.military.com/ASVAB/0,,ASVAB_MOS.html.

In 1964, the U.S. Congress passed the Civil Rights Act. Intended to bring about equality in hiring, transfers, promotions, compensation, access to training, and employment-related decisions, Title VII of the act made it unlawful to discriminate or segregate based on race, color, national origin, or gender in all terms and conditions of employment. Issued as an interpretation of Title VII in 1978, the Uniform Guidelines on Employee Selection Procedures recommended that employers analyze their hiring and promotion processes to determine whether their selection procedures (including the use of tests) were discriminatory. If the selection rate for any race, sex, or ethnic group was less than four-fifths (or 80%) of the selection rate for the group with the highest selection ratio, the selection process could be considered potentially discriminatory, because it results in adverse impact for the underselected groups. (Chapter 15 includes more detailed discussing of using psychological tests in compliance with the Civil Rights Act of 1964.)

Aptitude Testing in Education

During the 1970s, Americans noticed an apparent decline in SAT scores (Haney, 1981). National averages for the SAT between 1952 and 1963 stayed approximately the same despite the fact that 7% more students took the SAT during those years. However, between 1964 and 1970, national average scores began to decline significantly. By 1977, both the SAT math and verbal scores had declined (Dutch, 2010).

This time, the concern was not with how “intelligent” Americans were but rather with how much American students were learning in public schools. Between 1963 and 1975, the College Board reported that college-bound high school students answered approximately 5% fewer SAT questions correctly—a 60- to 90-scale point decline in aggregate SAT scores. As a result, the College Board and Educational Testing Service convened a special panel that concluded that a 14-year decline in average scores was due to two factors. First, more students were taking the SAT, and these students not only had weaker academic records but also were coming from more diverse backgrounds. Again, the implication was that the traditional test takers, middle- and upper-class White students, were more likely to make high grades. Second, the type of educational experience students had during the late 1960s and early 1970s had caused a decrease in performance on standardized tests. Among the reasons given for

a decline in educational experience were a “diminished seriousness of purpose and attention” and a “marked diminution in young people’s learning motivation” (Haney, 1981, p. 1026).

However, Berliner and Biddle (1995) state,

So although critics have trumpeted the “alarming” news that aggregate national SAT scores fell during the late 1960’s and the early 1970’s, this decline indicates nothing about the performance of American schools. Rather, it signals that students from a broader range of backgrounds were then getting interested in college, which should have been cause for celebration, not alarm. (p. 21)

Integrity Testing in Organizations

Yet another concern has been integrity testing. **Integrity tests** measure individual attitudes and experiences toward honesty, dependability, trustworthiness, reliability, and prosocial behavior (Society for Industrial and Organizational Psychology, 2012). Typically, integrity tests require test takers to answer questions about the following:

- Illegal use of drugs or engagement in unacceptable theft or criminal activities
- Opinions about illegal or inappropriate activities
- Personality or beliefs
- Reactions to theoretical and/or hypothetical situations (Eisenberg & Johnson, 2001)

There are two basic types of integrity tests. One requires individuals to respond to questions about previous experiences related to ethics and integrity. These tests are overt and include very straightforward questions. The other type requires individuals to respond to questions about their preferences and interests. These tests are more personality based and measure propensity to engage in unacceptable work behaviors. From the preferences and interests, inferences are drawn about how the individual may behave in the future. Both types are used by organizations to identify individuals who are likely to engage in inappropriate, dishonest, and antisocial behavior at work.

Employers have used integrity tests for many years both to screen job applicants and to keep existing employees honest. According to research, their use is justified. Organizations lose anywhere from \$10 billion (as reported by the American Management Association) to \$150 billion (as reported by the Federal Bureau of Investigation) annually due to workplace behaviors such as theft and shoplifting (Net Industries, 2012). Results of employee surveys reveal the following:

- 56% of working people admit they have lied to their supervisors.
- 41% say they have falsified records.
- 64% admit using the Internet for personal reasons during working hours.
- 35% have stolen from their employers, by their own admission.
- 31% abuse drugs or alcohol. (Profiles International, 2011, para. 8)

According to Profiles International (2011), a company that provides assessment instruments to organizations, over one-third of organizations declaring bankruptcy report being “stolen out of business” by their employees (para. 7). Further, the average shoplifter steals \$59, and the average employee caught stealing takes \$549! What do employees steal? If you define “employee theft” as the theft, use, or misuse of assets without permission, employees steal money, supplies, and merchandise or company property. Employee theft can also include the time employees are paid for time during which they did not work.

IN THE NEWS

Box 2.3 Can Your Personality Get You Hired or Fired?

In an ABC News report titled “Can Your Personality Get You Hired or Fired,” Tory Johnson (2006) discusses how organizations use personality tests and his experience taking two such tests: the California Psychological Inventory and the Myers-Briggs Type Indicator. Johnson’s article follows. What do you think?



Personality tests have been around for more than a century, but employers are using them now more than ever when hiring. The main reason: to select the best possible candidates and reduce turnover, which costs a company between a quarter to one-and-a-half times the departing worker’s salary.

Even though the word “test” implies pass or fail, there’s no such thing in personality assessments. There’s no right or wrong, no numerical score. Instead, these tools assess our “soft” skills—personality types, strengths, styles and preferences.

More than 2,500 types of personality tests are used today, and they generally fall into two distinct categories for employment purposes: those used for selecting and hiring new workers and those used for developing and advancing existing staffers.

Last week I took two of these tests: the California Psychological Inventory, which is popular in hiring because it helps predict how an employee might interact with other people, and the Myers-Briggs, which is the gold standard for assessing preferences and styles useful for worker development.

A confession: I was incredibly nervous before starting the online assessments. I was fearful of the unknown. What if the tests revealed weaknesses I wasn’t aware of? My mind wandered every which way.

But as I dived into the 350-plus questions between the two assessments, all of that fear dissipated. I very quickly realized it’s all based on my opinions, with no right or wrong.

CPP.com, a leading publisher and administer [*sic*] of many of these tests, gave permission to share some of the questions I had to answer on each of the two tests I took. These are by no means mini tests but rather an illustrative sample of items that appear in each test.

From Myers-Briggs Type Indicator Instrument, by Katherine C. Briggs and Isabel Briggs-Myers:

Are you inclined to:

- A) value sentiment more than logic, or
- B) value logic more than sentiment?

Do you prefer to:

- A) arrange dates, parties, etc., well in advance, or
- B) be free to do whatever looks like fun when the time comes?

Would you rather work under a boss who is:

- A) good-natured but often inconsistent, or
- B) sharp-tongued but always logical?

At parties do you:

- A) do much of the talking, or
- B) let others do most of the talking?

When you start a big project that is due in one week, do you:

- A) take time to list the separate things to be done and the order of doing them, or
- B) plunge right in?

Which one word in each of the following pairs appeals to you more?

- A) sensible
- B) fascinating
- A) imaginative
- B) realistic
- A) devoted
- B) determined

From CPI 260 Assessment, by Harrison G. Gough, Ph.D.:

Answer TRUE or FALSE as to how you feel each statement applies to you.

- I have a natural talent for influencing people.
- I always see to it that my work is carefully planned and organized.
- People often expect too much of me.
- It is hard for me to just sit still and relax.
- The idea of doing research appeals to me.
- I enjoy hearing lectures on work affairs.
- I read at least 10 books a year.
- I always try to consider the other person's feelings before I do something.
- I like parties and socials.

As you can see from the examples, it's impossible to paint a picture of someone based on just a few answers, which is why so-called mini assessments have absolutely no value. However, when you've answered a comprehensive assessment featuring multiple questions on similar topics, a pattern of strengths and styles will emerge. Sophisticated scoring systems are used to generate meaningful results, and certified interpreters are able to tell you what it all means and how to apply it to your career development.

“Faking Good”

Many people are quick to ask if it's possible to cheat or beat the assessments. The answer is no. In fact, the CPI has a built-in mechanism designed to catch a test-taker who is trying to do what's called faking good. By this I mean people who take every opportunity to paint themselves in an exceptionally positive light will likely be flagged. One true/false question along these lines is: “I have never deliberately told a lie.”

There are 25 questions that revolved around the same issue, and they are designed to get your honest answer, not just what you think makes you sound the best or most truthful. If you appear to be too good to be true, you'll likely be flagged by the test administrator.

It's best to be honest—not only for the employer's sake but really for yours too. Be true to yourself. If you don't get the job because of it, there's a good chance you wouldn't have been a good fit, and it's a blessing to know that before an offer is made.

A Critical Concern

The biggest criticism of personality tests stems from a fear that employers rely too heavily on them in making decisions. But every employer and interpreter I spoke to emphasized the importance of using these assessments as only one part of the decision-making process.

Think of it in terms of the SAT for college admittance. A student is so much more than a simple SAT score. The best colleges make decisions based on GPA, course loads, the high school profile, essays, recommendations and more, including the SAT score.

The same is true in hiring: Assessments are only one piece of a much more comprehensive process that includes interviews, role playing, recommendations and more. If you're asked to take a personality test by a prospective employer, ask how it will be used in the overall hiring process, and confirm that the information stays confidential. If you're passed over for a position and you believe it's because of your personality assessments, don't panic. Since assessments are specific to an individual employer and are measured against the position you're applying for and the company's culture and needs, the results do not follow you from position to position. Getting fired from a job sticks with your employment record, but these assessments do not.

SOURCE: Johnson (2006).

Although the use of integrity tests might be justified by alarming figures associated with employee theft, many individuals and labor groups oppose their use because they believe that integrity tests (a) are neither valid nor reliable and therefore falsely classify some honest people as dishonest, (b) are an invasion of privacy, and (c) have a different and more inhibiting effect on minorities, eliminating higher percentages of minorities than Whites from job opportunities (U.S. Congress, Office of Technology Assessment, 1990). In the early 1990s the APA expressed concern about the reliability and validity of such tests. After 2 years of research, an APA task force concluded that for most integrity tests, publishers have little information regarding whether the tests actually predict honesty. As a result, the APA urged employers to stop using those integrity tests for which little validity information was available (APA Science Directorate, 1991). Instead, the APA suggested employers rely on only those tests that have substantial evidence of their predictive validity. (Chapter 8 discusses predictive validity in more detail, and Chapter 15 provides more information on integrity testing in organizations.)

Although a history of controversy over the use of intelligence, aptitude, and integrity tests exists, today there is also an emerging controversy over the use of personality tests. For a review of some of this controversy, read In the News Box 2.3.

INTERIM SUMMARY 2.2 SOCIAL AND LEGAL IMPLICATIONS OF PSYCHOLOGICAL TESTING

- Even with their widespread use, psychological tests are not without their critics.
- Some individuals are very concerned that psychological tests are biased and do not result in the correct institutional decisions.
- Much of this controversy is focused on the use of intelligence, aptitude, and integrity tests.
- This controversy has influenced social movements, laws, and guidelines on how psychological tests should and should not be used.

Chapter Summary

By now, you should understand that psychological testing is important because psychological tests are used to make important decisions. Both individuals and institutions use the results of psychological tests to make decisions. Individual decisions are those made by the person who takes a test, and institutional decisions are made by others who use an individual's performance on a test to make a decision about that individual. Furthermore, institutions make decisions using either a comparative method or an absolute method. Comparative decisions involve comparing an individual's score with other people's scores to see who has the best score. Absolute decisions involve setting a minimum score that test takers must achieve to qualify.

Testing is also important because different professionals, in a variety of clinical, educational, and organizational settings, use psychological tests for many purposes. In educational settings, administrators, teachers, school psychologists, and career counselors use psychological tests to make a variety of educational decisions, including admissions, grading, and career decisions. In clinical settings, clinical psychologists, psychiatrists, social workers, and other health care professionals use psychological tests to make diagnostic decisions, determine interventions, and assess the outcomes of treatment programs. In organizational settings, human resources professionals and industrial/organizational

psychologists use psychological tests to make decisions such as whom to hire for a particular position, what training an individual needs, and what performance rating an individual will receive.

Even given their widespread use, psychological tests are not without their critics. Many people have been, and continue to be, concerned that psychological tests discriminate against certain racial and economic groups, resulting in fewer educational and employment opportunities for these groups. During the 20th century, this concern influenced social movements and resulted in legislatures passing laws and courts setting case law that determined how psychological tests can and cannot be used.

Engaging in the Learning Process

KEY CONCEPTS

After completing your study of this chapter, you should be able to define each of the following terms. These terms are bolded in the text of this chapter and defined in the Glossary.

- absolute decisions
- comparative decisions
- individual decisions
- integrity tests
- institutional decisions
- nature-versus-nurture controversy
- within-group norming

LEARNING ACTIVITIES

The following are some learning activities you can engage in to support the learning objectives for this chapter.

<i>Learning Objectives</i>	<i>Study Tips and Learning Activities</i>
<i>After completing your study of this chapter, you should be able to do the following:</i>	<i>The following study tips will help you meet these learning objectives:</i>
Describe different types of decisions that are made using the results of psychological tests.	<ul style="list-style-type: none"> • Make a list of all the psychological tests you have ever taken. Write down any decisions you made about yourself based on the results of each test. Write down any decisions others made about you based on the results of each test. Indicate whether others used a comparative method or an absolute method when evaluating your test score.
Explain which professionals use psychological tests, in what settings, and for what reasons.	<ul style="list-style-type: none"> • Schedule time to talk with three of the following professionals: a clinical psychologist, a career counselor, a secondary school administrator, a school psychologist, an industrial/organizational practitioner, a human resources director. Interview these professionals and find out what tests they use on a day-to-day basis, why they use these tests, and how they use the test scores to make decisions. Be prepared to share your findings with your class.

(Continued)

(Continued)

<i>Learning Objectives</i>	<i>Study Tips and Learning Activities</i>
Describe some concerns individuals have regarding the use of psychological tests as well as the social and legal implications of psychological testing, especially as they relate to intelligence, achievement, aptitude, and integrity testing.	<ul style="list-style-type: none"> • Find two news or journal articles: one discussing controversies in psychological testing prior to the 1980s and one discussing legal challenges associated with psychological testing after 1980. Compare and contrast the controversies discussed in each article. Be prepared to share your articles and findings with your class. • On a piece of paper, create three columns: Intelligence and Achievement, Aptitude, Integrity. Write as many of the social and legal implications as you can remember in each of the categories.

PRACTICE QUESTIONS

The following are some practice questions to assess your understanding of the material presented in this chapter.

Multiple Choice

Choose the one best answer to each question.

1. What type of decision is made when a high school administrator uses your test score to place you in a gifted program?
 - a. Absolute
 - b. Comparative
 - c. Individual
 - d. Institutional
2. Hector completed several interest inventories at the career center at his college. He used the results to decide on a college major. What kind of decision did Hector make?
 - a. Institutional
 - b. Individual
 - c. Comparative
 - d. Absolute
3. What method is an organizational leader using to make a decision when the leader continues to consider your job application because your score was one of the highest on a pre-employment test?
 - a. Absolute
 - b. Comparative
 - c. Individual
 - d. Institutional
4. A manager at XYZ Corporation administers an employment test to help determine which job candidate will be offered a job. The manager makes the decision by looking to see who performed at a minimum level. The manager used the test to make what kind of decision?
 - a. Individual
 - b. Absolute
 - c. Comparative
 - d. Normative
5. In educational settings, teachers, administrators, school psychologists, and career counselors use psychological tests for all EXCEPT which one of the following purposes?
 - a. Measure student learning
 - b. Award scholarships
 - c. Identify career interests
 - d. Plan treatment programs
6. In organizational settings, human resources professionals and industrial/organizational practitioners use psychological tests for all EXCEPT which one of the following purposes?
 - a. Make hiring decisions
 - b. Diagnose disorders

- c. Determine training needs
 - d. Evaluate employee performance
6. What has been a major concern of the general public regarding the use of psychological tests?
 - a. Test publishing companies make too much money selling psychological tests.
 - b. Psychological tests unfairly discriminate against certain racial groups.
 - c. Psychological tests are neither reliable nor valid.
 - d. Local and federal government regulation of psychological testing is too prevalent.
 7. What debate centers around whether people are born with their intelligence or acquire their intelligence during their lives?
 - a. Innate versus learned
 - b. Mature versus learned
 - c. Innate versus nurture
 - d. Nature versus nurture
 8. What test would you use if you were interested in determining whether individuals qualify for specific jobs in military branches of government?
 - a. Rorschach Inkblot Test
 - b. General Aptitude Test Battery (GATB)
 - c. Armed Services Vocational Aptitude Battery (ASVAB)
 - d. Leadership Practices Inventory (LPI)
 9. What solution was introduced because an examination of GATB scores showed that more Whites were being referred for jobs than were African Americans and Hispanics?
 - a. Ethnic norming
 - b. Situational norming
 - c. Within-group norming
 - d. Between-group norming
 10. What is the term used to describe when test takers' raw scores are compared with those of their own racial or ethnic group?
 - a. Ethnic norming
 - b. Situational norming
 - c. Race norming
 - d. Between-group norming
 11. What do integrity tests claim to measure?
 - a. Ability to perform a job
 - b. Personality
 - c. Individuals' ethics
 - d. Honesty

Short Answer/Essay

Read each of the following, and consider your response carefully based on the information presented in this chapter. Write your answer to each question in two or three paragraphs.

1. Describe the different types of decisions that are made using the results of psychological tests. Provide an example of each.
2. How might individuals use the results of psychological tests? How might organizations use the results of psychological tests?
3. What are the similarities and differences between comparative decisions and absolute decisions? Give examples of each.
4. Who uses psychological tests, and for what reasons do they use them?
5. What are some of society's concerns about intelligence, aptitude, and integrity testing?
6. How do past controversies over psychological testing compare to current controversies?
7. What concerns have been expressed regarding the use of the following personality tests: Rorschach Inkblot Test, Minnesota Multiphasic Personality Inventory (MMPI), and Myers–Briggs Type Indicator (MBTI)?

ANSWER KEYS

Multiple Choice

1. d	2. b	3. b	4. b	5. d
6. b	7. d	8. c	9. c	10. c
11. d				

Short Answer/Essay

Refer to your textbook for answers. If you are unsure of an answer and cannot generate the answer after reviewing your book, ask your professor for clarification.