fifth edition

Assessment in Counseling

A Guide to the Use of Psychological Assessment Procedures

Danica G. Hays

fifth edition

Assessment
in **Counseling**

A Guide to the Use
of Psychological
Assessment Procedures

Danica G. Hays

Assessment in Counseling

fifth edition

A Guide to the Use of Psychological Assessment Procedures

10 9 8 7 6 5 4 3 2 1

American Counseling Association

5999 Stevenson Avenue
Alexandria, VA 22304

Director of Publications • Carolyn C. Baker

Production Manager • Bonny E. Gaston

Editorial Assistant • Catherine A. Brumley

Copy Editor • Kimberly Kinne

Cover and text design by Bonny E. Gaston

For Chris—my partner, best friend, and colleague.

Thank you for making everything more meaningful.

Table of Contents

Section I
Foundations of Assessment in Counseling

Section II
Basic Statistical and Measurement Considerations

Section III
Initial Psychological Assessment

Section IV
Types of Assessment

Section V
The Assessment Report

Preface

The purpose of this book is to provide information about the various assessment procedures that are specifically relevant for practicing counselors. The book deals with the use of these assessment procedures in the counseling process and emphasizes the selection, interpretation, and communication of psychological test results and highlights the basic principles of psychological assessment. It stresses the importance of integrating assessment results with other information about the client. One primary assumption undergirds this text: Counselors engage in assessment practices *every day*, and these practices affect relationships, treatment decisions, and culturally responsive counseling. Furthermore, assessment involves both quantitative and qualitative indicators.

The book is not designed to be a comprehensive textbook or desk manual on the various assessment tools themselves. There are a number of excellent books that describe psychological tests and other assessment procedures in detail. It is expected that counselors will make use of such publications along with other resources as they evaluate assessment tools. As with the previous editions, the latest developments regarding those assessments commonly used by counselors and other mental health professionals are included. New to this edition are some innovative ways to integrate assessment into the counseling profession.

Some of the key features of the fifth edition of the text include the following:

- bolded key terms to facilitate comprehension of major concepts;
- chapter pretests ("Test Your Knowledge") to gauge previous learning;
- self-development activities, such as reflective exercises and class and field activities;
- "Tip Sheets," or practical, user-friendly information about major assessment concepts, issues, and practices;
- inclusion of practitioner voices of various assessment topics ("Assessment in Action");
- case examples that highlight assessment issues and score reports;
- sample assessment items with an expanded list of common assessment tools;
- coverage of the history of assessment, test access issues, cultural bias in assessment, high-stakes testing, qualitative assessment, and specialty areas of assessment and related standards;
- review questions and chapter summaries;
- sample assessment report; and
- common statistical formulas used in assessment.

The text is organized into five sections. Section I, Foundations of Assessment in Counseling, includes introductory concepts of assessment that are useful for conceptualizing measurement and statistical concepts and working with various types of assessment. The four chapters in this section include a discussion of basic assessment terms; the history of assessment; purpose and use of assessment; the assessment process related to selection, administration, interpretation, and communication; ethical, legal, and professional issues in assessment and related assessment standards; and multicultural assessment practices. Section II, Basic Statistical and Measurement Considerations, includes two chapters that address foundational knowledge in statistics and measurement. Specifically, the following concepts are discussed: scales of measurement; reliability, validity, and correlation; test development; measures of central tendency and variability; and raw score transformation. Section III, Initial Assessment in Counseling, includes two chapters related to common assessment tasks typically found at the beginning of the counseling relationship to gauge mental health and substance abuse symptoms. This section addresses the intake interview; mental status examination; several general screening inventories; specialized assessment of suicide risk, substance abuse, depression, anxiety, anger, self-injury, eating disorders, and attention-deficit/hyperactivity disorder; and use of the *Diagnostic and Statistical Manual of Mental Disorders*.

Section IV, Types of Assessment, is the largest section and includes six chapters. The section is devoted to specific classes of assessment, including intelligence, ability, career development, and personality. In this edition you will find expanded coverage in areas such as high-stakes testing, projective assessments, and interpersonal assessment involving intimate partner violence and child abuse. Furthermore, recent revisions in intelligence and ability assessment are discussed. Section V, The Assessment Report, provides one chapter that outlines general guidelines for communicating assessment findings to a client and other stakeholders as well as developing a research report. The text also includes several appendices: names and acronyms of commonly used tests with publisher contact information (Appendix A), web links for ethical guidelines and assessment standards (Appendix B), common statistical formulas (Appendix C), a sample assessment report (Appendix D), and an answer key for "Test Your Knowledge" items (Appendix E).

In graduate courses that cover the use of tests and other assessment procedures in counseling, information about the various tools is typically covered, but the actual use of psychological assessment procedures in counseling often must be learned through trial and error. This text should help remedy that situation by providing information to assist the counselor in choosing, administering, and interpreting assessment procedures as part of the counseling process.

Acknowledgments

I thank Carolyn Baker, American Counseling Association (ACA) director of publications, for her support and responsiveness throughout the writing and production process. I am also grateful for the work of Bonny Gaston, production manager, and the other ACA staff members who made this edition possible.

I appreciate the contributions of Albert B. Hood and Richard W. Johnson to the practice of psychological assessment in general and as authors of the previous editions of this text. The counseling profession has certainly been influenced by their countless achievements in research and practice, and I am humbled to build on their work in this fifth edition.

Finally, I am thankful for my students and mentors in the assessment world who remind me every day of the important role of assessment.

• • •

About the Author

Danica G. Hays, PhD, LPC, NCC is an associate professor of counseling and chair of the Department of Counseling and Human Services at Old Dominion University. She is a recipient of the Outstanding Research Award, Outstanding Counselor Educator Advocacy Award, and Glen E. Hubele National Graduate Student Award from the American Counseling Association as well as the recipient of the Patricia B. Elmore Excellence in Measurement and Evaluation Award and President's Special Merit Award from the Association of Assessment in Counseling and Education (AACE). Hays served as founding editor of *Counseling Outcome Research and Evaluation*, a national peer-refereed journal of the AACE, and is editor of *Counselor Education and Supervision*, a national peer-refereed journal of the Association for Counselor Education and Supervision. She served as president of the AACE in 2011–12. Her research interests include qualitative methodology, assessment and diagnosis, trauma and gender issues, and multicultural and social justice concerns in counselor preparation and community mental health. She has published numerous articles and book chapters in these areas and coauthored or coedited five books to date: *Developing Multicultural Counseling Competence: A Systems Approach*; *Qualitative Inquiry in Clinical and Educational Settings*; *Mastering the National Counselor Exam and the Counselor Preparation Comprehensive Exam*; *The ACA Encyclopedia of Counseling*; and *A Counselor's Guide to Career Assesment Instruments*.

• • •

section
I
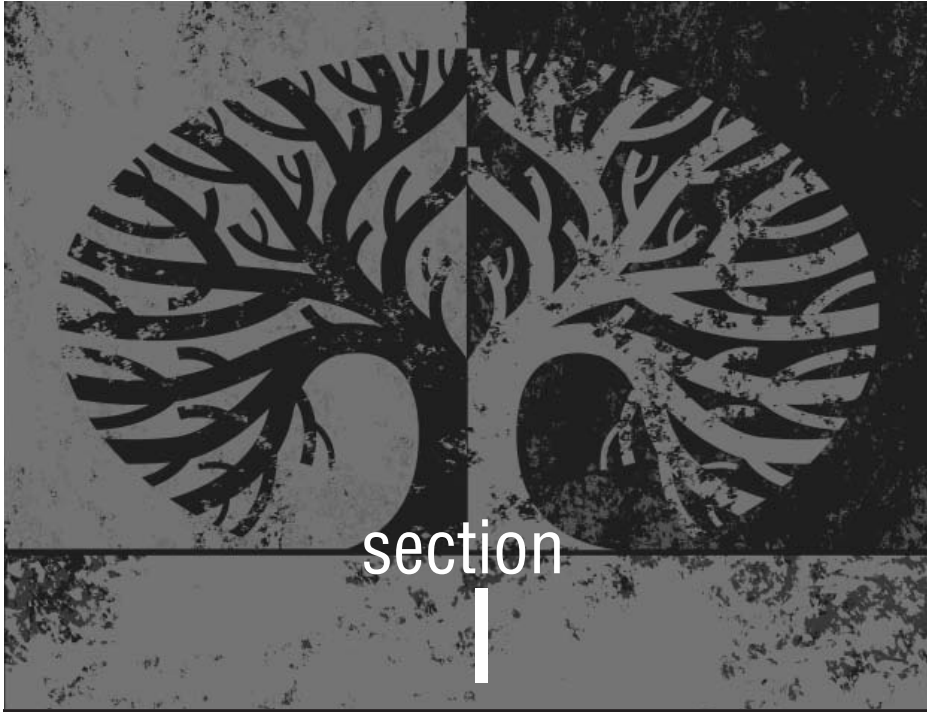
Foundations of
Assessment in Counseling

1

Use of Assessment in Counseling

What is assessment? What are the different ways counselors use assessment in the work they do? How did assessment become such an important part of counseling? In this chapter several key assessment terms are defined, and the purpose and uses of assessment are described. Then, a brief history of assessment is provided followed by a discussion of current attitudes toward assessment use. Finally, the chapter concludes with key questions and guiding principles of assessment in counseling.

 Test Your Knowledge

Respond to the following items by selecting T *for "True" or* F *for "False":*

☐ T   ☐ F   1. Assessment aids counseling by providing information for the client alone.
☐ T   ☐ F   2. *Assessment* and *test* are synonymous terms.
☐ T   ☐ F   3. Early group tests were used to assess intelligence and ability among World War I recruits.
☐ T   ☐ F   4. A problem-solving model is a useful method for conceptualizing the purpose of assessment.
☐ T   ☐ F   5. Personality assessment is the most significant area counselors are known for in assessment development.

Introduction to Assessment

Assessment is a part of everyone's daily lives. In any instance where someone has to make a judgment or solve a problem based on an outcome or information gained, assessment

is occurring. Individuals are recipients and participants of assessment data. Think back to your early memories of being assessed, tested, or evaluated in some way. Did it relate to a spelling or history test in school? Did it involve a report card you brought home? Were you being assessed for a disability or placed in a gifted program? Did you feel sad or anxious about something? Now, think of maybe more recent memories: taking a college or graduate entrance exam, discussing with a physician or counselor some symptom or issue you are experiencing, selecting a career path, interviewing for a job, even trying out a new recipe or working on a home improvement project, to name a few. No matter the memories—positive or negative—assessment occurs in various settings: schools, colleges, and universities; homes; health care settings; agencies; neighborhoods; communities; and so on.

It is not surprising, then, that assessment has always played an important part in counseling. From its inception, the field of counseling typically involved helping students with academic and career planning on the basis of test results. In recent years, the role of counseling (and the nature of assessment) has broadened to address a variety of concerns, such as self-esteem, shyness, personal growth, family and couple relationships, sexual identity, sexual abuse, cross-cultural communication, substance abuse, eating disorders, depression, anxiety, and suicidal ideation. Counselors also rely on assessment data for program planning and evaluation. Clients use assessment results to understand themselves better and to make plans for the future. The assessment process can be therapeutic in itself by helping clients to clarify goals and gain a sense of perspective and support.

Key Assessment Terms

There are many terms associated with assessment in counseling. In this section five key terms (i.e., *assessment, tests, measurement, variable,* and *psychometrics*) are presented. Throughout the text information on terms associated with these are outlined. Before defining these terms, it is important to define what the term *client* means throughout the text. A **client** may be an individual or group of individuals being assessed in various settings, such as counseling agencies, private practice settings, schools, colleges and universities, and career centers. A client can also refer to places or settings in general, such as in cases of program evaluation (e.g., a character education program). Finally, a client may be associated with objects or things such as dropout rates, divorce rates, violence, trauma, or neighborhoods. In essence, clients are people, places, or things.

Assessment

Assessment is an umbrella term for the evaluation methods counselors use to better understand characteristics of people, places, and things. Other terms used interchangeably in counseling to describe assessment are *appraisal* and *evaluation*. For most purposes, assessment can be conceptualized in terms of problem solving (Brown-Chidsey, 2005; Lovitt, 1998; Nezu & Nezu, 1993). The *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 1999) defines assessment as "any systematic method of obtaining information from tests and other sources, used to draw inferences about characteristics of people, objects, or programs" (p. 172). The first part of the definition ("any systematic method of obtaining information from tests and other sources") indicates that a broad range of evaluation methods—such as standardized tests, rating scales and observations, interviews, classification techniques, and records—may be used as a means of obtaining data about clients. The second part of the definition ("used to draw inferences about characteristics of people, objects, or programs") emphasizes the use of assessment data to help counselors understand their clients and the situations in which clients find themselves. Collectively, these two definition parts refer to a broad process of tool selection, administration and interpretation of data to provide a basis for forming

and testing hypotheses regarding the nature of a client's issues, and possible treatment approaches. The assessment process is discussed in more depth in Chapters 2, 6, and 15.

Some of the common assessment categories discussed in this text are intelligence (Chapter 9), ability (Chapter 10), career (Chapters 11 and 12), and personality (Chapter 13). These categories include both formal and informal assessment methods (see Chapter 2). Following are brief definitions of each category:

- **Intelligence assessment**: evaluation of cognitive abilities such as communication, reasoning, abstract thought, learning, and problem solving. Intelligence has been defined in many ways, although intelligence assessment is primarily measured through tests geared toward more traditional definitions.
- **Ability assessment**: assessment of acquired information (*achievement*) or an ability to acquire information (*aptitude*) about a particular subject matter or domain. Ability assessments are typically used for educational purposes, although some career and intelligence assessments may also be categorized as ability measures.
- **Career assessment**: measure of a client's career development process as well as the content domains of that process. Process-oriented variables include career readiness, concerns, planning, and maturity. Content domains involve career values and interests inventories. Career assessment can involve individual tools or more comprehensive assessment programs.
- **Personality assessment:** examination of individual attributes, types, and traits related to cognitions, emotions, actions, and attitudes. Personality assessment can be classified as structured (objective) or unstructured (projective).

As you can see from these descriptions, assessment categories are not fixed and can overlap one another.

Tests

A **test** is a systematic and often standardized process for sampling and describing a behavior of interest for individuals or groups. Tests can measure past, present, and/or future behavior or some reflection or feeling toward a behavior of interest. Tests can be interpreted in reference to a test taker's previous performance (**self-referenced**), some objective or criterion (**criterion referenced**), or that of a standardization sample (**norm referenced**). Standardization and test norms are discussed in more depth in Chapter 6.

Questionnaires and inventories, such as personality and interest inventories, elicit self-reports of opinions, preferences, and typical reactions to everyday situations. In practice, questionnaires and inventories also are often referred to as tests if they meet certain standardization criteria.

Tests are only one aspect of assessment. Assessment is a more comprehensive activity than testing by itself because it includes the integration and interpretation of test results and other evaluation methods. In sum, assessment involves judgments based on quantitative and qualitative descriptions of client data from a variety of sources.

Measurement

Measurement is a description of the degree to which a client possesses some characteristic. Traditionally, measurement deals with quantitative units, such as those associated with length (e.g., meter, inch), time (e.g., second, minute), mass (e.g., kilogram, pound), and temperature (e.g., Kelvin, Fahrenheit). In the physical sciences, measurement has been described as the actual or estimated magnitude of quantity relative to another (see Michell, 1997). The measurement concept has long been applied to the social sciences, such as when S. S. Stevens (1946) defined measurement as the assignment of numerals to objects or events according to some rule. These "rules" refer to scales of measurement (i.e., nominal,

ordinal, interval, and ratio; see Chapter 5). In addition, measurement in social sciences relates to providing data that meet some criteria, and thus tests are administered to assess the degree to which criteria are met.

Variable

Another key term is **variable**, which gets assigned a label through measurement. A variable refers to a construct or concept that can take on more than one value. Values can be qualitative or quantitative. For example, qualitative variables can include groupings such as gender, ethnicity, sports team, and hair color; they tend to involve **categorical variables**. Quantitative variables might include **continuous variables** (i.e., variables measured on some continuum), such as test scores, age, and rank. In assessment, you will encounter several types of variables: **independent variables** (preexisting variable or variable able to be manipulated that is assumed to influence some outcome), **dependent variables** (construct affected by the independent variable; also known as an outcome or response variable), and **extraneous variables** (a "noise" variable that impacts a dependent variable yet is unrelated to the assessment process—also known as a *confounding variable*).

Psychometrics

Psychometrics is the study of measurement technique and theory. Although a lengthy discussion is beyond the scope of this text, psychometricians have proposed common theories and techniques such as classical test theory, item response theory, Rasch modeling, factor analysis, and structural equation modeling. Classical test theory and its common concepts of measurement error, reliability, and validity are discussed in Chapter 5.

Purpose of Assessment in Counseling

Now that you have a basic understanding of the general terminology, let's take a look at how and why assessment is used in counseling. Assessment is beneficial in counseling because it provides information for both counselors and clients so they can understand and respond to client concerns as well as plan and evaluate programs. In addition, it can be therapeutic and can help clients understand both their past and present attitudes and actions as well as their plans for the future. Thus, assessments serve a diagnostic use, help to evaluate client progress, and are useful to improve or promote client awareness, knowledge, and skills. Gregory (2011) further cited several test uses: classification (i.e., program placement, screening, and certification), diagnosis and treatment planning, client self-knowledge, program evaluation, and research to guide counseling theory and technique development. Whichever purpose(s) counselors cite as the reason for assessment, it is important to convey this purpose to the client throughout the assessment process. That is, assessment should be part of the learning process for a client rather than something that is tacked on to counseling sessions.

Because performing an assessment is similar to engaging in problem solving, the five steps in a problem-solving model can be used to describe a psychological assessment model (Chang, D'Zurilla, & Sanna, 2004; D'Zurilla & Goldfried, 1971; D'Zurilla & Nezu, 1999). Depending on a client's problem-solving style, he or she will have varying levels of success in resolving a problem (Nezu, D'Zurilla, Zwick, & Nezu, 2004). Following is a brief description of the five steps involved (see Table 1.1 for specific ways the model relates to the assessment process):

1. **Problem Orientation.** This first step assesses how a problem is viewed (can be positive or negative) and requires the client to recognize and accept the problem. With completion of this step, the client and counselor can begin to approach it in a systematic fashion as indicated by the problem-solving model.

Table 1.1
Assessment and Problem-Solving Steps

Problem-Solving Step	Assessment Purpose	Counseling Example
Problem Orientation: stimulate counselors and clients to consider various issues	Almost any assessment procedure can be used to increase sensitivity to potential problems. Instruments that promote self-awareness and self-exploration can stimulate clients to cope with developmental issues before they become actual problems. Surveys of groups or classes can help counselors identify common problems or concerns that can be taken into account in planning programs for clients.	A counselor uses a needs assessment, such as an alcohol screening inventory, to identify areas of focus.
Problem Identification: clarify the nature of a problem or issue	Assessment procedures can help clarify the nature of the client's problem and ultimately strengthen communication and the overall counseling relationship as well as clarify goals. For example, screening inventories or problem checklists can be used to assess the type and the extent of a client's concerns. Personal diaries or logs can be used to identify situations in which the problem occurs. Personality inventories can help counselors and clients understand personality dynamics underlying certain situations.	A counselor can provide a diagnosis to classify a set of concerns or symptoms, such as the case of a relationship difficulty or an anxiety disorder.
Generation of Alternatives: suggest alternative solutions	Assessment procedures enable counselors and clients to identify alternative solutions for client problems, view problems from different angles, as well as stimulate new learning. For example, an assessment interview can be used to determine what techniques have worked for the client in the past when faced with a similar problem. Checklists or inventories (such as a study skills inventory or work skills survey) yield data that can be used to generate alternatives.	A counselor uses an interest inventory to suggest alternative career choices for a client. A counselor helps a client identify positive self-statements to create alternatives.
Decision Making: determine appropriate treatment for the client	Counselors use assessment materials to help clients weigh the attractiveness of each alternative and the likelihood of achieving each alternative. The likelihood of achieving different alternatives can be evaluated by expectancy (or experience) tables that show the success rate for people with different types of test scores or characteristics (Anastasi & Urbina, 1997). Balance sheets or decision-making grids enable clients to compare the desirability and feasibility of various alternatives (Howard, 2001).	A counselor uses a values clarification exercise to assess the attractiveness of various alternatives. A counselor uses a personality inventory to help select a client's intervention.
Verification: evaluate the effectiveness of a particular solution	Assessment procedures to verify success may include goal attainment scaling (Kiresuk, Smith, & Cardillo, 1994), self-monitoring techniques (Korotitsch & Nelson-Gray, 1999), the readministration of tests that the client completed earlier in counseling, client satisfaction surveys, and the use of outcome questionnaires (Wells, Burlingame, Lambert, & Hoag, 1996). In addition to serving as a guide for the counseling process, verification efforts also provide a means of accountability for the counseling agency.	Client feedback can be used to make changes to an intervention. A counselor can request a client self-monitoring exercise to assess maintenance of change.

2. **Problem Identification.** This step involves the counselor and the client attempting to identify the problem in as much detail as possible. A client is more likely to continue in counseling and to achieve positive outcomes if the counselor and client agree on the nature of the problem (Busseri & Tyler, 2004). Identification of the problem also aids in communication with others, such as referral sources, family, and friends.

3. **Generation of Alternatives.** In the third step, the counselor and client generate alternatives to help resolve the problem. Counselors use assessment procedures to assist clients in discovering strengths on which they can build to overcome difficulties or enhance development.

4. **Decision Making.** In this step clients anticipate the consequences of the various alternatives. According to classical decision theory, choice is a function of the probability of success and the desirability of the outcome (Horan, 1979). This equation emphasizes the importance of assessing both the likelihood of success of various alternatives and the attractiveness of those alternatives for the client. Clients will usually want to consider those alternatives that maximize the likelihood of a favorable outcome.

5. **Verification.** The counselor in this final step should discuss with the client how the client will know when the problem has been solved. This step requires that goals be clearly specified, that they be translated into specific behavioral objectives, and that the possibility for progress in accomplishing these goals be realistically viewed. Counselors are to verify the effectiveness of their interventions.

In understanding the purpose of assessments in counseling, it is also important to understand what purposes *do not* characterize assessments—particularly tests. Tyler (1984) highlighted several things tests do not measure:

- Tests cannot measure unique characteristics, only attributes common to many people.
- Individual assessment cannot be used to make group comparisons; counselors can only estimate how well an individual will function in a culture for which the assessment is appropriate. Tests are not suitable for comparing groups that are not identical.
- Because test scores are plotted on a distribution of scores, we tend to infer scores on the distribution ends as "good versus poor" (p. 48). To this end counselors often evaluate scores without examining the appropriate norms and without considering that highness or lowness of scores do not measure a client's worth.
- On a related note, a "good score" may measure some absence of pathology (such as in personality assessment) rather than universal attributes to which humans aspire.
- Tests cannot measure innate characteristics. Although there are biological components to some attributes (such as intelligence), beginning at birth these components interact with various environmental factors that further shape responses. Thus, even if counselors can assume that intelligence tests are free from cultural bias (which they are not), responses on intelligence tests are a combination of hereditary tendencies and individual responses to particular environments.
- Test scores are not final measurements of anything but outlets—in conjunction with multiple assessment sources—for facilitating client growth. Clients' high-stakes decisions should not be based solely on test scores.

To this list I add the following:

- Tests cannot measure all things equally. Some things, such as reaction time, may be easier to assess than others, such as intelligence or disability.
- Tests are not necessarily indicative of the totality of behaviors, attitudes, or skills. Tests are only one sampling of these areas and thus should be evaluated as such.
- Test results are not always useful. In fact, they are often misused in decision making and applied to individuals inappropriately. Concerns include the following:

(a) misuse with minority groups, who may differ significantly from the population for whom the test was developed; (b) use of tests to label or stereotype a person based on the test results; and (c) the disproportionate influence of tests in so-called high-stakes decisions, such as selection for college or employment. In some situations, too much emphasis may be placed on test results, often because of their quantitative or scientific nature; in other situations, pertinent test information may be disregarded, especially if it conflicts with an individual's personal beliefs or desires.

Please see the tip sheet at the end of the chapter for sound assessment procedures.

History of Assessment

Let's step back from how assessment is used (and should not be used) today and reflect on how counselors began using assessment in the first place. This section presents early key developments in intelligence, ability, interest, and personality assessment; most of these developments occurred from the late-19th to mid-20th century. After reviewing this brief history of assessment, perhaps you can understand why current assessment practices across various settings exist, how beneficial assessment can be to counselors and clients, what mistakes those who have administered tests have made, and why certain practices should be continually challenged and scrutinized. Table 1.2 provides a time line of major assessment developments.

Table 1.2
Key Historical Events in Assessment

Year	Event
2200 BC	Chinese tested aspiring public officials for work evaluations and promotion decisions
1879	Wundt founded the first psychological laboratory, conducting several experiments with brass instruments
1880s	Galton initiated the social science testing movement, measuring individual differences in sensory processes
1890	James Cattell coined the term *mental test*
1900	Esquirol and Seguin performed formalized intelligence assessment in medical communities
	College Entrance Examination Board (CEEB, now known as the College Board) created
1905	Binet–Simon scale developed; revised in 1908 and 1911
	Goddard misused test at Vineland Training School and Ellis Island
1916	Stanford–Binet Scale created
1917	Army Alpha and Beta tests, Woodworth Personal Data Sheet developed
1921	Rorschach Inkblot Test published
1923	Terman and colleagues develop the Stanford Achievement Test
1926	Scholastic Aptitude Test published by CEEB
	Strong Vocational Interest Blank created
1938	Buros Center for Testing developed the Mental Measurements Yearbook
	Thematic Apperception Test (TAT) created
1939–1950s	Wechsler Scales of Intelligence developed
1940s	Myers–Briggs Type Indicator published
1942	Minnesota Multiphasic Personality Inventory (MMPI) created
1947	Educational Testing Service created
1964	Civil Rights Act
1974	Family Educational Rights and Privacy Act
1975	Education of All Handicapped Children Act (Pub. L. No. 94-142)
1990	American With Disabilities Act (Pub. L. No. 101-336)
1995	Individuals With Disabilities Education Act Amendments
1996	Health Insurance Portability and Accountability Act (HIPAA)
2001	No Child Left Behind Act (Pub. L. No. 107-110)[a]

[a]Additional legislation concerning assessment is presented in Chapter 3.

The discussion focuses primarily on historical events from the mid-1800s until present day. The earliest form of testing dates back to Ancient China, where in 2200 BC (over 4,400 years ago!) the Chinese used a civil service testing program to assess, evaluate, place, and promote its employees. Every three years officials tested employees on five topics: civil law, military affairs, geography, agriculture, and revenue. The testing program was abolished in 1906 after several complaints and questions about its administration and utility, although the program influenced American and European civil service program placements in the 1800s. Let's jump 4,000 years later, when individuals began recording formal assessment procedures in the social sciences.

Developments in Individual Intelligence Assessment

In the mid to late 1800s, there was an increasing interest in studying individual human differences, particularly concerning intelligence. Charles Darwin's *Origin of Species,* with its focus on genetic variation and evolution and thus individual differences, was used as a case for testing human differences. The study of intelligence increased, given its links to discussions of evolution at the time. In the late 1800s, experimental psychologists—primarily Wilhelm Wundt (1832–1920), Sir Francis Galton (1822–1911), and James Cattell (1860–1944)—revolutionized the way intelligence and ability were measured. They focused on quantifiable measures of sensory processes (e.g., visual and auditory processes, reaction time) using brass instruments in human laboratories to indicate intelligence. Wundt, one of the founders of modern psychology, studied mental processes and was able to highlight that individual differences do exist (even though his interests were more in understanding general features of the psyche). Galton, Charles Darwin's half-cousin, was considered a prolific scholar and creator of several significant mathematical and scientific concepts, such as correlation, regression, and central tendency statistics; meteorology; fingerprinting; hearing loss; and heredity. He is considered the founder of eugenics, claiming genetics was the determinant of genius and mental competence differences. Galton is also referred to as the founder of mental tests; he demonstrated that individual cognitive differences do exist and can be measured. Although Galton's tests are now considered simplistic, in the 1880s and 1890s he tested over 17,000 individuals on physical (e.g., height, weight, head size, length of middle finger) and behavioral (e.g., hand squeeze tests, lung capacity, visual acuity, reaction time) domains to indicate intelligence (Forrest, 1995). Cattell, who studied with both Wundt and Galton, coined the term *mental test* and articulated 10 mental tests (presence of each indicates intelligence) similar to Galton's. Examples include strength of hand squeeze, rate of hand movement, degree of rubber tip on forehead pressure needed to cause pain, weight differentiation of identical-appearing boxes, reaction time for sound, and number of letters repeated upon hearing them.

Thus, in the early 20th century there was increased interest in what was called mental testing—now referred to as intelligence testing. Work in the medical community, where there was an increasing distinction between emotional problems and intellectual disabilities (i.e., mental retardation), set the stage for more formalized intelligence assessment. In Paris, France, two physicians—Jean Esquirol (1772–1840) and Edouard Sequin (1812–1880)—studied language use, identified various levels of verbal intelligence, and examined motor function in patients to initially conceptualize performance intelligence. Sequin was particularly instrumental in performance tests, with the development of the **Sequin form board**. This board is still used today in neuropsychological tests and involves fitting 10 blocks within designated slots on an upright board (DuBois, 1970).

The first intelligence test was developed by Alfred Binet (1857–1911), who as minister of public instruction in Paris introduced the 1905 **Binet–Simon scale** in collaboration with his Sarbonne colleague Theodore Simon (1972–1961). They were commissioned by the French government to assess children with intellectual disabilities. They developed a scale that contained 30 tasks and was designed to assist in educational placement; the scale relied on

a general factor of intelligence, versus lower level sensory processes, and it relied heavily on verbal ability (Goodenough, 1949). Sample tasks included following a movement with eyes, repeating three spoken digits, repeating a sentence of 15 words, putting three nouns (e.g., *Paris, river, fortune*) in a sentence, reversing the hands of a clock, and defining abstract words by distinguishing between them (e.g., *boredom* and *weariness*).

In 1908 Binet and Simon revised the scale and dropped many of the simpler tasks (used previously to classify those with severe intellectual disabilities) and added higher level tasks. The revised scale contained 58 tasks. The concept of *mental level* was developed for this scale, later referred to as *mental age*. Using a standardization sample of 300 children ages 3–13, Binet and Simon were able to calculate the number of items passed by a majority of children of a certain age. In the 1911 scale revision, mental level was further refined, whereas each age level had five associated tasks. Ultimately, mental age was compared against chronological age, and others suggested that an **intelligence quotient (IQ)** be developed.

The 1908 Binet–Simon scale was significantly misused, particularly in the United States at the Vineland Training School in New Jersey—a school for "feebleminded" children. Henry Goddard, the first American psychologist to translate the Binet–Simon scale, tested 378 school residents, classifying 73 as "idiots," 205 as "imbeciles," and 100 as "feebleminded." He then tested 1,547 "normal" children and noted that 3% of the sample could be classified as feebleminded. Goddard was invited in 1910 to Ellis Island to assess intelligence among immigrants. Through his assessments he increasingly emphasized that the rate of feeblemindedness was much higher among immigrant populations (Gould, 1996). Specifically, he found feeblemindedness for his small immigrant samples at the following rates: 83% of Jews, 80% of Hungarians, 79% of Italians, and 87% of Russians (see Goddard, 1917). These findings were used to make the case that feebleminded individuals were a threat to social order, and thus low IQ immigrants should be deported. Unfortunately, Goddard's writings heavily influenced immigration restrictions (Gould, 1996). Fortunately, Howard Knox developed several performance tests to be used with immigrants and was able to debunk some of Goddard's conclusions (see Figure 1.1). Knox's work highlighted the necessity of future intelligence tests containing performance or nonverbal parts (Richardson, 2003).

In 1916 Lewis Terman (1857–1956) and colleagues at Stanford University revised the scale (i.e., **Stanford Binet Scale**) and suggested that IQ be calculated by dividing an

Figure 1.1
Knox Administering Performance Tests to Ellis Island Immigrants, 1912–1916

individual's mental age by his or her chronological (actual) age and multiplying that fraction times 100. For example, if a 10-year-old child performs at the level of an 11-year-old (mental age), his IQ would equal 110. The Stanford Binet Scale was the gold standard in intelligence testing for several decades and is now in its fifth edition. In 1927 Charles Spearman (1863–1945) conceptualized a general and specific factor theory of intelligence, and in 1941, Raymond Cattell (1905–1998) coined the terms *fluid* and *crystallized intelligence*, influencing how future assessments were constructed and interpreted. Other developments, including the Wechsler scales (Wechsler, 1949, 1955), are discussed in Chapter 9.

Developments in Group Intelligence Assessment

Although individual intelligence tests had made significant contributions in a short time—including a beginning understanding of the misuse of testing—test administrators learned quite quickly that group testing was more time-efficient. The use of group tests became increasingly warranted when during World War I there was a need to screen new recruits. Robert Yerkes (1876–1956), while president of the American Psychological Association (APA), developed the first two group tests: the Army Alpha and Army Beta intelligence tests. The **Army Alpha** measured verbal ability, numerical ability, ability to follow directions, and knowledge of information. The **Army Beta** was a nonverbal counterpart to the Army Alpha and was used to evaluate illiterate or non-English-speaking recruits. During World War I over 1.5 million recruits were administered these tests for placement purposes. Figure 1.2 provides sample items from these tests.

Developments in Ability Assessment

Eventually group tests were used to assess more than just intelligence. Group aptitude tests were developed after World War II, as more specialized careers (e.g., flight engineers, pilots, navigators) required more stringent selection for flight schools, and previous intelligence tests (i.e., Army Alpha, Army Beta) were not sufficient (Goslin, 1963). In essence, it was clear that previously developed intelligence tests were limited because not all important job functions were covered.

With greater attention paid to education after World War I, ability testing flourished in public schools and higher education. The Army Alpha and Beta tests were released for public use and became the model for future ability tests. Edward Thorndike (1874–1949) spearheaded the development of several standardized achievement tests in public schools, including rating scales, spelling tests, arithmetic reasoning, and handwriting assessments, to name a few. These tests were distinguished from earlier ones by a more reliable administration format. The **College Entrance Examination Board** (CEEB) was established in 1900 to regulate group test use in colleges and universities. In 1926 the CEEB developed the first aptitude test for college admissions, the **Scholastic Aptitude Test** (SAT; Goslin, 1963). Functions of the CEEB were subsumed under the Educational Testing Service (ETS) in 1948, creator of modern-day assessments such as the **Graduate Record Examination** (GRE), **Law School Admissions Test** (LSAT), **Praxis**, and **Test of English as a Foreign Language** (ToEFL). In addition, Terman and his colleagues in 1923 developed a standardized achievement test, the **Stanford Achievement Test** (SAchT). So, as a student in the 1930s and 1940s, you certainly would have had significant exposure to testing!

Developments in Career Assessment

Upon reading about major developments in intelligence and ability testing, it may be evident to you that there was some focus on vocational assessment at the same time. Similar to how intelligence and ability assessments developed from societal needs (e.g., educational reform and placement, military screening), career assessment developed in response