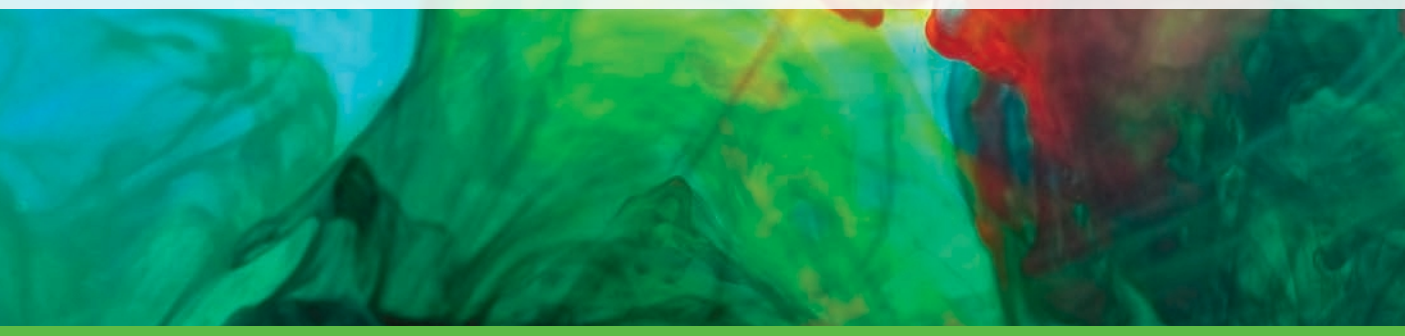
GLOBAL EDITION

Psychological Testing

History, Principles, and Applications

SEVENTH EDITION

Robert J. Gregory



ALWAYS LEARNING

PEARSON

Psychological Testing

Recently, Gardner (1998) has added three tentative candidates to his list of intelligences. These are naturalistic, spiritual, and existential intelligences. Naturalistic intelligence is the kind shown by people who are able to discern patterns in nature. Charles Darwin would be a prime example of such a person. Gardner believes that the evidence for this kind of intelligence is relatively strong. In contrast, spiritual intelligence (a concern with cosmic and spiritual issues in one's development) and existential intelligence (a concern with ultimate issues, including the meaning of life) are less well proved as independent intelligences. In general, the theory of multiple intelligence is compelling in its simplicity, but there is little empirical investigation of its validity.

STERNBERG AND THE TRIARCHIC THEORY OF SUCCESSFUL INTELLIGENCE

Sternberg (1985b, 1986, 1996) takes a much wider view on the nature of intelligence than most previous theorists. In addition to proposing that certain mental mechanisms are required for intelligent behavior, he also emphasizes that intelligence involves adaptation to the real-world environment. His theory emphasizes what he calls successful intelligence or "the ability to adapt to, shape, and select environments to accomplish one's goals and those of one's society and culture" (Sternberg & Kaufman, 1998, p. 494).

Sternberg's theory is called *triarchic* (ruled by three) because it deals with three aspects of intelligence: componential intelligence, experiential intelligence, and contextual intelligence. Each of these types of intelligence has two or more subcomponents. The entire theory is outlined in Table 5.5.

Componential intelligence, also known as analytical intelligence, consists of the internal mental mechanisms that are responsible for intelligent behavior. The components of intelligence serve three different functions. *Metacomponents* are the executive processes that direct the activities of all the other components of intelligence. They are responsible for determining the nature of an intellectual problem, selecting a strategy for solving it, and making sure that the task is completed. The metacomponents receive constant feedback as to how things are going in problem solving. Persons who are strong on the metacomponential aspect of intelligence are very good at allocating their intellectual resources.

TABLE 5.5 An Outline of Sternberg's Triarchic Theory of Intelligence

Componential (Analytical) Intelligence
Metacomponents or executive processes (e.g., planning)
Performance components (e.g., syllogistic reasoning)
Knowledge acquisition components (e.g., ability to acquire vocabulary words)
Experiential (Creative) Intelligence
Ability to deal with novelty
Ability to automatize information processing
Contextual (Practical) Intelligence
Adaptation to real-world environment
Selection of a suitable environment
Shaping of the environment

Source: Summarized from Sternberg, R. J. (1986). *Intelligence applied: Understanding and increasing your intellectual skills.* San Diego, CA: Harcourt Brace Jovanovich.

In a problem-solving study using novel forms of analogies, Sternberg (1981) found that higher intelligence is associated with spending relatively more time on global or higher-order planning, and relatively less time on local or lower-order planning. For example, consider this analogy problem:

Man: Skin:: (Dog, Tree):(Bark, Cat)

The examinee must choose the two correct terms on the right that will complete the analogy. (The correct choices are Tree and Bark.) Using reaction time measures for a series of such novel or nonentrenched problems, Sternberg (1981) found that persons of higher intelligence spend more time in global planning—forming a macrostrategy that applies to this and similar problems—than did persons of lower intelligence. Thus, a crucial aspect of intelligence is knowing when to step back and allocate intellectual effort instead of obtusely attacking a difficult problem.

Performance components are the well-entrenched mental processes that might be used to perform a task or solve a problem. These aspects of intelligence are the ones that are probably measured the best by existing intelligence tests. Examples of performance components include short-term memory and syllogistic reasoning.

Knowledge acquisition components are the processes used in learning. Sternberg has emphasized that in order to understand what makes some people more skilled than others, we must understand their increased capacity to acquire those skills in the first place. A case in point is vocabulary knowledge, which is learned mainly in context rather than through direct instruction. More-intelligent persons are better able to use surrounding contexts to figure out what a word means; that is, they have greater knowledge-acquisition skills. Their increased vocabulary results, in large measure, from their increased ability to "soak up" the meanings of words they see and hear in their environment. Thus, vocabulary is an excellent measure of intelligence because it reflects people's ability to acquire information in context.

The second aspect of Sternberg's theory involves experiential intelligence. According to the theory, a person with good **experiential intelligence** is able to deal effectively with novel tasks. Experiential intelligence is also known as creative intelligence. This aspect of his theory explains why Sternberg is so critical of most intelligence tests. For the most part, the existing tests measure things already learned by presenting tasks that the subject has already encountered. According to Sternberg, intelligence also involves the capacity to learn and think within new conceptual systems, not just to deal with tasks already encountered. A second aspect of experiential intelligence is the ability to automatize or "make routine" tasks that are encountered repeatedly. An example of automatizing that applies to most of us is reading, which is carried out largely without conscious thought. But any task or mental skill can be automatized, if it is practiced enough. Playing music is an example of an extremely high-level skill that can become automatized with enough practice.

The third aspect of Sternberg's theory involves contextual intelligence. **Contextual intelligence**, also known as practical intelligence, is defined as "mental activity involved in purposive adaptation to, shaping of, and selection of real-world environments relevant to one's life" (Sternberg, 1986, p. 33). This aspect of Sternberg's theory appears to acknowledge that human behavior has been shaped by selective pressures during our evolutionary history. Contextual intelligence has three parts: adaptation, selection, and shaping.

Adaptation refers to developing skills required by one's particular environment. Successful adaptation will differ from one culture to the next. In the pygmy cultures of Africa, adaptation might involve the ability to track elephants and kill them with poison-tipped spears. In the Western industrial nations, adaptation might involve presenting oneself favorably in a job interview.

Selection might be called niche finding. This aspect of contextual intelligence involves the ability to leave the environment we are in and to select a different environment more suitable to our talents and needs. Feldman (1982) has illustrated how selection can operate in the career choices of gifted children, thereby determining whether they are highly accomplished as adults. She followed up on the Quiz Kids who were featured in radio and television shows of the 1950s. These were extremely bright children by conventional standards, most with IQs of 140 and higher. A few became highly successful as adults. However, most of them led rather ordinary lives, devoid of the spectacular accomplishments that might have been predicted from their childhood precocity. Those who were most successful had found occupations highly suited to their abilities and interests. In sum, they had selected environmental niches that fitted them well. Sternberg would argue that the ability to select such environments is an important aspect of intelligence.

Shaping is another way to improve the fit between oneself and the environment, especially when selection of a new environment is not practical. In this application of contextual intelligence, we shape the environment itself so that it better fits our needs. An employee who convinces the boss to do things differently has used shaping to make the work environment more suited to his or her talents.

Sternberg (1993) has developed a research instrument based on his theory and has used the test to examine the validity of the triarchic approach. The Sternberg Triarchic Abilities Test (STAT) is unique in going beyond the typical questions that invoke analytical intelligence; the test includes creative and practical questions as well. For example, in one subtest examinees are presented with a map of an area, such as an entertainment park, and then must answer questions about navigating effectively through the area shown in the map (practical intelligence). In another subtest examinees are presented with verbal analogies preceded by incorrect, counterfactual premises (e.g., money falls off trees). Examinees must solve the analogies as though the counterfactual premises were true (creative intelligence). In factor-analytic studies of American, Finish, and Spanish samples, the triarchic model was a better fit to the data than the usual outcome of finding a single factor of general intelligence (Sternberg, Castejon, Prieto, Hautamaki, & Grigorenko, 2001).

Although Sternberg's triarchic theory is the most comprehensive and ambitious model yet proposed, not all psychometric researchers have rushed to embrace it. Detterman (1984) cautions that we should investigate the basic cognitive components of intelligence before introducing higher-order constructs that may be unnecessary. Rogoff (1984) questions whether the three subtheories (componential, experiential, contextual) are sufficiently linked. Other comments on the triarchic theory can be found in *Behavioral and Brain Sciences* (1984, pp. 287–304).

Whatever the final verdict on the triarchic theory of intelligence, Sternberg's insistence that intelligence has several components not measured by traditional tests rings true to anyone who has studied or administered these tests. He cites the case of a colleague who was asked to test a number of residents at an institution for those with mental retardation. These residents had just planned and successfully executed an escape from the security-conscious school, a feat requiring high levels of practical intelligence. Yet, when administered the Porteus Maze Test (Porteus, 1965), a standardized test reputed to involve planning ability, they could not solve even the simplest maze correctly. Sternberg (1986) has made it clear that intelligence just has too many components to be measured by any single test.

Topic 5B Individual Tests of Intelligence and Achievement

Individual intelligence testing is one of the major achievements of psychology since the founding of the discipline. In response to the success of the Binet-Simon scales in the early 1900s, psychologists developed and refined dozens of individual tests of intelligence patterned after this pathbreaking instrument. The explosive growth in group tests of intelligence, fostered by the enthusiastic acceptance of the Army Alpha and Beta tests during and after World War I, also provided impetus to the individual testing movement. Many contemporary individual tests of intelligence owe their lineage to Binet, Simon, and the Army testing programs.

The successful application of intelligence tests inspired educators and psychologists to look for ways to appraise the academic progress of students with school-based achievement tests. In turn, this led to the puzzling discovery that many children of normal or even superior intelligence lagged far behind in school achievement. From this discovery, the concept of learning disability gradually developed, and a whole new field of assessment was born.

The purpose of this topic is to provide an overview of noteworthy approaches to the testing of individual intelligence and achievement, and to introduce the reader to the essentials of learning disability assessment. However, an exhaustive survey of individual cognitive tests is simply beyond the scope of this or any other basic reference. New and revised tests appear practically every month, and thousands of new research findings are published every year. We have chosen to review tests that are widely used or that illustrate interesting developments in theory or method. Readers can find information on additional tests in the *Mental Measurements Yearbook* series, now published every two or three years by the Buros Institute.

ORIENTATION TO INDIVIDUAL INTELLIGENCE TESTS

The individual intelligence tests reviewed in this topic include the following:

- Wechsler Adult Intelligence Scale-IV (WAIS-IV)
- Wechsler Intelligence Scale for Children-IV (WISC-IV)
- Stanford-Binet: Fifth Edition (SB5)
- Detroit Tests of Learning Aptitude-4 (DTLA-4)
- Cognitive Assessment System-II (CAS-II)
- Kaufman Brief Intelligence Test-2 (KBIT-2)

Collectively, these instruments probably account for 95 percent of the intellectual assessments conducted in the United States.

The Wechsler scales have dominated intelligence testing in recent years, but they are by no means the only viable choices for individual assessment. Many other instruments measure general intelligence just as well—some would say better. Consider the implications of a now familiar

observation: For large, heterogeneous samples, scores on any two mainstream instruments (e.g., Wechsler, Stanford-Binet, McCarthy, Kaufman scales) typically correlate .80 to .90. Often the correlation between two mainstream instruments is nearly as high as the test–retest correlation for either instrument alone. For purposes of producing a global score, it would appear that any well-normed mainstream intelligence test will suffice.

But producing an overall score is not the only goal of assessment. In addition, the examiner usually desires to gain an understanding of the subject's intellectual functioning. For this purpose, the overall IQ is important, but there are instances in which the global score may be irrelevant or even misleading. To understand a referral's intellectual functioning, the examiner should also inspect the subtest scores in search of hypotheses that might explain the unique functioning of that individual. Of course, examiners need to undertake subtest analysis cautiously, armed with research-based findings on the nature and meaning of subtest scatter for the test in use (Gregory, 1994b).

If the examiner's goal is to understand intellectual functioning and not merely to determine an overall score, the differences between tests become quite real. Every instrument approaches the measurement of intelligence from a different perspective and yields a distinctive set of subtest scores. Furthermore, a test well suited for one referral issue might perform abysmally in another context. For example, the WAIS-IV performs admirably in the testing of mild mental retardation but contains too few simple items for the effective assessment of persons with moderate or severe developmental disability.

A central axiom of assessment is that the choice of a testing instrument should be based on knowledge of its strengths and weaknesses as they pertain to the referral question. Put simply, the skilled examiner does not blindly rely on a single test for every referral! Instead, the skilled examiner flexibly chooses one or more instruments in light of the perceived assessment needs of the examinee. Each of the tests discussed in this topic has its special merits and also its particular shortcomings. The test user must know these strong and weak facets in order to choose the instruments best suited for each unique referral.

THE WECHSLER SCALES OF INTELLIGENCE

Beginning in the 1930s, David Wechsler, a psychologist at Bellevue Hospital in New York City, conceived a series of elegantly simple instruments that virtually defined intelligence testing in the mid- to late twentieth century. His influence on intelligence testing is exceeded only by the pathbreaking contributions of Binet and Simon. It is fitting that we begin the survey of individual tests with a historical summary of the Wechsler tradition, followed by a discussion of individual instruments.

Origins of the Wechsler Tests

Wechsler began work on his first test in 1932, seeking to devise an instrument suitable for testing the diverse patients referred to the psychiatric section of Bellevue Hospital in New York (Wechsler, 1932). In describing the development of his first test, he later wrote, "Our aim was not to produce a set of brand new tests but to select, from whatever source available, such a combination of them as would meet the requirements of an effective adult scale" (Wechsler, 1939). In fact, the content of his scales was largely inspired by earlier efforts such as the Binet scales and the Army Alpha and Beta tests (Frank, 1983). Readers who peruse *Psychological Examining in the United States Army,* a volume edited by Yerkes (1921) just after World War I, might be astonished to discover that Wechsler purloined dozens of test items from this source, many of which have survived to the present day in contemporary revisions of the Wechsler tests. Wechsler was not so much a creative talent as a pragmatist who fashioned a new and useful instrument from the spare parts of earlier, discontinued attempts at intelligence testing.

The first of the Wechsler tests, named the Wechsler-Bellevue Intelligence Scales, was published in 1939. In discussing the rationale for his new test, Wechsler (1941) explained that existing instruments such as the Stanford-Binet were woefully inadequate

for assessing adult intelligence. The Wechsler-Bellevue was designed to rectify several flaws noted in previous tests:

- The test items possessed no appeal for adults.
- Too many questions emphasized mere manipulation of words.
- The instructions emphasized speed at the expense of accuracy.
- The reliance on mental age was irrelevant to adult testing.

To correct these shortcomings, Wechsler designed his test specifically for adults, added performance items to balance verbal questions, reduced the emphasis upon speeded questions, and invented a new method for obtaining the IQ. Specifically, he replaced the usual formula

$$\text{IQ} = \frac{\text{Mental Age}}{\text{Chronological Age}}$$

with a new age-relative formula

$$\text{IQ} = \frac{\text{Attained or Actual Score}}{\text{Expected Mean Score for Age}}$$

This new formula was based on the interesting presumption—stated in the form of an axiom—that IQ remains constant with normal aging, even though raw intellectual ability might shift or even decline. The assumption of **IQ constancy** is basic to the Wechsler scales. As Wechsler (1941) put it:

> The constancy of the I.Q. is the basic assumption of all scales where relative degrees of intelligence are defined in terms of it. It is not only basic, but absolutely necessary that I.Q.'s be independent of the age at which they are calculated, because unless the assumption holds, no permanent scheme of intelligence classification is possible.

Although Wechsler's view has been largely accepted by contemporary test developers, it is important to stress that the assumption of IQ invariance with age is really a statement of values, a philosophical choice, and not necessarily an inherent characteristic of human nature.

Wechsler also hoped to use his test as an aid in psychiatric diagnosis. In pursuit of this goal, he divided his scale into separate verbal and performance sections. This division allowed the examiner to compare an examinee's facility in using words and symbols (verbal subtests) versus the ability to manipulate objects and perceive visual patterns (performance subtests). Large differences between verbal ability (V) and performance ability (P) were thought to be of diagnostic significance. Specifically, Wechsler believed that organic brain disease, psychoses, and emotional disorders gave rise to a marked $V > P$ pattern, whereas adolescent psychopaths and persons with mild mental retardation yielded a strong $P < V$ pattern. Subsequent research demonstrated many exceptions to these simple diagnostic rules, and also helped refine the nature of these two major elements of intelligence. For example, verbal intelligence is now better known as verbal comprehension, and performance intelligence is more commonly recognized as perceptual reasoning. Nonetheless, the distinction between verbal and performance skills has proved useful for many purposes, such as studying brain–behavior relationships, and examining age effects on intelligence. Wechsler's armchair division of subtests into verbal and performance sections, even though refined and extended by others, continues to endure as a major contribution to contemporary intelligence testing (Kaufman, Lichtenberger, & McLean, 2001).

General Features of the Wechsler Tests

Including revisions, David Wechsler and his followers have produced more than a dozen intelligence tests in a span of about 70 years. A major reason for the continued success of these instruments has been the faithful adherence to the familiar content and format first introduced in the Wechsler-Bellevue. By sticking with a single successful formula, Wechsler and company ensured that examiners could switch from Wechsler test to another with minimal retraining. This was not only good psychometrics but also shrewd marketing insofar as it guaranteed several generations of faithful test users.