
UNIT 5 ASSESSMENT OF INTELLIGENCE*

Structure

- 5.1 Objectives
- 5.2 Introduction
- 5.3 History of Assessment of Intelligence
 - 5.3.1 The Binet-Simon Scale: The First Intelligence Test
 - 5.3.2 The Revised Scales and the Advent of IQ
- 5.4 Individual Tests
 - 5.4.1 Stanford-Binet Scale of Intelligence
 - 5.4.2 The Wechsler Scales
 - 5.4.3 The Kaufman Scales
 - 5.4.4 Das-Naglieri Cognitive Assessment System
- 5.5 Group Tests
 - 5.5.1 Advantages and Disadvantages of Group Testing
 - 5.5.2 Examples of Group Intelligence Tests
- 5.6 Verbal and Nonverbal Tests
- 5.7 Culture-Fair Tests
- 5.8 Issues in Intelligence Testing
- 5.9 Let Us Sum Up
- 5.10 Key Words
- 5.11 Answers to Self Assessment Questions
- 5.12 Unit End Questions
- 5.13 References
- 5.14 Suggested Readings

5.1 OBJECTIVES

After reading this Unit, you will be able to:

- Present an account of the history of intelligence testing;
- Explain the essential properties of a test;
- Know and differentiate between the individual and group intelligence tests;
- Explain the difference between verbal and nonverbal intelligence tests; and
- Discuss the relevance of culture fair tests.

5.2 INTRODUCTION

In the previous Units (Units 3 and 4), you have learned about the definition, nature and theories of intelligence. You must have noted that psychologists are not unanimous regarding the nature of intelligence. Even, there is not a single definition accepted by all the theorists. So, how do we go about assessing intelligence? Assessment of intelligence is one of the most highly researched topics in psychology and consequently, a number of intelligence tests have been developed, most of which are based on a specific theory of intelligence. Assessing intelligence has implications for a wide range of areas, specifically for educational achievement and career recruitment.

In this Unit we will discuss about the history of intelligence testing, the types of tests and then discuss about certain issues concerning it.

* Dr. Manaswini Dash, Assistant Professor, Department of Psychology, Utkal University, Bhubaneswar, Odisha

5.3 HISTORY OF ASSESSMENT OF INTELLIGENCE

If we look back to the history of contemporary testing of intelligence, we can know that it started in the nineteenth century with rise in interest in the care and treatment of mentally retarded (now known as intellectual disability, refer to Unit 3). In 1838, a French Physician, Esquirol pointed out that there are different degrees of mental retardation, varying along a continuum from normality to ‘low-grade idiocy’ (no longer used now). He attempted to classify the different degrees of retardation on the basis of individual’s linguistic ability, as he believed that individual’s use of language is the most dependable measure of his/her intelligence.

Another French physician, Seguin, in 1837, established the first school for the education of mentally retarded children. He was employing physiological method of training, i.e., exercises for sensory discrimination as well as motor control was given. Some of the procedures developed by Seguin were eventually incorporated into performance or nonverbal tests of intelligence. His Seguin Form Board test required the individual to insert blocks of different shapes into the corresponding depressions/spaces on a board as quickly and as perfectly as possible. This test is used even today.

English biologist, Sir Francis Galton (1883) believed that individual’s ability to perceive the difference among stimuli is associated with his judgment and intelligence. He devised a number of tests of sensory discrimination like the Galton bar, the Galton whistle and weights for measuring kinesthetic discrimination. He believed that these tests could be used for measuring intelligence. James McKeen Cattell was also of the same opinion as Galton.

Now let us know about the first scales constructed to assess intelligence.

5.3.1 The Binet-Simon Scale: The First Intelligence Test

In 1904, the French Government appointed Alfred Binet to develop a measure to identify children with notably below average intelligence for their age, so that special education can be given to them. With this objective, in 1905 Binet with the help of Theodore Simon developed the first intelligence test, which is popularly known as the first Binet - Simon scale.



Figure 5aF: Binet and Simon

Source: upload.wikimedia.org (Binet) and en.wikipedia.org (Simon)

This scale consisted of 30 problems measuring a variety of functions, such as judgment, comprehension and reasoning. Binet believed that these functions are the essential components of intelligence. The items were arranged in ascending order of difficulty. The items were administered to 50 normal children of 3 to 11 years of age, some mentally retarded children and some adults in order to determine the difficulty level of the items. In this scale, no precise method to calculate the total score was available.

5.3.2 The Revised Scales and the Advent of IQ

In 1908, Binet and Simon revised their 1905 scale. As the 1905 scale had been designed to diagnose children with significantly below average intelligence, many of the items of the test were very simple. In 1908 version, many of the very simple items were dropped and new items were added.

The major innovation of this later version was the introduction of the concept of mental level. The test was standardized on about 300 children of 3 to 13 years of age. Binet and Simon arranged the items according to age level. All the items typically passed by a particular age group were placed at that particular age level. For example, all the items passed by 80-90 per cent of the 3 year olds were placed in the 3-year level. Similarly all the items were ordered according to age level up to age 13. The scoring system consisted of a basal age (all the items of that age level were passed by the child) and credits for partial years, finally expressing the scores in terms of a “mental level”. Mental level of a child corresponded to the age of normal children with whose performance, he or she was equal. Thus, a child of seven years may perform at the level of nine years old. Soon after, in various translations of the scale, mental level was substituted by “mental age”.

Mental age was being compared with the chronological age to indicate the individual’s intellectual level. For example, if an 8 -year old is functioning at the mental level (or mental age) of a 6-year old, he/she is said to be retarded by 2 years. On the other hand, if an 8-year old child is functioning at the mental level (mental age) of a 10-year old child, he is said to be advanced or ahead by 2 years. William Stern (1912) coined the term IQ or intelligence quotient to denote the ratio of mental age to chronological age.

In 1911, the Binet-Simon scale was revised for the third time. More items were added to extend the scale upto the age group of 16 years. The American version of the Binet-Simon test was developed by Terman and Merrill at Stanford University in 1916. It is known as Stanford-Binet Intelligence scale. Terman suggested for multiplying 100 to intelligence quotient in order to avoid fraction.

$$IQ = (\text{Mental Age} / \text{Chronological Age}) \times 100$$

In fact, intelligence testing is one of the major achievements of psychology in the twentieth century. After the success of Binet- Simon scale, several tests of intelligence were developed and refined. These tests can be grouped under (a) individual and group tests, (b) verbal and non-verbal tests, and (c) performance tests. Thus, tests can be categorized based on administration as well as the nature of test items. Refer to the Figure 5bF below.

TYPES OF INTELLIGENCE TESTS

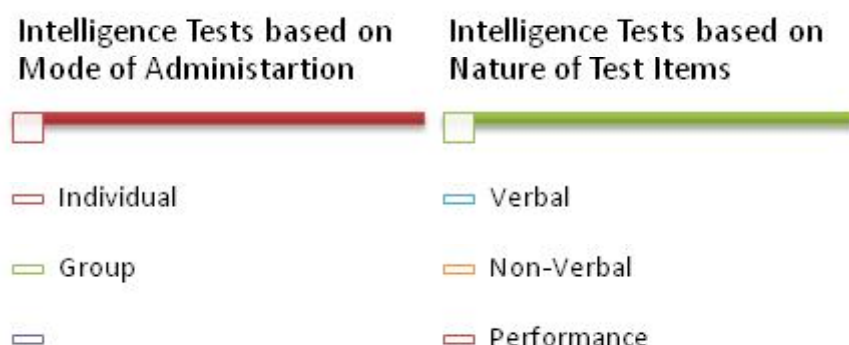


Figure 5bF: Types of Intelligence Tests

Intelligence tests are used in a variety of situations:

- 1) The most common use is in the context of educational setting. Due to their highly significant correlation with academic achievement, intelligence tests are very often used as tests of scholastic aptitude.
- 2) They are also used with normal adolescents and adults for educational and occupational counseling, and personnel selection.
- 3) Further, in clinical setting, they are used for identification and classification of persons with intellectual disability.

Here, let us see what constitutes a good test? When we need to use tests for educational, career or clinical purpose, there are certain considerations one needs to keep in mind. These are called the characteristics or properties of a test.

Properties of a good test

While choosing a particular test in a particular situation, one needs to consider the psychometric properties of the test. Three main properties are: standardization, reliability and validity.

An essential feature of any test is standardization. It implies two things: First, uniformity of procedures in administration and scoring of the test; and second, establishment of norms. Norm is the normal or average performance. Before using a test you should consider the characteristics of the standardization sample on which the norm has been established. Two other psychometric properties of a test are reliability and validity. Reliability is the consistency of scores obtained by a person when retested, be it with the same test, or an equivalent form of test, or by different examiners. Validity says whether the test measures what it intends to measure. You must have already read about properties of a good test in Unit 2 in the context of assessment of personality. You can refer back to it, to further your understanding.

Thus we need to use only such intelligence tests which have good psychometric properties. This is very important because otherwise the findings will not be correct and it may led to wrong information, labeling and stigma with regard to the individual tested.

Self Assessment Questions I

- 1) What do you mean by ‘mental level’?

.....
.....
.....

- 2) What is IQ?

.....
.....
.....

- 3) What is the meaning of test standardization?

.....
.....

5.4 INDIVIDUAL TESTS

Let us now be familiar with some intelligence tests. As we mentioned earlier, intelligence tests can be individual or group tests.

Some individual intelligence tests are described below.

5.4.1 Stanford-Binet Scale of Intelligence

You have already read about the Binet - Simon scale (1905, 1908, 1911) in the earlier section. This test was translated and adapted by Terman (1916) and his associates at Stanford University. This American version of the test included a number of altogether new items and some old items which were revised. The entire scale was re-standardized on an American sample of 1000 children and 400 adults. The fifth version of the Stanford – Binet scale (SB5) was released in 2003. The major changes introduced in each version of the test are given in the Table 5aT below.

Table 5aT: Major Developments in the Binet-Simon and Stanford-Binet Scale

Year	Test/Authors	Comments/ Developments
1905	Binet and Simon	Simple 30 item test
1908	Binet and Simon	Introduced the mental age concept
1911	Binet and Simon	Expanded to include adults
1916	Stanford-Binet (Terman and Merrill)	Used the concept of IQ
1937	Stanford-Binet-2 (Terman and Merrill)	First use of parallel forms (L and M)
1960	Stanford-Binet-3 (Terman and Merrill)	Modern item analysis method used
1972	Stanford-Binet-3 (Thorndike)	SB-3 re-standardized on 2100 persons
1986	Stanford-Binet-4 (Thorndike, Hagen, and Sattler)	Content coverage was broadened by 15 subtests
2003	Stanford-Binet - 5 (Roid)	Five factors of intelligence

(Source: Gregory, R. J. (2004). Psychological Testing)

The SB5 Model of Intelligence

The Stanford-Binet – fifth edition (SB5 model) of intelligence consists of 10 subtests for assessing the intellectual and cognitive abilities. It assesses five cognitive factors such as Fluid reasoning, Knowledge, Quantitative Reasoning, Visual Spatial Processing, and Working Memory. Each cognitive factor is measured by two subtests.

The earlier versions of the Stanford – Binet were yielding a composite IQ. But the SB5 provides three IQ scores (Full score IQ, Verbal IQ, and Nonverbal IQ), as well as five factor scores. The mean and SD of each score (IQ and Factor scores) are set at 100 and 15 respectively.

The SB5 is suitable for use with children from age of two years to adults of 85 years.

The scale was standardized on a sample of 4800 individuals of United States stratified by gender, ethnic, regional and educational levels based on the 2000 census. The reliability and validity are very strong. The reliability of the three IQ scores are in the .90s and that of the subtests range from .70 to .85 (Roid, 2002).

5.4.2 The Wechsler Scales

David Wechsler developed three scales, one designed for adults, one for school-age children, and one for preschool children. These scales are used as measures of general intelligence; at the same time, they are also useful in psychiatric diagnosis. Wechsler observed that brain damage, psychotic deterioration, and emotional difficulties may affect some intellectual functions more than others. Hence, he argued that individual's performance on different subtests should be analyzed in order to diagnose different psychiatric conditions. Therefore, his tests contain several subtests from verbal as well as nonverbal domain.



Figure 5cF: (a) David Wechsler, (b) WAIS - III

Source: (a) ru.wikipedia.org, (b)alchetron.com, Free Social Encyclopedia for the World

Initially Wechsler was interested in developing a scale to measure intelligence in adults. He went through the then available scales. He realized that in these tests, particularly in the Stanford-Binet,

- the test items were not suitable for adults,
- manipulation of words were over-emphasized,
- speed of performance was emphasized at the expense of accuracy,
- mental age norms that were used were not relevant to adult testing.

To overcome these shortcomings, Wechsler developed the 'Wechsler Bellevue Intelligence Scale', which

- was specifically for adults,
- added performance items to balance verbal items,
- reduced the emphasis on speed of performance, and
- replaced mental age with Deviation IQ.

Deviation IQ is the intellectual level of an individual compared to her/his age-mates.

Instead of using the usual formula, $IQ = \frac{\text{Mental Age}}{\text{Chronological Age}}$, he invented a new age-relative formula:

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

With this new formula, Wechsler assumed that IQ remains constant in spite of normal aging, even though the intellectual ability might decline. This assumption of IQ constancy is basic to the Wechsler scales.

In 1955, the Wechsler- Bellevue was revised into Wechsler Adult Intelligence Scale (WAIS).

Keeping the basic pattern the same in terms of form and content, Wechsler(1949) prepared the Wechsler Intelligence Scale for Children (WISC), which was meant to measure intelligence of children of 6 to 16 years of age. In WISC, many items were directly taken from the adult scale and easier items of the same type were added to each subtest. Another addition to the Wechsler family in 1967 was the Wechsler Preschool and Primary Scale of Intelligence (WPPSI), a downward extension of WISC, originally meant to make intellectual assessment of 4 to 6 and a half year olds. But now it has been revised to assess the intelligence of children aged 2 years 6 months through 7 years 7 months.

All these scales have undergone several revisions. Now the fifth version of the original Wechsler- Bellevue is available in the name of WAIS-IV (2012). WISC has also been revised for five times. The latest version WISC-V was published in 2014. After three revisions, WPPSI is in its current version, WPPSI-IV (2012).

All the Wechsler scales consist of a number of verbal and performance subtests. Some of the subtests are core tests and others are supplemental. These subtests together can provide a Verbal IQ score, a Performance IQ, and a Full Scale IQ score. In addition to IQ scores, they can yield four Index scores.

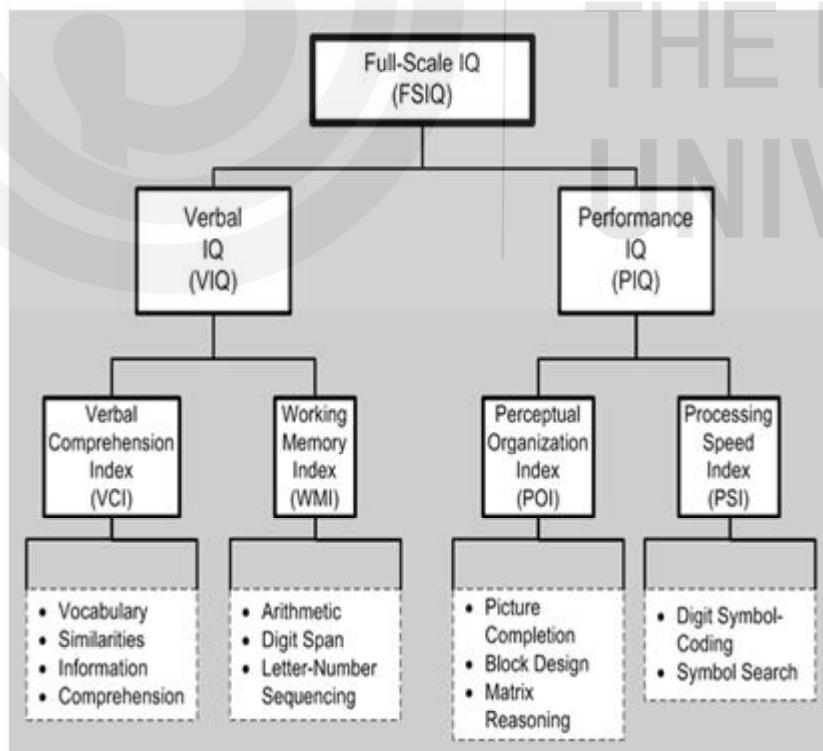


Figure 5dF: Subtests of Wechsler Scales

Source: en.wikipedia.org

General Features of the Wechsler Tests

- 1) Common subtests for different ages. The three scales, WAIS-IV, WISC-V, and WPPSI-IV, all have the same core subtests, such as: Vocabulary, Information, Symbol Search, Block Design, Matrix Reasoning, Picture Completion, and Comprehension. Once the examiner learns to administer any one of the core subtests, (such as information) on any of the Wechsler scale, can easily transfer the skill to the other Wechsler scales.
- 2) Multi-subtest approach enables the examiner to make intraindividual analysis of the strengths and weaknesses of the individual rather than just compute a single global score.
- 3) Verbal and a performance subtests constitute the Verbal and Performance Scales respectively. One can assess verbal comprehension and perceptual organization skills, working memory and processing speed separately with these subtests.
- 4) A common metric for IQ and Index scores, namely Verbal Comprehension, Perceptual Reasoning, Working Memory, and Processing Speed. The mean and standard deviation for IQ and Index scores are 100 and 15 respectively. This IQ is known as Deviation IQ. Moreover, the scaled scores on each subtest have a mean of 10 and Standard Deviation of 3. Scaled scores allow the examiner to analyze the relative strengths and weaknesses of the examinee in the skills measured by the subtests.

All the versions of the Wechsler tests have been standardized on large stratified samples. The samples have been carefully stratified on such variables as sex, race/ethnicity, educational level, and geographic region etc. as per the census figures. All the Wechsler scales have high reliability and validity. The respective manuals provide the information regarding reliability and validity of the tests.

5.4.3 The Kaufman Scales

Kaufman and Kaufman, in the 1980s and 1990s, developed three scales which are individually administered. These scales were designed to include developmentally appropriate tasks anchored in theories of intelligence.

The Kaufman Assessment Battery for Children (K-ABC, 1983) is in line with the information processing theory of Das (1984) and Luria (1966). The test comprises seven tests of Simultaneous Processing, three subtests of Sequential Processing and an Achievement Scale consisting of six subtests. The Simultaneous Processing subtests require the synthesis and organization of visuo-spatial information that can be surveyed as a whole. The Successive Processing subtests require serial or temporal arrangement of the separate bits of information as well as short-term memory. These two mental processing scales are similar to those used in neuropsychological examinations. The Achievement subtests tap tasks similar to those included in traditional aptitude and intelligence tests, rather than measuring the factual knowledge taught in school.

The battery gives four global scores: Simultaneous Processing, Successive Processing, Mental Processing Composite and Achievement. The mean and SD of all these scores are set at 100 and 15 respectively.

The second edition of the K-ABC (K-ABC-II) was published in 2004. K-ABC-II incorporates two distinct theoretical models, Luria's and that of Cattell-Horn-Carroll. Before testing, the examiner has to decide which model to follow.

The Kaufman Adolescent and Adult Intelligence Test (KAIT, 1993) is a measure of intelligence for ages 11 through 85 years. This test is based on the theoretical propositions of Horn and Cattell (1966) as well as Luria (1980) and Piaget (1972). The KAIT consists of two batteries: The Core Battery and the Expanded Battery and a brief Mental Status test. The former is composed of a Crystallized Scale and a Fluid Scale, each having three subtests. The Crystallized Scale is related to concepts acquired from schooling and acculturation, while Fluid Scale measures the ability to solve new problems. The Expanded Battery is used with individuals suspected of neurological damage.

Kaufman Brief Intelligence Test (K-BIT, 1990, 2004) is a quick screening measure for estimating the level of intellectual functioning in 4 to 90 year olds. It consists of one verbal and one performance subtest. The test yields three scores, i.e., (verbal, performance and composite). The scores are expressed in terms of deviation IQ units.

The psychometric properties of the Kaufman scales are satisfactory having adequate normative sample, and high reliability and validity.

5.4.4 Das-Naglieri Cognitive Assessment System (D-N CAS)

The D-N CAS was published in the late 1990s. It is an individually administered instrument for assessing the cognitive processing. This test has its theoretical background in the PASS model of intelligence advanced by Das, Naglieri, and Kirby (1994). The CAS comprises several verbal as well as nonverbal subtests to measure four basic cognitive processes, namely Planning, Attention, Simultaneous processing and Successive processing. These processes are believed to be involved in learning, but are independent of schooling.

The CAS is used with children of ages 5.0 to 17.11 years. From the psychometric point of view, the test is sound.

Self Assessment Questions II

1) What are the shortcomings of the Stanford-Binet scale?

.....
.....
.....
.....

2) Which processes are measured by the D-N CAS?

.....
.....
.....

3) Name the four index scores measured by the Wechsler scale.

.....
.....
.....

5.5 GROUP TESTS

As compared to the individual tests of intelligence, a group test is one that can be administered to more than one person at the same time. Thus a group intelligence test allows us for testing the intelligence of several persons, may be hundreds at a time, quickly and efficiently.

When the United States entered World War I in 1917, there was a pressing need for selecting one and half million candidates on the basis of their general intellectual level within very short time for recruitment to military services. The then available tests, since required manipulation of materials or oral responses from the examinee, could not be adapted to group administration. In this setting, under the direction of Robert M. Yerkes the Army Alpha and the Army Beta were developed. While the former was designed for general routine testing, the latter being a nonlanguage scale was used with illiterates and those who did not know English. Thus the Army Alpha and the Army Beta can be considered as models for group testing. They underwent a number of revisions. After the end of World War I, these tests were released for civilian use.

Features of Group Tests as different from Individual Tests

1) *Multiple-choice versus open-ended format*

Although open-ended questions asking for free responses were used in the early group tests, today most group tests employ multiple-choice items. This change in format has ensured uniformity and objectivity in scoring.

2) *Quick and objective machine scoring versus examiner scoring*

Since multiple-choice items are used, scoring takes less time because of the use of scoring templates. Even scoring can be far quicker if it is done by an optical scanning device. Computer scoring eliminates the chances of examiner errors and halo effects that may occur in scoring individual tests.

3) *Group versus individualized administration*

While administering a group test, the examiner only reads the instructions and enforces the time limits. Thus, her/his role is minimal having no scope for one-to-one interaction. In individual test administration, examiner rapport is an important requirement. The examiner needs to be trained and experienced in training.

4) *Control of item difficulty*

In individually administered tests, often there are entry rules and discontinuation rules. The examiner follows these rules to ensure that the test taker takes the items according to her/his ability level. In the group tests, no such starting and discontinuation rules are there.

5) *Application: Screening versus diagnosis and remedial planning*

Group intelligence tests are commonly used for screening the subjects for a particular course or a job. But traditional individual intelligence tests are used as an aid for the diagnosis of and planning remediation for children with learning difficulties.

6) *Very large versus not-so-large standardization sample*

Group tests, in general, are standardized on very large samples compared to individual tests.

5.5.1 Advantages and Disadvantages of Group Testing

Advantages

- 1) Group tests are designed primarily for mass testing. Simultaneously, they can be administered to a number of persons.
- 2) In group testing, examiner's role is simplified. On the other hand, individual tests like Stanford-Binet or Wechsler, the examiner has to have extensive training.
- 3) Scoring is more objective and simpler in comparison to individual tests.
- 4) As it is easier to obtain large amount of data with group tests, they provide better established norms.

Disadvantages

- 1) Since in group testing the examiner has less opportunity to establish rapport with the examinees, some people, particularly those having motivational problems or difficulty in following directions may score far below their true ability.
- 2) Group testing lacks flexibility, as all examinees are required to answer all the items. If there are too many easy or difficult items, the examinee may feel bored or frustrated respectively.
- 3) Invalid scores are difficult to be recognized.

5.5.2 Examples of Group Intelligence Tests

i) Multidimensional Aptitude Battery (MAB)

The Multidimensional Aptitude Battery (MAB; Jackson, 1984) is a paper-pencil equivalent of the WAIS-R. The test consists of 10 subtests, five each from verbal and performance categories. Although the subtests have the same name as the WAIS-R subtests, the test items are not same. The items within each MAB subtests are multiple-choice and are arranged in order of increasing difficulty. Like WAIS-R, this test also provides 10 subtest scores, as well as Verbal, Performance and Full Scale IQs. It is appropriate for people of 16 to 74 years of age. The psychometric properties are comparable with that of WAIS-R.

ii) Cognitive Abilities Test (CogAT)

The Cognitive Abilities Test (Lohman & Hagen, 2001) is a group-administered, standardized, norm-referenced test. This test has been designed to assess the basic skills that are prerequisite to classroom-based learning. It measures students' learned reasoning and problem-solving skills in three different areas: verbal, quantitative, and nonverbal. Each of these areas of cognitive abilities is measured by three subtests. Thus, there are a total of nine subtests.

CogAT is a multilevel battery consisting of a series of overlapping tests. There are two levels for students in kindergarten through grade 3 and eight levels covering grades 3 through 12. For each level, there are different sets of items. But adjacent levels contain some common contents. Because of the overlapping nature of test contents, proper assessment of students at both the extremes (i.e., poor and bright students) is possible.

The test provides age-based normalized standard scores for each battery with mean of 100 and standard deviation of 15. In addition, per centile ranks and stanines are also available. The reliability of the test is exceptionally good while the validity is also reasonably good.

iii) **Culture Fair Intelligence Test (CFIT)**

The Culture Fair Intelligence Test (Cattell, 1940; IPAT, 1973) is a nonverbal measure of fluid intelligence. It was designed to measure analytical and abstract reasoning ability in a manner so as to make it free from the effects of culture as far as possible. Originally the test was named as Culture Free Intelligence Test. Later the name was changed to Culture Fair Test as it was realized that cultural influences cannot be eliminated totally from any intelligence test.

The CFIT consists of three scales: Scale 1 is used with mentally retarded adults and children from four to eight years of age; Scale 2 is meant for use with adults with average intelligence and children aged eight to thirteen. High school and college students as well as adults with superior ability are administered the Scale 3. Each scale comprises four subtests: Series, Classification, Matrices, and Condition. Two equivalent forms, Form A and Form B are available for each scale.

Scale 1 requires considerable interaction between the tester and the test taker. Hence, in a strict sense, Scale 1 is not a group test. The other two scales are truly group tests of intelligence.

The scores are expressed in terms of normalized standard score IQs with mean 100 and standard deviation 16. Psychometrically the CFIT is a worthy test, but it needs to be revised.

iv) **Raven's Progressive Matrices (RPM)**

Raven's Progressive Matrices (Raven, 1938, 1986, 1992) is a nonverbal test of inductive reasoning, originally designed to measure Spearman's g factor. Spearman defined g as the "education of correlates". Education refers to the process of finding out the relationships based on the fundamental similarities among the stimuli. The test contains figures presented in a matrix form i.e., they are arranged in a pattern with rows and columns. The task of the test taker is to choose the missing part of the matrix from the alternatives given below it. The test taker, in order to perform the item correctly, must identify the recurring pattern of relationship among the figures.

Raven, in actuality, constructed three different instruments: Coloured Progressive Matrices for five to eleven year old children, Standard Progressive Matrices (1996 Edition) for individuals between the ages of 6 to 80 years, and Advanced Progressive Matrices for adults and adolescents with superior intelligence. Although the three forms of the test appear to resemble each other, the strategies required for solving the problem in each form are different.

The test is a nonverbal one. Use of language is restricted only to giving instructions to the test-taker. Even, if necessary, the test can be explained through pantomime also. Once the test-taker understands his task, he does not require language anymore for solving the problems. Therefore the RPM is best suited for testing persons who do not know English. Thus, culture's influence is kept at minimum.

The test can be administered individually or in groups. For normal adults and adolescents in their late teens the reliability coefficients (both split-half and test-retest) vary between .80s and .90s. But for very young subjects, RPM does not possess sufficient reliability. The predictive validity of the test against academic criteria is lower than that of usual verbal intelligence tests.

Self Assessment Questions III

1) Write the advantages of group tests of intelligence.

.....

.....

.....

.....

.....

5.6 VERBAL AND NONVERBAL TESTS

In the earlier sections, we described intelligence tests as grouped into individual and group tests based on administration. In this section, you will learn that intelligence tests can also be classified as verbal and nonverbal tests based on the nature of items in the test.

A verbal test, as the name suggests, is the one in which language is required for successful performance in it. Verbal intelligence is the ability to comprehend and solve language-based problems. The problems require the ability to analyze verbal information and language-based reasoning like verbal analogies. If we look back into the history, we can know that French physician, Esquirol (1838), attempted to diagnose and classify the mentally retarded on the basis of their language skills. Binet also overemphasized on language skills in his intelligence test. But psychologists soon began to realize that this test was inappropriate for people who do not know English and those who have limited language ability like illiterates, children and persons with speech and hearing impairments, due to its overemphasis upon verbal skills. Therefore, a number of nonverbal intelligence tests came into being. Though the Wechsler scales, Kaufman scales, Das-Naglieri test, Otis Quick-scoring Intelligence Test, Multilevel Ability Test etc. have a verbal component, they also include some nonverbal subtests.

Nonverbal tests of intelligence tap the ability to analyze visual information and solve problems without necessarily using words. The tasks may involve the ability to recognize visual sequences, understand and recognize the relationships between visual concepts and situations, as well as perform visual analogies.

The earliest nonverbal/performance test was the Seguin Form Board. Though several performance tests were developed in the early part of the twentieth century, two of the more known tests were the Kohs Block Design test (Kohs, 1920) and the Porteus Maze Test (Porteus, 1915, 1919). The Kohs Block Design test requires the individual to assemble painted blocks to match a pattern. The Porteus test consists of a graded series of mazes. This test is still available today, but is not in wide use.

The Army Beta test came into being out of the necessity to recruit the illiterate and non-English-speaking persons to the army. So obviously it was a non-language test consisting of a number of visual-perceptual and motor tests like tracing a path through mazes and saying the correct number of blocks shown in a three-dimensional drawing.

The Culture Fair Intelligence Test (Cattell, 1940; IPAT, 1973) and Raven's Progressive Matrices (Raven, 1938, 1986, 1992, 1995) are well-known nonverbal intelligence tests which are discussed in the previous section.

Self Assessment Questions IV

1) Name some nonverbal tests of intelligence.

.....
.....
.....
.....
.....

5.7 CULTURE FAIR TESTS

In 1910, Goddard was assigned the work of assessing intelligence of immigrants of Ellis islands. He noted that the immigrants, mostly constituted by Hungarians, Jews, Italians and Russians were having low IQs, and their low IQs were due to environmental deprivation. Goddard (1912, 1917) has pointed out that Non-white people are usually at a disadvantage as intelligence tests are biased towards White people. Critiques have highlighted that intelligence tests are biased towards the majority group in the society. The minority people, the disadvantaged group are usually at a disadvantage.

Between any two cultural groups there are cultural differences. Different cultural groups value, demand and thereby motivate different behaviour. Hence, for assessing individuals from different cultures, culture free tests are needed. Some of the earliest cross-cultural tests (Knox, 1914) were developed during 1910s. Usually cultures vary along the following parameters:

- i) Language – Different cultural groups speak different languages
- ii) Reading experience – Cultures are different with respect to their educational background. Illiteracy prevails in some cultures.
- iii) Speed – The tempo of daily life, motivation to hurry and the value attached to speed of performance vary widely among different cultural and ethnic groups, even within a nation.
- iv) Environmental experiences – People of different cultures experience different natural and social phenomena. For example, people of many cultures do not have an experience of snowfall. If they are asked about information relating to snowfall, they may be unable to answer.

The classic culture free tests earlier were attempting to control these cultural parameters. The authors try to make the test content culture free as far as possible. However, it is realized that impact of culture cannot be eliminated completely from these tests. Therefore the term ‘culture fair’ is used in place of ‘culture free’ tests.

Typical Culture Fair Instruments

- i) The Culture Fair Test (Cattell, 1940; IPAT, 1973) is a nonverbal test. We have already discussed about it under the section of ‘Group Tests’.
- ii) Raven’s Progressive Matrices (RPM -Raven, 1938, 1986, 1995) has also been discussed in the ‘Group Tests’ section. RPM frees itself from the effects of language and speed. The test contains abstract figures that are considered not to favour any particular culture.

- iii) The Leiter International Performance Scale-Revised (Roid&Miller, 1997) was originally published in 1940. This scale was developed for use with different ethnic groups of Hawaii, but later applied to several African groups. The distinctive feature of the test was almost elimination of verbal instructions and absence of any time limit. The test is individually administered. This test covers four domains of functions, namely, Reasoning, Visualization, Attention, and Memory.
- iv) Draw-a-Man Test developed by Goodenough (1926) and its revised version, the Goodenough-Harris Drawing Test (Harris, 1963) adopts a different approach. The test taker is asked to make a picture of a man. The test, instead of artistic skill, emphasizes on the child's ability to observe accurately and think conceptually. Credit is given for inclusion of body parts, clothing, proportion, and perspective etc. Credit points obtained by the test taker are converted to standard scores.

5.8 ISSUES IN INTELLIGENCE TESTING

In the previous sections we discussed some standardized intelligence tests which are quite sophisticated and widely used. However their widespread use has always been controversial. Critics of intelligence testing have raised a number of issues. Let us discuss some of the issues associated with intelligence testing.

- Intelligence tests have been found to be useful in predicting mainly academic achievement. Intelligence test findings can be used to help students find out their level in different mental abilities. This may help design further training and appropriate remedial programmes for the students. Given the benefits of intelligence testing, there is also the issue of labeling a child based on the test findings. A child who has not performed well on intelligence test may be labeled as dull or incompetent. Such a child may be stigmatized and discriminated in school as well as in family. The misuse of intelligence testing may do more harm than good to the child.
- There has been question raised with regard to the validity of intelligence tests also. It is pointed out that IQ tests measure factors other than intelligence like environmental factors such as quality of education, exposure to stimulation, prior knowledge of the information on the test and the individual's test-taking skills etc. If all these non-intelligent factors are responsible for individual differences in IQ scores, then it compromises the accuracy of IQ tests. In such cases, the use of IQ tests as part of diagnosis of intellectual disability and learning disability may not give the correct picture.
- Intelligence is not limited to IQ alone. The IQ tests measure only a part of overall competence of human beings. Success in life depends on several other aspects, such as creativity, social competence, practical problem solving abilities etc. Further, the traditional IQ tests may not recognize the other kind of intelligences, for example as proposed by Gardner.
- Some critics argue that most intelligence tests are biased against certain groups, particularly those who vary from mainstream Western society. Language and nature of the test items may create problems in estimating intelligence of people from non-white cultures. For instance, tribals have a rich source of indigenous knowledge, but they may not perform well on the standard tests of intelligence. The reason is that the tests do not reflect their context, knowledge and culture.
- Intelligence testing for infants and preschoolers has also been questioned for its sensitivity, accuracy and predictive value.

5.9 LET US SUM UP

Now that we have come to the end of this Unit, let us list all the major points that we learnt.

In the nineteenth and early part of twentieth century, interest in the care and treatment of mental retardation increased. Hence, the need for identifying the intellectually poor people was felt. Esquirol and Seguin had their own tests used for this purpose. However, the credit of the first intelligence test goes to Alfred Binet who was given the task of developing a measure to identify the retarded children. In 1905 Binet and Simon developed the first intelligence test, which is popularly known as the first Binet Simon Scale. In 1916, the American version of the Binet-Simon test was developed by Terman and his associates at Stanford University. The scores on the Binet test were expressed in terms of 'Mental Age' and later the expression, 'IQ' came into use to indicate the intellectual level of a child.

In order to address the inherent inconvenience in the interpretation of mental age and its use in finding out adult intelligence, David Wechsler talked of 'Deviation IQ'. He developed three scales, one designed for adults, one for school-age children, and one for preschool children.

Besides the Stanford-Binet and Wechsler scales, there are a number of other individual and group intelligence tests. An individual test is administered to one individual at a time and a group test is one that can be administered to more than one person at the same time.

Intelligence tests can also be classified as verbal and nonverbal tests. A verbal test is one in which language is required for successful performance in it. Nonverbal tests without necessarily using words tap the ability to analyze visual information to solve problems.

Psychologists acknowledged that intelligence tests developed for Western people are not suitable for use with nonwestern people. So they developed a number of culture fair tests for cross-cultural use.

Though intelligence tests have become popular, their widespread use is controversial. There are chances of misuse of IQ tests and if misused, they can be more harmful than useful.

5.10 KEY WORDS

- Mental age** : refers to the age at which an individual performs on the test items. For example, if a 6 year old child performs like a 9 year old child, then his mental age is 9 years.
- Intelligence quotient** : is a measure of intelligence expressed as the ratio of mental age divided by chronological age and multiplied by 100.
- Standardization** : refers to uniformity of procedures in administration and scoring of the test; and establishment of norms.
- Norm** : is the normal or average performance by the individuals of a particular age group.

- Binet Simon scale** : was the first intelligence test developed by Binet and Simon in 1905 to identify children with notably below average intelligence for their age, so that special education can be given to them.
- Wechsler scales** : consist of the intelligence scales developed by David Wechsler to assess intelligence of adults, school-age children and pre-primary school children.
- Deviation IQ** : refers to intellectual level of an individual compared to her/his age-mates. It has a mean of 100 and standard deviation of 15.
- Culture fair intelligence tests** : refer to tests which are not biased by cultural differences in the individuals taking the test.

5.11 ANSWERS TO SELF ASSESSMENT QUESTIONS

Self Assessment Questions I

- 1) 'Mental level' of a child corresponds to the age of normal children with whose performance he or she is equal. Soon after, in various translations of the Binet-Simon scale, mental level was substituted by 'mental age'.
- 2) IQ denotes the ratio of mental age to chronological age. Terman suggested for multiplying 100 to intelligence quotient in order to avoid fraction.
$$IQ = (\text{Mental Age} / \text{Chronological Age}) \times 100$$
- 3) Standardization implies two things: First, uniformity of procedures in administration and scoring the test, and second, establishment of norms.

Self Assessment Questions II

- 1) Shortcomings the Stanford-Binet scale:
 - the test items were not suitable for adults,
 - manipulation of words were over-emphasized,
 - speed of performance was emphasized at the expense of accuracy,
 - mental age norms used were not relevant to adult testing.
- 2) Planning, Attention, Simultaneous and Successive processing
- 3) Verbal Comprehension, Perceptual Reasoning, Working Memory, and Processing Speed.

Self Assessment Questions III

- 1) Advantages of group intelligence tests:
 - They can be simultaneously administered to a number of persons.
 - The examiner does not need extensive training for administering the test.
 - Scoring is more objective and simpler in comparison to individual tests.

- As it is easier to obtain large amount of data with group tests, they provide better established norms.

Self Assessment Questions IV

- 1) The Seguin Form Board, the Kohs Block Design test, the Porteus Maze Test, the Army Beta test, the Culture Fair Intelligence Test and the Raven's Progressive Matrices.

5.11 UNIT END QUESTIONS

- 1) Give a historical account of intelligence testing.
- 2) Describe some individually administered intelligence scales.
- 3) Discuss group intelligence tests.
- 4) What is a culture fair test? Give some examples of culture fair intelligence tests.

5.12 REFERENCES

Anastasi, A. & Urbina, S. (1997). *Psychological Testing*. Singapore: Pearson education.

Binet, a., & Simon, T. (1905). Methodesnouvelles pour le diagnostic du niveau intellectuel des anormaux. *AnneePsychologique, 11*, 191-244.

Cattell, J. M. (1890). Mental tests and measurements. *Mind, 15*, 373-380.

Cattell, R.B. (1940). A culture-free intelligence test. I. *Journal of Educational Psychology, 31*, 161-179.

Das, J. P. (1984). Simultaneous and successive processes and K-ABC. *Journal of Special Education, 18*, 229-231.

Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes: the PASS theory of intelligence*. Boston: Allyn& Bacon.

Esquirrol, J. E. D.(1838). *Des maladies mentalesconsiderees sous les rapports medical, hygienique, et medico-legal* (2 vols.). Paris: Bailliere.

Galton, F. (1883). *Inquiries into human faculty and its development*. London: Macmillan.

Goddard, H. H. (1912). *The Kallikak Family: A study in the heredity of feeble-mindedness*. New York: Macmillan.

Goddard, H. H. (1917). Mental tests and the immigrant. *Journal of Delinquency, 2*, 243-277.

Goodenough, F. (1926). *Measurement of intelligence by drawings*. New York: World Book Co.

Harris, D. B. (1963). *Children's drawings as measures of intellectual maturity: A revision and extension of the Goodenough Draw-a-Man Test*. San Diego, CA: Harcourt Brace Jovanovich.

Horn, J.L. &Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligence. *Journal of Educational Psychology, 57*,253-270.

- Jackson, D. N. (1984). *Multidimensional Aptitude Battery (MAB): Manual*. Port Huron, MI,: Sigma assessment Systems.
- Jackson, D. N. (1994). *Multidimensional Aptitude Battery (MAB): Manual*. Port Huron, MI: Sigma assessment Systems.
- Kaufman. A. S. & Kaufman, N. L. (1983a). *Kaufman Assessment Battery for children: Administration and Scoring Manual*. Circle Pines, MN: American Guidance Service.
- Kaufman. A. S. & Kaufman, N. L. (1983b). *Kaufman Assessment Battery for children: Interpretive Manual*. Circle Pines, MN: American Guidance Service.
- Kaufman. A. S. & Kaufman, N. L. (1993). *Kaufman Adolescent and Adult Intelligence Test: Manual*. Circle Pines, MN: American Guidance Service.
- Kohs, S. C. (1920). The Block-Design Tests. *Journal of Experimental Psychology*, 3(5), 357-376.
- Knox, H. A. (1914). A scale based on the work at Ellis Island for estimating mental defect. *Journal of the American Medical Association*, 62, 741-747.
- Lohman, D. F., & Hagen, E. P. (2001). *Cognitive Abilities Test (Form 6)*. Itasca, IL: Riverside.
- Luria, A. R. (1980). *Higher cortical functions in man* (2nd ed.). New York: Basic books.
- Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. *Human development*, 15, 1-12.
- Porteus, S.D. (1915). *The Porteus Tests*. Victoria Education department, Melbourne, Australia.
- Porteus, S. D. (1919). *The Porteus Tests — Vineland Revision*. Publication of the Training School at Vineland, N.J.
- Raven, J., Raven, J. C., & Court, J. H. (1995). *Manual for Raven's Progressive Matrices and Vocabulary scales-Section 1: general Overview* (1995 Edition). Oxford, England: Oxford Psychologists Press.
- Roid, G. H. (2003). *Stanford-Binet Intelligence Scales*, Fifth Edition (SB:V). Itasca, IL: Riverside Publishing.
- Roid, G. H. & Miller, L. J. (1997). *Examiner's manual: Leiter International Performance Scale-Revised*. Wood Dale, IL: Stoelting.
- Stern, W. (1912). *Die psychologische Methoden der Intelligenzprüfung*. Leipzig, Germany: Barth.
- Terman, L. M. (1916). *The measurement of intelligence*. Boston: Houghton Mifflin.
- Thurstone, L. L. (1938). Primary mental abilities. *Psychometric Monographs*, No. 1.
- Wallace Wallin, J.W. (1911). A Practical Guide for Administering the Binet-Simon Scale for Measuring Intelligence. *The Psychological Clinic*, 5, (1).
- Wechsler, D. (1949). *Manual for the Wechsler Intelligence Scale for Children*. New York: Psychological Corporation.

Wechsler, D. (2008). *WAIS-IV administration and scoring manual*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2012a). *Wechsler Preschool and Primary Scale of Intelligence – Fourth Edition (WPPSI-IV): Administration and scoring manual*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2014b). *Wechsler Intelligence Scale for Children-Fifth Edition administration and scoring manual*. San Antonio, TX: NCS Pearson.

Logsdon, A. (2018, August 28). *Importance of Verbal Intelligence*. Retrieved from <https://www.verywell.com/what-is-verbal-intelligence-2162167>.

5.13 SUGGESTED READINGS

Anastasi, A. & Urbina, S. (1997). *Psychological Testing*. Singapore: Pearson education.

Gregory, R.J. (2004). *Psychological testing: History, principles, and Applications*. Pearson education.

Kaplan, R. A. & Saccuzzo, D. P. (2013). *Psychological Assessment and Theory*. New Delhi: Cengage Learning.

Morgan, C.T., King, R.A., Weisz, J.R., & Schopler, J. (1986). *Introduction to Psychology*. New Delhi: Tata McGraw-Hill.