

12

READINGS

in

COUNSELLING PSYCHOLOGY



Edited by

E. M. HASSAN, Ph.D
S. E. OLADIPO, Ph.D
J. W. OWOYELE

© Department of Counselling Psychology

First Published 2008

All rights reserved. No part of this book may be reproduced in any form, electronic or mechanical including photocopy, recording or any information storage and retrieval system, without permission in writing from the publisher.

ISBN 978-065-980-3

Published by:

Department of Counselling Psychology,
College of Applied Education and Vocational Technology,
Tai Solarin University of Education,
Ijagun, Ijebu-Ode.

Printed by:

Silverprint Communications

237, Oba Adeboye Road, P.O.Box 186. Ijebu-Igbo, Ogun State.

TABLE OF CONTENTS

Preface.....	v
Foreword.....	vi
Notes on Contributors.....	vii
1. Behaviour Modification	
Prof. Sotonade, O.A.T	1
2. Techniques of Counselling	
Ewumi, A.M.	17
3. Adolescence Psychology	
Dr Oluwatimilehin, J.T.B.	27
4. Introduction to Physiological Psychology	
Adegbenga A. Onabamiro	51
5. Educational Psychology: An Introduction	
Owolabi Gabriel, M. A.	78
6. Human Resources Management and Counselling	
Dr Akinsanya Omolade Oluwatoyin	94
7. Introduction to Guidance and Counselling	
Dr Shobola, A. A	105
8. Group Counselling	
Dr Ajufo, B.I.	132
9. Conflicts in School System	
Adebanjo, M.A.	143
10. School Administration and the Guidance Counsellor	
Dr Ijaduola, K.O.	151
11. Introduction to Organisational Psychology	
Okubanjo, A.O.	160
12. Educational Guidance and Counselling	
Owoyele, J.W.	185
13. Multicultural Counselling	
Dr Hassan, E.M.	199
14. Addiction: What it is and how to Overcome it	
Dr Oladipo, S.E.	210
15. Theories of Counselling	
Dr Odeleye, D.A.	224
16. Group Dynamics	
Dr Owodunni, A.A.	233
17. Sex Education and Marital Guidance	
Adebiyi, F.A.	241

18	Language Disorders: Counselling Interventions to Facilitate the Study of English as a Second Language.	
	Oluwole Akinbode	263
19.	Scope, Nature and Uses of Psychological Tests	
	Adejumo, G.O	271
20.	Inferential Statistics in Educational	
	Dr Seyin, Adejoke	280
21.	Teaching Listening as a Receptive Language Skill: a Pre-requisite Knowledge for Counsellors with English as a Teaching Subjects in Secondary Schools	
	Eweniyi, I.T	310
22.	Personality and Social Development	
	Dr Pitan, O.O	320
23.	Introduction to Special Education	
	Oyeyemi, O.F	330
24.	Essential Elements in Learning Disabilities	
	Toyobo, O.M	360
25.	Emotion: Effect on Learning	
	Oke, O.R	380

SCOPE, NATURE AND USES OF PSYCHOLOGICAL TESTS

ADEJUMO, G. O.
Covenant University, Otta

INTRODUCTION

This chapter contains discussion on what a psychological test is and introduces you to some tests you might never have considered to be psychological tests. After exploring the history of psychological testing, it discusses the three defining characteristics of psychological tests and the assumptions one must make when using psychological tests. After discussing how tests are classified, it distinguishes four commonly confused concepts: psychological assessment, psychological tests, psychological measurement, and surveys. It concludes by sharing printed and online resources that are available for locating information about psychological testing and specific tests.

OBJECTIVES

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

1. Define a psychological test?
2. Trace the history of psychological testing from Alfred Binet and intelligence testing to the tests of today.
3. Describe the three characteristics that are common to all psychological tests and assumptions one must make when using psychological tests.
4. Describe the different ways one can classify psychological tests and differences among four commonly used terms that students often get confused: psychological assessment, psychological tests, measurement, and surveys.

WHAT IS TEST?

The word "test" refers to any means, (often formally contrived) used to elicit responses to which human behaviour in other contexts can be related. A test is designed to measure a particular body of knowledge, skills, abilities, or performances which are of interest to the test user.

There are 3 basic concepts in understanding what a test is:

1. A test focuses on a particular *domain*.
2. A test is a *sample* of behaviour, products, answers, or performances from the domain.
3. A test permits the user to make *inferences* about the larger domain of interest.

What is a Psychological Test?

Tests are defined broadly as psychological and educational instruments developed and used by testing professionals in organizations such as schools,

industries, clinical practice, counselling settings and human service and other agencies, including those assessment procedures and devices that are used for making inferences about people in the above-named settings.

However, psychological test can be defined as an instrument/process that requires an individual to perform some behaviour(s) used to measure a personal attribute, trait, or characteristic thought to be important in describing/understanding behaviour or predicting outcomes. Anastasi and Urbina (1997) defined a psychological test as an objective and standardized measure of a sample of behaviour

Three Defining Characteristics of Psychological Tests

1. *They must include a representative sample of behaviour:* All Psychological tests require the respondent/client to do something. Psychological tests are not exhaustive measures. It is a representative sample of the measured behaviour. There should be a clear connection between the test and the measured behaviour in a real world setting.

2. *Samples must be collected under standardized conditions:* The behaviour sample is obtained under standardized conditions. Each individual taking a psychological test should be tested under essentially identical conditions in seating arrangements, lighting conditions noise levels, interruptions, answering common questions .Standardization is vital because many test results are referential in nature i.e. the performance is measured relative to everybody else's performance. Standardization reduces between subject variability due to extraneous variables and it is easier to obtain with tests designed to be administered en masse.

3. *There must be rules for scoring:* Good standardized psychological tests must have a set of rules or procedures for scoring responses to a test.

Differences among Tests

1. Behaviour the test taker performs.
2. Attribute, trait, or characteristic measured or outcome predicted.
3. Content, format, and how administered.
4. How scored and interpreted.
5. Psychometric quality.

HISTORY OF PSYCHOLOGICAL TESTING

Circa 1000 BC. Chinese introduced written tests to help fill civil service positions

1850 The United States begins civil service examinations.

1890 James Cattell develops a "mental test" to assess college students. Test includes measures of strength, resistance to pain, and reaction time.

1905 Binet-Simon scale of mental development used to classify mentally retarded children in France.

1914 World War I produces need in U.S. to quickly classify incoming recruits. Army Alpha test and Army Beta test developed.

- 1916 Terman develops Stanford - Binet test and develops the idea of Intelligence Quotient
- 1920 - 1940 factor analysis, projective tests, and personality inventories first appear.
- 1941- 1960 vocational interest measures developed
- 1961- 1980 item response theory and neuropsychological testing developed
- 1980- Present: Wide spread adaptation of computerized testing. "Smart" tests which can give each individual different test items develop

Test Classification Methods

Tests can be classified as maximal performance, behaviour observation, or self-report; standardized or nonstandardized; objective or projective; by dimension measured; or by subject. All tests can be classified on a number of continuums

1. **Individual or Group test:** Indicates how the test is administered. Many versions of I.Q. tests are given in a one to one situation.
2. **Speed or Power Test:** Refers to whether any time constraints are built into the test. A classic **speed** test would contain many simple items and a strict deadline, while a classic **power** test contains no time deadline but very difficult items.
3. **Cognitive or Affective Test:** Achievement and aptitude tests (like the CUSAT) attempt to measure mental activity and are cognitive tests. The CUSAT is an aptitude test designed by Covenant University to assess the chance for academic success during the first year of university.
4. **Achievement Tests:** assess knowledge of information already learned.
5. **Aptitude Tests:** attempt to gauge whether a person is capable of learning a specific knowledge base.
6. **Affective tests:** are designed to assess interests, attitudes, and personal values of an individual. Most personality tests are considered affective tests.
7. **Objective or Non objective Scoring:** Objective scoring procedures are fully specified before grading begins so that anyone grading the test would calculate the same score for a particular set of answers. Objective scoring is one consideration in determining whether a test is classified as:
8. **Standardized versus Nonstandardized:** Refers to all aspects of creating, testing, and administering a psychological test. As you may well guess, standardized tests are seen as theoretically superior measuring devices. Most scholastic examinations or teacher-made tests are *nonstandardized*. *Standardized* tests have established *norms* to which one can compare an individuals performance. *Norms:* The normal distribution of scores on a standardized tests determined by the *test standardization group*.

Psychological Assessment, Tests, Measurement and Surveys

- Psychological assessment involves multiple methods for gathering data.
- Measurement is the assignment of numbers according to rules.
- A test is a measurement when the behaviour sample can be expressed as a numerical score.
- Surveys focus on group outcomes, and tests focus on individual differences.]
- Survey results are typically reported at the question level, and test results are typically reported as an overall score.

The Importance of Psychological Testing

- Individuals use tests to make important decisions
- Individual decisions are those made by the person who takes a test, and institutional decisions are those that others make as a result of an individual's performance on a test.
- Comparative decisions involve comparing people's scores with one another to see who has the best score, and absolute decisions involve seeing who has the minimum score to qualify

Who Uses Psychological Tests and for What Reasons

- In educational settings, administrators, teachers, school psychologists, and career counsellors use tests to make educational decisions, including admissions, grading, and career decisions
- In clinical settings, clinical psychologists, psychiatrists, social workers, and other health care professionals use tests to make diagnostic decisions, determine interventions, and assess the outcome of treatment programs
- In organizational settings, human resources professionals and industrial/organizational psychologists use tests to make decisions such as who to hire for a particular position, what training individuals need, and what performance rating an individual will receive

Assumptions of Psychological Tests

There are important assumptions we make in using psychological tests, these include:

- That a test measures what it says it measures,
 - That an individual's behaviour will remain stable over time so the test scores,
 - That individuals understand test items similarly,
 - That individuals can report accurately about themselves,
 - That individuals will report their thoughts and feelings honestly
 - That test score is equal to ability plus some error
- We can increase our confidence in many of these assumptions by following certain steps during test development

Test Development

A number of steps have been suggested in the measure development process (Churchill, 1979; Gerbing & Anderson, 1988). They are:

Step 1: The first step in test development involves defining the test universe, the target audience, and the purpose of the test. In determining the test universe: the developer prepares a working definition of the construct the test will measure and check the psychological literature to help with this. The target audience can be determined by identifying the characteristics of the people one expects to take the test. These include characteristics that will affect how the test taker responds to the test such as reading level; disability; motivation to answer honestly and language. The purpose i.e. what the test will measure must be stated e.g. vocational interest. These must include how the outcomes will be used. Will the scores be used to compare test takers (normative) or will they be compared to some achievement level (criterion)? Will the scores be used to predict some performance or make a diagnosis? Will the scores be used cumulatively to help prove/disprove a theory, or individually to provide information about the test taker? This type of information helps the user to determine if the test is appropriate for things such as group administration, paper-pencil or oral administration. It helps the user make a more informed decision about the test's usefulness.

Step 2: The next step in test development is to write out a test plan, which includes the construct(s) definition, the test format, the administration method, and the scoring method. In construct definition the developer must provide his construct with a concise definition. The definition should include: 1) An operationalized statement regarding observable and measurable behaviours. 2) Boundaries of the test domain. The content that the developer is testing. Include content that is not appropriate. Include an estimate (percentage) of how many questions are needed to sample the test domain. Test Format: Choose the test format (e.g. objective or subjective) and the type of questions the test will contain (e.g. multiple-choice, true/false, short answer, verbal questions and responses). Most tests use a consistent format. However, if you use different formats be sure to provide detailed instructions about each type of question. Specify Administration and Scoring Methods: Specify how the test should be administered and scored. Answer questions like: How will the test be administered? How long do test takers have to complete the test? Should the test be given in a group setting or to individuals? Does the test need to be scored by the publisher or administer? Is there a particular weighting for each question? What type of data will the test yield? The most common method for scoring is termed the cumulative model. This type of model states that the more the test taker responds in a given fashion, the greater the exhibition of the attribute being measured. The most common method gives one point for each measure of the attribute. The total accumulation of the 'one points' becomes the raw score. These tests typically yield interval-level data. Other scoring methods include: A) Categorical model: This type of method is used to place test takers into a given group. This type of model generally yields nominal data. B) Ipsative

model: The test taker's scores on various scales within the test are compared to each other and yield a profile of the individual. C) Note: All the above 3 models can be combined in any fashion on any given test.

Step 3: Develop Test Items

This step can be divided into five:

1. Concise definition of the test construct(s) and operationalizes each construct in terms of observable and measurable behaviours. The construct will be what the test will measure and its definition should have the following characteristics!

- It should define one specific thing (be unidimensional) to be measured. For example, a definition of construct homophobia (fear of gays or lesbians) should not include content related to construct sexual orientation (identification as heterosexual, lesbian, gay, bisexual).
- The definition should be inclusive of the entire domain of content of the construct. It should cover (but not necessarily name) every event associated with the occurrence of the construct.
- It should be specific enough that other constructs would clearly not be covered by the definition. For example, a definition of assertive behaviour should be worded in such a way that instances of aggressive behaviour or submissive behaviour would not be covered by the definition.
- The construct as defined should be amenable to being discussed in terms of more or less of it occurring or being present at a particular time.

2. The test developer chooses an objective or subjective format and the type of test question (for example, multiple choice, true/false, open-ended, essay). The format is based on the goal of the test. Item format may be objective or subjective. Response format may be forced or open ended. e.g. College students will be administered a 20-item paper-and-pencil questionnaire on procrastination in which they will respond to multichotomous levels of agreement for each item (e.g. strongly agree, agree, disagree, or strongly disagree). Response biases need to be considered since they may occur depending on the test format. There are several response styles (test taking attitudes) possible

- *Social desirability* - examinee tendency to answer items in a socially accepted way.
- *Acquiescence* - examinee tendency to consistently agree in response to items.

Randomizing position of the items or reversing the direction of the items minimizes response bias. Then the test developer specifies how the test will be administered and scored. Three models for scoring are:

- the *cumulative model*-Adding response score values to obtain an overall total score. e.g., the procrastination test is developed to obtain a total score by adding the items response values. The total score is proposed to indicate the level of procrastination (e.g., a higher score means a higher level of procrastination).
- the *categorical model*, and

- the *ipsative mode*: The test taker's scores on various scales within the test are compared to each other and yield a profile of the individual. Ipsativity comes from having to choose one item over another. For example 'which of the following describes you best? (a) I am outgoing. (b) I work hard'.

The scoring model determines the type of data (nominal, ordinal, or interval) the test will yield. After completing the test plan, the test developer is ready to begin writing the actual test questions and administration instructions.

3. **Generating items:** This is essentially a creative process where the test developer makes up as many items as possible that seem to relate to construct. In theory, one should "sample items" from the domain defined by the concept. In practice, for example in marketing research, *focus groups* are often utilized to illuminate as many aspects of the concept as possible. In educational and psychological testing, one commonly looks at other similar questionnaires at this stage of the scale design, again, in order to gain as wide a perspective on the concept as possible. After writing the test, the developer conducts a pilot test followed by other studies that provide the necessary data for validation and norming.

4. **Choosing items of optimum difficulty:** In the first draft of the test, test developer will include as many items as possible. He/she then administers this test to an initial sample of typical respondents, and examines the results for each item. First, He/she would look at various characteristics of the items, for example, in order to identify *floor* or *ceiling* effects. If all respondents agree or disagree with an item, then it obviously does not help us discriminate between respondents, and thus, it is useless for the design of a reliable scale.

Item difficulty is the proportion of respondents who agree or disagree with an item, or who answer a test item correctly. The larger the percentage getting an item right, the easier the item. The higher the difficulty index, the easier the item is understood to be (Wood, 1960). To compute the item difficulty, divide the number of people answering the item correctly by the total number of people answering item. The proportion for the item is usually denoted as p and is called item difficulty (Crocker & Algina, 1986). An item answered correctly by 85% of the examinees would have an item difficulty, or p value of .85, whereas an item answered correctly by 50% of the examinees would have a lower item difficulty, or p value of .50. The test developer may also look at the item means and standard deviations and eliminate those items that show extreme means, and zero or nearly zero variances.

Item discrimination is the ability of the item to differentiate those students with more knowledge from those with less. To calculate the discrimination index subtract the number of students in the lower group that got an item correct from those in the upper group, and divide by the number of students that made up the upper or lower group. An item discriminates positively if more students in the upper group got an item right than students in the lower group..

$$\text{Difficult} = \frac{U_c + L_c}{U_n + L_n} * 100$$

Group					
			*		
		A	B	C	D
Upper	4	5	0	6	
Lower	2	6	0	7	

Note: * denote correct response
 Item difficult: $(0 + 0)/30 = .00p$
 Discrimination Index: $(0-0)/15 = .00$

Group					
			*		
		A	B	C	D
Upper	0	0	15	0	
Lower	0	0	15	0	

Note: * denote correct response
 Item difficult: $(15 + 15)/30 = 1.00p$
 Discrimination Index: $(15-15)/15 = .00$

Using this steps

- (1) Score the tests
- (2) Arrange answer sheets from high to low
- (3) Separate answer sheet three subgroups
 - (a) upper 27%
 - (b) middle 46%
 - (c) lower 27%
- (4) Count and record responses per foil in upper group
- (5) Count and record responses per foil in lower group

Responses	A	B	C	D*	E	Omit	Diff.	Net D
Upper 27% = 50	28	2	1	19		0		19
Lower 27 = 50	24	8	1	17		0	36%	4%

$$D = 19 + 17/100 * 100 = 36\%$$

$$\text{Discrimination} = \frac{U_c + L_c}{U_n + L_n} * 100 \quad D = 19 - 17/19 * 50 = 4\%$$

Other analysis to be done include:

Average interitem correlation and item-total correlation. These provide information for increasing the test's internal consistency. Each item should be highly correlated

for increasing the test's internal consistency. Each item should be highly correlated with every other item measuring the same construct. In item response theory, the performance of each item is related to the test taker's ability on the construct being measured. The resulting item characteristic curve (ICC) is a graph of the probability of answering an item correctly given a certain level of ability. It combines item difficulty and discrimination. Item bias refers to an item being easier for one group than for another group. ICCs provide information on item bias by comparing group scores. Test developers use questionnaires and expert panels for a qualitative analysis of test items

Step 4 Revising the Test

Test items are dropped based on their consistency, difficulty, discrimination, and bias until a final form of the test is reached. If enough items were piloted, no items will be rewritten. Items that are rewritten must be piloted to be sure they now meet criteria for retention in the test. Other test components, such as the instructions, should be revised based on the results of a qualitative analysis in the form of questionnaires or interviews with test takers, subject matter experts, and testing experts

Step 5 Validation, Norms, Cut Scores, and the Test Manual

Standards for designing the validation study are similar to those for designing the pilot study, including using a representative sample of the target audience that is large enough to conduct the desired statistical tests. Test developers should follow professional code of ethics. Using more than one test site will provide evidence that the results generalize from site to site. The scores resulting from the validation study should not be used for decision making or evaluation of individuals. When the test user will make predictions from test results, a cross-validation study is important. Test developers expect the resulting validity coefficient to be lower, and the difference between the coefficients in the validation study and those in the cross-validation study is called shrinkage.

Differential validity results when tests have different validity coefficients for different groups. Single group validity means that a test is valid for only a specific group. Norms and cut scores can be developed from the validation data to provide test users with information for interpreting test scores. Cut scores can be determined empirically using the correlation between the test and an outside criterion or by a panel of expert judges

Finally, the developers compile the test manual, which has been in the process of development along with the test. The test manual includes the answer key, instructions for the administrator and test user, information on test development, validation and cross-validation, norms, and cut scores.

What makes a Good Test?

Psychologists evaluate psychological tests in terms of 'reliability' and 'validity'. Reliability relates to the dependability of test scores while validity refers to what the test is attempting to measure.

What is Validity of a Test?

Validity is the extent to which a test measures what it claims to measure. It is vital for a test to be valid in order for the results to be accurately applied and interpreted. Validity can be divided into *predictive validity*, *concurrent validity*, *content validity*, and *construct validity*. The first two of these may be considered together as *criterion-oriented* validation procedures.

What is Criterion-Related Validity?: A test is said to have criterion-related validity when the test is demonstrated to be effective in predicting criterion or indicators of a construct. There are two different types of criterion validity — predictive validity and concurrent validity.

Predictive validity: occurs when the criterion measures are obtained at a time after the test. Examples of test with predictive validity are career or aptitude tests, which are helpful in determining who is likely to succeed or fail in certain subjects or occupations

Concurrent validity: occurs when the criterion measures are obtained at the same time as the test scores. This indicates the extent to which the test scores accurately estimate an individual's current state with regards to the criterion. For example, on a test that measures levels of depression, the test would be said to have concurrent validity if it measured the current levels of depression experienced by the test taker. Also, concurrent validity is studied when one test is proposed as a substitute for another (for example, when a multiple-choice form of spelling test is substituted for taking dictation), or a test is shown to correlate with some contemporary criterion (e.g., psychiatric diagnosis).

Content validity: is established by showing that the test items are a sample of a universe in which the investigator is interested. Content validity is ordinarily to be established deductively, by defining a universe of items and sampling systematically within this universe to establish the test.

Construct Validation: Test has construct validity if it demonstrates an association between the test scores and the prediction of a theoretical trait. Intelligence tests are one example of measurement instruments that should have construct validity. It is involved whenever a test is to be interpreted as a measure of some attribute or quality which is not "operationally defined." Indices of reliability and validity range between 0 and 1. Good reliabilities are in the order of .75 and above, however, validities are usually somewhat lower and partly depend on the context in which the test is being used. Generally speaking, acceptable validities range between about .20 to about .50. Reliability is a prerequisite for validity since an unreliable test cannot validly measure anything.

Writing the Instructions for the New Test

The instructions for the test administrator should cover group or individual

administration, specific requirements for the location, required equipment, time limitations or approximate time for completion of the test, script for the administrator to read to test takers, and required training for the test administrator. The instructions on the test need to be simple, concise, and written at a low reading level. Complicated instructions for responding are likely to lead to confused test takers and an increased probability of response errors. The scoring instructions and test key ensure that each person who scores the test will follow the same process. The scoring instructions should explain how the test scores relate to the construct the test measures

Models of Clinical Assessment

Diagnoses: The information-gathering model represents the use of standardized tests to make diagnoses.

Intervention: The therapeutic model represents the use of tests as an intervention that provides new information for the client to use for self-discovery and growth.

Evaluating programme outcomes: The differential treatment model represents the use of tests for conducting research or evaluating program outcomes

Making Diagnoses Using the Interview and Structured Personality Tests

A diagnosis is the definition of a client's problem or disorder, and screening is the process of conducting a psychological assessment to arrive at a diagnosis. The clinical interview in which the practitioner observes and gathers information about the client is a primary diagnostic tool. There are three types of interviews: (a) the structured clinical interview, which has a predetermined set of questions and yields a score; (b) the nondirective clinical interview, in which the practitioner's questions follow up on the client's report of symptoms or problems; and (c) the semistructured interview, in which some questions are predetermined but the practitioner also asks follow-up questions based on the client's responses.

The practitioner who uses the nondirective approach risks three sources of bias: hypothesis confirmation bias, the self-fulfilling prophecy, and ethnocentrism. The nondirective interview may be more useful for an intervention (the therapeutic model) than for a diagnosis. Practitioners also use standardized personality tests, such as the Minnesota Multiphasic Personality Inventory-2 (MMPI-2), the NEO Personality Inventory, the California Personality Inventory (CPI), and the 16 Personality Factor Questionnaire (16PF), to make diagnoses

Projective Techniques

Projective techniques ask test takers to give meaning to ambiguous stimuli. Projective story telling requires test takers to tell a story about some visual stimuli such as pictures. In projective drawing, test takers draw and interpret their own pictures. In sentence completion tests, the assessor administers partial sentences, either verbally or on paper, and asks test takers to respond by

completing each sentence. A major weakness of most projective tests is a lack of evidence of traditional psychometric soundness such as reliability and validity. Therefore, the value of projective tests may be in their usefulness as interventions rather than as diagnostic or research instruments.

Neuropsychological Tests

Neuropsychology is a branch of psychology that focuses on the relation between brain functions and behaviour. Neuropsychologists use electrophysiological techniques such as the electroencephalogram (EEG), a continuous written record of brainwave activity, and event-related potential (ERP), a record of the brain's electrical response to the occurrence of a specific event. Developmental applications involve assessments that determine whether the client is developing normally. Two developmental tests for infants are biochemical assessment, which is an analysis of blood gases to determine the concentration of oxygen and carbon dioxide at the tissue level, and electrophysiological assessment, which involves monitoring vital signs such as heart rate and spontaneous electrical activity of the brain.

Neuropsychologists conduct neurobehavioral assessments such as eliciting various reflexes to assess the functioning and maturity of an infant's central nervous system. Neuropsychologists use tests that measure temperament or personality as well as intelligence, academic achievement, language, perception, and attention tests to assess social-emotional functioning in children. When treating the elderly, neuropsychologists take into account the normal changes in brain function that result from aging, including the fact that aging can exaggerate the psychological impact of medical conditions. Three major areas of psychopathological disorders – anxiety, depression, and schizophrenia – have been researched extensively by neuropsychologists.

Types of Psychological Test

Thousands of different tests have been developed to measure hundreds of different psychological concepts. Some have stood the test of time although others have not. Anything conforming to the above definition can be viewed as a test although they may also be referred to as "inventories", "questionnaires" or "blanks". Some tests, such as tests of intelligence or achievement, are made up of items with well defined correct answers; while others, such as personality inventories, do not have right or wrong answers, but are designed to capture a person's predispositions, tendencies and preferences.

Aptitude

Aptitude tests seek to measure candidate's qualities or traits (physical, social, conceptual, analytic, or practical). The outcome of the test will aim to assist in predicting the candidate's capacity (potential) to develop those competences/skills that are needed to perform a task/job/role well. The traits that we associate

with job performance must be definable in concrete, measurable terms.

Infant Development Scales

Early intervention can help children with handicaps reach their fullest potential, so it is important that handicapping conditions be identified as soon as possible. Several tests have been constructed to compare an infant's *developmental* level with the *expected* level for his or her age group. They identify children who are "at risk". The most commonly used infant scales include the following four:

- **Brazelton Neonatal Behavioural Assessment Scale:** This scale tests an infant's (1) neurological intactness; (2) interactive behavior (including control of motor functions such as putting the thumb in the mouth and remaining calm and alert in response to stimuli such as a bell, a light, and a pinprick); and (3) responsiveness to the examiner and need for stimulation. This test is administered during the newborn period only.
- **Bayley Scales of Infant Development:** These scales test mental abilities including memory, learning, and problem-solving behaviour; motor skills; and social behaviors such as social orientation, fearfulness, cooperation, and language-both receptive and expressive.
- **Gesell Developmental Schedules:** These schedules test for fine- and gross-motor skills; language behavior; adaptive behavior including eye-hand coordination, imitation, and object recovery; and personal-social behaviour including reaction to persons, initiative, independence, and play response.
- **Denver Developmental Screening Test:** This test is used to identify problems or delays that should be more carefully evaluated at a later time. It measures four areas: personal/social, fine-motor/adaptive, language, and gross-motor skills.
- **Preschool and School-age Intelligence Tests**
- **Stanford-Binet:** This test can be used with both preschool and school-age children and is usually given to children between the ages of two and eight. Examples of what is required include remembering where an object was hidden, building a four-block tower to match an existing tower, explaining the uses of common objects, and identifying by name pictured objects. Some items on the Stanford-Binet are culture-specific so the test is best suited for middle-class English-speaking children. Depending upon the child's age, the test requires vision, eye-hand coordination, hearing and speech.
- **Wechsler:** This test has separate forms for preschool and school-age children. The pre-school form is called the Wechsler Preschool and Primary Scale of Intelligence (WPPSI), and the school-age form is called Wechsler Intelligence Scale for Children-Revised (WISC-R). The WISC-R is the test most likely to be used to assess the cognitive functioning of a school-age

child. It has six verbal and six performance subtests.

- **Special Abilities Tests**

- In addition to IQ Tests, there are special abilities tests. Such tests can provide valuable information prior to a full-scale evaluation or can add to information obtained from an IQ test.

- **Bender Visual Motor Gestalt Test:** This test is used to assess visual perceptual skills and eye-hand coordination. The child is given nine geometric figures, one at a time, and asked to copy them.

- **Peabody Picture Vocabulary Test, Revised:** This test assesses familiarity with vocabulary words without requiring a child to speak. The child is shown four pictures at a time and must point to (or otherwise indicate) the one that corresponds to the word the tester says.

- **Wide Range Achievement Test Revised-revised (WRAT-R):** This achievement test provides a gross assessment of the child's math, reading recognition and spelling skills.

- **Key Math:** This test evaluates a child's grade level in the attainment of math skills including basic operations, time, money, fractions, etc.

Peabody Individual Achievement Test (PIAT): This test evaluates achievement in reading and spelling. The child's ability to recognize letters and words as well as his or her comprehension skills are assessed.

- **Kaufman-Assessment Battery for Children (K-ABC):** This battery of tests is designed to assess the child's style of information processing. This can be very helpful in understanding a child's learning style and can be used by the psychologist or teacher to plan teaching strategies. An achievement scale is also included in the test. The test has been researched widely with minority populations.

Personality Tests

Personality tests can be used to help determine a child's emotional state. Tests of this kind provide the child with hypothetical situations based on real life. The child's responses to these situations allow the psychologist to gather information about the unique features of the child's personality. The results of personality tests should be considered in conjunction with the observations of those familiar with the child (particularly parents) and the results of other psychological tests. Generally there are two types of tests: objective and projective.

Objective Tests: Objective tests are pencil-and-paper tests containing several hundred items designed to determine the child's predominant personality traits or behaviour. The best known objective test is the Minnesota Multiphasic Personality Inventory (MMPI), which was originally designed for use with adults but can be used also with adolescents. The Personality Inventory for Children

(PIC) is one of the few objective tests for younger children, but questions are answered by the parents or the results are based on the parents' view of the child's behaviour. In Nigeria, Adolescent Personal Data Inventory is available

Projective Tests: Projective tests provide the child with a stimulus-such as inkblots, a set of pictures, or incomplete sentences-with the idea that the child's responses will reveal his or her unique view of the world, including issues of concern and emotional needs. Another type of projective test provides instructions for the child to draw a picture, again with the idea that the drawing will reveal information about the child's inner self.

The Rorschach Test was the first inkblot test and the one still most commonly used. The Holtzman Inkblot Technique is another projective test that may be substituted for the Rorschach. The most common picture-story tests are the Thematic Apperception Test (TAT) and the Children's Apperception Test (CAT). Other picture story tests are the Michigan Picture Test, the Tasks of Emotional Development Test, the Blacky Pictures, and the Make-a-Picture-Story Test. Completion tests consist of a series of incomplete sentences or stories, which the child is asked to complete. One such test is the Rosenzweig Picture Frustration Study, which presents a frustrating situation in cartoon form. The statements of one character are left blank for the child to fill in.

Draw-a-person, Draw-a-man and Kinetic Family Drawing: Tests are also often used as projective tests. The child may be asked to draw pictures of a person, house, tree or a family. These pictures often reveal the child's feelings about himself and other important people in his or her life. Each of these tests can provide useful information about a child's needs and concerns. Each test, however, relies heavily upon the interpretations of a psychologist, and thus requires that he or she be well trained, experienced, and competent in using the particular technique. The tests must be interpreted cautiously and used only in conjunction with other sources of information about the child.

REFERENCES

- Ebel, R.L., & Frisbie, D.A. (1986). *Essentials of educational measurement*. Englewood Cliffs, NJ: Prentice-Hall.
- Gronlund, N.E., & Linn, R.L. (1990). *Measurement and evaluation in teaching* (6th ed.). New York: MacMillan.
- Henrysson, S. (1971). Gathering, analyzing, and using data on test items. In R.L. Thorndike (Ed.), *Educational measurement* (p. 141). Washington DC: American Council on Education.
- Millman, J., & Greene, J. (1993). The specification and development of tests of achievement and ability. In R.L. Linn (Ed.), *Educational measurement* (pp. 335-366). Phoenix, AZ: Oryx Press.
- Nunnally, J.C. (1972). *Educational measurement and evaluation* (2nd ed.). New York: McGraw-Hill.

- Pedhazur, E.J., & Schmelkin, L.P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Erlbaum.
- Popham, W.J. (1981). *Modern educational measurement*. Englewood Cliff, NJ: Prentice-Hall.
- Sax, G. (1989). *Principles of educational and psychological measurement and evaluation* (3rd ed.). Belmont, CA: Wadsworth.
- Wiersma, W. & Jurs, S.G. (1990). *Educational measurement and testing* (2nd ed.).
- Wood, D.A. (1960). *Test construction: Development and interpretation of achievement tests*. Columbus, OH: Charles E. Merrill Books, Inc