

**COURSE  
GUIDE**

**EGC 809  
PSYCHOLOGICAL TESTING**

**Course Team**      Dr. Aminu Kazeem Ibrahim (                    ) –  
NOUN



**NATIONAL OPEN UNIVERSITY OF NIGERIA**

© 2017 by NOUN Press  
National Open University of Nigeria  
Headquarters  
University Village  
Plot 91, Cadastral Zone  
Nnamdi Azikiwe Expressway  
Jabi, Abuja

Lagos Office  
14/16 Ahmadu Bello Way  
Victoria Island, Lagos

e-mail: [centralinfo@nou.edu.ng](mailto:centralinfo@nou.edu.ng)

URL: [www.nou.edu.ng](http://www.nou.edu.ng)

All rights reserved. No part of this book may be reproduced, in any form or by any means, without permission in writing from the publisher.

Printed 2016, 2017

ISBN: 978-058-400-X

| <b>CONTENTS</b>                           | <b>PAGE</b> |
|---|-------------|
| Introduction.....                         | iv          |
| Course Aims.....                          | iv          |
| Course Objectives.....                    | v           |
| Working through the Course.....           | v           |
| Course Materials.....                     | vi          |
| Study Units.....                          | vii         |
| Textbooks and References.....             | viii        |
| Presentation Schedule.....                | viii        |
| Assignment File.....                      | viii        |
| Assessment.....                           | viii        |
| Tutor-Marked Assignment.....              | .ix         |
| Final Examination and Grading.....        | ix          |
| Course Marking Scheme.....                | .ix         |
| How to Get the Most from this Course..... | xii         |
| Facilitators/Tutors and Tutorials.....    | xii         |
| Summary and Conclusion .....              | xii         |

## **INTRODUCTION**

The course EGC 809: Psychological Testing is a two credit unit course designed for Masters Students in Guidance and Counseling Programmes, of the National Open University of Nigeria. This Course Guide tells you briefly what the course is all about, the course materials you will need and how you can derive maximum benefit from the customised self-instructional materials. It also provides hints on your Tutor-Marked Assignment, details of which will be given to you at your study centre. Tutorial sessions, where you can seek clarifications on the course materials will be arranged for you at your centre. It is in your own interest to attend the tutorial sessions! The time and venue will be made known to you at your center.

The course is a three module course designed to equip the prospective counselors with the psychological testing concepts, kinds, uses and classifications of test validity and reliability standardisation and basic test construction principles including objective test items and finally test administration, scoring analysis and interpretation required to analyse, synthesise, diagnose, prognose and offer effective treatment (counseling). The first module introduces you to the Concept of Psychology, psychological testing; kinds, uses and classifications of psychological tests and relevance of psychological tests to counseling were clearly discussed. The second module deals with educational tests and measurement, test validity and reliability, types of validity and reliability, theory of reliability and methods of estimating tests reliability and principles of test construction and the third module deals with tests standardisation, uses of strong vocational interest blank, kuder preferential blank, achievement test and intelligence test, uses of personality test, problems associated with the use of personality tests and ways of overcoming the problems associated with the uses of personality tests. The ability to get the right test or measurement at the right time is seen as an essential ingredient in counseling process.

At the end of this course; Psychological Testing you are expected to be able to identify, adopt/adapt, modify, develop or construct an effective tests that will help you in your counseling profession.

## **COURSE AIMS**

The aim of this course is to expose you to the details of psychological testing. This aims will be achieved by:

- Introducing you to psychology, psychological testing; educational testing and measurement.
- Understanding the relevance of psychological testing to counseling
- Teaching you the kinds, uses and classifications of test
- Introducing you to principles of test construction with emphasises on objective test items
- Differentiating between psychological, education and counseling tests
- Understanding the nature and functions of testing with types and uses of aptitude test.

## **COURSE OBJECTIVES**

On successful completion of this course you should be able to:

- define psychology and identify branches psychology
- explain the concept of psychological testing
- enumerate the relevance of psychological testing to counseling
- identify the kinds, uses and classifications of test
- define the concept of test validity, reliability and standardisation
- list and explain types of validity and reliability
- calculate the reliability estimate of a test
- apply the principles of test construction to the construction of objective test items
- compare the psychological, educational and counseling tests
- discuss the uses of strong vocational interest Blank, Kuder Preferential Blank, Achievement, Intelligence and Personality tests
- identify problems associated with the use of psychological tests
- discuss ways by which problems associated with personality testing can be overcome.

## **WORKING THROUGH THE COURSE**

You are required to spend a lot of time and commitment to studying this course; you are required to read books, journals, conference proceedings and other materials. You will need to find out more about the examples given to illustrate certain points. You will need to spend three hours on each unit of this course.

We advise that you avail yourself of the opportunity of the tutorial sessions, where you benefit from clarifications from your course facilitators, and the interaction with your colleagues.

## COURSE MATERIALS

Major components of the course are:

1. Course Guide
2. Study Units
3. Textbooks
4. Assignment File
5. Presentation Schedule

## STUDY UNITS

There are three modules in this course divided into 15 study units

### Module 1 Overview of Psychological Testing

- Unit 1 Introduction to Psychology
- Unit 2 Psychological Testing
- Unit 3 Kinds, Uses and Classification of Psychological Tests
- Unit 4 Distinctions between Psychological, Educational and Counselling Tests.
- Unit 5 Relevance of Psychological Tests to Counseling

### Module 2 Test Development

- Unit 1 Educational Testing and Measurement
- Unit 2 Tests Validity and Reliability
- Unit 3 Types of Validity and Reliability
- Unit 4 Theory of Reliability and Methods of Estimating Test Reliability
- Unit 5 Principles of Test Construction and Its Application to the Construction of Objective Test Items

### Module 3 Psychological, Educational and Counseling Tests

- Unit 1 Tests Standardisation
- Unit 2 Uses of Strong Vocational Interest Blank, Kuder Preferential Blank, Achievement Tests, and Intelligence
- Unit 3 Uses of Personality Tests
- Unit 4 Problems Associated with Personality Tests
- Unit 5 Ways of Overcoming Problems Associated with Personality Tests

Each unit consists of three hours works and includes the specific objectives, direction for study, reading materials and other sources. The units consist of Self-Assessment Exercise questions, which will assist you to test what you have read in the course material. The self-assessment exercise helps you to assess your progress and reinforce your understanding of the material. These exercises at the end of each unit together with the tutor-marked assignment will assist you in achieving the stated objectives at the beginning of each unit.

### **TEXTBOOKS AND REFERENCES**

These textbooks are necessary for you. You should try to buy them yourself.

Abdulsalami, A. S. (2005). *Research Methods in Education*. Nigeria: Stirling- Horden Publishers Ltd.

Muhammed Y. B. (1998). *Measurement and Evaluation in Education*. Kano: Debis Co. Press & Publishing Company Ltd.

Saka, A. O. Y. (1996). *A Collection of Essays on Guidance and Counseling*. Kano: Debis Co. Press & Publishing Ltd.

Sani, S. (2008). *Understanding Guidance and Counseling*. Zaria: Ahmadu Bello University Press Ltd.

Robert, J. G. (2003). *The History of Psychological Testing*. America Psychological. Standards for Education and Training in Psychological Assessment.

Position of the Society for Personality Assessment – An Official Statement of the Board of Trustees of the Society for Personality Assessment. *Journal of Personality Assessment*, 87, pp.355 – 357.

Aminu, K. I. & Gali, S. (2012). *Guidance and Counseling: Theories and Practice in Schools*. Kano: Albarka Printing Ltd.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999).

Standards for Educational and Psychological Testing. Washington, DC: American Educational Research Association.

## **PRESENTATION SCHEDULE**

Dates for prompt completion and submission of your Tutor-Marked Assignment (TMAs) and attendance of tutorials will be reflected in your course materials. *You should remember to submit all assignments at the stipulated date and time.* You should work as scheduled, and do not lag behind in your work!

## **ASSIGNMENT FILE**

You will find details of works you are to submit to your tutor for marking in your assignment file. The marks you obtain for your assignments will count towards the final mark you obtain for this course. Further information on assignments will be found in the assignment file itself and in the assessment section of this course guide. Each unit of this course has at least two assignments. These assignments are meant to help you understand the course and assess your progress.

## **ASSESSMENT**

There are three aspects to the assessment for this course. The first is in form of the Self-Assessment Exercise, the second are the Tutor-Marked Assignments, and the third is the written examination/final examination. You are advised to attempt the exercises with all sincerity as that will assist you greatly. In tackling the assignments, you are expected to apply the information, knowledge and techniques gathered during the course. The assignments must be submitted to your tutor facilitator for formal assessment in accordance with the deadlines stated in the presentation schedule and the assignment file. The work you submit to your tutor for assessment will account for 30% of the total course score. At the end of the course, you will sit for a final or end of course examination of about three hours duration. This examination will form 70% of the total score for the course.

## **TUTOR-MARKED ASSIGNMENT**

The tutor-marked assignment (TMA) is a continuous assessment component of your course. You are required to submit at least four (4) TMAs before you are allowed to sit for the end of course examination. The TMAs would be given to you by your facilitator, and would be returned to you after they have been graded.

Assignment questions for the units in this course are contained in the assignment file. You will be able to complete your assignment from the information and materials contained in your study unit and references. However, it is desirable to demonstrate that you have read and researched into other references; this will give you a wider view point and a deeper understanding of the subject. Ensure that each assignment reaches your facilitator before the deadline in the presentation schedule and assignment file. If for any reason you cannot complete your work on time, contact your facilitator before the assignment is due to discuss the possibility of extension. Extension sought after the due date, may not be granted.

## **FINAL EXAMINATION AND GRADING**

The end of course examination for this course will be about three hours, and will account for 70% of the total course score. The questions will be fashioned after the self-assessment exercise and tutor-marked assignments that you have previously encountered during your course. All areas of the course will be examined. The time for the examination is not fixed, but you will be given adequate notice of the examination.

## **COURSE MARKING SCHEME**

### **Assessment**

### **Marks**

Assignment 1 – 4    The best three (3) of the four assignments will be rated based on 10% each making a total of 30% of the course marks.

Final Examination  
marks

70% of overall course

Total

100% of course marks

## **HOW TO GET THE MOST FROM THIS COURSE**

In distance learning, the study units replace the university lectures. This is one of the great advantages of distance learning; you can read and work through specially designed study materials at your own pace, and at a time and place that suits you. Think of it as reading the lecture instead of listening to the lecturer. The study units tell you when to read, your text materials and recommended books for your further reading. You are provided exercises to attempt at appropriate point in time, just as a lecturer might give in a classroom situation.

Each of the study units follows a common format. The first item is an introduction to the subject matter of the unit, and how a particular unit is integrated with other units, and the course as a whole. Next to this is a set of learning objectives. These objectives state the mental tasks you should be able to accomplish by the time you have completed the unit. These learning objectives are therefore, meant to guide your study. The moment a unit is finished, you must go back and check whether you have achieved the objectives. If this is made a habit, you will significantly improve your chances of passing and transferring the course.

The main body of the unit guides you through the required reading from other sources; this either from your references or a reading section. The following is a practical strategy for working through the course: If you run into any trouble, telephone your tutor or visit the study centre nearest to you. Remember, your tutor's job is to help you. When you need assistance, do not hesitate to ask your tutor to provide it.

### **Read this Course Guide thoroughly**

Organise a study schedule. Design a 'Course Overview' to guide you through the course. Note the time you are expected to spend on each unit and how the assignments relate to the units. Important information, e.g. details of your tutorials, and the date of the first day of the semester is available at the Centre. You need to gather all the information into one place, such as your diary or a wall calendar. Decide on whatever method you choose, and write in your own dates and schedule of work for each unit.

Once you have created your study schedule, do everything to stay faithful to it. The major reason why students fail is that they lag behind in their course work. If you get into difficulties with your schedule, please, let your tutor know before it is too late for help.

Turn to Unit 1, and read the instruction and the objectives for the unit. Assemble the study materials. You will need your references and the unit you are studying at any point in time.

As you work through the unit, you will know what sources to consult for further information.

Visit your study centre whenever you need up-to- date information. Before the relevant due dates (about four weeks before due dates), visit your study centre for your next required assignment. Keep in mind that you will learn a lot by doing the assignments carefully. They

have been designed to help you meet the objectives of the course and, therefore, will help you pass the examination. Submit all assignments as at when due.

Review the objectives for each study unit to confirm that you have achieved them. If you are not sure about any of the objectives, you can move to the next unit. Study unit by unit through the course, and try to space your study so that you can keep to the schedule.

When you have submitted an assignment to your tutor for marking, do not wait for its return before starting the next unit. Keep to your schedule. When the assignment is returned, pay particular attention to tutor's comments, both on the tutor-marked assignment form and also the written comments on the ordinary assignments.

After completing the last unit, review the course and prepare yourself for the final examination. Check that you have achieved the unit objectives (listed at the beginning of each unit), and the course objectives (listed in the Course Guide).

## **FACILITATORS/TUTORS AND TUTORIALS**

There are 15 hours of tutorial provided in support of this course. You will be notified of the dates, times and venue of these tutorials, as well as the name, and phone number of your facilitator, as soon as you are fixed in a tutorial group.

Your tutor or facilitator will mark and comment on your assignments. She/he keeps a close watch on your progress, so as to render necessary assistance when required. You mail your tutor-marked assignment to your tutor before the scheduled date (at least two weeks are required). They will be marked by your tutor and returned to you as soon as possible.

Do not hesitate to contact your facilitator by telephone, e-mail, and discuss your problems for necessary assistance.

The following might be circumstances in which you would find help necessary.

Contact your tutor if:

- you do not understand any part of the study units of the assigned reading.
- you have difficulty with the self-test or exercises.
- you have a question or problem with an assignment or with the grading of an assignment.

You should try your best to attend the tutorials. This is the only chance for a face-to-face contact with your course facilitator, and to ask questions which are answered instantly. You can raise any problems encountered in the course of your study. To derive maximum benefit from course tutorials prepare a list of questions before the tutorial session. You will learn a lot by your active participation in the discussion.

## **SUMMARY AND CONCLUSION**

This course is expected to enrich your scholarship on psychological testing in terms of concept of psychological testing, meaning of psychology, educational testing, measurement and relevance of psychological testing to counseling. It entails kinds, uses and classifications of test, test validity, reliability and standardisation of tests, with principles of test construction by placing emphasises on objective test items. Distinctions between psychological, education and counseling tests, nature and functions of testing, uses of aptitude tests such as strong vocational interest Blank, Kuder Preferential Blank, Achievement test, Intelligence test, personality test and the problems associated with personality testing. The modules and units have been carefully organised to enable you achieve the course objectives and it is hoped that the various assessment strategies and resources will provide the necessary support you need for a successful completion of this course. It is expected that you find this course motivating as we wish you success as you begin your relationship with NOUN.

**MAIN  
COURSE**

| <b>CONTENTS</b> |   | <b>PAGE</b> |
|-----------------|---|-------------|
| <b>Module 1</b> | <b>Overview of Psychological Testing..... ..</b>  | <b>1</b>    |
| Unit 1          | Introduction to Psychology..... ..  | 1           |
| Unit 2          | Psychological Testing..... ..   | 6           |
| Unit 3          | Kinds, Uses and Classification of Psychological Tests..... ..   | 11          |
| Unit 4          | Distinctions between Psychological, Educational and Counselling Tests..... ..                                   | 18          |
| Unit 5          | Relevance of Psychological Tests to Counseling....  | 21          |
| <b>Module 2</b> | <b>Test Development .....</b>   | <b>24</b>   |
| Unit 1          | Educational Testing and Measurement..... ..   | 24          |
| Unit 2          | Tests Validity and Reliability..... ..  | 31          |
| Unit 3          | Types of Validity and Reliability..... ..   | 34          |
| Unit 4          | Theory of Reliability and Methods of Estimating Test Reliability..... ..  | 47          |
| Unit 5          | Principles of Test Construction and Its Application to the Construction of Objective Test Items..... ..         | 50          |
| <b>Module 3</b> | <b>Psychological, Educational and Counseling Tests..</b>  | <b>61</b>   |
| Unit 1          | Tests Standardisation..... ..   | 61          |
| Unit 2          | Uses of Strong Vocational Interest Blank, Kuder Preferential Blank, Achievement Tests, and Intelligence..... .. | 64          |
| Unit 3          | Uses of Personality Tests..... ..   | 68          |
| Unit 4          | Problems Associated with Personality Tests..... ..  | 71          |
| Unit 5          | Ways of Overcoming Problems Associated with Personality Tests..... ..   | 74          |

## **MODULE 1      OVERVIEW OF PSYCHOLOGICAL TESTING**

|        |   |
|--------|---|
| Unit 1 | Introduction to Psychology  |
| Unit 2 | Psychological Testing   |
| Unit 3 | Kinds, Uses and Classifications of Psychological Tests                |
| Unit 4 | Distinctions between Psychological, Educational and Counselling Tests |
| Unit 5 | Relevance of Psychological Tests to Counselling                       |

### **UNIT 1      INTRODUCTION TO PSYCHOLOGY**

#### **CONTENTS**

|     |   |
|-----|---|
| 1.0 | Introduction  |
| 2.0 | Objectives  |
| 3.0 | Main Content  |
| 3.1 | Meaning of Psychology                               |
| 3.2 | Branches of Psychology                              |
| 3.3 | Methods of Scientific Investigation in Psychology   |
| 3.4 | Relevance of Psychology to Guidance and Counselling |
| 4.0 | Conclusion  |
| 5.0 | Summary   |
| 6.0 | Tutor-Marked Assignment                             |
| 7.0 | References/Further Reading                          |

#### **1.0      INTRODUCTION**

Counseling is much more than providing assistance to the clients. To be effective counselors acquire a body of knowledge and a set of skills that help them accomplish a behavioural change in the client and meet the challenges of the clients. Where does this knowledge come from? One source is the discipline of psychology. Counselors and psychologists are interested in the same things: behaviour and changes in behaviour. Counselors must deal with behaviour in the counseling process, attempting to bring about changes in behaviour in accord with counseling goals (immediate, long-range or process counselling goals). This ascertains makes the meaning, braches, methods and relevance of psychology, to be introduced in this course.

## **2.0 OBJECTIVES**

At the end of this unit, you should be able to:

- define psychology
- identify the branches of psychology
- describe methods of scientific investigation in psychology
- explain the relevance of psychology to guidance and counseling.

## **3.0 MAIN CONTENT**

### **3.1 Meaning of Psychology**

Psychology is described as the study of behaviour and the processes involved in changing it (Rosser, 1982). Human being contains a complex network of abilities, potentialities, competencies, propensities, attitudes, feelings, thoughts and past experiences that can be determined through psychological testing for the purpose of change or modification.

Psychology according to Chauhan (1979) is defined as a science of behaviour. This definition contains science and behaviour. Science is generally defined as a systematic body of knowledge which may be verified at any time by any person or number of persons under certain conditions. Science tries to find out those conditions and factors which determine or cause the occurrence of a particular event through the process of experiments, guided observations and testing. Behaviour as defined by J. B. Watson, an American psychologist, is an action which can be seen and observed in an objective way. Behaviour may be internal or external stimulation which can be measured or tested in an objective way.

Psychology deals with concepts of learning, memory, perception, motivation and personality traits. All these concepts lead to behaviour. Behaviour in this context may mean anything that we do such as thinking, talking, sneezing, laughing, sleeping and even loving (Orji, 1997). Psychology as an ancient discipline was part of philosophy.

### **3.2 Branches of Psychology**

Psychology in a sense appears to be an umbrella word because, psychology is used in almost all spheres of human life to understand and improve the existing conditions. There are many branches of psychology, as psychology

turns its attention to human, we have human psychology therefore, and branches of psychology include:

1. Clinical psychology
2. Industrial psychology
3. Developmental psychology
4. Physiological psychology
5. Social psychology
6. Counselling psychology
7. Military psychology
8. Experimental psychology
9. Genetic psychology
10. Community psychology
11. Abnormal psychology
12. Educational psychology
13. Animal psychology
14. General psychology
15. Sport psychology
16. Occupational psychology

### **3.3 Methods of Scientific Investigation in Psychology**

Psychology did not emerge as a science until the later part of the nineteenth century (Rosser, 1986). Prior to that time, psychological questions were considered the province of philosophy. Philosophy and psychology still has a close link, and both focus on the same general issues that is, the nature of man. Today however, the two disciplines have come to use different methods. Philosophy depends on a purely logical method, psychology on a scientific method. This method is an approach to the answering of question to the gathering of information, and to the drawing of conclusions. Borrowed from the older sciences, it has proven itself a useful technique, one based on demonstration and testing of phenomenon. Specifically, methods used in the study of psychology are:

- Case history or case study method
- Experimentation
- Longitudinal and cross sectional method
- Observation
- Correlational etc. All these methods follows a laid down steps of:
  - Awareness of the problem
  - Location and definition of the problem
  - Collection of data

- Formulation of hypotheses
- Testing the hypotheses
- Verification of the hypotheses
- Making a valid conclusion. These follow the strategies of professional counselors in many situations.

### **3.4 Relevance of Psychology to Counselling**

The relevance of psychology to counseling are:

- Psychology provides the counselor with the technique of understanding human behaviour. This understanding is used by the counselor to help counselees solve the problems they have brought before him. For example, if a brilliant and well-behaved child suddenly begins to perform poorly in his academic work and becomes defiant, an investigation by the counselor may reveal that the strange behaviour is meant to achieve some goal which may not be known even to the child himself.
- Psychologists have made important contributions to counseling in the form of data and theories such as behavioural and cognitive theories
- Psychotherapy researches are in line with and applied to counseling practices
- Psychology also assists the counselor to understand developmental characteristics. For example, the pre-school child behaves differently from the child in lower or upper school level. The major characteristics, needs, tasks and accomplishments of the different stages are utilised for counseling delivery
- Psychology provides tools for the assessment of abilities and progress of clients
- Psychological testing are used in many counseling processes to diagnose, prognose, predict and so on of client related problems

### **4.0 CONCLUSION**

It may therefore be said that psychology is to the counselor while physiology and anatomy are to the physician. Practicing counselors need an understanding of the psychology before comprehending the concept of psychological testing. Without it counseling treatment may be superficial.

## 5.0 SUMMARY

Psychology is defined as the science of behaviour and has many branches including counselling psychology. Scientific methods are used in the study of psychology that involves testing and application of psychological testing. Relevance of psychology to counseling was also outlined.

## 6.0 TUTOR-MARKED ASSIGNMENT

1. In your own experience, what is psychology?
2. Identify any five branches of psychology.
3. Discuss any five relevance of psychology to the counselor?

## 7.0 REFERENCES/FURTHER READING

- Adeloye, T. (2005). *Fundamental Educational*. Kano: Debis – Co Press and Publishing Limited.
- Orji, A. S & Anikweze, C. M. (1997). *Adolescent Psychology*. Oweri: Peace Press Limited.
- Roser, R. A. (1982). Information Use by Pre School Children in Altruistic Decision-Making: An Exploratory Investigation of Donating Behaviour. *Journal of Genetic Psychology*. 141, pp. 19 -27.
- Saka, A. O. Y. (1996). *A Collection of Essays on Guidance and Counseling*. Kano: Debis-Co Press and Publishing Limited.

## UNIT 2 PSYCHOLOGICAL TESTING

### CONTENTS

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Meaning of Psychological Test
  - 3.2 The Field of Psychological Testing
  - 3.3 Principles of Psychological Testing
  - 3.4 Psychological Test Security
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

### 1.0 INTRODUCTION

The history of psychological test can be traced to the imperial examination system in China. The imperial examination is an early form of psychological testing; it was used to assess candidates based on their proficiency in topics such as civil law and fiscal policies. The ancient Chinese game known in the Western countries as the tangram was used to evaluate a person's intelligence, along with the game Jiulianhuan or nine linked rings. The game being one of the earliest psychological tests in the world was used to assess a person's flexibility and creativity of thinking. Modern psychological testing began in France in the 19th century. It assisted in separating mental retardation from mental illness and reducing the neglect, torture, and ridicule heaped on both mentally retarded and mentally illness persons. For counselors therefore, psychological testing can be used and it is being used for many similar psychological or personality problems.

### 2.0 OBJECTIVES

At the end of this unit, you should be able to:

- explain the meaning of psychological testing
- outline the scope (field) of psychological testing
- identify the principles guiding the uses of psychological tests
- discuss the implications of psychological test security on their adoption.

### **3.0 MAIN CONTENT**

#### **3.1 Meaning of Psychological Test**

Psychological test refers to an objective and standardised measure of an individual's mental and/or behavioural characteristics. A psychological test is also an instrument designed to measure unobserved constructs known as latent variables. In nature, psychological tests are typically, but not necessarily, a series of tasks or problems that the client has to solve. Psychological tests can strongly resemble questionnaires that are also designed to measure unobserved constructs, but differ in that psychological tests ask for a respondent's maximum performance. For psychological test to be useful, it must be both valid and reliable (validity and reliability are discussed in later units). In many situations, psychological assessment is similar to psychological testing but psychological assessment usually involves a more comprehensive assessment of the individual. Psychological assessment is a process that involves checking the integration of information from multiple sources, such as tests of normal and abnormal personality, test of ability or intelligence, tests of interests or attitudes, information from interviews, occupation and so on, before a valid and reliable judgment can be made.

#### **3.2 The Field of Psychological Testing**

The field or scope of psychological testing is characterised by the use of samples of behaviour used to assess psychological construct(s). Examples of psychological constructs used in psychological testing include cognitive and emotional functioning of a given individual or client. Psychometrics is the technical term for the science behind psychological testing. It was Francis Galton that coined the terms psychometrics and eugenics and also developed a method for measuring intelligence based on nonverbal sensory – motor tests. By samples of behaviour, one means observations of an individual performing task that have usually been prescribed beforehand, which often means scores on a test. These responses are often means scores on a test. These responses are often compiled into statistical tables that allow the evaluator or guidance counselor to compare the behaviour of the individual being tested to the responses of a 'norm group'. The understanding of psychological testing must be based on strong formulation of the discipline of psychology.

### **3.3 Principles of Psychological Testing**

The principles of conducting psychological testing include:

1. **Standardisation:** All steps including procedures in conducting psychological tests must base on consistency and under the same environment so as to achieve the same testing performance or results from the group or individuals being tested.
2. **Objectivity** psychological tests scoring must be objective that is free from subjective judgments, free from all biases, based on the fact that the same results or performance can be obtained on the test by anybody.
3. **Test Norms:** Psychological tests must have the norms that specifies the average test score within a large group of people where can be compare to the results of others by establishing a point of comparison or frame of reference
4. **Reliability:** Psychological tests results must be the same results that will be obtained after multiple testing
5. **Validity:** A psychological test being designed and administered must measure what it is intended to measure.

### **3.4 Psychological Tests Security in Practical**

In practical terms, a lot of psychological tests are generally not available to the public users including the counselors, but rather, have restrictions that come from publishers of the tests and from psychology licensing boards that prevent the disclosure of the tests with the information about the interpretation of the results.

Publishers of many of these tests also consider both copyright and matters of professional ethics to be involved in protecting the secrecy of their tests, and they sell tests only to people who have proved their education and professional qualifications to the developer's satisfaction.

Sometimes purchasers are legally bound from giving test answers or the tests themselves out to the public unless permitted under the test developer's standard conditions for administration of the tests.

The International Test Commission (ITC) is an international association of national psychological societies and test publishers. This association publishes the international guidelines for Test use, which prescribes to "protect the integrity" of the tests by not publicly describing test techniques and by not "coaching individuals" so that they might unfairly influence

their test performance. The implication of this is that counselors who are interested after coming across an interested psychological test must legally obtain permission from the publisher for their use.

#### **4.0 CONCLUSION**

Knowledge of the discipline of psychology is crucial to the effective application of psychological tests by the counselors.

#### **5.0 SUMMARY**

A psychological test is an objective and standardised measure of an individual's mental and/or behavioural characteristics. The field of psychological testing is characterised by the use of samples of behaviour used to assess psychological construct(s). Examples of psychological constructs used in psychological testing include cognitive and emotional functioning of a given individual or client. There are principles underlying the use of psychological tests, these principles are standardisation, objectivity, test norms, reliability and validity. In addition, most of the psychological tests are not generally available to the public users except through legal permission from the developer and regulated by the international association of national psychological societies.

#### **6.0 TUTOR-MARKED ASSIGNMENT**

1. In your own experience, what do you understand by psychological testing?
2. Discuss the field of psychological testing.
3. Identify and discuss any three principles guiding the uses of psychological tests.

#### **7.0 REFERENCES/FURTHER READING**

American Educational Research Association, American Psychological Association & National Council on Measurement in Education (1999). *Standard for Educational and Psychological Testing*. Washington, D. C.: American Educational Research Association.

Robert, J. G. (2003). *The History of Psychological Testing*. American Psychological.

Mellenbergh, G. J. (2008). Surveys. In H. J. Ader & G. J. Mellenbergh (Eds.). (with contributions by D. J. Hand). *Advising on Research*

*Methods: A Consultant's Companion.* pp. 183 – 209. The Netherlands: Johannes Van Kessel Publishing.

Mellenbergh, G. J. (1989). Item Bias and Item Response Theory. *Internal Journal of Educational Research.* 13 (2) pp. 127 – 143.

## **UNIT 3 KINDS, USES AND CLASSIFICATION OF PSYCHOLOGICAL TESTS**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Kinds or Types of Psychological Test
  - 3.2 Classification of Psychological Tests
  - 3.3 Tests Classification by the Use to which they are put
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

### **1.0 INTRODUCTION**

One of the major instruments used in counseling process is a test. Test is also the most common formal instrument for measuring psychological traits of individual. Psychological tests to counselors are like a thermometer, a Geiger counter and such similar instruments to a physicist. This unit presents kinds, uses and classifications of psychological tests.

### **2.0 OBJECTIVES**

At the end of this unit, you should be able to:

- identify kinds or types of psychosocial tests
- discuss the uses of psychological test
- distinguish the various categories of psychological test.

### **3.0 MAIN CONTENT**

#### **3.1 Kinds or Types of Psychological Test**

There are several types of psychological test, namely:

- Achievement tests
- Aptitude tests
- Intelligence tests
- Diagnostic tests

- Performance tests
  - Public safety employment tests
  - Sexological tests
  - Interest tests
  - Direct observation tests
  - Personality tests
  - Attitude tests
  - Neuropsychological tests
  - Infant and preschool assessment
1. **Achievement Test:** These are tests constructed for the purposes of measuring how much learning of a course objective has taken place or achieved by the group or individual after a given time.
  2. **Aptitude Test:** They are tests constructed for the purposes of measuring the potential or ability of the individual or group to attain or achieve certain attribute or learning at a later stage. Aptitude tests are psychological tests to measure specific abilities, such as mechanical or clerical skills. Sometimes these tests must be specially designed for a particular job, but there are also aptitude tests available that measure general chemical and mechanical aptitudes. Such as Minnesota Clerical Test.
  3. **Intelligence Tests:** These tests may be aptitude tests or achievement tests. For example IQ achievement tests purport to be measures of intelligence, while achievement tests are measures of the use and level of development of use of the ability. IQ (or cognitive) tests and achievement tests are common norm-referenced tests.
  4. **Diagnostic Tests:** These tests are commonly constructed and Administered on individual or group for the purposes of investigating or identifying certain personality problem areas in order to determine needed remedial or counseling intervention.
  5. **Performance Tests:** These tests are designed to measure individual or group physical demonstration of certain abilities and skills. It is an ongoing performance assessment of abilities and skills in a given task
  6. **Public Safety Employment Tests:** These tests are constructed for vocations within the public safety i.e. fire service, law enforcement,

corrections, emergency, medical services. Industrial and organisational psychology tests requires for initial employment and advancement throughout the ranks are under these categories of test. For example, the National Firefighter Selection Inventory – NES and the Integrity Inventory – IT

7. **Sexological Tests:** The field of sexology provides different psychological evaluation devices in order to examine the various aspects of the discomfort, problem or dysfunction regardless of whether they are individual or relational ones. Although, there are limited number of tests that are specifically meant for the field of sexology.
8. **Interest Tests:** These tests are psychological tests designed to measure a person's interests and preferences. The tests are used primarily for career counseling interest tests may include items on daily activities from among which applicants select their preferences. The rationale is that if a person exhibits the same pattern of interests and preferences as people who are successful in a given occupation, then the chances are high that the person taking the test will find satisfaction in that occupation.
9. **Direct Observation Tests:** These types of tests are usually constructed and conducted with families in laboratory settings, home, or with children in a classroom. In most cases, the purposes of these tests are clinical for example, to establish a pre-intervention baseline of a client aggressive behaviour or to observe the nature of a parent-child interaction in order to understand a relational disorder. Example of Direct observational tests are the Parent-Child Interaction Assessment – 1 (PCIA) and the Mac-Arthur Story Stern Battery (MSSB).
10. **Personality Test:** This type of test is constructed to measures personality attributes or traits. The tests are commonly described as either objective tests or projective tests. Although, the terms “objective test” and “projective tests” have come under criticism such in the Journal of Personality Assessment. The recent descriptive are “rating scale or self-report measures of personality” and “free response measures of personality”.
11. **Attitude Tests:** Attitude tests measures an individual or group feelings about an event, person or object. Typically used attitude

tests are Thurstone Scale or Likert Scale so as to measure specific attitudinal items.

- 12. Neuropsychological Tests:** These tests consist of specifically designed tasks used to measure a psychological function known to be linked to a particular brain structure or pathway. These tests can be used in a clinical context to assess impairment after an injury or illness known to affect neuro-cognitive functioning. In research, neuro-psychological tests can be used to contrast neuro psychological abilities across experimental groups.
- 13. Infant and Preschool Assessment:** These tests are designed for children ages birth to six years of age. The tests measure reflexes developmental milestones to sensory and motor skills, language skills, and simple cognitive skills. The assumption behind the tests is that infant and preschool aged children have limited capacities of communication. Examples of these tests include standard – Benet Intelligence Scales, for Early Childhood, McCarthy Scales of Children’s Abilities, Infant Ability Preschool Intelligence and School Readiness, Gesell Developmental Schedules etc.

### 3.2 Classification of Psychological Tests

There are several broad categories of psychological tests, namely:

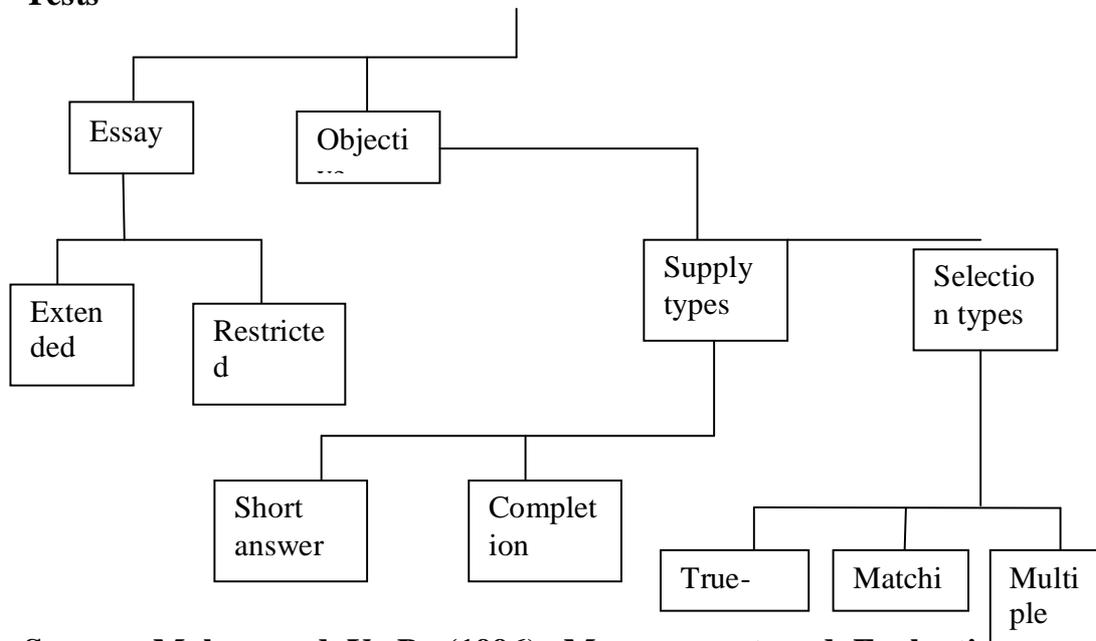
1. Tests classification by the nature of the personality trait being measured. These classification is of two types:
  - a. Tests of maximum performance tests under this classification attempt to determine just how well a person can perform in a specified task. The person is encouraged to do his best and his response is compared with a key prepared by experts which distinguishes between correct and incorrect responses. Examples of test of maximum performance are: (i) aptitude tests (ii) ability tests, and (iii) achievement tests
  - b. Tests of typical performance. This classification of tests tries to get an index of an individual’s usual or typical behaviour as he goes about his day-to-day life. In this case, instead of encouraging the individual to do his best, he is expected to report how he would normally behave in a specified situation. The individual’s responses are not judged right or wrong, but interpreted as truly or not. Test of typical performance include personality tests and interest inventories.

2. Tests classification by how the personality trait is being measured  
 Classifications under this are:

- Objective vs. Subjective (Essay) tests
- Paper and Pencil vs. Performance tests
- Speed vs. Power test
- Individual vs. Group tests
- Language vs. Non-language nor verbal vs non-verbal tests
- Teacher-made tests vs. Standardized tests

a. Objectives tests are tests that cover a wide area of content and are structured in such a way that they can easily be scored, even by machines. There are different forms of objective tests such as multiple-choices, true-false, short-answer matching. The objective tests are also classified into supply types (short answer, completion) tests and selection types (True-False, Matching and Multiple Choice) tests, on the other hand the essay type tests, these tests tries to measure a candidate' knowledge analysis and organisational ability. Essay type tests are not difficult to set of essay type tests include extended response and restricted response type tests. Both objective and essay type tests are examples of paper and pencil tests.

**Table 1: Schematic Presentation of Essay and Objective Type Tests**



**Source: Muhammed Y. B. (1996). Measurement and Evaluation in Education**

- b. Paper and pencil tests are meant to be answered using paper and pencil while performance tests are tests that requires the candidate to physically demonstrate certain abilities and skills
- c. Speed tests are tests that measures a candidate's ability to think fast and accurately while power tests allows candidates to be given enough time to complete all the items on the test before judgment
- d. Individual tests are tests administered to candidates individually but group tests are administered to a number of candidates in a group
- e. Language or verbal tests are tests that measure language fluency and require the candidate to speak-out while non-language or non-verbal tests require non-verbal responses or answers from the candidates.
- f. Teacher-made and standardised tests.

### **3.3 Tests Classification by the Use to which they are put**

This is the third dimension in the classifications of tests. Tests are classified under this category based on the function of testing. For example, placement tests results are used for placement or promotion purposes.

## **4.0 CONCLUSION**

Guidance counselors should realise that knowledge of the kinds, uses and classification of psychological tests will assist them greatly in the selection and utilisation of different types of psychological tests.

## **5.0 SUMMARY**

Different types of psychological tests were identified and discussed. Uses of psychological tests in line with psychological test types were also discussed. The unit also outlined and discussed various classifications of tests as with nature of trait being measured; how the trait is measured and the functions of the trait being measured.

## **6.0 TUTOR-MARKED ASSIGNMENT**

1. Discuss any five uses of guidance psychological tests to the counseling profession.
2. Identify the major classifications of tests according to your own experience.

## 7.0 REFERENCES/FURTHER READING

- Muhammed, Y. B. (1996). *Measurement and Evaluation in Education*. Kano: Debis-Co Press and Publishing Limited.
- Ipaye, T. (1982). *Continuous Assessment in Schools*. Ilorin: University Press.
- Gipps, C. (1990). *Assessment: A Teacher Guide to the Issues*. London: Hodderand Stoughton.
- Abdulsalami, A. S. (2005). *Research Methods in Education*. Nigeria: Stirling Horden Publishers (Nig.) Ltd.
- Yusuf, K. (2009). *Elements of Educational Measurement and Evaluation*. Zaria: Tamnaza Publishing Company Limited.

## **UNIT 4 DISTINCTIONS BETWEEN PSYCHOLOGICAL, EDUCATIONAL AND COUNSELLING TESTS**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Description of Psychological, Educational and Counselling Tests
  - 3.2 Distinctions between Psychological, Educational and Counselling Tests
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

### **1.0 INTRODUCTION**

Tests are mostly developed by test experts, teachers and counselors for different uses. Counselors may adopt, adapt or modify the intended test or may develop their own but in so doing all processes for developing a test must be followed. However, counselor must note that differences exist between psychological, educational and counseling tests. This unit examines these differences.

### **2.0 OBJECTIVES**

At the end of this unit, you should be able to:

- describe psychological tests
- discuss educational tests
- explain counseling tests.

### **3.0 MAIN CONTENT**

#### **3.1 Meaning of Psychological Test**

A psychological test is an objective and standardised measure of an individual's mental and/or behavioural characteristics. A psychological test is an instrument designed to measure unobserved constructs, also known as latent variables. Psychological tests can strongly resemble questionnaires,

which are also designed to measure unobservable constructs, but differ in that psychological tests ask for a respondent's maximum performance whereas a questionnaire asks for the respondent's typical performance.

Educational tests are teacher made tests and standardised measures of learning outcomes. Educational tests are designed to measure what learners have benefited after a given learning outcome. They may be objective items or essay designed to measure the success or failure of a given educational programme. Educational tests usually vary from one context to the other. Educational tests usually followed planned or stated objectives of instruction. A school administrator may construct and administers a test to a large number of candidates and from their performances decides which candidates should be administered.

Counseling tests are psychological or educational tests for counseling purposes. Counseling tests may be adopted, adapted, modified. Psychological or educational tests or developed by counselor for counseling uses. Broadly speaking, tests that serves counseling purposes are called counseling tests.

### **3.2 Distinctions between Psychological, Educational and Counselling Tests**

The following illustrates the relationships and differences between psychological, educational and counseling tests:

1. Psychological tests such as interest tests assess a person's interests and preferences. These tests are used primarily for career counseling. In the classroom the test measures learner's academic choices
2. Psychological tests provides the basis for teachers and counselors in measuring or classifying intelligence levels of clients and students
3. The roles of psychological and educational test outcomes is to provide guidance and counseling to individual concern
4. Some tests such as interest tests, Rorschach Inkblot test, Thematic Apperception tests are purely for counseling and psychotherapy
5. Psychological tests are useful in developing therapeutic rapport between the counselor and the client, also useful in creating inferences to follow-up with other methods. These can also be applied in motivating the learner by the teacher, toward learning tasks.
6. Psychological tests of aptitude are used by teachers and counselors to measure specific abilities in clients and learners, such as mechanical or clerical skills

7. One major distinction between the three categories of tests is that psychological tests is an objective and standardised measure of an individual's mental and or behavioural characteristics; educational tests are combinations of teacher – made tests and standardised test and counseling tests are tests used for the benefits of counseling processes.

#### **4.0 CONCLUSION**

It is difficult to say that, there are no relationships between psychological, educational and counseling tests. This is because, both psychological, educational and counseling tests focused on people, clients or learner's personality.

#### **5.0 SUMMARY**

Descriptions were made of psychological, educational and counseling tests. Distinctions between psychological, educational and counseling tests were discussed based on types of psychological tests.

#### **6.0 TUTOR-MARKED ASSIGNMENT**

Using relevant test examples, distinguish between psychological, educational and counseling tests.

#### **7.0 REFERENCES/FURTHER READING**

<http://go.speedbit.com/search.aspx?gusesof+strong+vocational+interest+blank%2ct> ...Retrieved on 10/25/2013

Ochri, O. (2003). *Some Psychological Tests for Identification of Children with Behaviour Problems*. Lagos: Babs Oladele Printing Press Limited.

Sambo, S. (2008). *Understanding Guidance and Counseling*. Zaria: Ahmadu Bello University Press.

## **UNIT 5 RELEVANCE OF PSYCHOLOGICAL TESTS TO COUNSELLING**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Relevance of Psychological Tests to Counselling
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

### **1.0 INTRODUCTION**

Psychological test is one that gives fairly stable and predictable result that can be depend upon during counselling process. This is one of the reasons that makes counsellors to uses personality tests in many circumstances. This unit dicusses revelance of psychological tests to counselling.

### **2.0 OBJECTIVE**

At the end of this unit, you should be able to:

- identify and discuss the roles of psychological tests to counselling.

### **3.0 MAIN CONTENT**

#### **3.1 Relevance of Psychological Tests to Counselling**

The following are relevance of Psychological tests in counselling:

- a. Psychological tests are used for diagnosing or identifying personality problems, dysfunction and learning difficulties so that appropriate remedial steps can be taken.
- a. Psychological tests such as aptitude tests are used for predicting students or clients' future in a specific field such as language-fluency, music, academic specialisations
- b. Tests are used for selecting candidates into either vocations or educational programmes
- c. Tests are used to promote and improve learning

- d. The feedbacks provided by psychological tests are used for self-assessment and counseling purposes.
- e. Psychological tests results are need for research purposes
- f. Psychological tests such as neuro-psychological tests are used in a clinical context to assess impairments.
- g. Sexological tests are used to examine the various aspects of discomfort, problem or dysfunction.
- h. Interest tests are used for the assessment of a person's or group interests and preferences
- i. Aptitude tests are used for measuring specific abilities skills
- j. Attitude tests are used for the assessment of individual or group feelings about an event, person or object
- k. Infants and preschool assessment are used for the assessment of reflexes, developmental milestone, sensory and motor skills, language skills and simple cognitive skills in infants and preschool aged children ages birth to six years of age.
- l. Direct observation tests are used to establish a pre-intervention baseline of a child's hyperactive or aggressive classroom behaviour or to observe the nature of a parent child interaction in order to understand a relational disorder
- m. Public safety employment tests are used by public safety organisation and industrial such as national fire fighter to tests initial employment and advancement throughout the ranks
- n. Personality test measures and outcomes are useful in developing therapeutic rapport between counselor (therapist) and the client. In addition, they are also useful in creating inferences to follow-up with other therapeutic methods.

#### **4.0 CONCLUSION**

Counselors need to understand the various uses, purposes and different application of the varieties of psychological tests.

#### **5.0 SUMMARY**

The unit discusses various uses and purposes of psychological tests by counselors and other practioners such as psychotherapists, researchers and others.

#### **6.0 TUTOR-MARKED ASSIGNMENT**

Discuss any five uses of psychological tests to the counselors?

## **7.0 REFERENCES/FURTHER READING**

Muhammad, Y. B. (1996). *Measurement and Evaluation in Education*.  
Kano: Debis-Co Press and Publishing Company.

John, P. K. (1988). *Educational Research Methodology and Measurement*.  
New York: Pergamon Press.

**MODULE 2      TEST DEVELOPMENT**

- Unit 1      Educational Test and Measurement
- Unit 2      Test Validity and Reliability
- Unit 3      Types of Validity and Reliability
- Unit 4      Theory of Reliability and Methods of Estimating Test Reliability
- Unit 5      Principles of Test Construction and their Application to the Construction of Objective Test Items

**UNIT 1      EDUCATIONAL TEST AND MEASUREMENT****CONTENTS**

- 1.0      Introduction
- 2.0      Objectives
- 3.0      Main Content
  - 3.1      Meaning of Educational Test
  - 3.2      Meaning of Measurement
  - 3.3      Meaning of Evaluation
  - 3.4      Relationships between Test, Measurement and Evaluation
- 4.0      Conclusion
- 5.0      Summary
- 6.0      Tutor-Marked Assignment
- 7.0      References/Further Reading

**1.0      INTRODUCTION**

In Nigeria like any other society, education is considered a means of effecting national development and enhancing rapid social and behavioural change. These aims are also targeted under counseling either in a group or on individual basis. Education is also considered as a means of enabling individuals to attain a means of sustenance, live a 'good life' and eventually contributes positively to his society.

Evaluation is an integral part of counseling and thus, knowledge of test or assessment techniques is an indispensable tool for any counselor anywhere. The roles of the school counselor with the introduction of continuous assessment in our educational system have made this requirement even more apparent. Unfortunately, a very significant proportion of the school counselors lack the necessary competence in testing and evaluation techniques. This unpleasant situation has made this aspect of the programme to mandatory.

## **2.0 OBJECTIVES**

At the end of this unit, you should be able to:

- discuss educational test
- define measurement
- explain evaluation
- state the differentiate between test, measurement and evaluation.

## **3.0 MAIN CONTENT**

### **3.1 Meaning of Educational Test**

Test is a systematic procedure for comparing the performance of an individual with a designated standard of performance. Test is also a series of tasks designed and administered to measure a person's or client's behaviour at a given time. In the process of designing a test, the series of tasks are systematically planned and arranged to elicit the desired behaviour. For example, in a test of typical performance, the tests attempt to get an index of an individual's usual or typical behaviour as he goes about his day-to-day life. In such measures, rather than urge the individual to do his best, he is required to report how he would normally behave in a specific situation such as in group or in a given event.

One important aspect of tests is that, the performance of individual in a test is meaningless unless a comparison is made with a fixed criterion of minimum proficiency of the individual group performances.

### **3.2 Meaning of Measurement**

Measurement refers to the process of assigning numerals to objects or events according to agreed upon rules. It is also the process of assignment of numbers to persons, subjects or events according to certain established rules. In education, measurement is a quantitative description of test scores; one of the most frequently used instruments for measurement. From these definitions, only random assignment of numbers will fail to qualify as measurement. For example, when we say a client scores 50% in an achievement test of Biology we are measuring his performance in that subject.

#### **Levels of Measurement**

A fundamental assumption in measurement is that anything that exists in some amount can be measured. Traits may be measured on the basis of the

type of information which different operations of measurement yield. By direct observation one can rank order a group of student from the tallest to the shortest according to height. On the other hand, he may use a ruler to record the height of each student in feet and inches. These two operations clearly lead to different information. While the former procedure permits statements such as student 'A' is taller than student 'B' the later operation permits statements of how much taller or shorter one student is than another. Differences along these lines serve as basis for measurement of variables and traits, the scale level of the measurement being determined by the nature of the information made available by the measuring operation used. Generally, one of the four sets of rules is used which results in four different types of measurement scales, namely: nominal, ordinal, interval and ratio scales.

- a. The nominal scale: This is simplest form of quantification and serves only to differentiate objects or persons. When used at the nominal level numbers, serve identification purpose only. For example, numbers used for houses along a particular street only serve to differentiate the houses from one another. Nominal numbers used at this level of measurement cannot be subjected to the arithmetical operation of addition, subtraction, multiplication or division.
- b. The ordinal scale: Measurement at this level places persons or objects according to relative positions only. The essential requirement at this level of measurement is that there are some procedures for determining each object being measured whether it has more, the same or less of the attribute in question. Ranking is the only permissible operation as differences between adjacent numbers are not necessary equal. For example, when we rank order a group of students by height, the tallest student might be 60 inches, the next tallest 48 inches, and third student might be 42 inches. Note that the difference between the first and the second student is 12 inches, while the difference between the second and the third student is 6 inches. Ordinal scale only shows which objects have more of the variable than other objects.
- c. The interval scale: Measurement at ordinal level has all the qualities of nominal and ordinal scales. In addition, interval scale provides for equal intervals from an arbitrary origin. Equal differences in the number represent equal differences in the attribute being measured. But when measurement is made without an absolute zero it is not legitimate to multiply or divide the values recorded. For example, it would be possible to say that student A is taller than student B and given the quality of intervals and those yielding similar

measurements of 40 and 28 inches is the same. But it cannot be said that the student who is 24 inches higher than the table is twice as tall as the student who is 12 inches higher than the table. If the table were 30 inches high, the actual heights will be 54 and 42 inches, respectively.

- d. The ratio scale: This most refined all forms measurement scales. Ratio scale possesses all the features of nominal, ordinal and interval scales and in addition it provides for true zero—that is a situation where we have complete absence of the attribute being measured. All manner of mathematical operations such as addition, subtraction, division and multiplication can be performed.

### 3.3 Meaning of Evaluation

The term evaluation refers to the systematic process of collecting, analysing and interpreting the outcome of measurement. The value judgment we make on the basis of the outcome of our measurement is called evaluation. Evaluation may be quantitative description of individual or group performance or qualitative descriptions, this is where measurement only identifies amount of the test score but evaluation lays the amount against criteria in order to make value judgment about the observed amount.

#### The Purposes of Evaluation

The main purpose of evaluation is to determine how well the objectives of instruction have been attained. Educational test is the most frequently used form of measurement in education with applications to counseling process. All purposes of tests were classified into three main headings, namely: prognosis or prediction diagnosis and research (Bichi, 1996).

#### 1. Prognosis or Prediction

Prognosis or prediction is the one that comes to mind when we think of intelligence tests given to school children to predict their academic performance. Similarly to this is the application of tests to select individuals for vocation or career. Counselors and teachers with school and teachers with school administrators uses tests to make decisions such as: which candidate should be admitted, which students should be grouped together, which student should repeat the year which student should receive certificate, which career choice fit a student etc. Though, less obviously a problem of prediction all the questions above and many similar ones that could be asked have implications for prediction. Thus, whenever a test is used to predict future performance, or to select individuals for training or employment, or to determine the extent of attainment after a period

of training or counseling treatment their function is that of making a prediction of future performance.

## 2. **Diagnosis Purpose**

The major use of tests (psychological, counseling and educational tests) is that of diagnosis. Diagnosis is more elaborate form of prediction. Evaluation uses diagnostic procedures to search for underlying causes of some learning difficulty, such as reading difficulty. In this process, diagnostic tests may be required.

## 3. **Research Purpose**

The third use of evaluation is in the area of research. Persistent educational problems or individual group problems that are unresolved despite conventional efforts may call for the use of research techniques to discover the underlying causes but this must be dictated by the outcome or interpretations of evaluation.

## **General Principles of Evaluation**

Some of the principles of evaluation are:

- a. Identify what is to be evaluated
- b. Select appropriate technique such as test that will serve the purpose of evaluation
- c. Comprehensive evaluation must be based on varieties of evaluation techniques
- d. Evaluation techniques have other strengths and weaknesses
- e. Selection of evaluation techniques must be based on the directions of decisions to be made. Evaluation process is used to determine a person's abilities that is, what the person can do such as in academic achievement, aptitude and skill. In addition, evaluation process also used to determine social behaviour such as interpersonal relationship, interest attitudes, social adjustment and so on.

## 3.4 **Relationships between Educational Test, Measurement and Evaluation**

- a. Measurement is the process of assigning numerals to objects, events or individual exhibited behaviour or trait on the other hand, evaluation is the value judgment that we make on the basis of the outcome of measurement. For example, if we say Aisha's score of 70 in a given achievement test is good, the achievement test is an

- example of educational test, the score of 70 is measurement and judgment of 'good' is the evaluation.
- b. Educational tests are yardstick for measurement and evaluation is the interpretation of the measurement
  - c. Evaluative judgment are sometimes based entirely on measurement obtained from test scores more often, however, evaluative judgment are based on a composite of measurement such as tests, field trip reports, and research projects.
  - d. The word 'assessment' is often used as a composite term to refer to the processes of test, measurement and evaluation.
  - e. Since traits are hypothetical and observable, they cannot be measured directly, we must posit that there are some observable features related to the underlying variable, the trait, and then tap the trait indirectly by measuring the observable feature with the use of appropriate test. The observable feature and the trait are not the same thing of course, but we must presume they are related. The observable feature is some subset of behaviours; the trait is some quantitative characteristic that can be validly judged based on the outcome of measurement. Counselors therefore measure personality traits, including intelligence, indirectly rather than directly and then pass judgment.

#### **4.0 CONCLUSION**

Educational test are integral part of guidance and counseling and, this knowledge of the meaning of test and its associate of measurement and evaluation is an indispensable tool for any counselor to practice the profession effectively.

#### **5.0 SUMMARY**

Educational test is a series of tasks designed and administered to measure a person's or client's behaviour at a given time. Measurement is the process of assigning numerals to objects or events according to agreed upon rules. Evaluation refers to the systematic process of collecting, analysing and interpreting the outcome of measurement. Purposes of evaluation (Assessment) are prognosis, diagnosis and research. Relationships between educational tests, measurement and evaluation can be illustrated with the case of counselor using test to measure personality trait and interpreting the result so as to offer counseling intervention.

## **6.0 TUTOR-MARKED ASSIGNMENT**

1. Briefly write short notes on the following:
  - Educational test
  - Measurement
  - Evaluation
2. Discuss five importance of test, measurement and evaluation to you as a counselor

## **7.0 REFERENCES/FURTHER READING**

Bidi, M. Y. (1996). *Measurement and Evaluation in Education*. Kano: Debis-Co Press & Publishing Company Limited.

Abdulsalami, A. S. (2005). *Research Methods in Education*. Nigeria: Stirling – Horden Publishers (Nig.) Ltd.

Rosenthal, T. L. & Zimmerman, B. J. (1978). *Social Learning and Cognition*. New York: Academic Press.

## **UNIT 2 TEST VALIDITY AND RELIABILITY**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Meaning of Test Validity
  - 3.2 Meaning of Test Reliability
  - 3.3 Characteristics of Test Validity and Reliability
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

### **1.0 INTRODUCTION**

Psychological tests are designed to provide information that when interpreted will lead to making practical decisions. The criteria that guide a guidance counselors or teachers in selecting good psychological tests that will measure an intended trait are validity and reliability

### **2.0 OBJECTIVES**

At the end of this unit, you should be able to:

- discuss test validity and test reliability
- state differentiates between test validity and test reliability.

### **3.0 MAIN CONTENT**

#### **3.1 Test Validity**

Test validity refers to the extent to which a test measured what it purported to measure. In another view, test validity refers to the extent to which the results of an evaluation procedure serve the particular uses for which they are intended. For a test to be valid it must possess the ability to measure precisely what it sets out to measure. In addition, for validity to occur there must be evidence to support the specified interpretation of the test results

### **3.2 Test Reliability**

Test reliability refers to the accuracy or precision of a measuring instrument. The consistency with which the test measures what is expected to measure. A test is reliable when subsequent measurements give the same or approximately the same numerical status or results of the trait being measured. Test reliability also is internally consistent or gives consistent results over time, across raters, etc.

### **3.3 Characteristics of Test Validity and Test Reliability**

#### **3.3.1 Characteristics of Test Reliability**

- i. ability or power of the test to measure the trait purported to measure
- ii. ability or power of the test result procedure serve the particular uses for which it is intended
- iii. ability of the test items to indicate specifically the trait or aspect of measurement in a step by step approach
- iv. ability of the test items to cover each of the sub-segment of the trait under investigation
- v. ability to provides the consistency that makes reliability possible

#### **3.3.2 Characteristics of Test Validity**

- i Tests must provide the consistency that will makes validity possible
- ii Ability to indicates the level of confidence that can be placed on the test results
- iii. Ability to produce the same or approximately the same test results by different raters
- iv Ability of the test results to be interpreted by interest users

### **4.0 CONCLUSION**

Test validity and reliability are the two basic qualities desired of good psychological tests. A psychological test that is valid must also possess the quality of reliability and the other way round.

### **5.0 SUMMARY**

Test validity is the extent to which a test measure precisely what is designed to measure. A test that is designed to observe trait will only be valid when it observe trait but not interview trait. Test validity is the consistency with which the test measures what is expected to measures by

providing the same or approximately the same results overtime. Characteristics of test validity and reliability were also discussed.

## **6.0 TUTOR-MARKED ASSIGNMENT**

1. Define test validity and test reliability in your own understanding
2. A test that is valid and reliable must have distinct characteristics. Briefly explain these characteristics

## **7.0 REFERENCES/FURTHER READING**

Tunji, A. (2008). *Fundamentals of Educational Psychology*. Kano: Debis – Co. Press and Publishing Company Limited.

Ellis, L. (1994). *Research Methods in the Social Sciences*. Madison, Wisconsin: WCB Brown & Beachmark.

Kerlinger, F. N. (1986). *Foundations of Behavioural Research* (3rd ed.). New York: Holt, Rinehart and Wiston.

## **UNIT 3    TYPES OF VALIDITY AND RELIABILITY**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Types of Validity
  - 3.2 Types of Reliability
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

### **1.0 INTRODUCTION**

Test validity is an indicator of some abstract concept is valid to the extent that it measures what it purport to measure and test reliability is the consistency with which it measures whatever it measures. In the process of ensuring that a test is valid or reliable there are methods or techniques that are used. These techniques or methods are referred to as types of validity and reliability of a test.

### **2.0 OBJECTIVES**

At the end of this unit, you should be able to:

- list and explain types of test validity
- state and explain types of test reliability.

### **3.0 MAIN CONTENT**

#### **3.1 Types of Test Validity**

The following approaches are ways of establishing the validity of a test. Since, it is not the test that is being validated, but an interpretation of data arising from a specified procedure:

##### **A    Content Validity**

The process of context validity focuses upon the extent to which the content of an indicant correspond to the content of the theoretical concept it is designed to measure. For example, the self-concept indicates focused upon

the same content as the conceptual definition of self-concept. In a simple description, content validity, refers to the extent to which a test measures domain of tasks under consideration.

Establishing content validity involves specifying the domain of content for the concept and constructing and selecting indicants that represent that domain of content. This, sampling is the essence of content validity. For example, a counselor might expect his client to correctly explain the meaning of the one hundred and fifty words that describes self-concept, because explain the meaning of one hundred and fifty words in a single test would take too long, the counselor might select fifteen words to represent the one hundred and fifty words. If the client correctly explains the meaning of the one hundred and fifty words that describes self-concept, because explaining the meaning of one hundred and fifty words in a single test would take too long, the counselor might select fifteen words to represent the one hundred and fifty words. If the client correctly explains the meaning of seventy percent of the fifteen words, the counselor would like to generalise and claim that the client would also be able to correctly explain the meaning of the seventy percent of the one hundred and fifty words. For the counselor to achieve this, the sample test must constitute a representative sample of the one hundred and fifty indicants of self-concept. Developing content validity requires first, carefully defining in behavioural terms the content areas to be measured. Second, breaking these areas into categories representing all major aspects of the area and thirdly, constructing a sufficient number of test items or indicants in each category. One of the best options to do this is by drawing a table of specification. Table of specification is defined as a 2 –way chart showing subject matter content on one part and the instructional objectives on the second part. Below is a sample of table of specification:

**Table 2.1: Table of Specification with Relative Emphasis in Percent given to Content Areas and Objectives for a Test in Behavioral Science**

| Content areas | Objectives          |                         |                       |       |
|---------------|---------------------|-------------------------|-----------------------|-------|
|               | Knowledge indicants | Comprehension indicants | Application indicants | Total |
| Meaning       | 3                   | 4                       | 5                     | 12    |
| Types         | 8                   | 6                       | 4                     | 18    |
| Importance    | 5                   | 5                       | 5                     | 15    |
| Differences   | 5                   | 5                       | 5                     | 15    |
| Total         | 21                  | 20                      | 19                    | 60    |

Table of specification is used to judge the content validity of a test, assists the counselor or tests placing undue emphasis on certain area to the

detriment of other area either in the content or in the objective area. The table is mostly used when constructing a new test. Content validity can also be determined by (a) comparing the test items with the objectives of the test, and (b) item analysis result.

#### B. Criterion – Related Validity

This is also called empirical validity. Gronlund (1981) cited in Muhammed (1996) defines criterion – related validity as the extent to which test performance is related to some other valued measure of performance. Criterion-related validity is divided into two types namely:

- i. Concurrent validity which is obtained when data from test scores are obtained at the same time with data from some outside criterion. For example, a test of theory of computer might be used to estimate current skill in the operation of computer as determine by the observer. If candidate who scored high on the test tend to show greater skills in computer operation, then the observer can claim a reasonable concurrent validity for his test of theory of computer.
- ii. Predictive Validity which is the second type of criterion – related validity. In predictive validity, a criterion is identified and success in this test is used to predict future status. For example, a high score in theory of computer may be a predictive for future operational of computer.

Another way of obtaining concurrent and predictive validity is to compare the scores. For example, in the case of concurrent validity the two sets of data to be compared are obtained simultaneously, while in the case of predictive validity obtaining the two sets of data can be spearman-Rank Order correlation coefficient or Pearson product moment correlation coefficient.

For the Spearman-Rank order

$$\text{Formula: } r = 1 - \frac{6\sum D^2}{N(n^2-1)}$$

For the Pearson Product moment formula:

$$r = \frac{N\sum XY - (E_x)(E_y)}{\sqrt{nE_x^2 - (E_x)^2} \sqrt{nE_y^2 - (E_y)^2}}$$

The Spearman – Rank correlation coefficient is a non-parametric statistics and the Pearson-Product moment correlation coefficient is a parametric statistics.

Illustration of the Spearman-Rank correlation coefficient computation is shown in example 1:

Example 1:

Step 1: Arrange the two pairs of data for each client in X and Y column

Step 2: Rank clients from 1 to the last number for each set of data in  $X_r$  and  $Y_r$

Step 3: Calculate the difference (D) in ranks by subtracting  $Y_r$  from  $X_r$ .

Step 4: Square each difference in rank to obtain difference squared  $D^2$

Step 5: Sum or add the squared difference  $D^2$  to obtain  $ED^2$

Step 6: Substitute the formula as thus.

**Table 2.2: Computation of Spearman Rank Order Correlation Coefficient**

| X  | Y  | $X_r$ | $Y_r$ | D  | $D^2$       |
|----|----|-------|-------|----|-------------|
| 71 | 61 | 9     | 7     | -2 | 4           |
| 62 | 62 | 6     | 8     | +2 | 4           |
| 69 | 55 | 7     | 5     | 02 | 4           |
| 45 | 41 | 2     | 2     | 0  | 0           |
| 53 | 52 | 4     | 4     | 0  | 0           |
| 78 | 67 | 10    | 10    | 0  | 0           |
| 37 | 40 | 1     | 1     | 0  | 0           |
| 48 | 43 | 3     | 3     | 0  | 0           |
| 60 | 58 | 5     | 6     | +1 | 1           |
| 70 | 63 | 8     | 9     | +1 | 1           |
|    |    |       |       |    | $ED^2 = 14$ |

$$\begin{aligned}
 r &= 1 - \frac{6 \times 14}{10(10^2 - 1)} \\
 &= 1 - \frac{84}{10(99)} \\
 &= 1 - \frac{84}{990} \\
 &= 1 - 0.085 \\
 &= 0.92
 \end{aligned}$$

Example 2: Using Pearson product moment correlation coefficient formula:

$$r = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2} \sqrt{n\sum Y^2 - (\sum Y)^2}}$$

Illustration of the Pearson Product Moment Correlation coefficient computation is shown below:

Step 1: Arrange the two pairs of data for each client in X and Y column

Step 2:  $\sum X$  = addition of the X data

Step 3:  $\sum Y$  = addition of the Y data

Step 4:  $\sum X^2$  = addition of the squared X data

Step 5:  $\sum Y^2$  = addition of the squared Y data

Step 6: Exy = addition of the Products of paired X and Y data

Step 7: N = number of paired products

Step 8: Substitute the formula as thus:

**Table 2.3: Computation of the Pearson Moment Correlation Coefficient**

| X       | Y       | X <sup>2</sup>         | Y <sup>2</sup>         | XY          |
|---------|---------|------------------------|------------------------|-------------|
| 71      | 61      | 5041                   | 3721                   | 4331        |
| 62      | 62      | 3844                   | 3844                   | 3844        |
| 69      | 55      | 4761                   | 3025                   | 3795        |
| 45      | 41      | 2025                   | 1681                   | 1845        |
| 53      | 52      | 2809                   | 2704                   | 2756        |
| 78      | 67      | 6084                   | 4489                   | 5226        |
| 37      | 40      | 1369                   | 1600                   | 1480        |
| 48      | 43      | 2304                   | 1849                   | 2064        |
| 60      | 58      | 3600                   | 3364                   | 3480        |
| 70      | 63      | 4900                   | 3969                   | 4410        |
| X = 593 | Y = 542 | X <sup>2</sup> = 36737 | Y <sup>2</sup> = 30246 | EXY = 33231 |

$$r = \frac{10 \times 33231 - (593)(542)}{\sqrt{10 \times 36737 - (593)^2} \sqrt{10 \times 30246 - (542)^2}}$$

$$\frac{332310 - 321406}{\sqrt{367370 - 351649} \sqrt{302460 - 293764}}$$

$$\frac{10904}{\sqrt{15721} \sqrt{8696}}$$

$$= \frac{10904}{125.38 \times 93.25}$$

$$= \frac{10904}{11691.685}$$

$$= 0.93$$

The counselor can now interpret the calculated or obtained alpha value of Spearman – rank order or Pearson Product Moment Correlation with the grouping of results interpretation of correlation tests with the Best and Khan (1989) table of a value interpretation of correlation coefficient:

**Table 2.4: Interpretation of Correlation Coefficient**

| Coefficient of r | Relationship |
|------------------|--------------|
| 00 – 20          | Negligible   |
| 20 – 40          | Low          |
| 40 -60           | Moderate     |
| 60 – 80          | High         |
| 80 – 100         | Very high    |

Source: Best and Khan (1989). Research in Education

- a. **Construct Validity:** Construct validity is the third way of establishing the validity of a psychological test. Construct is psychological quality that we assume to exist and used to explain some aspect of behaviour. Examples of construct include verbal ability, reasoning ability, intelligence, creative thinking, interpersonal relation skills or ability. Whenever test performance is interpreted on the basis of a certain construct, it implies we are referring to construct validity.

Construct validity therefore, is the extent to which test performance can be interpreted in terms of certain psychological constructs are hypothetical that is, each construct carries with it a number of associated meanings relating to how a person who possessed the specified traits would behave in a given situation. If we say a person is under emotional stress, for example, we would expect him to behave or react in a certain ways in various situations. This, establishing construct validity for a test involves: (1) identifying the constructs presumed to account for test performance (2) stating the hypotheses regarding test performance from the theory underlying each construct and (3) verifying the hypotheses empirically. Suppose one wish to verify the claim that a new test he constructed measures emotional stress. From what is known about 'emotion and stress' he might make the following predictions:

- a. Emotional stress individuals will score higher on the test than non-emotional stress individuals
- b. Scores on the test will correlate positively with scores on trait-factional test
- c. Scores on the test will be positively related to counselor's therapist's or psychologist's rating of emotional stress
- d. Scores on the test will be little influenced by other non-emotional stress situations.
- e. Scores on the test will correlate positively with scores on other emotional stress related tests.

After stating the predictions, then each of the predictions will be tested one by one. If each yields a positive result, then the combined evidence lends support to the claim that the test measures emotional stress. If any of the predictions is not confirmed, for example, the scores of the test are considerably influenced by other non-emotional stress situation, then we must conclude that: either:

- i. The test does not measure emotional stress, or
- ii. There is something wrong with the theory
- iii. The experimental design failed to test the hypothesis properly.

This implies that a decision must be made as to which of the three conditions has occurred. It should be noted that, establishing construct validity for a test challenges both the theory and the test simultaneously. It is important to keep in our minds that a test is not valid in and of itself, but rather valid for a particular purpose, thus, a test may be valid for one purpose, but not valid for another purpose.

Evidence for construct validation can be gathered with the use of the following methods:

- a. Study of the test taking process: this involves analysis of the test taking process such as examining the test items to determine what factors or trait they appear to measure, client taking the test might be asked to think aloud to see whether the items call for the intended reasoning.
- b. Group Difference: This involves comparing the scores of the groups that are known to differ. For example, persons of different age groups will be expected to perform differently on any test of intellectual ability, emotional stressed individuals
- c. Changes in performance: this method depends on the nature of the construct in question. Some test scores are expected to vary or remain stable after a given period. Scores on an intelligence test will remain stable even with direct teaching; scores on an anxiety scale or test will be expected to vary from one testing occasion to another.
- d. Correlation with other tests: The scores of any particular test will be expected to positively correlate with the scores of another test that measures the same trait and also scores of a given test will have low or no correlation with scores of a test measuring different trait.

In the process of establishing construct validity of psychological tests, some variables may still influence the construct therefore; the test constructor

should understand these variables so as to control them from contaminating the test result. Some of these variables are:

- a. Unclear directions: This is when test directions are not clear or are not clear as possible, to the clients or testees. They may not be sure of what to do or how to record the answers causing unnecessary interpretations.
- b. Complex or awkward wording: if the language or grammar and sentence structure are too difficult for the client(s) taking the test.
- c. Inappropriate level of item difficulty: when test items are too easy or too difficult the validity will be low. In addition, if test items are too easy most clients will obtain high scores and if test items are too difficult most clients will obtain low scores thereby making validity to be unrealistic.
- d. Poor construction of test items: if tests are poorly constructed items may give unintended clues. For example, grammatical inconsistent with the neither stem, other alternative nor may include extreme adjectives like “always” or “never” or the item may be ambiguous to the point of confusing the testees.
- e. Improper arrangement of items: if for example, difficult test items are placed early in the arrangement of the test items, testees may spend too much time on them and this may prevent them from reading items they could easily answer, they may also lose hope, become frustrated and finally makes the test result to be unrealistic.
- f. Test items being too short: this is a situation when the items on the test refused to cover all the content or aspect of the behaviour being measured, this will affect the validity of the test being measured.
- g. Identifiable pattern of answers: This is a situation where arrangement of correct answers or responses follow the same pattern (e.g. A, B, C, D, or AAAA or B BB B) will enable clients or testees to guess answers.

Variables a – g are factors in the test itself that can influence test validity. Other factors that can influence test validity is in test administration and scoring that include the manner in which a test is administered and scored such as lack of sufficient time to complete the test items, cheating, extra help given to testees in difficulty, un-conducive physical and psychological environment like noise, heat or cold, insufficient space and lighting, mis-scoring or wrong totaling, halo effect. Factors in clients or testees response that can affect test validity are emotional disturbances, test anxiety, poor motivation and responses etc. a response set is a consistent tendency to follow a certain pattern in responding to test items. E.g. some clients will respond ‘thru’ when they do not know the answer to a true-false item,

whereas other clients will respond 'false', the nature of the group being tested such as age, sex, ability level, cultural background of the group may also affect test validity.

In view of the above variables, test constructor need to device an effective mechanism for controlling or reducing them so as to have an acceptable level of validity from our tests.

### **Types of Test Reliability**

Test Reliability is the ability or power of a test to produce the same or approximately the same numerical status in a consistent manner to the object or person or behaviour being measured. Other terms used to describe reliability of a test are dependability, stability, consistency, predictability and accuracy. Types of test reliability also form methods of estimating or establishing test reliability. Although, psychometricians believed that it is not possible to compute the true reliability of any test (to be discussed under theory or reliability in later unit), however, there are four procedures for estimating test reliability. They are:

1. **Test-Retest Method:** The method defines variable under investigation in terms of random fluctuation in performance from one testing occasion to another. That is, error is defined as 'anything that leads a client or testee to get a different score on one testing than he obtained on another testing? Error in this way makes reliability obtained to be referred to as coefficient of stability. Test procedure is the administration of the same test twice on the same group of testees with a given time interval between the first and second administration of the same test. At the end of the administration, the two results are then correlated. The correlation coefficient provides an estimate of the test reliability. If the result are stable, testees who obtained high scores on the first test should also obtain high scores on the second test and those with low test scores.

The problem with test-retest method procedure is the time interval between the two testing sessions. If the interval is short the testees will remember some of their answers from the first test to the second and the constancy of the results will be inflated and if the time interval is long, there will be actual changes in the testees, and the testing procedure with the stability of the results will be underestimated. However, the test retest approach sees this problem as error. Therefore, alternative approach or method is next!

2. **Equivalent – Forms Method:** This method is also called parallel or alternate form method. The method involves the use of two different, but equivalent forms of the test items. The two equivalent forms of the test are administered to the same group of testees in close succession and the resulting scores are correlated. The correlation coefficient provides a measure of equivalence. When tests are administered this way, it provides a measure of equivalent and stability in addition to taking care of all possible sources of variation in the test scores, i.e. stability of testing procedures, constancy of testees characteristics being measured and representativeness of the sample tasks included in the test. The major weakness of the equivalent forms method is that, it is difficult to construct two tests that are truly equivalent. Two test measuring the same traits in the same way. In addition it calls real fluctuation in the trait 'error'. The alternative to this approach is next!
3. **Split-Half Method:** The procedures in the split-half method are that, a single form of a test is administered to a group of testees and then divided into two halves for scoring purposes; the two sets of scores are correlated. A usual method of splitting the test into two halves is to take the odd numbered items in one half and the even numbered items in the other half. Since test reliability is a function of test length, the Spearman Brown formula is used to estimate the reliability of the scores based on the full length test. The computation is:

$$\text{Reliability on full test} = \frac{2 \times \text{reliability on } \frac{1}{2} \text{ test}}{1 + \text{reliability on } \frac{1}{2} \text{ test}}$$

Computation example:

Assuming the correlation coefficient between the two halves of a test is .7:

$$\begin{aligned} \text{Reliability on full test} &= \frac{2 \times .70}{1 + .70} \\ &= \frac{1.40}{1.70} \\ &= 0.83 \end{aligned}$$

The test reliability coefficient computed through the split-half method is called measure of internal consistency. Beside this method is:

4. **Kuder-Richardson Method:** This is another way of thinking about variable error in the term of inconsistency of performance of the

testees on the items within a test. It is a measure of internal consistency. Richard-Kuderson Formula 21 can be used to obtain internal consistency of a test. When using this formula, one should note that it is more concerned with the homogeneity of the items, that is whether each item measures the same characteristic or not. In this formula only one single test is administered and does not require splitting the test responds into half for scoring purposes. For example, using the formula to estimates reliability is thus:

$$KR\ 21 = \frac{K}{K-1} \left\{ 1 - \frac{K-M}{KS} \right\}$$

K = the number of items in the test

M = the mean of the test scores

S = the standard deviation of the test scores

Computational illustration using KR 21:

**Table 2.5: Sample of Data for Calculating the Mean and Standard Deviation**

| X      | X <sup>2</sup>        |
|--------|-----------------------|
| 10     | 100                   |
| 8      | 64                    |
| 8      | 64                    |
| 9      | 81                    |
| 7      | 49                    |
| 8      | 64                    |
| 6      | 36                    |
| X = 56 | ∑X <sup>2</sup> = 458 |

$$\sum x = 56$$

N

$$K = 7$$

$$M = \frac{\sum x}{N} = \frac{56}{7} = 8$$

N 7

$$S = \frac{\sqrt{\sum x^2 - 8^2}}{N}$$

$$= \frac{\sqrt{458 - 8^2}}{7}$$

$$= \frac{\sqrt{65.43 - 64}}{7}$$

$$= \sqrt{1.43}$$

$$S^2 = 1.195$$

In the process of estimating the reliability of a test, some factors may affect the results. Some of these factors are:

1. Procedures used to estimate test reliability usually produces slight different results. For example, parallel test usually results in low reliability estimate, test re-test given to group of testees and split-half Kuder-Richardson result in serious over-estimate if applied to speed test.
2. Test length that is, the length of a test. In practice, the longer the test, the higher the reliability. A lengthier test provides a more representative sample of the behaviour under investigation. Lengthier tests also have content validity. However, this effect holds true only when the items added are of the same quality as the items that are already in the test. Adding very simple or very difficult items in the test does not add anything to the reliability of the test.
3. Range of abilities has influence on test reliability based on the heterogeneity of the analysis group. In practice, the more heterogeneous the group, the higher the estimate of reliability.
4. Testing environment such as light, verification influences estimate of test reliability
5. Poor motivation on the part of testees also affects reliability estimate
6. Subjectivity of the rater may also influence test reliability estimate

#### **4.0 CONCLUSION**

Test validity and reliability are among the qualities of a good psychological educational or counseling test therefore, counselors need to familiarise themselves with their types and roles toward adopting, adapting or constructing a test.

#### **5.0 SUMMARY**

Different types of test validity were identified and discussed. In the discussion, it was pointed out that, types of test validity are also forms the methods of estimating or establishing validity of a test. Reliability types were also identified and discussed, explanation on how types of test reliability can be used to estimate reliability of a test were given

## **6.0 TUTOR-MARKED ASSIGNMENT**

1. Outline types of test validity and reliability you know.
2. Discuss the process involved in using types of test validity and reliability in test estimations.

## **7.0 REFERENCES/FURTHER READING**

Yoloye, (1998). Federal Ministry of Education, Science and Technology. *A Handbook on Continuous Assessment*. Lagos: Heinemann Education Books.

Gupta, & Gupta, V. (2007). *An Introduction to Statistical Methods*. (23rd ed.). New Delhi: Vikas Publishing House.

Spiegel, M. R. & Stephens, I. J. (2008). *Theory and Problems of Statistics*. (4<sup>th</sup> ed.). New York: McGraw-Hill.

Muhammed, Y. B. (2004). *Introduction to Research Methods and Statistics*. Kano: Debis-Co Press and Publishing Company.

## **UNIT 4 THEORY OF RELIABILITY AND FACTORS THAT INFLUENCES, ESTIMATED TEST RELIABILITY**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Theory of Reliability
  - 3.2 Factors that Influences Test Reliability
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

### **1.0 INTRODUCTION**

A reliable psychological test is one that gives fairly stable and predictable result that can be depended upon. Although, all measurements contain variable errors. Variable errors are described as those errors arising from accidents and inaccuracies due to many causes such as, incorrect test wording, directions, distractions, ill-health, fatigue and others. Variable errors carries from person to person and also for a given person each time he takes the test. These are some of the reasons why illustration of the theory and factors that influences estimate test reliability will be appropriate at this point.

### **2.0 OBJECTIVES**

At the end of this unit, you should be able to:

- explain in detail the assumptions of the reliability theory
- identify and discuss factors that influences the estimated test reliability.

### **3.0 MAIN CONTENT**

#### **3.1 Theory of Reliability**

Reliability Theory presume that, every measurement contains some amount of error. Thus, if we measure the length of a table for example, the

score we observe is made up of two components: a – true component b – error component.

We never know the true length nor do we ever know the error scores. The true length is what we would get if the measurement were perfect, that is, no error at all. Now if we measure the same table several times it is possible to observe a different score each time we measure the table. If the length of the table remains the same, any variation in observed length represents an error in measurement. If the variation is small the error is small and if the variation is large the error is large. Thus, the amount of variation is an index of the size of error in measurement. If instead of only one table we measure the length of several tables of different length the total variance cannot be said to be all error since part of the variance is due to actual differences in the length of the tables. Thus, the total variation is made up of two components: true score variance and error variance. Since we cannot attribute all the variation to error, it is reasonable to ask what proportion of the observed differences is attributable to true or actual difference in table length. If this proportion is large then the measurement reflects true differences and if the proportion is small, a large part of what has been observed is error. This proportion of the amount of variation in true score to total variation is termed reliability. In terms of test, reliability is defined as 'the ratio of the true score variance to the observed variance'. From this expression, it can be seen that if the measurement involves nothing, but error (e.g. if all the tables have the same true length and any variation in observed score is entirely attributable to error and reliability is zero. On the other at all = 0 the ratio become = 1. This, reliability varies from 0 to 1, taking a value of zero when measurement involves nothing, but error and value of one when there is no variable error at all in the measurement.

### **3.2 Factors that Influences Test Reliability**

In test practices when computing reliability coefficients we must be aware of the factors that may possibly influence the size of value we get. These factors are:

## **4.0 CONCLUSION**

Counselors needs to understand the contribution of reliability theory toward identifying different factors that affect the estimate of test they use in counseling profession.

## **5.0 SUMMARY**

Theory of reliability explains possible ways by which error can be made in the process of estimating test reliability. Two types of error were identified as true-error component and error component. The true-error refers to the errors arising from the actual measurement and the error components are those error committed unintentionally. The discussions on the theory later extend to various factors that may influence the estimated reliability of a test.

## **6.0 TUTOR-MARKED ASSIGNMENT**

1. Explain in detail the assumptions of reliability theory
2. List and discuss five variables that may influence the size of reliability of a test

## **7.0 REFERENCES/FURTHER READING**

- Muhammad, Y. B. (1996). *Measurement and Evaluation in Education*. Kano: Debis-Co Press and Publishing Company.
- John, P. K. (1988). *Educational Research Methodology and Measurement*. New York: Pergamon Press.

## **UNIT 5 PRINCIPLES OF TEST CONSTRUCTION AND THEIR APPLICATIONS TO THE CONSTRUCTION OF OBJECTIVE TEST**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Principles of Test Construction
  - 3.2 Construction of Objective Tests
  - 3.3 Procedures for Standardising a Test
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

### **1.0 INTRODUCTION**

Developing a valid and reliable test involves careful planning based on laid down principles. Counselors may adopt, adapt or modify an already developed test based on legal permission or may develop their own. Whatever is the case, test construction or development must follow laid down processes known as principles of test construction.

### **2.0 OBJECTIVES**

At the end of this unit, you should be able to:

- list and explain the principles of test construction
- apply the principles of test construction to the development of objective test items.

### **3.0 MAIN CONTENT**

#### **3.1 Principles or Processes of Test Construction**

Counselor constructing or developing a test for use must follow the laid down processes. These processes or principles are broadly classified into three. These are:

- a. Formulation or stating of objectives for the desired test;
- b. Construction of a test blueprint or table of specification; and
- c. Writing of test items and organising tests.

Formulation or stating of objectives is one of the most important tasks expected of a test constructor. The first prerequisite to the tasks of objective formulation is to have a set of data that describes the behaviour in part, information about the nature of client problems, and personality traits such as data on intelligence, physical characteristics, dominant interest, vocational ambitions, etc.

Objectives should be specific and measurable. It should clearly state the behaviour expected and which is to be measured. Objectives should be stated in a way that any interested party can tell whether or not it can be achieved, measured and tested by the uses of one or other of his senses.

Three criteria have been identified for stating objectives. They are:

- Each statement of an objective should indicate both the desired behaviour and the type of situation in which it occurs;
- Objectives should be stated in terms of desired client behaviour rather than counselor behaviour;
- Objectives should be formulated to a level of specificity.

The first criterion is on the client behaviour or trait to be measured. How the behaviour can be demonstrated based on situation. The second criterion is on what the client is to do, how he is think, to feel and to act. The task expected of the counselor here is to arrange the situation so that the clients have an opportunity to behave in the ways specified by the objectives. It should be noted that client desirable behaviour is the goal, and counselor technique is only a means to achieving the goal. Attention to this criterion tends to enrich the nature of the test activities selected by the counselor. In addition, client behaviour is the primary target to be assessed, and attention to this criterion facilitates the assessment process. Assessment based on what the client problem look like is virtually a pre-requisite to assessment of the power of the test. A third criterion in the statement of objectives is designed to control the statement are genuinely helpful in assessment or measurement. In the process of objectives statement, general objectives are of few types, but there are hundreds of specific objectives subsumed for even one reasonably substantial trait. The criterion seeks to mediate between the general and the specific in terms of the use to be made of objectives. Therefore, the third criterion is: objectives should be formulated to a level of specificity such that it is possible to readily infer some test

items appropriate for assisting clients achieve each objective and also to devise means of assessing the achievement of the measurement objectives.

This criterion assumes that the counselor, or whoever is formulating the objective, understands the subject matter and the nature of appropriate behaviour or problem. The criterion further asserts that under these conditions the counselor should be able to state actual behaviour or problem characteristics or features for the client. In achieving this, it is better to use specific active verbs such as indicate, count, specify, state, yes, no etc. so that the client can easily interpret or translate what the test item is trying to find out or measure.

### **Construction of a Test Blueprint or Table of Specification**

The test table of specifications or test blueprint is a two-dimensional table; it contains the objectives of the behaviour or problem under investigation in one direction and the content that is the behaviour/traits or behaviour problem in another. The major decision to be made is the number of items to be included in the test. This decision is based upon the type of test format to be used (objective or essay) and the time available for the test. The rule is that the number of items increases as the amount of time available also increases.

The table of specification is shown in Table 1, but brought forward for illustration

**Table 2.5: The Early Stage of the Table of Specialisations**

| Content areas | Objectives  |               |               | Total |
|---------------|-------------|---------------|---------------|-------|
|               | Knowledge i | Comprehension | Application i |       |
| Meaning       |             |               |               |       |
| Types         |             |               |               |       |
| Importance    |             |               |               |       |
| Differences   |             |               |               |       |
| Total         |             |               |               | 60    |

At this phase the table of specialisations indicates that there are four content areas to be measured and that it is intended that knowledge, comprehension and application areas will be measured by the test. The table contained a total of 12 empty cells. The 12 cells come from the fact that there are 4 content areas and 3 objectives. This implies that  $4 \text{ multiply by } 3 = 12$ . Each cell is the intersection of the content area and an objective.

In each of these cells will be indicated the number of test items to be set in that content area to measure the indicated objectives. The next task is to

decide how many of the 60 items will be assigned to each of these 12 cells. The decision is based on the degree of importance attached to each cell. For instance, what percentage of the 60 items will be knowledge items and what percentage will be set in content area one?

The answers are determined by the special circumstances of the individual cases. One practical indicator of the importance attached to a content area is the value of the content on the client behavioural change. Table 8 indicates that in this particular case knowledge is given 21 %, comprehension is given 20% and application is given 19%.

**Table 2.6: The Final Stage of the Table of Specialisations**

| Content areas | Objectives  |               |               |       |
|---------------|-------------|---------------|---------------|-------|
|               | Knowledge i | Comprehension | Application i | Total |
| Meaning       | 3           | 4             | 5             | 12    |
| Types         | 8           | 6             | 4             | 18    |
| Importance    | 5           | 5             | 5             | 15    |
| Differences   | 5           | 5             | 5             | 15    |
| Total         | 21          | 20            | 19            | 60    |

In table 2.6 the distribution is as follows: 21%, 20% and 19% to knowledge, comprehension and application. It should be noted that adjustment can be made to increase or decrease the percentages allocated to the content areas depending on the grand total that the test constructor intended. The next level is the writing of test items.

### Writing of Test Items and Organising Tests

Writing of test items and organising them together to produce an acceptable test requires skill, patience and constant practices. After completion of the table of specifications that is, the preparation of the situations to be presented to the client to elicit the desired behaviour. The following guidelines are generally recommended:

1. Counselor developing a test should always construct more items than the expected numbers in the table of specifications
2. The advantage of constructing more numbers of items required is that it allow for the discarding items that are not very good
3. Excessive window dressing should be avoided. One should not put any more in an item than is necessary for the purpose of that item.
4. The use of long and involved statements should be avoided because such statements makes clients waste time thinking out the elements that are important in answering the item

5. Degree of accuracy required from the client should be specify. This can be done either in the item or in the directions for a series of items
6. Use of extraneous cues should be avoided by checking carefully for information a client can use to get an item right even though he does not have the abilities the item is measuring.
7. Avoid giving clues to one item in the statement of another
8. Negative statements and double negative statements should be avoided. Such statements are easily missed by clients particularly in speed test.

### **Organising and Administrating Tests**

In organising and administrating an already prepared test items, there are two major considerations in setting up the test, namely:

1. The ease with which the client can understand what he/she is to do, and where and how is to record his responses or answers
2. The ease with which the counselor will be able to locate and score the responses or answers. To achieve these effectively, the counselor can group items of the same type together, number all items consecutively.

From the first item to the last, arrange each sub-division of the test so that the easier items come before the more difficult ones, items should be legible, precede each group of items with simple and clear statements telling how and where the client is to indicate his answers, make client work from a separate copy of the test, have all necessary materials on hand such as erasers, pencils, provide the best physical surroundings with safety, adequate light, ventilation, desk space and make sure that there is no distracting sound or object. If speed is not a factor in the objective covered by the test, enough time should be allowed so that client has an opportunity to attempt all the items.

### **3.2 Using the Knowledge of the Principles of Test Construction to Develop an Objective Test Items**

There are two commonly used types of test items, namely objective and essay test items. The objective test items are mostly used in counseling process. This is because, it is easy to score, it ensures objectivity and caters for all categories of clients irrespective of their educational background.

### **Objective Tests**

The term objective simply refers to the scoring procedure employed rather than the manner in which the responses or answers are made.

Constructing an acceptable objective test items requires the counselor to:

1. Identify the objectives of the test that is, what and what the test wishes to measure
2. Have comprehensive knowledge of the content (behaviour or trait) area to be measured, and
3. Have knowledge of the mental processes of the client or clients whom the test will be administered to
4. Take decision on the objective form of the test such as short-answer items completion question items, true-false or alternative-response test items, matching exercise, and multiple-choice items
5. Have knowledge of the strength and weaknesses of these categories of objective test items

The procedure for analysing clients' answers or responses to objective test items is called Item Analysis. This procedure is used to determine the functioning effectiveness of each item. The characteristics of item analysis are that:

- a. It determines the adequacy of the item. In a norm-referenced test, it is used to determine whether the items adequately discriminate between high and low achievers. In a criterion referenced test, it will determine whether the item adequately measures the effect of the service provided by the counselor.
- b. It determine whether the test items were of appropriate level of difficulty
- c. It determines when distracters are used whether they were effective

### **Types and Procedures of Item Analysis**

Item analysis is of four types, namely:

1. Difficulty level of the item (difficulty index)
2. The discriminatory power of the item
3. The effectiveness of the distracters
4. Effectiveness of counseling intervention.

Each of this type of item analysis has procedures for obtaining them. For example:

- a. Calculating the difficult index involves the following procedure:

$$\text{Difficulty index} = \frac{U + L}{T}$$

Computation using the above formula:

- ❖ Rank all the test scores from highest to the lowest score
- ❖ Divide the group of scores into three: for example, if the scores are 60, the first 20 scores with the highest total scores and the next 20 with the lowest total scores
- ❖ The 20 middle scores will be set aside, they will not be used in the analysis
- ❖ Calculate the difficulty index using the formula:  

$$= \frac{U + L}{T} \times 100$$
  - U = the number of clients in the upper group who get the item correct
  - L = the number of clients in the lower group who get the item correct, and
  - T = the total number of clients in the upper and lower who participated in the test.

Assuming the 60 clients, 20 highest scoring clients will be referred to as upper group and the next 20 lowest scoring clients will be referred to as lower group. The two groups will be used for item analysis. For example, 17 clients from the upper group got the item correct and 15 from the lower group got the item correct, the difficulty level will be:

$$= \frac{17 + 15}{40} \times 100$$

$$= \frac{32}{40} \times 100$$

$$= 0.8 \times 100$$

$$= 80\%$$

The estimated value of difficulty index can be influenced by the item being too easy or very difficult. For example, too easy item makes difficulty index close to 100%, very difficult item makes difficulty index close to zero. Item that are extremely too difficult are too easy should be discarded. Low difficulty index can be improved and average level of difficulty index can be selected for future use.

**b. Calculating the Discriminating Power of the Item**

Discriminatory power of an item refers to the extent an item can discriminate between the highest and lowest achievers. The procedure for discriminating power makes the highest achievers as upper group and the lowest achiever as the lower group. If an item discriminates in a positive direction more clients in the upper group will get the item correct than the lower group; this is known as positive discriminating power and the result is usually in a discriminating index of 1.00. Negative discriminatory power on the other hand is obtained in an item that more clients in the lower group than the upper group got the item correct. The result of negative discriminatory power is usually with an index of -1.

Any Item that has no discriminatory power usually have an index of .00 indicating that the same number of client in the upper and lower group got the item correct.

How to calculate the discriminatory power involves:

$$\text{Formula: } \frac{U - L}{\frac{1}{2} T}$$

U = the number of clients in the upper group who get the item correct

L = the number of clients in the lower group who get the item correct, and

$\frac{1}{2} T$  = the total number of clients in the upper and lower who participated in the test.

$$\begin{aligned} \text{Thereby: } & \frac{17 - 15}{20} \\ & = \frac{2}{20} \\ & = 0.10 \end{aligned}$$

It should be noted that zero is usually added to the obtained answer or value of discriminatory power of item.

When interpreting the discriminatory power the following criteria are used:

1. 1.00 = positive discriminatory power
2. -1.00 = negative discriminatory power
3. 0.00 = No discriminatory power

Test items with positive discriminatory power are selected for future use, those with negative discriminatory power are improved for future use and those with no discriminatory power are discarded.

### c. Calculating the Effectiveness of the Distracters

Following our explanations on the interpretations of discriminatory power of test items that a positive discriminatory power item is accepted for future use, but if negative discriminatory power is obtained then, series of questions may be raised particularly on the effectiveness of the distracters. Further analysis of the item can assist the test developer to improve the item.

The ground rule for determining the effectiveness of the distracters is through inspection. Hills, (1976) cited in Yusuf (2009) presented a desirable general pattern of examinee's response in a multiple-choice items as follows:

- a. Every alternative (A, B, C, D) must be chosen by at least one examinee
- b. More examinees in the upper group than in the lower group should answer the correct choice
- c. More examinees in the lower group than in the upper group should answer each distracter
- d. More than half of the examinees should answer the correct choice

**Table 2.7: Determining the Effectiveness of Distracters**

| Alternatives     | A | B | C | D  |
|------------------|---|---|---|----|
| Upper group = 20 | 0 | 2 | 2 | 17 |
| Lower group = 20 | 0 | 1 | 3 | 15 |

In the table above, the correct answer is D as such the item discriminates in a positive direction since 17 in the upper group and 15 in the lower group got the item correct. Alternative B is poor distracters because it attracted more examinees from the upper group than the lower group. Alternative or option A is absolutely inadequate because it attracted no examinee. Alternative C is a good distracter because it attracted more examinees from the lower group than the upper group. Moreover, if an estimate of the effectiveness of each distracter is desired, same formula used for calculating discriminatory power can be employ.

For illustration:

1. Alternative or option A:  $\frac{0 - 0}{20} = 0.0$
2. Alternative B:  $\frac{2 - 1}{20} = 0.05$
3. Alternative C :  $\frac{2 - 3}{20}$

$$20 \\ = 0.05$$

Alternatives A and B are ineffective and alternative C must be improved for future use.

#### d. Calculating the Effectiveness of Counselling Intervention

The procedure for analysing the effectiveness or efficacy of counseling intervention is to construct a test, administer the test before intervention (pretest) so as to establish the presence of a behaviour problem in a client or group of clients and re-administer the test after administering the counseling intervention (post-test) and then compare the results of pre-test and post-test. Item analysis may be used. Item analysis measures the effects of intervention by emphasising the extent to which each stage of intervention or technique had effect on the client targeted behaviour. Inspection of intervention by intervention, comparison will provide information about the effectiveness of the intervention strategies A chart will be prepare with the list of the intervention strategies across the top and the list of clients' names by the side of the chart. The clients correct response to the effectiveness of interventions will be scored (+) and incorrect responses will be scored (-) for each student or a student with exposure to different interventions on both pre and post test periods. For example, the table below contains an illustration:

**Table 2.8: Clients Responds on the Effectiveness of Counselling Interventions at Pre and Post Test periods**

| Clients   | Counseling Interventions |   |   |   |   |   |   |   |
|-----------|--------------------------|---|---|---|---|---|---|---|
|           | 1                        |   | 2 |   | 3 |   | 4 |   |
| Hope      | -                        | + | + | - | + | + | - | + |
| Happiness | -                        | + | + | - | + | + | - | + |
| Funmilayo | -                        | + | + | - | + | + | - | + |
| Aminu     | -                        | + | + | - | + | + | - | + |
| shade     | -                        | + | + | - | + | + | - | + |

Interpretation of the responses:

1. Intervention Strategy 1: Two clients rated correctly the first intervention and four clients after the intervention. This can be said to be a fairly good intervention
2. Intervention Strategy 2: All clients responded correctly before intervention and incorrectly after intervention. This can be judged as ineffective intervention

3. Intervention Strategy 3: All the clients responded correctly before and after intervention. The intervention can be said to lack ability in alleviating the clients problem
4. Intervention Strategy 4: All the clients responded correctly before and after the intervention. This can be judged to be very effective intervention.

#### **4.0 CONCLUSION**

For the counselor to develop or construct an effective test it requires a detail understanding of the principles and applications of test construction. The counselor also needs to liaise with experts in test construction followed by constant practices.

#### **5.0 SUMMARY**

Principles known as processes of test construction were outlined and discussed. The construction of objective test items. Following the discussion of the principles of test construction were also discussed. Procedures for calculating and analysing clients' responses or answers to objective test items were illustrated.

#### **6.0 TUTOR-MARKED ASSIGNMENT**

1. List and explain the principles of test construction
2. Identify and discuss guidelines necessary for the development of objective test items
3. Discuss the criteria required for test standardisation

#### **7.0 REFERENCES/FURTHER READING**

Yoloye, (1998). Federal Ministry of Education Science and Technology. *A Handbook on Continuous Assessment*. Lagos: Heineman Educational Books.

Muhammad, Y. B. (1996). *Measurement and Evaluation in Education*. Kano: Debis-Co Press and Publishing Company.

Yusuf, K. (2009). *Elements of Educational Measurement and Evaluation*. Zaria: Tamaza Publishing Company.

Aminu, K. I. & Gali, S. (2012). *Guidance and Counselling: Theories and Practice in Schools*. Kano: Albarka Publishing Company Limited.

## **MODULE 3      TEST      STANDARDISATION      AND PERSONALITY TESTS PROBLEMS**

|        |   |
|--------|---|
| Unit 1 | Test Standardisation  |
| Unit 2 | Uses of Strong Vocational Interest Blank, Kuder Preferential Blank, Achievement Tests, Intelligence Tests and Personality Tests |
| Unit 3 | Problems Associated with Personality Tests  |
| Unit 4 | Strategies of Overcoming the Problems Associated with Personality Tests   |
| Unit 5 | Problems Associated with Personality Testing and Ways of Overcoming the Problems  |

### **UNIT 1      Tests Standardization**

#### **CONTENTS**

|     |                                     |
|-----|-------------------------------------|
| 1.0 | Introduction                        |
| 2.0 | Objectives                          |
| 3.0 | Main Content                        |
| 3.1 | Procedures for standardizing a test |
| 4.0 | Summary and Conclusion              |
| 5.0 | Tutor-Marked Assignment             |
| 6.0 | References/Further Reading          |

#### **1.0      INTRODUCTION**

Counselors may adopt, adapt or modify an already developed test based on legal permission or may develop their own. Whatever is the case, test construction or developers must follow laid down processes to standardize the tests. This unit outlines the criteria for standardizing tests.

#### **2.0      OBJECTIVES**

At the end of this unit, you should be able to:

- List the criteria for tests standardization.
- Explain the criteria for standardization of tests.

### **3.0 MAIN CONTENT**

#### **3.1 Test Standardization**

Standardization of a test refers to the qualities desired in a test. Some of these qualities have been identified and fully discussed in previous units. But for the purpose of clarifications the qualities are:

- ❖ Test validity
- ❖ Test reliability, and
- ❖ Test usability

Test validity is defined in terms of the extent to which a test measured what it purported to measure but contemporary use of test validity refers to the extent to which the results of an evaluation procedure serve the particular uses for which they are intended.

Test Reliability refers to the accuracy or precision of a measuring instrument that is, the consistency with which the instrument measures what it is expected to measure.

Test Usability: This comes after validity and reliability. It is the third requirement for test standardization. The term usability is an umbrella term under which all practical considerations before selecting or using a test are discussed. These include test being easy to administer, time required for administration of the test, ease of scoring and interpretation, cost of development and testing. All these factors are very important for consideration before selecting, adopting, adapting or developing a test.

### **4.0 SUMMARY AND CONCLUSION**

#### **4.1 SUMMARY**

The criteria for tests standardization refers to the qualities expected of a good test. These qualities include validity, reliability and usability.

#### **5.0 CONCLUSION**

For the counselor to develop or construct an effective test it requires a detail understanding of the qualities of a good test. The counselor also needs to liaise with experts in test construction followed by constant practices.

## **6.0 TUTOR-MARKED ASSIGNMENT (TMA)**

1. List the criteria used for standardizing a test?
2. Explain the criteria for standardizing a test?

## **7.0 REFERENCES/FURTHER READING**

- Yoloye, (1998). Federal Ministry of Education Science and Technology. A handbook on continuous assessment. Lagos. Heineman Educational Books.
- Muhammad, Y. B. (1996). Measurement and Evaluation in Education. Kano. Debis-Co Press and Publishing Company.
- Yusuf, K. (2009). Elements of Educational Measurement and Evaluation. Zaria. Tamaza publishing Company.
- Aminu, K. I. and Gali, S. (2012). Guidance and Counselling: Theories and Practice in Schools. Kano. Albarka Publishing Company Limited.

## **UNIT 2 USES OF STRONG, VOCATIONAL INTEREST BLANK, KUDER PREFERENTIAL BLANK, ACHIEVEMENT TEST AND INTELLIGENCE TEST**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Uses of Strong Vocational Interest Blank
  - 3.2 Uses of Kuder-Preferential Blank
  - 3.3 Uses of Achievement Test
  - 3.4 Uses of Intelligence tests
- 4.0 Summary and Conclusion
- 5.0 Tutor-Marked Assignment
- 6.0 References/Further Reading

### **1.0 INTRODUCTION**

Data collection is an integral part of counseling process. The general order of contemporal scientific methods which the counselors follow attests to this claim. Data collection techniques entail the determination of the information needed and of how to obtain it. Some of the uses of the various tests that can be used in counseling process are discussed under this unit.

### **2.0 OBJECTIVES**

At the end of this unit, you should be able to:

- Explain the uses of
- Strong vocational interest blank
- Kuder Preferential Blank
- Achievement test
- Intelligence test.

### **3.0 MAIN CONTENT**

#### **3.1 Strong Vocational Interest Blank**

Strong vocational interest Blank is an example of interest inventories. The blank ask the client to indicate her/his preference for various occupational areas such as mechanical, scientific, artistic, persuasive and social science:

Uses of Strong Vocational Interest Blank are:

1. It is used to measure client personality interest toward a targeted object, individual or event
2. It is used to determine client level of likes and dislikes, such as the kinds of activities she/he prefers to engage
3. Being a non-projective test personality, it is used to describe the personality structure of different group
4. It also used to measure client self that is client assess his/her feelings about himself/herself
5. In a correlational study, it was used to correlate intelligence of different individuals

#### **Kuder Preferential Blank**

Kuder Preferential Blank is an example of non-projective test similar to Strong Vocational Interest Blank. In many experimental studies it have been used to correlate different personality trait, such as intelligence. Uses include:

1. To correlate clients intelligence level
2. Describe personality structure of various groups of clients
3. The format of the test is sometimes used as checklist where an individual checks items that best characterized a trait.
4. It is also used to measure feelings of clients toward a targeted trait.

#### **Uses of Achievement Tests**

Achievement tests have been defined as those tests that measure the current status of individuals with respect to proficiency in given areas of knowledge or skill. Some examples of achievement tests are:

- The California Achievement Tests
- Standard Ford Achievement Tests
- Iowa tests of basic skills
- Metropolitan Achievement test
- Sequential tests of educational progress

**Uses of achievement tests include:**

1. Achievement tests are used to measure achievement in reading, language and Arithmetic
2. Achievement tests are used to measure presence of basic skills in client
3. Achievement tests such as metropolitan achievement test measures general academic achievement in vocabulary, reading and numerical ability
4. Achievement tests are used to determine the extent to which an individual benefited from a course of instruction
5. Achievement has been used in researches as dependent variables especially in areas involving instructional processes.

**Intelligence Tests**

Intelligence tests are psychological tests. There are different definitions of intelligence. For example, Spearman defines intelligence as the ability to develop relations and correlates, Terman considers intelligence to be the ability to carry on abstract thinking, Thorndike sees intelligences as the power of good responses from the point of view of truth and fact, Woodrow defines intelligence as the capacity to acquire capacity, etc.

Alfred Binet in the early 1900s developed the first standard measurement of intelligence in Paris. This was motivated by his interest in predicting the academic success of school children. The test was later revised by Lewis Terman of Stanford University for use with children in the United States. This was released in 1916 as the Stanford-Binet Intelligence Test. Since, then, a number of different tests of intelligence have been developed. The most re-known intelligence tests are:

- Binet Intelligence test
- Stanford – Binet Intelligence test
- Wechsler-Intelligence tests for pre-school, primary, children and adults.

**Uses of intelligence test are:**

1. Intelligence tests are used to measure the construct in a manner it is theorized
2. Intelligence tests are constructed with varieties of items that measures different abilities – verbal, numerical, spatial, etc
3. Intelligence tests are used to compare individual performances on the tests against some criteria, norm or standard

4. Intelligence tests are also used to identify intelligent level of individual for classifications
5. Intelligence tests assists the counselor to identify appropriate counseling intervention for client

#### **4.0 SUMMARY AND CONCLUSION**

##### **4.1 SUMMARY**

The unit introduces the concepts strong **vocational interest blank, kuder preferential blank, achievement tests and intelligence tests.**

##### **5.0 CONCLUSION**

Projective personality tests are utilized to a great extent by clinical psychologists, therapists and guidance counselors because the techniques require special skills for use (administration, scoring and interpretation).

##### **6.0 TUTOR-MARKED ASSIGNMENT (TMA)**

Explain the uses of the following tests by counsellots:

- a. strong vocational interest blank
- b. kuder preferential blank
- c. achievement tests
- d. intelligence tests?

##### **7.0 REFERENCES/FURTHER READING**

Kerlinger, F. N. (1986). *Foundation of Behavioural Research* (3<sup>rd</sup> Ed.) New York. Holt, Rinehart and Winston.

Adeloye, T. (2006) *Fundamentals of Educational Research*. Kano: Debis-Co. Press and Publishing Limited.

Frankel, J. R. & Wallen, N. E. (2000). *How to Design and Evaluate Research in Education*. Boston. McGraw Hill

## **UNIT 3 PERSONALITY TEST**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Uses of Personality tests
- 4.0 Summary and Conclusion
- 5.0 Tutor-Marked Assignment
- 6.0 References/Further Reading

### **1.0 INTRODUCTION**

Data collection is an integral part of counseling process. The general order of contemporal scientific methods which the counselors follow attests to this claim. Data collection techniques entail the determination of the information needed and of how to obtain it. Many of the uses of the various tests in counselling falls under personality tests

### **2.0 OBJECTIVES**

At the end of this unit, you should be able to:

- Explain the uses of personality tests.

#### **Personality Tests**

Psychological measures of personality are often described as either objective tests or projective tests, rating scale or self-report measures. For the purpose of measurement, Kerlinger (1986) defined personality as the organization of the traits of the individual. A trait is viewed as a characteristic of an individual revealed through recurring behaviours in different situations. Personality is an important variable (dependent or independent) in counseling processes. Personalities tests are broadly categorized tests are broadly categorized into two: projective and non-projective tests.

The projective personality tests are constructed in an attempt to do away with the problems of non-projective techniques – “conscious dishonesty of respondents or clients” or “socially acceptable responds”. Examples of projective tests are:

- The Roschach Inkblot Tests, and
- The Thematic Apperception tests

The Non-Projective tests are self-report measures or tests that usually ask an individual to respond to a series of questions or statements about his personality. Examples of non-projective tests are:

- Personality inventories like:
  - a. The Minnesota Multiphasic Personality Inventory
  - b. IPAT Anxiety Scale
  - c. Piers –Haris Children’s Self-Concept Scale
  - d. Kuder Preference Record
  - e. Rotters Locus of Control Scale

#### **Uses of Personality tests:**

1. Personality tests are used to measure individual traits and predisposition
2. Personality tests are used to assess individual feeling about himself/herself
3. Personality tests are used as checklist where individual check items that best characterized his/her trait in question
4. Personality tests are greatly used by clinical psychologists and counselor to presume their views when interacting with clients
5. Personality tests are adopted to a great extend in counseling researches

## **4.0 SUMMARY AND CONCLUSION**

### **4.1 SUMMARY**

The unit introduces the concepts of Personality tests. Uses of these tests were also discussed.

### **5.0 CONCLUSION**

Guidance counsellors, psychologists, and therapists uses tests with a great caution particularly the projective personality tests because the techniques require special skills for use (administration, scoring and interpretation).

### **6.0 TUTOR-MARKED ASSIGNMENT (TMA)**

Explain the uses of personality yests by counsellots?

## **7.0 REFERENCES/FURTHER READING**

Kerlinger, F. N. (1986). *Foundation of Behavioural Research* (3<sup>rd</sup> Ed.) New York. Holt, Rinehart and Winston.

Adeloye, T. (2006) *Fundamentals of Educational Research*. Kano: Debis-Co. Press and Publishing Limited.

Frankel, J. R. & Wallen, N. E. (2000). *How to Design and Evaluate Research in Education*. Boston. McGraw Hill

## **UNIT 4 PROBLEMS ASSOCIATED WITH PERSONALITY TESTS**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Problems associated with personality tests.
- 4.0 Summary and Conclusion
- 5.0 Tutor-Marked Assignment
- 6.0 References/Further Reading

### **1.0 INTRODUCTION**

Personality tests are instruments designed to measure an individual traits and predispositions. Personality tests are many – as many personality traits of an individual that can be outlined and defined. It was observed that tests such as personality tests provide inexact and limit measurement. As improved sampling and statistical methods developed, much controversy regarding the utility and validity of personality tests. Some of these problems are discussed in this unit.

### **2.0 OBJECTIVES**

At the end of this unit, you should be able to:

- Discuss five problems or limitations of personality tests
- Provides time strategies that can be used to overcome the problems associated with personality tests

### **3.0 MAIN CONTENT**

#### **3.1 Problems/Limitations of Personality Tests**

The following are researched to be problems of personality tests. Some of the problems are stated in the American Journal of Personality Assessment:

1. There are series of doubts about the theoretical assumptions behind projective personality testing. This problem started in the second half of the 1900s. This doubts lead to less use of the tests compared to the first half of the 1900s.

2. Personality tests typically, projective tests are being criticized of more time consuming to administer, score and interpreted.
3. The reliability and validity of personality tests are controversial. Experts argued that clinical judgment are used instead of norms and statistics to evaluate people's characteristics has convinced many that projective personality tests are deficient and unreliable (results are too dissimilar each time a test is given to the same person).
4. Projective personality tests also require special skills for use in administration, scoring and interpretations. Many practitioners (counselor, psychologist and psychotherapists) lack these skills and expertise required for the test.
5. Another general problem in tests including personality tests is that no tests provide exact measurement rather infer in their measurement
6. Personality tests provides limited measurement, approximate rather than an accurate measure of personality
7. Many of the personality tests are culturally biased and only creates artificial testing situations.
8. Many personality tests are influenced by a variety of factors such as hunger, fatigue, worry, anxiety, excitement conditions and influences test performance

Problem of personality tests also include improper interpretation of test scores or results. The interpreter sometimes placed too much faith in the test results and draw unwanted conclusions regarding the exactness of the test scores.

## **4.0 SUMMARY AND CONCLUSION**

### **4.1 SUMMARY**

Problems or limitations of personality tests were discussed in details. The next unit will look at various strategies that can be employed to reduce the problems.

### **5.0 CONCLUSION**

Upon all the problems observed with personality tests, many practitioners (counselors, psychotherapists, researchers and others) continue to rely on personality tests. Some testing experts (e.g. Cohen, Anastasi) suggest that personality tests results can be useful in developing therapeutic rapport, creating inferences to follow-up with other methods.

## **6.0 TUTOR-MARKED ASSIGNMENT (TMA)**

1. Discuss any five problems or limitations of personality tests

## **7.0 REFERENCES/FURTHER READING**

Robert, J. G. (2003). *The History of Psychological testing*. America Psychological American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and Psychological testing*.

Washington, DC: American Educational Research Association.

*Standards for Education and Training in Psychological Assessment*.

*Position of the Society for Personality Assessment – An Official Statement of the Board of Trustees of the Society for Personality Assessment. Journal of Personality Assessment, 87, 355 – 357* Aminu, K. I. & Gali, S. (2012). *Guidance and Counselling: Theories and Practice in Schools*. Kano. Albarka Printing Limited.

## **UNIT 5    WAYS OF OVERCOMING THE PROBLEMS ASSOCIATED WITH PERSONALITY TESTS**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Problems associated with personality tests.
- 4.0 Summary and Conclusion
- 5.0 Tutor-Marked Assignment
- 6.0 References/Further Reading

### **1.0 INTRODUCTION**

Personality tests are instruments designed to measure an individual traits and predispositions. Personality tests are many – as many personality traits of an individual that can be outlined and defined. It was observed that tests such as personality tests provide inexact and limit measurement. This unit examines some of the strategies that can be used to reduce problems associated with uses of personality tests.

### **2.0 OBJECTIVES**

At the end of this unit, you should be able to:

1. Provides some strategies that can be used to overcome the problems associated with the uses of personality tests

### **3.0 MAIN CONTENT**

#### **3.1 Ways of Overcoming the Problems Associated with Personality Tests**

The following strategies will assist the users of personality tests in guiding against its shortcomings:

1. The users of personality tests need to thoroughly understand the nature, validity, reliability and the established norms of the tests and the nature of the client or group to be tested.

2. Before using any personality test, the user should understand the person's or client's growth and development, such as the influences of home environment, vocabulary level, and reading ability.
3. The users should consider their effectiveness and skills before using any of the tests. The user should be honest in asking himself/herself of their capability in using the tests
4. Interest and consent of the client should be considered before using the tests
5. The users should not place too much emphasizes on the potentiality of the tests, but should combine different technique before interpretations
6. Users should not use first tests they find that appears to measure what they want, but search for more similar tests
7. Users of personality tests should search for available or attached tests manual or information that explains how the test can be used
8. In case of not commercial available test(s), users should seek the permission of the author before use. Try to acknowledge source
9. Users should conduct a pilot study to reestablish the quality estimates of the selected test(s).
10. Search, select and use culture free and culture fair tests

#### **4.0 SUMMARY AND CONCLUSION**

##### **4.1 SUMMARY**

Suggested ways and strategies of overcoming the problems associated with personality tests were discussed in the unit.

##### **5.0 CONCLUSION**

Upon all the problems observed with personality tests, many practitioners (counselors, psychotherapists, researchers and others) continue to rely on personality tests..

##### **6.0 TUTOR-MARKED ASSIGNMENT (TMA)**

1. Discuss ways by which problems of personality tests can be overcome?

## 7.0 REFERENCES/FURTHER READING

Robert, J. G. (2003). *The History of Psychological testing*. America Psychological American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and Psychological testing*.

Washington, DC: American Educational Research Association.

*Standards for Education and Training in Psychological Assessment*.

Position of the Society for Personality Assessment – An Official Statement of the Board of Trustees of the Society for Personality Assessment. *Journal of Personality Assessment*, 87, 355 – 357 Aminu, K. I. & Gali, S. (2012). *Guidance and Counselling: Theories and Practice in Schools*. Kano. Albarka Printing Limited.