

INFLUENCE OF CONTENT VALIDITY ON SECONDARY SCHOOL STUDENTS' ACADEMIC ACHIEVEMENT IN IMO STATE, NIGERIA

J. I. Nwaogazie

Department of Mathematics, Pre-degree Unit, Imo State University, Owerri

E-mail: nwaogaziejoel@gmail.com

Abstract

This study investigated the influence of content validity on secondary School Students' academic achievement in Imo state, Nigeria. The purpose of the study was to find out the extent to which content validity affects the academic achievement of senior secondary school students. A sample of four hundred and twenty students were randomly selected from fifteen thousand, five hundred senior secondary two students of 2009/2010 academic session in Imo state. The data used were collected from the students' scores using the researcher's constructed instrument called mathematics achievement and psychometric indices instrument (MAAII). The reliability index of the instrument was 0.9. The data were analyzed using mean and standard deviation for the research questions while the hypothesis was tested using Z-ratio at 0.05 level of significance. The null hypothesis was rejected. Based on the findings, it is recommended that teachers should ensure that they adequately cover the subject syllabus and that all examination test items should be drawn from all the content areas the students were taught. It was also recommended that teachers should be exposed to the techniques of constructing achievement tests that contain acceptable content validity indices through workshops, seminars and in-service training to update their knowledge. Finally, suggestions were made for further research.

Keywords: Content validity, academic, achievement, secondary school

Introduction

In recent years, there has been a loud cry that the academic performance of secondary school students as drastically reduced. The number of students that pass at credit level and above in external examination in Mathematics and English language is usually less than fifty percent of the total number of students that enroll to take the examinations. It is also observed that some students who pass the examinations set by their classroom teachers fail or make lower scores in external examinations in the same subjects.

In fact, the issue of fallen standard of education has become a household one in Nigeria. Newspapers and news media are full of the same

outcry. For instance, Elekwe (2010) wrote that two hundred thousand, three hundred (200, 300) students failed in two thousand and nine (2009) WAEC examination and that only thirty-one percent of the candidates who sat for the 2009 November/December West African Senior School Certificate Examination obtained credits in English and Mathematics, including three other subjects. If this trend is not checked, our country's dream of becoming one of the 20 technologically developed countries of the world by the year 2020 will not be achieved. Also, the poor academic achievement is evidenced by the result of WASCE over a 10 year period published by the National Bureau of statistics as shown in the following table

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Total entry	64062	75668	64337	102310	107890	93950	101952	105485	114927	124902
% pass at credit level and above	4	0	1	2	1	8	4	3	7	3
	11.15	18.81	32.81	36.55	34.50	36.91	33.97	38.20	41.12	46.75

Adewumi (1982) wrote on the effects of lack of qualified teachers on the teaching of

mathematics. This in his opinion is a major contributor to the fallen standard of education or

low academic achievement of students. Idu (1983) wrote that the factors responsible for the fallen standard include lack of qualified teachers, lack of commitment on the part of teachers, indiscipline on the part of pupils, etc.

Based on the fact that no study has been carried out on content validity indices in relation to academic achievement, the researcher was motivated to undertake a study on the influence of content validity in academic achievement of secondary school students in Imo State. Senior secondary school students in public schools in Imo State were used for this study.

Validity of an instrument is defined by Nwankwo (2006) as a measure of the extent to which an instrument measures what is supposed to measure. Validity refers to the results obtained with a research instrument and not the instrument itself. Iwuji (1997) wrote that the validity of a test is the degree of accuracy with which the test measures what it is intended to measure. The degree of validity may be low, moderate or high. Nworgu (1992) wrote that validity of a test refers to the extent to which a test measures what it is supposed to measure and nothing also. Therefore, validity of a test is dependent on the purpose. This means that a test which is valid for one purpose may not be valid for another. For instance, a test which is valid for assessing achievement in SS II mathematics will not necessarily be valid for SS III mathematics.

Types of test validity

Four main types of validity were discussed in Onunkwo (2002) and they are content, face, construct and criterion – related validities. All these validity types are not established and appropriate for all the instruments. For instance, content validity is appropriate for achievement test, while predictive validity can be established.

Content validity: Nkwocha (2004) sees content validity as the extent to which test items consist of a representative sample of the subjects matter and objectives the test is designed to measure. Gronlund (1969) in Nkwocha (2004) opines that content validity is demonstrated by showing how well the content of the test samples of the class of situations or subject matters about which conclusions are to be drawn. Content validity of

a test involves adequate coverage of the subject matter and behavioral objectives in a particular syllabus by an instrument. Therefore, Gronlund (1985) in Onunkwo (2002) described content validity as the degree to which the items of an instrument measures a representative of the subject matter content and the instructional objectives.

The content or subject matter are the topic to be treated while instructional objectives are the changes to be sought in the students. Instrument of measurement such as test should provide results which are representative of topics and behaviors we want to determine. Therefore, Onunkwo (2002) defines content validity as the art of testing all that the students are supposed to have studied i.e. topics and behavioral objectives and not just other things. A test blue print or table of specifications is employed to ensure a systematic coverage of the topics and the instructional objectives in order to produce content valid test. The table of specifications is a two-way chart with the topics listed along the left column and the objectives along the top row. This ensures easy computation of the weight of each topic or objectives as shown in the test. Furthermore, they are the kinds of learning to be tested, and the relative importance of individual topic and objectives. Content validity is also called curricula, rational or logical validity and it is most suited for achievement tests since achievement tests are designed to measure how well the examinee has mastered a specific course of the study.

Face validity: This is not really a true estimate of content validity but it gives a quick idea about the content validity of a test. Nkwocha (2004) sees face validity as the extent to which a test superficially looks like a test on the subject it intends to measure. This type of validity helps in sustaining the motivation of the users of the test. A mathematics test should contain mathematics symbols. Also, Horrocks and Schoonover (1968) opined that face validity is the extent to which a test looks as if it would measure what it is intended to measure, and therefore, refers to the facial appearance of the test. Therefore, it is concerned of how experts and students consider any instrument-whether the items, vocabulary, instructions, etc. appears too difficult for what

JSSI students will take or too much like what primary five pupils instead of SS II students for who it was intended.

Face validity of an instrument is achieved by giving copies of that instrument to specialist in the area as well as experts in measurement and evaluation. These experts should be requested to vet items of the instrument in terms of clarity of words, language difficulty, and relevance to the course content etc. the final version of the instrument must incorporate their recommendations. Students feel happy, motivated when they are administered with instruments which appeal to them in terms of difficulty and relevance.

Construct validity: Construct is used to refer to those abstract psychological traits such as emotion, anxiety, aggression, intelligence, etc. Thorndike and Hagan (1977) in Onukwo (2002) opined that construct validity refers to the accuracy with which an instrument describes an individual in terms of some psychological traits.

Also, Eze (1992) sees construct validity of a test as the extent to which the test measure a psychological construct or trait which it is supposed to measure. Similarly, Nkwocha (2004) defines construct validity as how well a test actually measures the psychological or theoretical construct it is designed to measure. The extent to which performance in a test reveals possession of a specified psychological construct is also construct validity. Some of the methods used in estimating construct validity of measuring instruments are internal consistency, age differentiation and correlation with other tests. The developers of such test are Stanford Binet scales; Wechsler scales use age differentiation as the major criterion in validating them. To show whether their scores exhibited a progressive increase with advancing age, these tests were matched against student's chronological ages. The reason for this was based on the idea that since abilities increase with age during childhood, it implies that test scores should also show such increase if the test is valid. Therefore, if test scores do not increase with age, it shows that the test is not a valid measure of ability it was designed to measure.

The method of internal consistency entails that students' scores in sub test is correlated with their scores in the total test. This is homogeneity test. Any sub test that correlates poorly with the total test is discarded because it does not measure the same general traits as the total test. As the sub test correlate more with the total test, the degree of homogeneity of test will increase and the degree of homogeneity of a test is its degree of construct validity.

Another technique involves the correlating of students' scores in a new test with their scores in an easier similar test and if the coefficient of correlation is high it indicates that the new test measure the same general area of behaviors as the earlier one. The degree to which the new test correlates with the older one is its degree of construct validity.

Criterion-related validity: Two sets of scores are of interest in criterion related validity. These scores are the predictor scores and the criterion scores. Students' scores in a test that intend to predict a particular achievement, aptitude or behavior are the predictor scores. The UTME scores or continuous assessment scores may be useful to predict student university achievement or achievement in the JSSCE/SSCE. The UTME score and continuous assessment scores are the predictor scores while the university or SSCE achievement score are the criterion scores.

Gronlund (1968) states that "criterion related validity is demonstrated by comparing the test scores with one or more external variables considered to provide a direct measures of the characteristics or behaviour in question". (p167). Salvia and Yesseldyke (1978) define criterion related validity as "the extent to which a person's score on a criterion measure can be estimated from those persons score" (p99). Borg and Gull (1979) sees criterion related validity as the ability of a test to measure an individual's behaviour on some other variable called a criterion. There are two types of criterion related validity namely predictive and concurrent validity.

Predictive-criterion-related validity: When an individual's score in test is used to predict his future performance in another test which measures a similar criterion, the test is said to

have predictive validity. Salvia and Yesseldyke (1978) state that predictive validity estimated tries to verify whether a person's test score allows an accurate estimation of that persons' score on a criterion measure administered some time in future. The correlation coefficient estimated of the two sets of score is regarded as the predictive validity index. Here, the second test that is correlated with the first one is given after a long time interval.

To determine this type of validity, a test is first administered to candidates entering into job or academic institution. The score of this test are called predictor score. These students, are followed up and at the end of the programme, their achievement or performance scores called criterion scores are correlated with the predictor scores using Pearson r. The higher the predictive validity index, the more effective the test is in predicting the test in question.

This type of validity is mostly applied in intelligence tests, aptitude tests, attitude and interest tests. Really, all tests used in selecting candidates into armed forces, business industry, education, etc. or in predicting achievement or performance should have high predictive validity.

Concurrent validity: This is another type of criterion related validity. It is indispensable for tests used in diagnosis of existing status. This is the most appropriate tests used in most psychological tests. A test is said to have concurrent validity when there is a correlation between the scores obtained by the group examiners in the test and score they got in the criterion test which they took within the same period. It should be noted that these two tests to be correlated are taken concurrently. The correlation coefficient obtained is called concurrent validity index.

Concurrent validity and predictive validity are similar in that each matches the predictor scores against the criterion scores. However, in the concurrent validity, both criterion scores and predictor scores are obtained at approximately the same time while in predictive validity, the criterion scores are obtained long after the predictor scores have been got. Also, predictive

validity is relevant for the use in predicting future achievement or performance while concurrent validity is relevant for tests used in the diagnosis of existing issues.

Factors affecting validity of measuring instruments: The following factors reduce the validity of measuring instruments; unclear instruments, use of irrelevant technical terms, poor construction of items, ambiguous statements, inappropriate of terms, non coverage of the content, difficulty level of the instrument, improper arrangement of items and poor arrangement of answers options. Nkwocha (2004) opines that factors which affect test validity are, item coverage, irrelevant items, ambiguity, item difficulty, organization of items, test duration, test civility and personal factors.

Validity coefficient: Onunkwo (2002) sees validity coefficients as the degree of performance of concept. Validity of measuring instruments is interpreted by using coefficient and such coefficient is called validity coefficient. The predictor scores are correlated with the criterion scores when computing the validity coefficient. The correlation statistics that can be used to compute the validity coefficient include Pearson r, biserial correlation, etc. The computed validity coefficient should be high enough in order to be accepted. Annastasi (1961), wrote that it must be high enough and significant beyond 0.05 or 0.01 level of significance.

The concept of academic achievement / achievement test

Academic achievement can be described as the quality and quantity of learning (changes) that has taken place in the learner in the school after being exposed to a course of instruction during the teaching and learning experience within a specified period. Bell (2009) defined academic achievement as the ability to study and remember facts and to be able to communicate your knowledge verbally or on paper. It further refers to how students deal with, or accomplish different tasks given to them by their teachers. In educational institutions, success is measured by academic achievement, or how well a student meets standards set out by local government and the institution itself as career competition grows even fiercer in the place of work. The importance

of students doing well in schools has caught the attention of parents, legislators and government departments alike (Bell 2009).

Nwana (1982) described achievement as how much the pupils have learned and that the test used to determine the achievement of pupils (students) is generally referred to as achievement test. All the various tests set by test experts, the national and international examination bodies are called teacher-made achievement test and standardized achievement tests respectively. Ukwuije (1993) pointed out that an achievement test is a test designed to find out how much a student or client has achieved after a specific instruction or training in a course or school subject. Score, grades, marks are used to denote academic achievement. Similarly, Nwana (1982) opined that an achievement test is any test given to determine how much the pupils have learned. This implies that an achievement tests gives an index of the students mastery of the subject matter taught them. Achievement tests made by classroom teachers are called teacher-made achievement test while those constructed by psychometricians are called standardized achievement tests.

Nkwocha (2004) defines a standardized achievement test as a test designed by experts from different areas of external candidates following an accepted quality standard for test construction and which sets a specified uniform procedures for its administration, scoring and grading. Ukwuije (1993), it was written that a standardized achievement test has its steps for administration, scoring, interpretation, systematically mapped out in the tests manual and remains the same irrespective of who the tester is, when and where the testing is done. He further pointed out that scores of standardized tests are usually reported in standard scores like Z-scores, percentile and stannine scores. Standardized achievement and teacher-made achievement tests can take the form of essay or objective and can equally be administered in writing or rally. Achievement tests can be used for diagnostic, prognostic, motivational, promotional, placement, certification, and evaluation etc, purposes.

Furthermore, different types of achievement test exist as put forward by different scholars. Onunkwo (2002) writing on types of achievement tests include survey tests; speed tests; power tests; readiness tests; diagnostic tests; placement tests; formative tests; summative tests; norm-referenced tests, and criterion-referenced tests. Similarly, in Ukwuije (1993), the following types of achievement test exist: survey tests batteries; survey test in specific areas; diagnostic tests and prognostic tests or readiness tests.

The purpose of the study was to investigate if content validity affects academic achievement and specifically, to find out the extent to which content validity affects the academic achievement of senior secondary school students.

The understated research question guided the conduct of the study: To what extent does content validity affect academic achievement of senior secondary school students?

The study tested one hypothesis at (P<0.05).

Method

The study used the analytic descriptive survey design. The population consisted of fifteen thousand senior secondary two (SS2) students of 2009/2010 in Imo State. Four hundred and twenty SS2 students were random sampled for the study. The sample size determined was arrived at by using the yaro yamen's formula as shown in the following formula.

$$n = \frac{N}{1+N(e)^2}$$

Where n = samples size sought
e = level of significance and N=population size.

Collection of data for the study was done using researcher's developed mathematics achievement and psychometric indices instrument (MAAPII). The instrument is a multiple choice objective test with options A to E and the students were expected to tick or circle the right option. The MAAPII contains 50 objective test items and it has three sections. Section A contains 20 validity items. The first 10 items have high content validity indices; while 11 to 20 are items with

low content validity indices. The content validity index of items with high content validity is one (1) while the range of values of items with low content validity is 0.25 to 0.5.

Administration and scoring of the instrument

Copies of MAAPI were administered directly to the testees by the researcher with the help of research assistants. The instrument was so named because the items that make up this achievement test are of different levels of psychometric indices required to achieve the objectives of the study. Instructions guiding the answering of the items were explained to the testees. Copies of the answered question papers were collected from the examinees on the spot. For the purpose of scoring the instrument, each item a student got right was qualified 1 point. The total score of each student for each sub-section is got by adding the scores for all the items in the sub-section. The maximum score for each sub-section is 10.

Determination of the content validity

The content validity of the test items were determined using test blue print and judgment by two content specialists and the data obtained from this were used to compute an index of item – objective congruence as recommended by Rovineli and Hambleton in Onunkwo (2002). The item – objective congruence is computed by the formula

$$I_{10} = \frac{(M-1)(S_o - S'_o)}{2N(M-1)}$$

Where; I_{10} = index of item – objective Congruence for item I and objective O

M = number of objectives

N = number of content specialists

S_o = sum of the rating assigned to objective O

S'_o = sum of rating assigned to all other objectives except objective O

Reliability of the instrument

Reliability of the MAAPI was determined by adopting test – retest method which is a measure of temporal stability. The instrument was administered to the same students twice. The gap between the first and second administration was seventeen days. Their individual total scores for the two administrations were gotten. Then, the value of the reliability coefficient was arrived at using Pearson Product Moment formula.

Data analysis technique

The mean and standard deviation were used to answer the research question so as to determine the difference between their mean scores while the hypothesis was tested using Z-test at 0.05 level of significance. Z – computed was compared with table Z at 0.05 level of significance (1.96) under N-2 degrees of freedom for a non-directional test. The null hypothesis was accepted if the calculated Z was less than the table Z while the null hypothesis was rejected if the calculated Z was greater than table Z.

Research question: To what extent does content validity affect academic achievement of senior secondary school students?

To answer this research question, the means and standard deviations of the students’ achievements based on items with high and low content validity indices were computed and the result of the analysis are presented in table.

Table 1: Mean achievement and standard deviation based on high and low content validity indices.

Variables	$\sum fx_1, \sum fx_2$	N	Means (X_1, X_2)	SD ₁
High C.V (X_1)	1462		3.48	1.64
Low C.V (X_2)	919	420	2.19	1.63

Table 1 shows that the mean academic achievement of students tested on items with high content validity indices is 3.48 while their mean academic achievements on items with low content validity indices is 2.19. This result shows that low and high content validity indices affect

academic achievement of students. The mean academic achievement of the students tested with items that have high content validity indices is greater than their mean academic achievement on items with low content validity indices.

Hypothesis

Ho: There is no significant difference between the mean academic achievement of students tested on items with high content validity indices

and their mean academic achievement on items with low content validity indices.

To test this hypothesis, a Z-test of two sample means was applied and the results of the analysis are presented in table 2.

Table 2: Mean achievement, standard deviation and Z-test based on high and low content validity indices.

Variables	$\sum fx_1, \sum fx_2$	N	(X_1 X_2)	SD_1 SD_2	Z_{crit}	Z_{cal}	Alpha-level	df	Result
High C.V (X_1)	1462		3.48	1.64					
Low C.V (X_2)	919	420	2.19	1.63	1.96	6.04	0.05	418	Significant

It is clearly seen from Table 2 that 420 students tested on items with high content validity had a sum ($\sum fx_1$) of 1462, mean (X_1) of 3.48 and standard deviation (SD_1) of 1.64 while on items with low content validity indices they had a sum ($\sum fx_2$) of 919, mean (X_2) of 2.19 and standard deviation (SD_2) of 1.63. The table has also shown that the calculated Z-ratio between the two variables is 6.04 while the table value of Z is 1.96, the alpha-level is 0.05 while the degrees of freedom is 418. Since the obtained or calculated Z-ratio (6.04) is greater than the critical value of Z (1.96) at 0.05 level of significance under 418 degrees of freedom for a non-directional test, the first null hypothesis is rejected. That is, there is a statistically significant difference between the mean academic achievement of students tested on items with high content validity indices and their mean academic achievement on items with low content validity indices.

two variables (6.04) is greater than the critical value of Z (1.96) at 0.05 level of significance under 418 degrees of freedom, the null hypothesis is rejected. That is, there is a significant difference between the mean academic achievement of students tested on items with high content validity indices and the mean academic achievement on items with low content validity indices.

This result showed that the students had a better achievement mean when tested with items with high content validity indices than when tested on items with low content validity indices.

Discussion of results

Summary of results

Research question: To what extent does content validity affect academic achievement of senior secondary school students?

The findings of the study are summarized below:

Hypothesis: There is no significant difference between the mean academic achievement of students tested on items with high content validity indices and their mean academic achievement on items with low content validity indices.

The mean academic achievement of students tested on items with high content validity indices is 3.48 while their mean academic achievement on items with low content validity indices is 2.19. This shows that secondary school students in Imo State had greater achievement mean when tested with high content validity than when tested on items that have low content validity indices. Also, since the calculated Z-ration between the

The result shows that the mean academic achievement score and standard deviation of students tested on items with high content validity indices are 3.48 and 1.64 respectively while their mean academic achievement score and standard deviation on items with low content validity indices is 2.19 and 1.63 respectively. The result also shows that the calculated Z-ration between the two variables is 6.04 while the critical value of Z is 1.96 at 0.05 alpha level under 418 degrees of freedom. Since the calculated Z-ratio (6.04) is greater than the critical value of Z (1.96) at 0.05 level of significance under 418 degrees of freedom, the

null hypothesis is rejected. This indicates that there is a significant difference between the mean academic achievement of students tested on items with high content validity indices and their mean academic achievement on items with low content validity indices.

This result means that the academic achievement mean of students on items with high content validity indices is greater than their mean academic achievement on items with low content validity indices. The result that the mean academic achievement of students tested on items with high content validity indices is greater than their mean academic achievement on items with low content validity indices is expected and not surprising. This is due to the fact that these items were drawn from the students' scheme of work and that they might have been taught the topics from where the items were drawn. The lower mean academic achievement of students on items with low content validity indices could be attributed to the fact that some of items were drawn from topics outside the students' scheme of work and also they might not have been familiar with the techniques of solving or answering such items. Other factors that could have contributed to the lower mean academic achievement of the students are organization of items, test duration, test civility and personal factors. The result of this study is in agreement with the earlier position of Iwuji (1997) that a student's failure in a test may be partly or wholly due to factors inherent in the test itself as well as circumstances surrounding its administration and scoring. One of the factors inherent in the test is item validity index.

Implications of results

The implications of the result of a study are the logical consequences which those results have for education (Nwana, 2005). The findings made in this study have implications for students, teachers, school administrators, the government and other educational providers.

The study found out that the mean academic achievement score of senior secondary school students tested on items with high content validity indices was greater than their mean academic achievement on items with low content

validity indices. This finding implies that teachers could raise the academic achievement of students in mathematics by ensuring proper coverage of the subject syllabus. There is also the implication that the students' academic achievement is affected negatively if they are tested on items with low content validity. The result further suggests that some students do not take the study of mathematics very serious while others take it serious. Also, the result of the study suggests that more school administrators ensure that their teachers cover the scheme of work in mathematics while few of them pay lip service to the supervision of their teachers.

Recommendations

Based on the result this study, the following recommendations were made:

1. Teachers should ensure that they adequately cover the subject syllabus and that all their test or examination questions should be drawn from all the content areas of the subjects students were taught.
2. Examination questions should be typed and should contain multiple choice and essay types.
3. School administrators should ensure that their teachers teach their students regularly and cover the subject scheme of work for the academic session.
4. Activities that can disrupt normal teaching/learning encounter should not be encouraged.
5. Parents should provide their children with the relevant text-books and ensure that their children do their assignments and work hard academically.

Limitations of the study

Every study has aspects in which it falls short of the ideals which the researcher has established or which the researcher is not even aware. According to Nwana (2005), it is a mark of intellectual honesty for a researcher to admit that these shortcomings did exist and to give an account of the way they have manifested. This study has few limitations.

Subject mortality: Some of the students used for first administration were absent during the second administration. This affected the number of students whose pair of scores were used to compute the reliability index of the instrument. Even though there was such limitation, the study has some credit to itself. For instance, the sample used (420) is far more than the minimum sample required from the population used, Yaro Yamens' formula (Baridam in Nwaogazie 2000). The study was able to determine the extent to which low and high content validity indices affect academic achievement of senior secondary school students in Imo State.

Finally, the findings are believed to be valid and reliable because the statistical tools used for data analysis tested for significant difference and would have taken care of such limitations.

References

- Adewumi, D.O. (1982). The effect of lack of qualified teachers on the teaching of mathematics in Nigeria. *Journal of Science Teachers Association of Nigeria* 2, 14-30.
- Anastasi, A. (1961). *Psychological testing*. (6th ed) New York: Macmillan Company.
- Bell, M. (2009). Defined Academic Performance. Online at http://www.chow.com/about474075_define_academic_performance.htm.retrieved on 08/09/2009.
- Borg, W.R. & Gull, M.D (1979). *Educational Research: An Introduction*. New York: Longman.
- Elekwe, T. (2010, January 15). WAEC Exam Result, Daily Champion, P. 1,6.
- Eze, D.N (1992). Reliability and validity. In B.G. Nworgu, *Measurement and evaluation: theory and practice*. Awka. Hallman Publishers.
- Gronlund, N.E. (1968). *Reading in Measurement and Evaluation, Education and Psychology*. New York: Collier Macmillan.
- Horrocks, J.E & Schoonover, T.I. (1968). *Measurement for Teacher*. Ohio: A bell and Howell Co.
- Idu; C. (1983, June 12). Falling Standards of Education: Factors Responsible. *New Nigeria*, P.3.
- Iwuji, V.R. (1997). *Measurement and Evaluation for Effective Teaching and Learning*. Owerri, Joe Mankpa's prints.
- Nkwocha, P.C (2004) *Measurement and Evaluation in the Field of Education*. Versatile Publishers. Owerri P.123.
- Nwana, O.C. (1982). *Educational Measurement for Teachers*, Walton: Thomas Nelson and son Ltd.
- Nwana, O.C (2005). *Introduction to Education Research*. Ibadan: Heinemann. Educational Books (Nig) Plc.
- Nwaogazie, J.I (2000). Relationship Between Continuous Assessment Scores and JSSCE Final Score in Imo State. Unpublished Masters Thesis. University of Port Harcourt.
- Nworgu, B.G (1992) *Educational Measurement and Evaluation, Theory and Practice*. Nsukka, Hallman Publishers.
- Onunkwo, G.I (2002). *Fundamental of Education Measurement and Evaluation*, Owerri: Cape Publishers International Ltd.
- Salvia, J. & Ysselydke J.E (1978). *Assessment in Special and REMEDIAL Education*. Boston: Houghton Muffin Company.
- Thorndike, R. & Hagan, E. (1977), *Measurement and Evaluation in Psychology and Education*. New York: John Wiley & Sons, Inc.
- Ukwuije, R.P (1993). *Appraisal Techniques in Guidance and Counseling*. Port Harcourt. CAPIIC Publishers.