Research on Humanities and Social Sciences
ISSN (Paper)2224-5766 ISSN (Online)2225-0484 (Online)
Vol.5, No.7, 2015 

# Analysing the Psychometric Properties of Mathematics in Public Examinations in Nigeria

MOYINOLUWA, TOYIN D.
Department of Educational Foundations,Faculty of Education, Nasarawa State University, Keffi
(NERDC, P.M.B 91, GARKI, ABUJA)

**Abstract**
This paper investigated the psychometric properties of Mathematics examinations conducted by four examination bodies (NABTEB, NECO, JAMB and WAEC) to establish the quality of the items presented to secondary school students for of the purpose of certification and placement. The study employed descriptive survey research design to seek for information from a segment of the population of senior secondary students to make generalization on all Nigerian students adopting the multistage stratified sampling technique. From each of the 6 geo-political zones, 2 states were randomly selected to obtain a total of 12 states. One co-educational Federal Government College was purposively sampled from each state, plus three public schools, one private secondary school and one technical college were selected from each state. Thirty SS3 students selected, using the systematic random sampling technique in each of the sampled schools that have been presenting students for public examinations. The test batteries were past question papers developed in four common subjects offered in public examinations as listed above. The statistical tools employed to provide empirical answers to the research questions are: calculation of difficulty indices using relevant formula; split half correlation, Distractor power estimate, and Discrimination indices using relevant formula, also quantitative validations of test were conducted. A fairly high proportion of the test items have appropriate difficulty index i.e. within ranges 0.25 – 0.75. Validity of test batteries used was fairly good. Other psychometric characteristics were found to be generally acceptable. It is recommended that, for easy tests, attention should be given to higher order skills required of graduate of secondary schools and coverage of most aspects of the syllabi prescribed by examination bodies. It is also recommended that the present standard of examinations in Nigeria should be sustained and improved upon.
**Keywords:**Psychometric properties, Public Examinations, Distractor Power, Discriminative Index, Reliability, Validity.

## Introduction

Over the years, performance of students in public examinations has not been encouraging. This has been a source of concern to government, educators, parents and the general public. Studies have shown that the reasons for the dismal performance could be attributed, among others, to the poor state of education in the Country, low quality teaching staff, nature of the subjects, inadequate preparation of students for examinations as well as cut in education budgets leading to shortages of facilities and equipment needed for effective teaching and learning.

Public examinations in Nigeria, particularly the Secondary School Certificate Examinations, have been afflicted by examination malpractices and mass failure. The current Assessment and Evaluation procedure is not effectively implemented and the situation is worsened by lack of efficient and committed supervision and monitoring of education process (Moyinoluwa, 2015).

In any case, since the senior secondary school certificate examinations were set, conducted, scored and graded by bodies external to the schools, it is possible that some of the viruses that account for poor performance could be related to the external examination bodies. Actually, several other factors were observed to be responsible for students' mass failure in public examinations, some blamed the home, society and parents for failure to inculcate discipline and learning habits in their children.

Yet, others blamed the Government for failure to provide human and material resources to facilitate good teaching and learning, some blamed the teachers for failure to inculcate the necessary knowledge, skills and behavior to students and also the students themselves for refusal to learn, while majority of others blamed WAEC or NECO for their failure to measure adequately the knowledge, skills and behaviour learned by students. Whatever the factors are, there is clearly the need to get something done urgently in order to create the type of assessment and examination procedures that is deserved, as it may be a silent cause for the poor performance level in the SSCE.

According to Adegun (2004), the basic qualification for admission into any higher institution in Nigeria is the Senior School Certificate issued by the WAEC, NECO and / or NABTEB. At the end of secondary school education, students are expected to sit for examinations conducted by the West African Examinations Council (WAEC), National Examinations Council (NECO) and National Business and Technical Examination Board (NABTEB).

The certificates awarded by these examination bodies are officially recognized in Nigeria as equivalent.

 

The certificates obtained by candidates that passed the examinations set by any of the three bodies could be used to obtain employment in the appropriate cadres of public service and/or in private companies and corporations. Indeed, some candidates often merge the grades they obtained from sittings in any two of the examinations.

More importantly, the quality of the certificates offered by any of these examination bodies determines the admissibility or otherwise of the possessors into institutions of higher learning both in Nigeria and abroad. A minimum of five credits passes in relevant subjects in these examinations ordinarily qualifies a candidate to seek admission into the tertiary institutions of his/her choice after satisfying the requirements for the University Matriculation Examinations conducted by the Joint Admissions and Matriculation Board (JAMB).

Psychometric characteristics of examinations refer to certain attributes inherent in tests upon which an assessment of candidates is based. These characteristics include the facility and difficulty indices, the discrimination index, the power of distractors, validity and reliability indices. It is perhaps worth mentioning that these attributes of a test are most often ignored. The Dictionary of Statistical Terms (Marriott, 2002) defined item analysis as analysis based on the responses to individual questions.

Oshkosh (2005) considers Item Analysis as probably the most important tool to increase test effectiveness. It is a scientific way of improving the quality of tests, and test items in an item bank. An item analysis provides three kinds of important information about the quality of test items.

-*Item difficulty* - a measure of whether an item was too easy or too hard, also called facility index.

-*Item discrimination* - a measure of whether an item discriminated between candidates who knew the material well and candidates who did not.

-*Effectiveness of alternatives* - determination of whether distractors (incorrect but plausible options) tend to be chosen by the less able students and not by the more able candidates.

Allen and Yen (2008) observed that the CTT model are procedures for determining item difficulty index, discrimination power and the distractor power of options for understanding and improving the reliability of psychological tests. Right from the middle of the second half of the 20$^{th}$ Century, item response theory developed alternative approaches for understanding, interpreting and administering psychological tests. Difficulty, Discrimination and chance factors are calculated and item characteristics identified for calibration and adaptability to administration of tests by use of computed.

Schnitzer (2008) has identified two major controversies associated with testing. They stem from two interrelated issues, namely: technical shortcomings in test design and ethical problems in interpretation and application of results. It is probable that some technical weaknesses exist in the tests developed and administered by the WAEC, NECO NABTEB and JAMB. Therefore, it is advisable that their results be viewed as only one kind of information about the test takers. Most criticisms of testing arise from the overvaluation of and inappropriate reliance on test results in making major life decisions.

For instance, many psychologists would agree that using tests to bar youngsters from educational opportunities as is the case with JAMB tests is unethical. The argument is that as long as unequal learning opportunities exist, they will continue to be reflected in test results. Apparently, the current study might yield findings that would serve as reference points for monitoring and refining ethical standards and public policy recommendations regarding the use of tests and public examinations. These previous studies had assumed that the examinations and the operations of the bodies handling them are normal and adequate. This assumption is not oblivious of the fact that the general level of performance is not independent of the quality of the measuring instruments and the social circumstances in which the measurement process has been undertaken. To this end, there is dire need to assess and analyze the psychometric properties of mathematics in public examinations conducted by WAEC, NECO, NABTEB and JAMB in Nigeria.

**Objectives of the Study**

1. Difficulty or facility index of the test items used by WAEC, NECO, NABTEB and JAMB for SSCE Mathematics.
2. Compute the discrimination power of each item used by the examining bodies in the respective subjects.
3. Distractor Power of the Options
4. Reliability coefficient (estimate of Internal Consistency) of the test instruments.
5. Content validity of the test instruments

**Research Questions**

1. How appropriate are the difficulty indices of test items used?
2. What is the discrimination power of the test items used by the examining bodies in Mathematics?
3. To what extent is the content of SSCE test items used in Public Examinations valid?
4. How functional are the distractor powers of the test items options in these Examinations?
5. What is the reliability coefficient (estimate of Internal Consistency) of each of the tests conducted by various Examination bodies?

 

## Methodology

### Research Design

The study employed a survey research design. Descriptive (Survey research) design was employed, since the data involved in the study were collected from the source without any manipulation. This Survey approach was considered most appropriate because the study sought information from a small segment of the population to make a generalization for all Nigerian senior secondary III students. The study focused on validation of test instruments with the aim of obtaining data from sampled respondents for establishing the Psychometric Characteristics of Public Examinations. Randomized block distribution of tests to the candidates was used.

### Population, Sample and Sampling Techniques

The population under investigation includes all SS III students at the final stage of their preparation for WAEC and NECO, SSCE as well as NABTEB papers. The representative samples were obtained from the students that wrote WAEC, NECO and NABTEB examinations in 2010 academic year. A multistage stratified sampling technique was adopted. First, the country was considered on the basis of the existing six geopolitical zones.

All Six zones were involved in the study. 2 states were selected randomly from each geopolitical zone using the 'dip hart' procedure to obtain a total of 12 states involved in the study. One co-educational Federal Government College was purposively sampled from each selected states of the federation, making 12 FGCs in the study. Additional three public Secondary Schools, 1 Private Secondary School and one technical college were selected from each geo-political zone.

The Public Schools were selected taking into consideration adequate representation of location (Urban &Rural); School ownership and School type (Boys only & girls only). Special attention was given to identification of Schools that take NABTEB examinations (technical college) since the examinations are taken by special candidates. The population under investigation includes all SS III students at the final stage of their preparation for WAEC and NECO, SSCE as well as NABTEB papers. The representative samples were obtained from the students that wrote WAEC, NECO and NABTEB examinations in 2010 academic year.

Thirty SS3 students was selected from each schools using systematic random sampling technique in the sampled secondary schools from the six geo-political zones of Nigeria. Making a total of 2,160 respondents randomly selected from the selected states. Only secondary schools that have been presenting students for public examinations were considered.

### Instrumentation

Test batteries were the main instruments used in this study and they were selected school subjects. Specifically, the test batteries were past question papers developed in Mathematics as offered in public examinations conducted by WAEC, NECO, NABTEB and JAMB. The instrument for the study was administered on the sample students in the SS III class in the selected schools from the 12 selected States for the study. Seventy-two (72) specialist subject teachers were recruited as Field Research Assistants to help with the test administration and invigilation.

### Data Analysis

The following statistical tools were employed to provide empirical information/answer to each research question as appropriate.

- Calculation of difficulty indices using relevant formula
- Split half correlation.
- Distractor power estimate
- Discrimination indices using relevant formula

Also quantitative validations of tests were conducted using the following criteria.

- Content validity of tests by year and examination body.
- Spread of items in relation to syllabus and number of items in test
- Structure of items in respect of coherence, tenses specificity, ambiguity e.t.c.
- Relation of test items to emphasis specification of the curriculum.

## RESULTS AND DISCUSSION

### Research question 1

How appropriate are the difficulty indices of test items used?
This question was addressed with the information in table.3.1

 

**Table 3.1: Number of Items with Appropriate Difficulty Index by Exam Body, Year and Subject.**

| | | % Item Goodness | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | WAEC | | | | | NECO | | | | | NABTEB | | | | | JAMB | | | | |
| Subject & Year | | < DT | DT | >DT | Items with Inappropriate Difficulty | (%) | < DT | DT | >DT | Items with inappropriate Difficulty | (%) | <DT | DT | >DT | Items with inappropriate Difficulty | (%) | >DT | DT | >DT | Items with Inappropriate Difficulty | (%) |
| | 2008 | 3 | 47 | 0 | 3 | 6.00 | 20 | 40 | 0 | 20 | 33.33 | 32 | 18 | 0 | 32 | 64.00 | 25 | 25 | 0 | 25 | 50.00 |
| MATHS | 2009 | 9 | 39 | 2 | 11 | 22.00 | 20 | 40 | 0 | 20 | 33.33 | 25 | 25 | 0 | 25 | 50.00 | 16 | 34 | 0 | 16 | 32.00 |

Difficulty \index (DT) is given as $0.25 \leq DT \leq 0.75$ = Appropriate
$DT < 0.25$ and $DT > 0.75$ = Not Appropriate

The acceptance range of difficulty was widened a little from the middle course of 0.4 - 0.6 to 0.25 – 0.75. All items having difficulty index lower or higher than this acceptance range were considered to be inappropriate. As shown on table 3.1, NABTEB has the highest percentage of items with inappropriate difficulty index as 0.64% of the items in 2008 in multiple choice items in Mathematics, The lowest number of items with inappropriate difficulty index was recorded by WAEC which had the highest percentage of 22 in 2009 Mathematics, with 2008 Mathematics recording the lowest. For NECO the 2008 and 2009 multiple choices test in Mathematics has the highest percentage of inappropriate difficulty with 33.3% The highest number of items with inappropriate difficulty index in the JAMB papers was recorded in 2008 Mathematic (50%)

**Research question 2**

What is the discrimination power of the test items used by the examining bodies in the respective subjects? This question was addressed with information in table.3.2

**Table 3.2: Discrimination Power of Test Items By Exam Body, Year and Subject**

| | | % Item Goodness | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | WAEC | | | | NECO | | | | NABTEB | | | | JAMB | | | |
| Subject & Year | | DI≥0.35 | 0.0≤DI<0.35 | SUM | % | DI≥0.35 | 0.0≤DI<0.35 | SUM | % | DI≥0.35 | 0.0≤DI<0.35 | SUM | % | DI≥0.35 | 0.0≤DI<0.35 | SUM | % |
| | 2008 | 35 | 15 | 50 | 70.00 | 31 | 29 | 60 | 51.67 | 9 | 41 | 50 | 18.00 | 16 | 34 | 50 | 32.00 |
| MATHS | 2009 | 28 | 22 | 50 | 56.00 | 23 | 37 | 60 | 38.33 | 11 | 39 | 50 | 22.00 | 17 | 33 | 50 | 34.00 |

Discrimination Index (DI) is given by: $DI \geq 0.35$ = Good
$0.0 \leq DI < 0.35$ = Poor

The acceptance range of discrimination power of multiple choice items is 0.4 and above. For this study, the lower level was extended to 0.35.it was observed that WAEC had the highest percentage 70% of items that discriminate well between the high and lower achievers in 2008 mathematics test. Others ranged between 48.33% to 58.33%. Among the certificate examinations NABTEB came in the rear with the highest Percentage in 2008 Mathematics.

**Research question 3**

To what extent are the test items used in Public Examinations valid?

This research question was addressed using content validity. The test items in each of the subjects and years of interest in this study were compared with the content specification of the syllabi used by each examination body under scrutiny with the help of subject experts. Four examination bodies were considered, viz: NECO, WAEC, NABTEB and JAMB. It was discovered that essentially all the four bodies have the same structure of syllabus and the objective questions were set vis –a- vis their syllabi.

The aim of this work was to determine whether the objectives of the syllabi were attained in terms of coverage and contents of the tests. The essay type questions were equally examined in the same manner to determine whether the questions were properly structured. The reports are documented as follows:

**MATHEMATICS**

Table 3.3 shows the content distribution of test items in mathematics across the examinations in 2008 and 2009 years of consideration

 

**Table 3.3: DISTRIBUTION ANALYSIS OF THE PUBLIC EXAMINATIONS OBJECTIVE TEST ITEMS.**

| A | WAEC 2008- 2009 | | NECO 2008 - 2009 | | NABTEB 2008 - 2009 | | JAMB 2008 - 2009 | |
|---|---|---|---|---|---|---|---|---|
| a) Number Bases | 2 | 1 | 1 | 1 | 1 | 1 | 3 | 4 |
| b) Fractions decimals and approximation | 2 | | 3 | 3 | | 1 | 2 | 2 |
| c) Indices | 3 | 3 | 1 | 2 | 1 | 2 | 1 | 1 |
| d) Logarithms | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 |
| e) Sequence | | | 1 | 1 | 3 | 2 | 1 | 1 |
| f) Sets | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 2 |
| h) Positive and Negative integers. Rationales | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| i) Surds | | 1 | 0 | 0 | 0 | 0 | 1 | |
| j) Ratio, Proportion and Rates | 2 | 1 | 3 | 2 | | | 1 | |
| k) Variation | 1 | 1 | 3 | 2 | 2 | 1 | 1 | |
| l) Percentages | | 2 | 2 | 1 | 3 | 1 | 1 | 2 |
| **B. ALGEBRAIC PROC** a) Algebraic Expression | 2 | 2 | 1 | 0 | 1 | 1 | 0 | 1 |
| b) Simple operations on Algebraic expressions. | 5 | 2 | 2 | 3 | 4 | 3 | 0 | 1 |
| c) Solution of linear equations | 1 | 3 | 1 | 3 | 1 | 2 | | |
| d) Change of subject of formula | 2 | 1 | 1 | 2 | 1 | | 4 | 1 |
| e) Quadratic equations | 1 | 1 | 1 | 2 | 1 | 3 | 2 | |
| f) Graphs of linear and quadratic | 2 | | 1 | 1 | 1 | 1 | 1 | |
| g) Linear inequalities | 3 | 3 | 1 | 1 | | | 2 | 3 |
| h) Algebraic Fractions | 1 | 2 | 2 | | 2 | 1 | | |
| **C. MENSURATION** a) Lengths and perimeters | | 2 | 4 | 1 | 5 | 7 | 2 | 0 |
| b) Areas | 1 | 2 | 2 | 1 | 1 | 4 | | |
| c) Volumes | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 |
| **D. PLANE GEOMETRY** a) Angles at a point | 2 | 4 | 1 | 3 | 4 | 1 | 1 | 0 |
| b) Angles and intercepts on parallel lines | 1 | | 1 | 1 | 1 | | | 1 |
| c) Triangles and at her polygons | 5 | 3 | 3 | 2 | 2 | 5 | 1 | 2 |
| d) Circles | 1 | 5 | 1 | 4 | 4 | 1 | | 2 |
| e) Construction | | | 2 | | 1 | 1 | | |
| f) Loci | | | 1 | 1 | | 1 | 2 | 1 |
| **E. TRIGONOMETRY** a) Sine, Cosine and tangier of angle | 1 | 2 | 2 | 2 | 1 | 2 | 2 | 1 |
| b) Angles of elevation and depression | | 2 | 2 | 0 | 0 | 0 | 0 | 1 |
| c) Bearings | 2 | | 1 | 2 | 0 | 0 | 0 | 0 |
| **F. STATISTIC AND PROBABILITY** a) Statistics | 4 | 2 | 9 | 7 | 2 | 3 | 5 | 6 |
| b) Probability | 1 | 3 | 2 | 4 | 4 | 2 | 4 | 4 |
| **G. VECTORS** a) Vectors in plane | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **H. MATRICES AND DETERMINANTS.** | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 2 |
| **I. COORDINATE GEOMETRY** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **J. FALCULUS** | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 2 |
| **TOTAL:-** | 50 | 50 | 60 | 60 | 50 | 50 | 50 | 50 |

**WAEC 2008 and 2009**

Detailed analysis shows that the multiple choice test items are well structured to meet most of the objectives of the syllabus. No ambiguity was observed. On content analysis, the questions are not evenly spread. In 2009, there were no questions on such parts of the syllabus, as A (b), e), B(f), D(b,e,f); E (b,c) as shown in the table attached. For 2008, no question on surd, construction loci, angles of elevation and depression.

**NECO 2008 AND 2009**

Structurally, the questions are very good. The sequence followed the pattern of the topics in the syllabus. On content analysis, some few areas were not included. Such areas include A(h,i); B(a,i); D(e) in 2009 and A(i) in 2008, as shown on table 3.3.

 

**NABTEB 2008 AND 2009**

The structure is good and met most of the objectives of NABTEB Mathematics syllabus. It was observed that items used by NABTEB were different from that of WAEC and NECO. The questions were so designed to meet the technical and trade needs of the candidates. In 2009, questions were not set in A(i,j), B(d,g); D(b), E(b,c), G(a,b). For 2008, no questions in A(b,I,j,k); E(b,c) vectors in a plane is in the syllabus but not effected in the 2 samples years, as shown on the tables 3.3.

**JAMB 2008 and 2009**

The structure of the questions is good and the spread was fairly adequate. Thus, the content of the syllabus were well covered except for the area of shapes. It touched some specifics in most section of the syllabus.

**Research question 4**

How functional are the distractor powers of the test items options in these examinations?

This research question was addressed using Tables 3.4 & 3.5

**Table 3.4: Functionality Coefficient of Multiple Choice Test Items by Examination Body, Subject and Year**

| Subject & Year | | Correlation Coefficients | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | WAEC | % | NECO | % | NABTEB | % | JAMB | % |
| MATHS | 2008 | 11.00 | 5.50 | 44.00 | 14.67 | 10.00 | 5.00 | 1.00 | 0.50 |
| | 2009 | 11.00 | 5.50 | 28.00 | 9.33 | 12.00 | 6.00 | 0.00 | 0.00 |

As shown in Table 3.4, JAMB observed the lowest percentage of items that showed less than 5% functionality coefficient in Mathematics of 2009 with 0.0% and 2008 Mathematics with 0.5%. WAEC and NECO test items options with poor functionality indices when compared with JAMB and NABTEB.

A critical observation at the choices made of the candidates in respect of options available to each multiple choice items in the 32 batteries considered indicated that there were a number of distractor that were not good as none of the testees in the upper and lower one quarter chose them. Table 3.4 shows the distribution of the number of bad distractor by subject, year and examination body.

**Table.3.5: Frequency of Bad Distractor By Subject, Year and Examination Body.**

| Subject & Year | | Number of bad distractor | | | |
|---|---|---|---|---|---|
| | | WAEC | NECO | NABTEB | JAMB |
| MATHS | 2008 | 1 | 6 | 0 | 0 |
| | 2009 | 1 | 2 | 2 | 0 |

Bad distractors are the options that were not chosen by any of the testees in both the lower and upper groups. The frequency of bad distractors should be seen from the perspective of the total number of options developed and used by the various examining bodies. For instance, 1 option out of the 200 used in the 2008 Maths paper by WAEC was bad. These imply an error level of 0.005.

As shown on Table 3.5 JAMB had the lowest frequency of bad distractions. No bad distractor was observed in the 2008 and 2009 Mathematics, NECO recorded the highest number of bad distractors in Mathematics in its 2008 paper, and NECO recorded 2 bad distractors in 2009.

**Research question 5**

What is the reliability coefficient (estimate of Internal Consistency) of each of the tests?

This research question was treated using the result in table 3.6.

**Table 3.6: Reliability Indices By Exam Body, Year and Subject**

| Subject & Year | | Correlation Coefficients | | | |
|---|---|---|---|---|---|
| | | WAEC ( r ) | NECO( r ) | NABTEB( r ) | JAMB( r ) |
| MATHS | 2008 | 0.86 | 0.79 | 0.46 | 0.37 |
| | 2009 | 0.82 | 0.87 | 0.41 | 0.43 |

Reliability is adjudged to be Good when r = 0.80

Acceptable when r is between 0.7 and 0.79

Poor when r < 0.7

Table 3.6 shows that coefficient of internal consistency which is a measure of the degree of reliability of the test batteries varies from year to year and from one examination body to another. On the average WAEC and NECO exams have better indices of reliability in all the subjects and the years under review while NABTEB and JAMB had relatively low coefficients with few outliers. The experience of NABTEB is worst compared with other public examination bodies under studied. NECO's 2009 paper had the highest reliability coefficient, followed by WAEC 2008 and NABTEB had the lowest with 0.41.

**Summary of Findings**

1. It is observed that WAEC, NECO, NABTEB and JAMB developed tests covering not less than 80% of their syllabi content.

 

2. Generally, a fairly high proportion of the test items have appropriate difficulty index i.e within the range of 0.25 – 0.75.
3. Functionality of the distractors is best in JAMB multiple choice items and least in NECO but in general terms, the distractors showed a fairly good combination of options.
4. Nineteen out of thirty-two papers have acceptable reliability index while thirteen were considered not reliable.

**Recommendations**

- ✓ Although more than 80% coverage was observed, the examination bodies should develop and use items in the remaining about 20% of the syllabi content in order to ensure total learning.
- ✓ As much as possible all forms of ambiguity and bias should be avoided in the development of items.
- ✓ For essay tests, attention should be given to higher order skills required of graduates from secondary schools and coverage of most aspects of the syllabi prescribed by examination bodies.
- ✓ Examination bodies should ensure that difficulty and discrimination of items as well as functionality of multiple choice test option fall within the acceptable limits for certification at the school certification level.
- ✓ The present standard of the examinations conducted by the Public Examination Bodies in Nigeria should be sustained and improved upon.

**Conclusion**

This study covered the entire Country in its sampling but focused on Mathematics and the most popular subject among students. The observation shows that the validity of test used by the examining bodies for certification at the secondary school level is fairly good. Other psychometric characteristics of the tests were found to be generally acceptable Areas needing attention were identified and recommendations for improvement made. There is however, the need to extend this exercise to all school subjects and to be carried out more often for the good health of the nation's education at the secondary school level.

**References**

Allen, M.J. and Yen, W.M (2002) Introduction to Measurement Theory. Long Groove, II Wareland Press.

Alonge, E. I. (1989). Measurement and Evaluation in Education and Psychology. Ado-Ekiti: Adebayo Publishing Company (Nig.) Ltd.

Anikweze, C. M. (2009). Simplified Approach to Educational Research. Enugu: SNAAP Press (Nig.) Ltd

Anikweze, C. M. (2010). Measurement and Evaluation for Teacher Education, 2nd Edition. Enugu: SNAAP Press (Nig.) Ltd

Becker, L. A. (1999). Reliability and Validity Part II Black, P. & William, D. (October 1998). "Inside the Black Box: Raising Standards through Classroom Assessment"; Phi Beta Kappan. Retrieved on 23/9/09 from http://www.pdkintl.org/kappan/kbla9810.htm.

Bond, L. A. (1996). Norm- and criterion-referenced testing. Practical Assessment, Research & Evaluation, 5(2). http://PAREonline.net/getvn.asp?v=5&n=2. Retrieved 23/9/ 2009

Chase, C.I. (1998). Measurement for Educational Evaluation. London: Addison - Wesley Publishing Company.

Ebel, R.L. (1991). Essentials of Education Measurement, 5th Edtion. Englewood Cliffs, New Jersey: Prentice Hall Inc.

Harbor-Peters, V.F. (1999). Noteworthy Points in Measurement & Evaluation. Enugu: Snaap Press Ltd.

Lal, J. P. (2007). Educational Measurement and Evaluation. New Delhi: Anmol Publications PVT Ltd.

Marriott, F.H.C. (2002). A Dictionary of Statistical Terms, 5th edition. London: Longman Scientific and Technical for the International Statistical Institute.

Nwana, O.C. (2007). Textbook on Educational Measurement and Evaluation. Owerri: Boma Way Publishers.

Oshkosh, U. W. (2002). Item Analysis, Testing Services Support; http://www.uwosh.edu/testing/facultyinfo/itemanalysis.php

Schnitzer, P. K. (2008). "Psychological Testing." Microsoft® Encarta® 2009 [DVD]. Redmond, WA: Microsoft Corporation.

Stiggins, R.J. (1994). Student-Centered Classroom Assessment. New York: Charles Merrill Publishing Company.

Thissen, D. and Orlando, M. (2005). Item Response Theory for items Second in two Categories. In Tussen, and Wiver, H. (Eds) Test Scoring Mahwah, N.S. Lawrence Erlbaum Associates Inc.

The IISTE is a pioneer in the Open-Access hosting service and academic event management. The aim of the firm is Accelerating Global Knowledge Sharing.

More information about the firm can be found on the homepage:
http://www.iiste.org

## CALL FOR JOURNAL PAPERS

There are more than 30 peer-reviewed academic journals hosted under the hosting platform.

**Prospective authors of journals can find the submission instruction on the following page:** http://www.iiste.org/journals/ All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Paper version of the journals is also available upon request of readers and authors.

## MORE RESOURCES

Book publication information: http://www.iiste.org/book/

Academic conference: http://www.iiste.org/conference/upcoming-conferences-call-for-paper/

### IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digtial Library , NewJour, Google Scholar