

## VALIDITY AND RELIABILITY OF 2017 SENIOR SCHOOL CERTIFICATE MATHEMATICS ITEMS AMONG STUDENTS IN ILORIN, NIGERIA

---

<sup>1</sup>OLADELE, J. I., <sup>2</sup>ADISHI, G. O., & <sup>3</sup>OWOLABI, H. O.

<sup>1&2</sup>Department of Social Sciences Education, University of Ilorin

<sup>3</sup>Department of Adults and Primary Education, University of Ilorin  
[gpadishi@gmail.com](mailto:gpadishi@gmail.com)

---

### Abstract

*This study analyzed the validity and reliability of Senior School Certificate Examination (SSCE) Multiple choice and Essay items in Mathematics. The study adopted a descriptive research design of survey type. The population comprised all Public Senior Secondary School students in Ilorin. Stratified random sampling technique was used to stratify the three local government areas. Simple random sampling technique was used to select 600 SSS 3 students from 6 public senior secondary schools. Purposive sampling was used to select SS 3 students from the selected schools. The instruments used were 2017 WAEC, NECO and NABTEB Mathematics items. The tests were critically examined to ascertain the content validity with respect to syllabus coverage and the levels of thinking of the items. Item analyses were carried out to calculate the content validity levels and reliability levels. The results revealed that the items of WAEC, NECO and NABTEB are in relation to the NERDC Curriculum. The levels of thinking of the items are suitable for the test takers. The results revealed that the items of the examinations were valid with NECO items having the highest content validity level in terms of syllabus coverage. NECO, WAEC and NABTEB multiple choice items were also reliable with reliability levels of 0.89, 0.96, 0.96 respectively. NABTEB essay items had the highest reliability level of 0.76. It was concluded that the tests are reliable and valid. Based on these findings, it was recommended that WAEC and NABTEB should increase the number of multiple-choice questions so as to cover a larger scope of their syllabus. The feedback from an analysis made on tests items should be made available to the item writers so that those areas making the items less valid and reliable will be improved upon.*

**Key words:** Content analysis, Validity, Reliability, Testees, Mathematics items.

### Introduction

The value and functionality of any educational system lie in its ability to actualize the goals of education. Evaluation, therefore, is useful in ascertaining the level of achievement of the objectives of an educational programme by learners, or determining the worth, quality or quantity of something in such a way that valid, unbiased and value judgement is given at the end of the process. According to Manga & Mangal (2008), Evaluation in the educational sector is a systematic process of

determining the extent to which instructional objectives are achieved by students. It includes both qualitative and quantitative descriptions of students' behaviour and value judgment concerning the adequacy and desirability of that behaviour. To achieve the stated educational objectives, learners must be evaluated adequately. A core aspect of evaluation that poses great concern is the content of the measuring tool. The content of the measuring tool is a vital tool in evaluation as it provides an opportunity for ascertaining the presence or measuring the levels of certain traits and characteristics of the testees.

It is worthy of mention that education is balanced when it is hinged on the three domains of learning. Bloom's taxonomy is a set of three hierarchical models used to classify educational learning objectives into levels of complexity and specificity. The three lists cover the learning objectives in cognitive, affective and sensory domains. The cognitive domain list has been the primary focus of most traditional education and is frequently used to structure curriculum learning objectives, assessments and activities (Bloom, Engelhart, Furst, Hill & Krathwohl, 1956).

According to Wass, Van der Vleuten, Shatzer and Jones (2001), test blueprint is a guide used for creating a balanced test items in an instrument which consists of a list of topics (with specified weight for each) that should be tested on an examination. The first step for writing test items is to have a blueprint (table of specifications). Blueprinting is the planning of the test against the learning objectives of a course or competencies essential to a specialty. According to Offor (2001), the purpose of examination is to ascertain the extent, to which students have achieved educational objectives in a particular course. Onwuka (1981) also opined that Educational objectives are expressed in terms of knowledge (cognitive domain), attitude (affective domain) and practical/motor skills (psychomotor domain). It becomes needful to examine the contents of tests, in terms of test items included, to see if they perform the role(s) for which they are intended. The constituent items of a test need proper scrutiny before administering them to the test takers.

A good test should be a fair representation of the content of the subject matter being tested. The test should cover all areas of the subject matter in the right weightings and proportions. Test could be in objective or essay form. Multiple choice test type is a typical example of objective tests. It is regarded as the most applicable, flexible and useful type of objective test items. They require students to select the answer from a number of possible alternatives. Multiple choice tests items are widely acclaimed as most reliable because of consistency in scoring the test as well as its fairness to all students (Osunde, 2009). On the other hand, essay tests normally require the writing of a series of logically arranged sentences on a topic. It affords test takers the opportunity of displaying their overall understanding of a topic and demonstrate their ability to think critically, organize their thoughts, and be creative and original. Moyinoluwa (2015) suggested that in designing 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.

Test contents of an item refer to the topical aspects that are contained in a test. It has to do with the extent to which a given test covers the list of topics in a particular subject matter and the levels of thinking of those individual items in the examination testing tool. Test contents play a pivotal role in the performance of the examinees. This therefore brings to bear, the importance of the validity and reliability of the test items. For this to be maintained and re-enforced, the weightings, emphasis on certain topical areas, distribution of items amongst the topics with respect to its relevance and quantity must be given topmost priority. Psychometric characteristics of a test refer to certain attributes inherent in tests upon which an assessment of examinees is based. These characteristics include the difficulty indices, the discrimination index, the power of distracters, validity and reliability indices. Reliability and validity are the two main psychometric properties a test must possess for it to be adjudged as standardized. According to Enyi (2002), if a test possesses other qualities but lacks validity, the test will be considered as not being useful.

Reliability of a test refers to the degree to which a test measures accurately and consistently yielding comparable results when administered a number of times (Akuezuilo&Agu, 2003). Establishing validity from the point of psychometric principles, on the other hand entails making the warrant explicit, examining the network of beliefs and theories on which it relies and testing its strength and credibility through various sources of backing. It requires determining conditions that weaken the warrant, exploring alternative explanation for good or poor performance and feeding them back into the system to reduce inferential errors (Mislevy, Wilson, Ercikan&Chudowsky, 2001). It is worthy of mention that the validity and reliability of test items can be ascertained through item analysis among others. According to Oshkosh (2005), item analysis is 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.

There are different evaluation agencies in Nigeria; the main ones being the West African Examinations Council (WAEC), the National Examination Council (NECO) and the National Business and Technical Examination Board (NABTEB). According to Nworgwu (1992), these evaluation agencies were set up to promote education, co-ordinate educational programmes, and to control and monitor the quality of education in the institutions of learning, with the essence of organizing public examinations in order to provide uniform standards to all examinees, irrespective of the type or method of instruction they have received.

Mathematics is a useful tool in the society, more so in the present technology age. No wonder Mathematics is a compulsory subject at primary and secondary school levels, though, not all the students are expected to become Mathematicians, but because of its application in everyday life (Oladele, 2004). Obodo, (1997) recognized that the persistent poor performances of students in mathematics especially at secondary levels of education (junior and senior) and the continued low enrolment of students in mathematics and science related disciplines in tertiary education have become sources of worry to both students and parents in this country. A recent analysis

presented by Anaduaka and Okafor, (2013) showed that students' performance has not yet significantly improved. The percentage of students that passed the subject at credit level from 2007 – 2011 still fell between 30 and 47 percent except in 2008 that the percentage got up to approximately 57%.

Test contents of equivalent examining bodies like WAEC, NECO and NABTEB have been studied, to ascertain if their test items can be assumed to be truly equivalent or the test items of one is more valid or reliable. This has stimulated the interests of stakeholders to a certain examining body. Moyinoluwa (2015) concluded that NABTEB 2008 and 2009 Mathematics questions have good structure and met most of the objectives of NABTEB Mathematics syllabus. Moyinoluwa (2015) also noted that WAEC 2008 and 2009 multiple choice test items are well structured to meet most of the objectives of the syllabus with no ambiguity and on content analysis, he noted that the questions are not evenly spread.

Adeniran (2000) concluded that NECO is inferior to WAEC in all standards; this result contrasted sharply with the findings of Moyinoluwa (2015) who employed the internal consistency method in estimating the reliability levels of NABTEB, NECO, JAMB and WAEC Mathematics papers and concluded that NECO's 2009 paper had the highest reliability coefficient. Bandele and Adewale (2013) submitted that WAEC, NECO and NABTEB Mathematics Achievement Examinations are highly reliable and valid. Kolawole (2007) argued that there is no significant difference in the reliability of NECO and WAEC multiple choice items in Mathematics.

All these research works have created a researchable gap, part of which the study intends to fill and also to add to the existing knowledge which is one of the goals of research by carrying out a research on Validity and Reliability of Senior School Certificate Mathematics Items.

### **Research Questions**

This study was sought to answer the following research questions with respect to the purpose of this study.

- a. What is the validity of the test contents of WAEC, NECO and NABTEB Mathematics papers (Multiple-choice and Essay tests)?
- b. What are the content validity levels of WAEC, NECO and NABTEB Mathematics papers (Multiple-choice and Essay tests)?
- c. What are the reliability indices of WAEC, NECO and NABTEB Mathematics papers (Multiple-choice and Essay tests)?

### **Methodology**

The research design for this study was the descriptive research design of survey type. The descriptive research design was adopted because it suited this study as data were collected through the test items and no variable was manipulated. This conforms to the definition given by Daramola (2006), that descriptive research design



describes the characteristics of a given population or areas of interest without manipulation.

The population for this study comprised all Public Senior Secondary School students in Ilorin. However, the target population for this study was all the 2017/2018SSS 3 students in Public Senior Secondary Schools in Ilorin. The population was divided into three strata, which are the three local government areas in Ilorin. The stratified random technique was used so as to have a sample that is highly representative of the population being studied. There are 79 public secondary schools in Ilorin metropolis. The three local government areas in Ilorin are Ilorin East Local Government Area with 30 public senior secondary schools, Ilorin South Local Government Area with 21 public senior secondary schools and Ilorin West Local Government Area with 28 public senior secondary schools (Kwara State Ministry of Education, 2017). From each local government area, the researcher then selected eight public senior secondary schools which made a total of 24 senior secondary schools that participated in the study.

SSS 3 students were chosen from each school using the Purposive random sampling technique. This technique was used because the students were expected to have covered a major part of the School Certificate Mathematics Syllabus. Their status placed them in the best position to respond to the tests prepared for this research work since they were preparing for their SSCE Mathematics examinations. The Simple Random Technique was used to select twenty five (25) students in each school. This brought the sample for the study to a total of six hundred (600) participants.

The instruments used for data collection in this study were WAEC, NECO and NABTEB 2017 Mathematics papers. These three Mathematics Papers were adopted in their original form. WAEC 2017 Mathematics papers have 50 test items with four options each in the multiple-choice test. The essay test is divided into two sections (Sections A and B). NECO 2017 Mathematics papers have 60 test items with five options each in the objective test. The essay test (Paper II) is divided into two parts (Parts I and II). NABTEB 2017 Mathematics papers have 50 test items with four options each in the objective test. The essay test (Paper II) is divided into three sections, viz: Sections A, B and C.

The statistic used was descriptive. Item analysis was carried out to ascertain the validity and reliability of the test items. The test contents were critically examined to ascertain the content validity with respect to syllabus coverage and levels of thinking of all the test items by the researcher.

The responses of the respondents were sub-divided into three levels; “zero” for low performance, “one” for average performance and “two” for high performance. Analysis of the objective tests was done on two levels of performance. “Zero” was awarded to testees who did not get a test item right and “one” was awarded to testees who got a test item right. With these levels of scores, the reliability index was calculated. Kuder-Richardson and Cronbach's alpha were used to determine the reliability coefficient of the Objective test items and Essay test items respectively. The

validity of the test items was done based on the level of syllabus coverage of the individual examining bodies. The responses of the respondents were computed using the Microsoft Excel Package.

## Results

**Research Question One:** What is the validity of the test contents of WAEC, NECO and NABTEB Mathematics papers (Objective and Essay tests)?

The core themes of the Nigerian Educational Research and Development Council (NERDC) curriculum in Mathematics are Numerical processes, Algebraic processes, Geometry and mensuration, Geometry and trigonometry, Statistics, Probability and Introductory calculus. These topics are necessary foundations for WAEC, NECO and NABTEB examinations syllabus.

The topics were written against the levels of thinking (Remembering, Understanding, Applying, Analyzing, Evaluating and Creating) in the Bloom's taxonomy revised by Anderson & Krathwohl (2001). The breakdown of the topics under these core themes are appended to this research work.

**Table 1:** Levels of thinking of the Multiple-choice Test Items in 2017 WAEC Mathematics

S/N	Contents	Remembering	Understanding	Applying	Analysing	Evaluating	Creating	Total
1	Number and numeration		1,6	4,5,10,33,38,43,46	3,7,34			12
2	Algebraic processes		42,49	11,8,9,2,12,41	13,37		40	11
3	Geometry and mensuration	18		14,15,16,17,35,36,39,44,48,50				11
4	Geometry and trigonometry		25	23,24,19,20,21,22,27	26			9
5	Statistics		29	28,30,31,47				5
6	Probability			32,45				2
7	Introductory calculus							0
<b>Total</b>		1 (2%)	6 (12%)	36 (72%)	6 (12%)	0 (0%)	1 (2%)	50

Table 1 shows the level of thinking of the objective test items in 2017 WAEC Mathematics. From table 1, it can be observed that out of the 50 items, 1 (2%) item fell under remembering; 6 (12%) items, understanding; 36 (72%) items, applying; 6 (12%) items, analysing; 0 (0%) item, evaluating; 1 (2%) item, creating. The results revealed that most of the items fell under applying while no item fell under evaluating.

**Table 2:** Levels of thinking of the Essay Test Items in 2017 WAEC Mathematics

S/N	Contents	Remembering	Understanding	Applying	Analysing	Evaluating	Creating	Total
1	Number and numeration			1a,4,5b,6b, 11a,12b, 13b		6a	5a	9
2	Algebraic processes			1b,2a,7a,7c, 12a,13aii	13ai		7b	8
3	Geometry and mensuration			3b,9aii,9b,1 1b	9ai			5
4	Geometry and trigonometry			2b,3a,10b	10a,			4
5	Statistics			8a,8bi,				2
6	Probability			5c,8bii,				2
7	Introductory calculus							0
<b>Total</b>		0 (0%)	0 (0%)	24 (80%)	3 (10%)	1 (3.3%)	2 (6.7%)	30

Table 2 shows the level of thinking of the essay test items in 2017 WAEC Mathematics. From table 2, it can be observed that out of the 30 items, 0 (0%) item fell under remembering; 0 (0%) item, understanding; 24 (80%) items, applying; 3 (10%) items, analysing; 1 (3.3%) item, evaluating; 2 (6.7%) items, creating. The results revealed that most of the items fell under applying while no item fell under remembering and understanding.

**Table 3:** Levels of thinking of the Multiple-choice Test Items in 2017 NECO Mathematics

S/N	Contents	Remembering	Understanding	Applying	Analysing	Evaluating	Creating	Total
1	Number and numeration			5,7,8,10,12,13, 14,15,18	1,2,3,4,6,1 1			15
2	Algebraic processes		26,27,28,29	9,19,21,22,25,3 0,33	20		24	13
3	Geometry and mensuration			17,37,38,39,40, 41,42,43,44, 46				10
4	Geometry and trigonometry			23,31,34,35,36, 45,47,48			32	9
5	Statistics			16,49,50,51,52, 54		53		7
6	Probability			55,56,57				3
7	Introductory calculus			58,59,60				3
<b>Total</b>		0 (0%)	4 (6.7%)	46 (76.7%)	7 (11.7%)	1 (1.7%)	2 (3.3%)	60

Table 3 shows the level of thinking of the objective test items in 2017 NECO Mathematics. From table 3, it can be observed that out of the 60 items in NECO Mathematics, 0 (0%) item fell under remembering; 4 (6.7%) items, understanding; 46 (76.7%) items, applying; 1 (1.7%) item, analysing; 1 (1.7%) item, evaluating; 2 (3.3%) items, creating. The results revealed that most of the items fell under applying while no item fell under remembering.

**Table 4:** Levels of thinking of the Essay Test Items in 2017 NECO Mathematics

S/N	Contents	Remembering	Understanding	Applying	Analysing	Evaluating	Creating	Total
1	Numerical processes			3a,3b,6,7a,7b,9b	1			7
2	Algebraic processes			9a,11a,11c			2,11b	5
3	Geometry and mensuration			4,10a,10b				3
4	Geometry and trigonometry							0
5	Statistics			5,12a,12b				3
6	Probability							0
7	Introductory calculus				7c,8a		8b	3
	<b>Total</b>	0 (0%)	0 (0%)	15 (71.4%)	3 (14.3%)	0 (0%)	3 (14.3%)	21

Table 4 shows the level of thinking of the essay test items in 2017 NECO Mathematics. From table 4, it can be observed that out of the 21 items, 0 (0%) item fell under remembering; 0 (0%) item, understanding; 15 (71.4%) items, applying; 3 (14.3%) items, analysing; 0 (0%) item, evaluating; 3 (14.3%) items, creating. The results revealed that most of the items fell under applying while no item fell under remember, understanding and evaluating.

**Table 5:** Levels of thinking of the Multiple-choice Test Items in 2017 NABTEB Mathematics

S/N	Contents	Remembering	Understanding	Applying	Analysing	Evaluating	Creating	Total
1	Numerical processes	21,25	1,9,24	3,12,16,22,30,38,41,43	19,40,49			16
2	Algebraic processes	15,26	11	2,10,13,14,20,42,44,45,46,48				13
3	Geometry and mensuration	37	4	8,23,27,28,31,32,50				9
4	Geometry and trigonometry			5,6,7,29,47	39			6
5	Statistics	34	35,36	33				4
6	Probability			17,18				2
7	Introductory calculus							0
	<b>Total</b>	6 (12%)	7 (14%)	33 (66%)	4 (8%)	0 (0%)	0 (0%)	50

Table 5 shows the level of thinking of the objective test items in 2017 NABTEB Mathematics. From table 5, it can be observed that out of the 50 items, 6 (12%) item fell under remembering; 7 (14%) item, understanding; 33 (66%) items, applying; 4 (8%) items, analysing; 0 (0%) item, evaluating; 0 (0%) item, creating. The results revealed that most of the items fell under applying while no item fell under evaluating and creating.



**Table 6:** Levels of thinking of the Essay Test Items in 2017 NABTEB Mathematics

S/N	Contents	Remembering	Understanding	Applying	Analysing	Evaluating	Creating	Total
1	Numerical processes			2a,2b,3a,7a,7b,8a	1a,3b,8b			9
2	Algebraic processes			1b	5a,5b			3
3	Geometry and mensuration			4,9a,11			6	4
4	Geometry and trigonometry			9b,10b				2
5	Statistics							0
6	Probability			10a				1
7	Introductory calculus							0
	Total	0 (0%)	0 (0%)	13 (68.4%)	5 (26.3%)	0 (0%)	1 (5.3%)	19

Table 6 shows the level of thinking of the essay test items in 2017 NABTEB Mathematics. From table 6, it can be observed that out of the 19 items, 0 (0%) item fell under remembering; 0 (0%) item, understanding; 13 (68.4%) items, applying; 5 (26.3%) items, analysing; 0 (0%) item, evaluating; 1 (5.3%) item, creating. The results revealed that most of the items fell under applying while no item fell under evaluating and creating.

**Research Question Two:** What are the content validity levels of WAEC, NECO and NABTEB Mathematics papers (Objective and Essay tests)?

The content validity levels of WAEC, NECO and NABTEB Mathematics papers was ascertained by examining the topical contents of the tests and comparing them with the Mathematics syllabi of the three examining bodies.

**Table 7:** Number of topics covered in WAEC Syllabus from the questions in 2017 WAEC Mathematics papers (Objective and Essay tests).

Examining body	Number of topics covered	Total number of topics	Percentage Coverage
WAEC	28	36	77.78%

Table 7 reveals the number of topics covered in WAEC Syllabus from the questions in 2017 WAEC Mathematics (Objective and Essay tests). It can be seen from the table that there are thirty six (36) topics outlined in the syllabus. From these topics, twenty eight (28) topics were covered in the exam. This brings the percentage coverage of the topics in the syllabus to be 77.78%.

**Table 8:** Number of topics covered in NECO Syllabus from the questions in 2017 NECO Mathematics papers (Objective and Essay tests).

Examining body	Number of topics covered	Total number of topics	Percentage Coverage
NECO	25	30	83.3%

Table 8 reveals the number of topics covered in NECO Syllabus from the questions in 2017 NECO Mathematics (Objective and Essay tests). It can be seen from the table that there are thirty (30) topics outlined in the syllabus. From these topics, twenty five

(25) topics were covered in the exam. This brings the percentage coverage of the topics in the syllabus to be 83.3%.

**Table 9:** Number of topics covered in NABTEB Syllabus from the questions in 2017 NABTEB Mathematics papers (Objective and Essay tests).

Examining body	Number of topics covered	Total number of topics	Percentage Coverage
NABTEB	32	44	72.73%

Table 9 shows the number of topics covered in NABTEB Syllabus from the questions in 2017 NABTEB Mathematics (Objective and Essay tests). It can be seen from the table that there are forty four (44) topics outlined in the syllabus. From these topics, thirty two (32) topics were covered in the exam. This brings the percentage coverage of the topics in the syllabus to 72.73%.

**Research Question Three:** *What are the reliability indices of WAEC, NECO and NABTEB Mathematics papers (Objective and Essay tests)?*

Reliability of each WAEC, NECO and NABTEB Mathematics Objective and Essay papers was carried out using the Kuder-Richardson Formula 20 (KR-20) and Cronbach's alpha respectively. The reliability obtained was then summarized and presented in Tables 10 and 11.

**Table 10:** Reliability of Mathematics Multiple-choice Test (WAEC, NECO & NABTEB).

Examination	Number of items	Reliability
WAEC	50	0.96
NECO	60	0.89
NABTEB	50	0.96

Table 10 shows the comparison of the reliability indices of 2017 WAEC, NECO and NABTEB Mathematics papers (Objective test). From table 10, it can be observed that WAEC Mathematics paper had a reliability index of 0.96; NECO Mathematics paper had a reliability index of 0.89 while NABTEB Mathematics paper had a reliability index of 0.96. The results revealed that WAEC and NABTEB Mathematics had equal reliability index with NECO Mathematics being the least of them.

**Table 11:** Reliability of Mathematics Essay Test (WAEC, NECO & NABTEB).

Examination	Number of items	Reliability
WAEC	13	0.61
NECO	12	0.71
NABTEB	11	0.76

Table 11 shows the comparison of the reliability indices of 2017 WAEC, NECO and NABTEB Mathematics papers (Essay tests). From the table, it can be observed that WAEC Mathematics paper had a reliability index of 0.61; NECO Mathematics paper had a reliability index of 0.71 while NABTEB Mathematics paper had a reliability index of 0.76. The results showed that among these three examinations, NABTEB Mathematics paper had the highest reliability index of 0.76.

### Discussion of the Findings

From the findings of this study, the levels of thinking of most of the items in WAEC, NECO and NABTEB Mathematics fell on 'Applying' in the revised Bloom's taxonomy. This shows that most of the items are application-based which is most suitable for Olevel external examinations.

It was discovered from this study that the content validity level of WAEC Mathematics (Objective and Essay tests) in terms of syllabus coverage was 77.78%; NECO was 83.3% while NABTEB was 72.73%. From the foregoing, NECO had the highest content validity level in terms of syllabus coverage. This may be due to the fact that NECO has the highest number of items in the objective test compared to WAEC and NECO. From this finding, only the content validity in terms of syllabus coverage of NECO agreed with the findings of Moyinoluwa (2015) who opined that WAEC, NECO, NABTEB and JAMB had not less than 80% of their syllabi content. The findings of this work on the level of coverage of NABTEB Mathematics (Objective and Essay tests) of the syllabus agreed with the findings of Moyinoluwa (2015) who concluded that NABTEB 2008 and 2009 Mathematics questions have good structure and met most of the objectives of NABTEB Mathematics syllabus. The findings of this work on syllabus coverage of WAEC also agree with the findings of Moyinoluwa (2015) who noted that WAEC 2008 and 2009 multiple choice test items are well structured to meet most of the objectives of the syllabus with no ambiguity.

This study also revealed that the reliabilities of the three examinations were high with WAEC and NABTEB having same reliability coefficient of 0.96; and NECO with a reliability coefficient of 0.89. This result, however for WAEC, may be as a result of the longer number of years in test development and administration. The researcher used the quality criterion suggested by Cohen, Manion and Morrison (2008) which suggests that reliability level is acceptable if it is 0.67 to interpret the reliability obtained. This finding varied with the findings of Ahmed (2015) who opined that NECO had the highest reliability coefficient amongst the three

examination bodies. The result of NECO having a high reliability coefficient of 0.89 also varied with the submission of Adeniran (2000) who concluded that NECO is inferior to WAEC in all standards. NECO cannot therefore, be adjudged substandard since it equally has a high reliability coefficient. The result of this research work agrees with the findings of Bandele and Adewale (2013) who submitted that WAEC, NECO and NABTEB Mathematics Achievement Examinations are highly reliable and valid.

The reliability levels of WAEC, NECO and NABTEB Essay Tests are 0.61, 0.71 and 0.76 respectively. NABTEB had the highest reliability levels of 0.76 amongst the three examinations. This may be due to the fact that NABTEB has been trying to maintain the same level of equivalence with WAEC and NECO for it to be continuously relevant. However, no research work known to the researcher has been done on the reliability levels of essay tests of WAEC, NECO and NABTEB.

### **Conclusion**

The content validity of these three tests is with no doubt as the outcome of the study showed that the three examination bodies have used tests that would yield valid and reliable data. Amongst the three examination bodies, NECO had the highest validity in terms of syllabus coverage. NABTEB had one of the highest reliabilities in the objective test and the highest reliability in the essay test. The results show that no one examination has it all in terms of having the highest item indices or being more valid and reliable.

### **Recommendations**

Owing to the findings from this study, it is hereby recommended that the three examination bodies should intensify their efforts on syllabus coverage as this will boost the confidence of stakeholders especially the test takers in the syllabus of the three examination bodies. WAEC and NABTEB can increase the number of objective test items from fifty (50) to sixty (60) so as to allow for more coverage of the syllabus. Test developers should improve on the modalities of setting essay tests as the reliability levels of the essay test items is relatively lower compared to the multiple-choice test items as revealed in this study.



## References

- Ahmed, T. A. (2015). Comparative analysis of psychometric properties of multiple-choice Test items in senior secondary school certificate examination in economics. *Unpublished M.Ed. Thesis, Department of Social Sciences, Faculty of Education, University of Ilorin, Ilorin.*
- Akuezuilo, E. O. & Agu, N. (2003). *Research and statistics in education and social sciences, methods and applications*. Awka: Nuelcenti Publishers & Academic Press Ltd.
- Anaduaka, U. S. & Okafor, C. F. (2013). Consistent poor performance of Nigerian students in mathematics in senior secondary certificate examination (SSCE): What is not working? *Journal of Research in National Development* 11 (1).
- Anderson, L. W.; Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.
- Bandeale, S. O. & Adewale, A. E. (2013). Comparative analysis of the item difficulty levels of WAEC, NECO and NABTEB mathematics achievement examinations. *Mediterranean Journal of Social Sciences*, 4 (2), 761-765.
- Bloom, B. S.; Engelhart, M. D.; Furst, E. J.; Hill, W. H. & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. New York: David McKay Company.
- Cohen, L.; Manion, L. & Morrison, K. (2008). *Research methods in education* (6th ed.). London & New York: Routledge Taylor & Francis Group. 133 – 164.
- Daramola, S.O. (2006). *Research and statistical methods in education for students and researchers in tertiary institution*. Ilorin: Bamitex Printing and Publishing.
- Enyi, G. S. (2002). *Introduction to measurement and evaluation for universities, polytechnics and colleges of education: A Professional Approach*. Enugu: Hamson Pal.
- Kolawole, E. B. (2007). A comparative analysis of psychometric properties of two Nigerian examining bodies for senior secondary schools mathematics. *Research Journal of Applied Science*. 2 (8): 913-915.
- Mahmud, J. A. (2012). Analysis of psychometric characteristics of economics objective test items in joint promotion examination, kwarastate. *Unpublished M.Ed. Thesis, Department of Social Sciences, Faculty of Education, University of Ilorin, Ilorin.*
- Mangal, S. K. & Mangal, U. (2008). *Teaching of social studies*. Asoke K. Ghosh publishers.
- Mislevy, R. J., Wilson, M.R.; Ercikan, K.; & Chudowsky, N. (2001). *Psychometric principles in student assessment*. Dordrecht, the Netherlands: Kluwer Academic Press.
- Moyinoluwa, T.D. (2015). Analysing the psychometric properties of mathematics in public examinations in Nigeria. *Research on Humanities and Social Sciences* 5 (7). 2015.

- Nworgu, B. G. (1992). *Educational measurement and evaluation: Theory and practice*. Nsukka. Hallman Publishers.
- Obodo, G. C. (1997). *Principles and practice of mathematics education in Nigeria*. Enugu General Studies Division, ESUT.
- Offor, E. I. (2001). Content validity of senior secondary school certificate examination (SSCE) chemistry. Faculty of Education, Imo State University, Owerri. *Unpublished M.Ed Thesis*.
- Oladele, O. (2004). Improving the teaching and learning of mathematics in secondary schools. *Oyo State Mathematics Conference Paper*.
- Onwuka, U. (1981). *Curriculum development for Africa*. Africana FEP Publishers Ltd. Nigeria.
- Oshkosh, U. W. (2005). Item analysis, testing services support; <http://www.uwosh.edu/testing/facultyinfo/itemanalysis>
- Osunde, A. (2009). Essay and multiple choice tests: bridging the gap. *Workshop papers on multiple choice test item writing procedures for academic staff*, University of Ilorin, Ilorin, on Monday 4th Monday, 2009. pp. 14 - 24.
- Wass, V.; Van der Vleuten, C.; Shatzer, J. & Jones, R. (2001). *Assessment of clinical competence*. Lancet; 357 (9260): 945–9.