



Analysis of the Psychometric Properties (Internal Consistency and Stability Coefficients) of WAEC and NECO Economics Multiple Choice Questions

Nwaneri, Zita Chika & Ursula. N. Akanwa

Department of Science Education, Michael Okpara University of Agriculture, Umudike

Corresponding Author: raphaelbucas@gmail.com,

Abstract

This study focused on the psychometric properties (internal consistency and stability) of the WAEC and NECO Economics multiple choice questions. The study seeks to determine which of the two examination bodies is conducting examination with better psychometric properties in Economics. The study was guided by two research questions and two null hypotheses. The ex-post facto research design was adopted. The sample for the study consisted 850 students drawn from a population of 8,213 Senior Secondary School Three students from public schools who enrolled for the External Senior School Certificate Examination conducted by the West African Examination Council and National Examination Council in 2017/2018 to 2019/2020 Academic Sessions in Abia State. The instrument for data collection were the Economics Multiple Choice past questions on the West African Examination Council and National Examination Council between 2017/2018 and 2019/2020. The data used were collected through direct administration of the instrument to the respondents by the researcher. The item analysis was based on IRT. The findings of the study revealed that the Economics multiple choice examination questions constructed by WAEC between 2017/2018-2019/2020 academic sessions had a significantly better internal consistency and stability, when compared to the Economics multiple choice examination questions constructed by NECO between 2017/2018 and 2019/2020 academic sessions. It was thus recommended that examination bodies like West African Examination Council, National Examination Council, teachers, and psychometricians should be encouraged to adopt IRT in developing test items used in measuring students' ability especially in Economics since it was designed to overcome the limitations of the Classical Test Theory (CTT).

Key Words: Psychometric Properties, Internal Consistency, Stability, West African Examination Council, National Examination Council, Economics, Multiple Choice Questions.

Introduction

Education is an instrument for societal development. It remains the tool for sustainable growth and development of any society. Education is the process of facilitating learning or the acquisition of knowledge, skills, values, belief and habits. Education can take place in formal and informal settings. The extent to which the students have acquired education in the formal setting could be determined through test which may be formative and summative evaluation. The development of human and material resources for rapid development of any nation depends on the availability and quality of education a nation provides for its citizenry. Thus, educational institutions which provide education for the citizens of the nation are the hub for the training of human beings towards achieving the goals of socio-economic development in the nation (Emeikwu, 2011).

A sound education system is the bedrock of development. The quality of education is measured by assessment of its process and products. Bandele and Adewale (2013) pointed out that assessment is an integral part of the education system which includes internal (school) and external (public) examinations and continuous assessment. One of the most commonly used assessment tools is the test (Monday, 2012). Tests have always been an important part of the school system that even the habitual absentees normally turn up in school and present themselves for tests or examinations. Testing has been fully accepted in most modern societies as the most objective means of obtaining information for decision making in schools, industries and government establishments. Therefore, tests are used for admission, recruitment, promotion, placement, evaluation, guidance,

research, and teaching purposes among others (Emeikwu, 2011).

Hence, test is an important instrument for measuring assessment. Test may be called a tool which is used to measure a particular characteristic of an individual or a group of individuals. Nkemakolam (2009) defines test as an instrument for systematic measure of a sample of behaviour or a systematic procedure for observing a person's behaviour and describing it with the aid of a numerical scale or category system. Test can also be defined as an instrument administered to the testees for determining some previously identified objectives in the individual (Ukwuije, 2013).

Test in line with the mode of preparation can be categorized into two types which are teacher-made achievement test and standardized achievement test. The teacher-made test according to Ukwuije (2012), are tests constructed by the classroom teacher to assess pupils/students and can be either of essay (short or extended) or the objective format (true/false, multiple choice type, completion, arrangement and matching) formats. Nkemakolam (2009) argued that these teacher-made tests do not possess high psychometric properties of validity, reliability, standard error of measurement, item difficulties, distractor effectiveness and discriminatory index. The items or questions are such that tests or examinations tend to focus mainly on the part of the curriculum or course content of the subjects that have been covered by the teacher with the pupils/students. Standardized tests are tests constructed or written by specialists or test construction experts known as psychometricians (Nkemakolam, 2003). These tests have norm, manuals and cover wide content areas. According to Nkemakolam (2009), a standardized test is a test that is administered and scored in a consistent or "standard", manner.

Furthermore, standardized tests possess psychometric properties such as validity, reliability, standard error of measurement, item difficulty, item discrimination index (Ukwuije, 2012). These are tests that are constructed and administered by Professional and Government bodies such as Joint Admissions and Matriculation Board (JAMB), West African Examination Council (WAEC) National Examination Council (NECO), SAT, UTME, National Business Technical Examination Board (NABTEB) and State Ministries of Education (that administer State Common Graduating/Entrance Examination) among others (Monday, 2012). In addition, these standardized test items are believed to be of high technical quality; largely because they have been developed by experts, tried out experimentally (that is

pre-tested) and selected on the basis of item difficulty, discriminating power and relationship to a clearly defined and rigid set of specification.

Tests, whether they are teacher-made or standardized tests are classified as essay and objective test. According to Joshua (2012), essay tests are tests that call on the test takers to supply the answers, arrange and organize their answers in a way that makes it run into pages. Essay tests are the test that measure the non-structured types of learning such as creative writing, critical thinking, problem solving, imaginative and organizational ability (Anikweze, 2010). Objective tests are tests that require very short or brief answers (sometimes just a letter of the alphabet) that are very specific and precise (Joshua, 2012). The two main types of objective tests are the recall and recognition (Calmorin, 1994). The recall types are categorized as simple recall and completion. The recognition type is categorized as alternate response, multiple choice, matching type, rearrangement type, analogy and identification. The Objective test is the most popular type of test used in most parts of the world since education became liberalized and school enrollment started to experience explosion (Anikweze, 2010).

There are different varieties of objective tests which are the supply items, selection items and rank order items; but the most popular type is the selection items which consist of the true/false, matching, multiple choice items, pictorial and completion. According to Joshua (2012), the most popular type in the secondary school is the multiple choice question which comprises best answers, correct answer, comparison variety and negative variety.

The present study seeks to explore psychometric properties, particularly the internal consistency and stability coefficients of West African Examination Council (WAEC) and National Examination Council (NECO) which are responsible for the award of Senior School Certificate since they conduct standardized tests. For the purpose of the present study, the focus is to compare the psychometric properties (internal consistency and stability coefficients) of the Economics Multiple Choice items conducted by WAEC and NECO.

Generally, Economics has been defined as the science which studies human behaviour in relationship between ends and scarce means which have alternative use (Osadebe, 2014). Economics is a social science subject which students offer at the senior secondary level, as one of the electives or group of subjects expected to be studied at the Senior Secondary School (SSS) level under the new National Policy on Education (FRN, 2013).

During these three years in the Senior Secondary School, the students are taught by their Economics teachers who are meant to cover the Senior School Certificate Examination Curriculum, reflecting these objectives of Economics. Hence, the students are examined at every level of the Senior Secondary School with the teacher made test, until the end of the third year when they are finally assessed with standardized tests which are supposed to possess psychometric properties which are conducted by external examination agencies like West African Examination Council (WAEC), National Examination Council (NECO) and National Business and Technical Examination Board (NABTEB) (Ukwuije, 2012).

The Senior School Certificate Examination (SSCE) Economics is written in two parts. Paper 1 is multiple choice questions while paper 2 consists of essay items. In the West African Senior School Certificate Economics Examination paper 1, candidates are expected to answer five (5) questions from paper 2 and fifty (50) multiple choice questions from paper 1 with four (4) options lettered A to D. The same thing is applicable to National Examination Council (NECO)'s conducted Senior Secondary School Certificate Economics Examination paper 2; except that its paper 1 consists of sixty (60) multiple choice or objective tests item questions with five (5) options to an item lettered A to E. (Osadebe, 2014). However, the quality of each test item used for assessment is very crucial. The quality of a test item used in testing is determined by the psychometric properties of the test which includes the validity, reliability, difficulty index, and discrimination power, standard error of the measurement and distracter effectiveness.

Reliability refers to the consistency of the scores obtained. That is, how consistent the scores are for each individual from one administration of an instrument to another and from one item to another (Kinyua & Okunya, 2014). Reliability is a measure of how stable, dependable, trustworthy and consistent a test is in measuring the same thing each time. Afolayan (2012) defined reliability of a test as the degree of consistency or stability which the test exhibits. Therefore, the study is interested in re-establishing the reliability of the Economics objective test of the Senior Secondary School Certificate Examination conducted by WAEC and NECO in order to find out if both are equivalent or differ significantly.

There are very few researches that have been conducted on psychometric properties of a test due to the technicality of this type of research work. A study was carried out by Obinne (2011) to investigate the psychometric analysis of two major examinations in Benue State, Nigeria: NECO and WAEC in Biology.

Findings from the analyzed data revealed that, there were significant differences in the standard error of measurement of Biology examination conducted by NECO and WAEC in 2000, 2001 and 2002. The finding implied that Biology Examination conducted by NECO has smaller SEM (high reliability) than those of WAEC.

Another study was carried out by Bandele and Adewale (2013) on the comparative analysis of the reliability and validity coefficient of WAEC, NECO and NABTEB Mathematics Examination in Ekiti State. The result revealed that the reliability coefficients of WAEC, NABTEB and NECO were 0.89, 0.89 and 0.77 respectively which was computed using the Kuder-Richardson 20. Hence there was no significant difference in the reliability coefficient of WAEC and NABTEB. Mathematics Examinations, as indicated by the t-cal of 0.32 and 0.57 respectively while that of NECO Examination was significantly different.

Furthermore, another study was carried out by Kolawale (2007) to compare the psychometric properties of WAEC and NECO Mathematics multiple choice items so as to ascertain whether the two tests are equivalent tests. The study was carried out in Ekiti State. The findings revealed that the WAEC and NECO mathematics multiple choice terms are significantly equivalent in terms of reliability since the r-calculated is 0.75; this shows that there is high relationship between WAEC and NECO and there is no significant difference between their means. This means that the two papers are equivalent hence reliable.

The study sought to:

1. establish the difference in the mean internal consistency coefficients of Economics multiple choice examination questions of May/June WAEC and June/July NECO examination between 2017/2018-2019/2020 academic sessions.
2. determine the difference in the mean stability coefficients of Economics multiple choice examination questions of May/June WAEC and June/July NECO examination between 2017/2018-2019/2020 academic sessions.

The following research questions were raised for the study;

1. What is the difference in the mean internal consistency coefficients of Economics multiple choice examination questions of May/June WAEC and June/July NECO examination between 2017/2018-2019/2020 academic sessions?

2. What is the difference in the mean stability coefficients of Economics multiple choice examination questions of May/June WAEC and June/July NECO examination between 2017/2018-2019/2020 academic sessions?

The following null hypotheses were tested at 0.05 level of significance.

H₀₁: There is no significant difference in the mean internal consistency coefficients of Economics multiple choice examination questions of May/June WAEC and June/July NECO examinations between 2017/2018 - 2019/2020 academic sessions.

H₀₂: There is no significant difference in the mean stability coefficients of Economics multiple choice examination questions of May/June WAEC and June/July NECO examinations between 2017/2018 - 2019/2020 academic sessions.

Methodology

The ex-post facto research design was adopted. The accessible population for this study consisted of all the 8,213 Senior Secondary School Three students from public schools who enrolled for the External Senior School Certificate Examination conducted by the West African Examination Council and National Examination Council in 2017/2018 to 2019/2020 Academic Sessions in Abia State from a total of 248 Public Senior Secondary Schools (Abia State Secondary Education Management Board, 2018). The sample for the study consisted of 850 students selected from 23 secondary schools across 8 local government areas in the three Education Zones.

The sampled students were students who had enrolled for Economics at both the West African Senior School Certificate Examination and National Examination Council Senior School Certificate Examination in the 2017/2018 to 2019/2020 Academic Sessions. The sample of the Education Zones, L.G.A, schools and students represented 100%, 47.05%, 10.48% and 10.35% respectively. The multi stage sampling process involving purposive and simple random sampling techniques were adopted for the study. The instruments are the Economics past questions of the West African Examination Council and National Examination Council between 2017/2018 and 2019/2020. The Economics Multiple Choice Questions were used since they were the ones for which the Psychometric properties being studied could be obtained. The instruments which are pulled questions from WAEC and NECO past question papers for the years 2017/2018 to 2019/2020 were administered to the students by trained research assistants who were experienced teachers of Economics subject in the selected schools under the supervision of the researcher. All the 850 sets of the copies of the instruments were returned, which gave a 100 percent return rate. The research questions were answered using the item analyses while the hypotheses were tested using the independent simple t-test at 0.05 level of significance. The null hypotheses of significant difference were accepted if the calculated t-value had a p-value which is equal or greater than critical or alpha value and degree of freedom at 0.05 but rejected if the p-value was less than the critical or alpha value and degree of freedom at 0.05 level of significance.

Results

The results are presented in tables as follows:

Table 1: The mean Internal consistency coefficients of Economics multiple choice examination questions of May/June WAEC and June/July NECO examination between 2017/2018-2019/2020 academic sessions

Years	WAEC		Years	NECO	
	No of items	K-21		No of Items	K-21
2020	50	0.78		60	0.67
2019	50	0.83		60	0.71
2018	50	0.81		60	0.69

The data on Table 1 revealed that WAEC in 2018 had an internal consistency value of 0.81 which is highly reliable, in 2019 WAEC had an internal consistency value of 0.83 which indicates highly reliable and in 2020 it had an internal consistency

value of 0.78 which also implied high reliability. The internal consistency of WAEC over 2018, 2019 and 2020 were regarded as highly reliable since Kerian (2015) regarded internal consistent values of over 0.70 as highly reliable. In a similar way, data on

Table 4.1 revealed that NECO in 2018 had an internal consistency value of 0.69 which is moderately reliable, in 2019 it had an internal consistency value of 0.71 which indicates highly reliable and in 2020 it had an internal consistency value of 0.67 which also implied moderate reliability. The internal consistency of NECO in 2018 and 2020 were

regarded as moderately reliable since Kerian (2015) regarded internal consistent values of 0.40 -0.69 as moderately reliable. Hence, in comparative analysis, the Economics multiple choice examination questions of May/June conducted by WAEC had a higher internal consistency(reliability) than that conducted by NECO.

Table 2: The t-test analysis of the significant difference in the mean internal consistency coefficients of Economics multiple choice examination questions of May/June WAEC and June/July NECO examinations between 2017/2018-2019/2020 academic sessions.

Examination	No of items	Mean	S.D	DF	t.cal	P-value	Decision
WAEC	50	0.81	0.41	108	3.68	0.000	Rejected
NECO	60	0.68	0.38				

The data on table 2 showed a t-cal value of 3.68 and P-value of 0.000. The calculated t-value of 3.68 in respect of the difference in the mean internal consistency coefficients of Economics multiple choice examination questions of May/June WAEC and June/July NECO examinations between 2017/2018-2019/2020 academic sessions is greater than t-critical value of 1.96 with degree of freedom of 108 at 0.05 level of significance. Therefore, the null

hypothesis of no significant difference in the mean internal consistency coefficients of Economics multiple choice examination questions of May/June WAEC and June/July NECO examinations is rejected. Consequently, the mean internal consistency coefficients of Economics multiple choice examination questions of May/June WAEC and June/July NECO examinations differed significantly between 2017/2018-2019/2020 academic sessions.

Table 3: The mean stability coefficients of Economics multiple choice examination questions of May/June WAEC and June/July NECO examination between 2017/2018-2019/2020 academic sessions

WAEC			NECO		
Years	No of items	R	Years	No of Items	R
2020	50	0.84	2020	60	0.61
2019	50	0.80	2019	60	0.67
2018	50	0.83	2018	60	0.70

The data on table 3 revealed that WAEC in 2018 had a stability value of 0.83 which is highly reliable, in 2019 it had a stability value of 0.80 which indicates high reliability and in 2020 it had a stability value of 0.78 which also implied high reliability. The stability of WAEC for 2018, 2019 and 2020 were regarded as highly reliable since Kerian (2015) regarded stability values of over 0.70 as highly reliable. In a similar way, data on Table 4.2 revealed that NECO in 2018 had a stability value of 0.70 which is moderately reliable, in 2019 it had a stability value

of 0.67 which indicates moderate reliability and 2020 it had a stability value of 0.61 which also implied moderate reliability. The stability of NECO in 2018 and 2020 were regarded as moderately reliable since Kerian (2015) regarded internal consistent values of 0.40 -0.69 as moderately reliable. Hence, in comparative analysis, the Economics multiple choice examination questions of May/June conducted by WAEC had a higher stability (reliability) than that conducted by NECO.

Table 4: The t-test analysis of the significant difference in the mean stability coefficients of Economics multiple choice examination questions of May/June WAEC and June/July NECO examinations between 2017/2018-2019/2020 academic sessions.

Examination	No of items	Mean	S.D	DF	t.cal	P-value	Decision
WAEC	50	0.88	0.54	108	4.75	0.000	Rejected
NECO	60	0.67	0.39				

The data on table 4 showed a t-cal value of 4.75 and P-value of 0.000. The calculated t-value of 4.75 in respect of the difference in the mean stability coefficients of Economics multiple choice examination questions of May/June WAEC and June/July NECO examinations between 2017/2018-2019/2020 academic sessions is greater than t-critical value of 1.96 with degree of freedom of 108 at 0.05 level of significance. Therefore, the null hypothesis of no significant difference in the mean stability coefficients of Economics multiple choice examination questions of May/June WAEC and June/July NECO examinations is rejected. Consequently, the mean stability coefficients of Economics multiple choice examination questions of May/June WAEC and June/July NECO examinations differed significantly between 2017/2018-2019/2020 academic sessions.

Discussion of Findings

The analysis in Table 1 revealed that WAEC in 2018 had an internal consistency value of 0.81 which is highly reliable, in 2019 had an internal consistency value of is 0.83 which indicates highly reliable and in 2020 had an internal consistency value of is 0.78 which also implied high reliability. The internal consistency of WAEC over 2018, 2019 and 2020 were regarded as highly reliable since Kerian (2015) regarded internal consistent values of over 0.70 as highly reliable. In a similar way, data on table 1 do revealed that NECO in 2018 had an internal consistency value of 0.69 which is moderately reliable, in 2019 had an internal consistency value of is 0.71 which indicates highly reliable and in 2020 had an internal consistency value of is 0.67 which also implied moderate reliability. The internal consistency of NECO in 2018 and 2020 were regarded as moderately reliable since Kerian (2015) regarded internal consistent values of 0.40 -0.69 as moderately reliable. Hence, in comparative analysis the Economics multiple choice examination questions of May/June conducted by WAEC had a higher internal consistency (reliability) than that conducted by NECO. The analyses in Table 2 also showed that the

Economics multiple choice examination questions constructed by WAEC between 2017/2018-2019/2020 academic sessions had a significantly better internal consistency when compared to the Economics multiple choice examination questions constructed by NECO between 2017/2018-2019/2020 academic sessions. These findings agree with the findings of Bandele and Adewale (2013) whose findings showed that WAEC, NECO and NABTEB mathematics achievement Examinations are highly reliable and valid.

The analysis in Table 3 revealed that WAEC in 2018 had an internal consistency value of 0.81 which is highly reliable, in 2019 had an internal consistency value of is 0.83 which indicates highly reliable and in 2020 had an internal consistency value of is 0.78 which also implied high reliability. The internal consistency of WAEC over 2018, 2019 and 2020 were regarded as highly reliable since Kerian (2015) regarded internal consistent values of over 0.70 as highly reliable. In a similar way, data on table 4.3 do revealed that NECO in 2018 had an internal consistency value of 0.69 which is moderately reliable, in 2019 had an internal consistency value of is 0.71 which indicates highly reliable and in 2020 had an internal consistency value of is 0.67 which also implied moderate reliability. The internal consistency of NECO in 2018 and 2020 were regarded as moderately reliable since Kerian (2015) regarded internal consistent values of 0.40 -0.69 as moderately reliable. Hence, in comparative analysis the Economics multiple choice examination questions of May/June conducted by WAEC had a higher internal consistency (reliability) than that conducted by NECO. These findings agree with the findings of Bandele and Adewale (2013) whose findings showed that WAEC, NECO and NABTEB mathematics achievement Examinations are highly reliable and valid. The data on Table 4 also indicated that The Economics multiple choice examination questions constructed by WAEC between 2017/2018-2019/2020 academic sessions had a significantly better stability when compared to the Economics multiple choice examination questions constructed by

NECO between 2017/2018-2019/2020 academic sessions.

Conclusion

The findings of the study revealed that the Economics multiple choice examination questions constructed by WAEC between 2017/2018-2019/2020 academic sessions had a significantly better internal consistency when compared to the Economics multiple choice examination questions constructed by NECO between 2017/2018-2019/2020 academic sessions. Also, the Economics multiple choice examination questions constructed by WAEC between 2017/2018-2019/2020 academic sessions had a significantly better stability when compared to the Economics multiple choice examination questions constructed by NECO between 2017/2018-2019/2020 academic sessions.

Recommendations

Based on the findings of the study, the researcher made the following recommendations.

1. The psychometricians and Test development experts should organize workshops to train teachers on test usage based on IRT model for their students.
2. The examination bodies and teachers should be encouraged to adopt (IRT) in developing test items used in measuring students' ability in economics.

References

- Bande S. O. & Adewale A. E. (2013). Comparative analysis of the reliability and validity coefficients of WAEC, NECO and NABTEB Constructed Mathematics Examination
- Calmorin, K. & Laurentina, S. (1994). *Educational Research, Measurement and Evaluation*, DUBLIN: Assessment book store, Inc.
- Emaikwu, S.O. (2011) Issues in test item bias in public examinations in Nigeria and implications for Testing. *International Journal of Academic Research in Progressive Education and Development*. 1, (1):175-187
- Kinyua, W. & Okuya, T. (2014). Validity and Reliability of teacher made test: case study of year 11 physics in Nyahururu District of Kenya. *African Education Research Journal* 2(2), 61 – 71.
- Monday T.J (2012). *Fundamentals of Test and Measurement in Education* Calabar: University of Calabar Press.
- Nkemakolam, E.O. (2009). *Measurement in Education: An introduction*: Enugu Jibaco Publishers

- Obinne (2011). A psychometric analysis of two major examinations in Nigeria: *International Journal of Education Science*, 3(2): 137-144
- Osadebe, P.W. (2014). Construction of economics achievement test for assessment of students. *World Journal of Education*. 14(2), 58 – 64.
- Ukwuije, R.P.I. (2012). Educational assessment: A sine qua non for quality education. 83rd inaugural lecture. University of Port Harcourt, Port Harcourt.