PSYCHOMETRIC PROPERTIES OF WEST AFRICAN EXAMINATION COUNCIL (WAEC) O'LEVEL MULTIPLE CHIOCE PAST QUESTION PAPER ON ECONOMICS FOR THE YEAR 2013

BY

DR. FREDRICK EKENE ONAH
DEPARTMENT OF LIFE SCIENCE EDUCATION
IMO STATE UNIVERSITY, OWERRI

AND

CELESTINE CHIBUIKE JIWUEZE
DEPARTMENT OF PHYSICAL SCIENCE EDUCATION
IMO STATE UNIVERSITY, OWERRI

Abstract

This study deals with the psychometric analysis of WAEC. The objective was to examine the difficulty, discrimination and distracter indices of SSCE Multiple Choice Past Question Papers in Economics for 2013. An attempt was made to find out whether difficulty, discrimination, and distracter indices meet the required standard using Classical Test Theory (CTT) item analysis to analyze the result. The design is indeed descriptive survey. Owerri Municipal Council, Imo State, Nigeria was the study area. The population for the study comprised all year three (SSIII) senior secondary school students who enrolled for May/June 2013 Economics senior secondary school certificate examination of WAEC in Owerri Municipal Council of Imo State. The sample for the study was one hundred (100) students. Stratified random sampling technique was used to get this sample. WAEC 2013 multiple choice Economics questions were the instruments for the study. Item analysis using SPSS computer programme was carried out. This implies that difficult level of items set by WAEC is relatively moderate. The result also indicates that discrimination and distracter indices of these two examinations are not in conformity with required standard. It is recommended that experts in the field of measurement and evaluation should be trained and employed by WAEC to write and /or vet tests set by WAEC to ensure a proper test construction standard.

Introduction

Evaluation is a process of seeking, obtaining and quantifying data with a view to making value judgment about objects, events or their characteristics. Conceptualized in this way, evaluation includes measurement and more. It is both quantitative and qualitative.

Educational assessment, which is the focus in this paper, has been defined in different

Psychometric Properties of West African Examination Council (WAEC) O'level
Multiple Choice Past Question Paper on Economics for the Year 2013

28

ways by different people. Some of these definitions, even by experts in the field, tend to be at variance with one another to the extent that Nworgu (2010) in a keynote address to the Annual Conference of the Association for Educational Assessment in Africa referred to the situation as 'a drift to babelism'. Notwithstanding this apparent lack of consensus among experts on a precise definition of educational assessment, we shall, in this discourse use the term to mean a systematic process of gathering data from a variety of sources in order to understand, describe and improve learning. This conception of educational assessment implies that:

· it encompasses both measurement and evaluation;
· decisions or judgments are made based not only on one single measurement or source of data but on multiple measurements or sources of data. In that sense, assessment is broader than either measurement or testing.

Summative- Assessment of learning (AOL)

Summative assessment or assessment of learning is that assessment carried out to determine what students have been able to learn at the end of a given lesson, unit, programme or period of schooling. It's uses, includes the fact that it helps to:

· Measure what students have learnt at the end of a unit,
· Promote students
· Ensure they have met required standards for certification on completion of a level or period of schoo
· Select students for entry into further education

Other uses of summative assessment have been identified to include sorting and ranking students. Test is one such tool for evaluation.

Test consists of a set of uniform questions or tasks to which a student is to respond independently and the result of which can be treated in such a way as to provide a quantitative comparison of the performance in different students (Nworgu, 2003). Testing is a fundamental part of teaching and learning process used not only as a basis for ranking students at the end of the teaching-learning process but to guide teaching, and aid in the development of curriculum, as well as in the assessment of needs, learning difficulties, level of mastery and differences among students. Based on the level of performance criterion, there are three types of tests, namely: the General Mental Ability Test, Separate Ability Test and Achievement Test. The practical relevance of these tests and their testing is largely dependent on their levels of reliability, validity, difficulty, discrimination and distracter.

Senior Secondary Certificate Examination is a standardized test designed to measure learner's achievement as stipulated in the National Policy on Education. This certificate examination is administered by the public examination bodies and provides uniform standards to members of the society irrespective of the type of education received. They are concerned with large scale testing programme. The scores of the standardized test administered are used to assess the performance of the students with the aim to identify students who have successfully completed a certain level of education,

ways by different people. Some of these definitions, even by experts in the field, tend to be at variance with one another to the extent that Nworgu (2010) in a keynote address to the Annual Conference of the Association for Educational Assessment in Africa referred to the situation as 'a drift to babelism'. Notwithstanding this apparent lack of consensus among experts on a precise definition of educational assessment, we shall, in this discourse use the term to mean a systematic process of gathering data from a variety of sources in order to understand, describe and improve learning. This conception of educational assessment implies that:

· it encompasses both measurement and evaluation;
· decisions or judgments are made based not only on one single measurement or source of data but on multiple measurements or sources of data. In that sense, assessment is broader than either measurement or testing.

Summative- Assessment of learning (AOL)

Summative assessment or assessment of learning is that assessment carried out to determine what students have been able to learn at the end of a given lesson, unit, programme or period of schooling. It's uses, includes the fact that it helps to:

· Measure what students have learnt at the end of a unit,
· Promote students
· Ensure they have met required standards for certification on completion of a level or period of schoo
· Select students for entry into further education

Other uses of summative assessment have been identified to include sorting and ranking students. Test is one such tool for evaluation.

Test consists of a set of uniform questions or tasks to which a student is to respond independently and the result of which can be treated in such a way as to provide a quantitative comparison of the performance in different students (Nworgu, 2003). Testing is a fundamental part of teaching and learning process used not only as a basis for ranking students at the end of the teaching-learning process but to guide teaching, and aid in the development of curriculum, as well as in the assessment of needs, learning difficulties, level of mastery and differences among students. Based on the level of performance criterion, there are three types of tests, namely: the General Mental Ability Test, Separate Ability Test and Achievement Test. The practical relevance of these tests and their testing is largely dependent on their levels of reliability, validity, difficulty, discrimination and distracter.

Senior Secondary Certificate Examination is a standardized test designed to measure learner's achievement as stipulated in the National Policy on Education. This certificate examination is administered by the public examination bodies and provides uniform standards to members of the society irrespective of the type of education received. They are concerned with large scale testing programme. The scores of the standardized test administered are used to assess the performance of the students with the aim to identify students who have successfully completed a certain level of education,

Psychometric Properties of West African Examination Council (WAEC) O'level
Multiple Choice Past Question Paper on Economics for the Year 2013

30

Where U= Number in the upper one-third (1/3) group who chose the option

    L= Number in the lower one-third (1/3) group who chose the option

    N= Number of testees in either the upper or lower one-third (1/3) group options that have positive distractor indices are preferred.

All these add up to the psychometric properties of a test. The development of achievement, ability, aptitude, interest and personality tests is generally a multi-step process that can follow one of two distinct measurement frameworks. These are usually called the Item Response Theory (IRT) and Classical Test Theory (CTT) measurement strategies (Macdonald and Paunanen 2002). In education, psychometricians apply CTT in order to achieve tasks such as developing and refining examination items, maintaining banks of items for examinations and equating for the successive versions of examination (for example, to allow comparisons between results over time).

After due consideration of these theories, this work is based on classical test theory (CTT) because an item can be more and less difficult than another item at the same time. This cannot happen in item response theory (IRT) because of the invariance property and this makes it difficult to publish IRT analyses because accepted procedure constantly changes.

This study was necessitated by public complaint and conception of the superiority of the test items of one examination over the other. The major objective of the study was to determine the difficulty, discrimination and distracter indices of Economics examination conducted by WAEC. Thus, the public will be enlightened on the interpretation of students' results from the examination conducted by WAEC. Presently, the performances of students in the examination conducted by WAEC are interpreted based on the sum of their total scores which is typical of Classical Test Theory (CTT). WAEC needs to consider the psychometric properties of tests in taking decisions on the observable performance of candidates in order to improve upon test construction, administration and analysis.

Senior School Certificate Examination is a standardized summative test designed to be taken at the end of six years programme in secondary school for certification. When this certificate is obtained, it serves as a pre-requisite for gaining admission into tertiary institution and for job opportunity. In some cases some people want WAEC status-quo to be maintained as the only reliably existed SSCE certificate. To what extent does the test instrument administered by WAEC satisfy the criteria recommended by the experts? These question need urgent answer and form the basis for the study.

Purpose of the Study

The purpose of the study is to evaluate the psychometric properties of WAEC; specifically the objectives of the study are as follows:

I)   To determine the difficulty index of the individual test items as administered by WAEC.

make available assessment for certificate placement, and select who is qualified for the next level in educational system. Nworgu (2003) posits that validity of a test refers to the extent to which a test measures what it is supposed to measure and nothing else. Reliability of an instrument means the consistency with which that instrument measures that traits, characteristics, constructs, it was designed to measure
.

Item Difficulty Index: According to Onunkwo (2002), item difficulty is defined as the percentage or proportion of persons answering each item correctly. Hence it gives an estimate of the proportion of students likely to answer an item correctly in subsequent testing periods, assuming the original group of student from which the index was computed is identical to the present group taking the test. Betz and Walsh (1990) defined item difficulty statistically as the percentage of persons who respond correctly to an item. The higher the percentage of people who answer an item correctly, the easier the item is considered to be, and vice versa. Item difficulty levels are usually expressed as P- values. An ideal item should have a p-value of 0.50 but realistically it could range from 0.40 to 0.70. It can be calculated

using the formula

$$P = \frac{U+L}{2N}$$

Where U= the number of testees in the upper one-third (1/3) of the group who passed the item.

L= the number of testees in the lower one-third (1/3) of the group who passed the item.

N= Number of testees in either the upper or lower one-third (1/3) of the group who took the test

Item Discrimination Index: This index according to Nworgu (2003) answers the question: Does the item distinguish between the bright and dull students? The discrimination index is a measure of the extent to which the item discriminates between the bright and dull students. An ideal item should have a d-value of +1.00 but realistically it could range from +0.30 to +1.00. It can be calculated using the formula

$$d = \frac{U-L}{N}$$

Where U= the number of testees in the upper one-third (1/3) of the group who passed the item.

Results and Discussion

Research Question One

What proportion of difficulty index of the individual scores of the items of WAEC fell with the realistic range from

Range	WAEC	
	F	Proportion
0.1 – 0.3	0	0
0.4 – 0.7	76	0.76
0.8 – 1.0	24	0.24
Total	100	1.00

Psychometric Properties of West African Examination Council (WAEC) O'level
Multiple Choice Past Question Paper on Economics for the Year 2013

32

Table 1: shows the proportion of difficulty index of the individual scores of the items of WAEC for different range and its proportion. The result shows a *higher* proportion in the difficulty level with 0.4-0.7. The result indicates that WAEC passed 76% difficulty level while 24% were relatively cheap and needed to be improved upon.

Research Question Two

How do the individual scores of the items set by WAEC discriminate from 0.3-1.0?

Table 2: Descriptive Analysis of Item Discrimination in WAEC.

Ranges	WAEC	
	F	**Proportion**
0.0-0.2	8	0.08
0.3-1.0	50	0.50
-Ve	42	0.42
Total	100	1.00

The data on table 2 contains the proportion of discrimination index of the individual scores of the items set by WAEC for different range and its proportion. The above analysis shows that 50% of the items are good while 50% are bad in terms of discrimination index and needed to be improved upon.

Research Question Three

What proportion of distractor index scores of the individual test options fell within the realistic range of -1 to + 1

Table 3: Descriptive Analysis of Item Distracter in WAEC.

Ranges	WAEC	
	F	**Proportion**
-ve	55	0.55
+ve	42	0.42
0	3	0.03
Total	100	1.00

Research Question Three

What proportion of distractor index scores of the individual test options fell within the realistic range of -1 to + 1

Table 3: Descriptive Analysis of Item Distracter in WAEC.

Table 3 shows the proportion of distracter index of the individual test options of the items set

Discussion of the result

The data generated for answering this research question 1 is contained in **table 1**, it clearly shows the range drawn from the frequencies and percentages of descriptive analysis of item difficulty in WAEC. Analysis of result in **table 1** showed that 76% of test items set by WAEC passed the realistic standard of difficulty level index ranging from 0.4 to 0.7. The finding is in line with the view of Onunkwo (2002) who opined that item difficulty is the percentage or proportion of persons answering each item correctly. It gives an estimate of the proportion of students likely to answer an item correctly in subsequent testing periods, assuming the original group of student from which the index was computed is identical to the present group taking the test.

The data generated in **table II,** showed that 50% of test items set by WAEC passed the realistic standard for discrimination index. This finding is in line with the view of Nworgu (2003) who states that discrimination index is a measure of the extent to which the item discriminates between the bright and dull students. While in **table III**, the analysis of result

showed that WAEC individual test options possess a poor distractor index and this would encourage guess work among students' testees. From the result of the study, WAEC would clearer understanding of their performance in test construction and be appropriately guided from now on. The finding is in line with the view of Nworgu, (2003) who see distractor index as that which finds out how many students choose the various options and more especially, whether the options appeal more to the dull students than the bright ones.

Conclusion

From the results of this research study, one would conclude that; Test items set by WAEC do not meet the required standard for measuring the difficulty level analysis. Result showed that items set by WAEC are relatively moderate. Furthermore, the result showed that WAEC test items lacks good discrimination index. Finally, the result showed that WAEC individual test options lack good distractor index, such that differences in performance might be due to the variation in the examination standard or due to the individual testee. The public needs to be convinced that the examinations conducted by WAEC are relatively equal standard with other

examination bodies. It is to say that there are certain qualities which a measuring instrument is expected to possess. Inadequate evaluation instruments, lack of training among teachers and some basic school infrastructural facilities lead to high dilapidation of educational system, but when these are being checked and improved by the authorities concerned, it will help to improve the standard of education.

Recommendation

On the basis of the findings of this study, the researchers recommend.

1. That expert in the field of measurement and evaluation should be trained and employed by WAEC to write and /or vet tests set by WAEC to ensure a proper test construction standard.

2. Inexperienced teachers should be sent on refresher courses, seminars and workshops on the best way of constructing multiple choice items.

3. Multiple choice items should be well constructed by providing plausible distractors so as to reduce the students' guessing abilities that could be achieved by the construction of multiple choice items.

4. Teachers and other test constructors should be addressed to adopt multiple choice test formats to a wide variety of approaches by using additional materials such as graphs, charts and diagrams.

Summary

The study investigates psychometric analysis of WAEC and examined the difficulty, discrimination and distracter indices of SSCE Multiple Choice Past Question Papers in Economics for 2013. An attempt was made to find out whether difficulty, discrimination, and distracter indices meet the required standard using Classical Test Theory (CTT) item analysis to analyze the result. The design is indeed descriptive survey. Owerri Municipal Council, Imo State, Nigeria was the study area. The population for the study comprised all year three (SSIII) senior secondary school students who enrolled for May/June 2013 Economics senior secondary school certificate examination of WAEC in Owerri Municipal Council of Imo State. The sample for the study was one hundred (100) students. Stratified random sampling technique was used to get this sample. WAEC 2013 multiple choice Economics questions were the instruments for the study and the item analysis was carried out using SPSS computer programme.

References

Allen, M.J., & Yen, W. M. (2002). *Introduction to Measurement Theory. Long Grove, IL: Waveland Press*

Baker, F. (2001). The Basics of Item Response Theory. E R I C Clearinghouse on Assessment and Evaluation, University of Maryland, College Park, MD. Available free with software included from IRT at Edres.org Betz, N.E SC & Walsh B.W (1990); Test and Assessment (2nd ed.).

examination bodies. It is to say that there are certain qualities which a measuring instrument is expected to possess. Inadequate evaluation instruments, lack of training among teachers and some basic school infrastructural facilities lead to high dilapidation of educational system, but when these are being checked and improved by the authorities concerned, it will help to improve the standard of education.

Recommendation

On the basis of the findings of this study, the researchers recommend.

1. That expert in the field of measurement and evaluation should be trained and employed by WAEC to write and /or vet tests set by WAEC to ensure a proper test construction standard.

2. Inexperienced teachers should be sent on refresher courses, seminars and workshops on the best way of constructing multiple choice items.

3. Multiple choice items should be well constructed by providing plausible distractors so as to reduce the students' guessing abilities that could be achieved by the construction of multiple choice items.

4. Teachers and other test constructors should be addressed to adopt multiple choice test formats to a wide variety of

approaches by using additional materials such as graphs, charts and diagrams.

Summary

The study investigates psychometric analysis of WAEC and examined the difficulty, discrimination and distracter indices of SSCE Multiple Choice Past Question Papers in Economics for 2013. An attempt was made to find out whether difficulty, discrimination, and distracter indices meet the required standard using Classical Test Theory (CTT) item analysis to analyze the result. The design is indeed descriptive survey. Owerri Municipal Council, Imo State, Nigeria was the study area. The population for the study comprised all year three (SSIII) senior secondary school students who enrolled for May/June 2013 Economics senior secondary school certificate examination of WAEC in Owerri Municipal Council of Imo State. The sample for the study was one hundred (100) students. Stratified random sampling technique was used to get this sample. WAEC 2013 multiple choice Economics questions were the instruments for the study and the item analysis was carried out using SPSS computer programme.

References

Allen, M.J., & Yen, W. M. (2002). *Introduction to Measurement Theory. Long Grove, IL: Waveland Press*

Baker, F. (2001). The Basics of Item Response Theory. E R I C Clearinghouse on Assessment and Evaluation, University of Maryland, College Park, MD. Available free with