Psychometric Properties of English Language in Junior Secondary School Certificate Examination Questions in Rivers State

Josephine Nkem Bosede Joe-Kinanee[i]
Department of Educational Psychology, Guidance & Counselling
Faculty of Education
University of Port Harcourt, Nigeria.
jjoekinanee@gmail.com

Ijeoma Margaret Opara[ii]
Department of Educational Psychology, Guidance & Counselling
Faculty of Education
University of Port Harcourt, Nigeria.

……..

Abstract

The study verified the psychometric properties of English Language in Junior Secondary School Certificate Examination (JSSCE) questions from 2014–2016 academic sessions. Five research questions were answered; it adopted the evaluation and descriptive survey designs. A sample of 1,068 was randomly drawn from 8 Local Government Areas out of a population of 65,054 junior secondary school 3 students in the 280 public and approved 397 secondary schools in the 23 Local Government Arears in Rivers State. Instruments were past objective items in 2014 to 2016 JSSCE scripts for English Language; JSS 1 to 3 curriculum and syllabus were used to generate data. Kuder-Richardson 20, percentages, formulae for calculating item analyses were used. Out of sample of 1,068 JSS 3 students, 1,035 (97%) did the test. The findings were that the JSSCE English Language 2014 - 2016 lacked content validity; the co-efficient of internal consistency was high among items in each test; all the tests had good difficulty, discrimination indices and distracter indices criteria. JSSCE English Language 2014 had the highest internal consistency, the best difficulty and discrimination indices, indicating that it was the best test among the three tests. The faults in the tests could be caused by unfamiliarity of the questions to the students, test questions could have been too difficult relative to the ability of the students, some items were ambiguous, not clearly written. It was recommended among others, that experts in measurement and evaluation should be used to carry out the rigorous process of item analyses by determining item difficulty, discrimination and distracter indices, not by hand-picking items from past questions. These would bring great improvement in the academic performance in English Language in Junior Secondary Schools in Rivers State.

Keywords: Psychometric Properties, English Language, Validity, Reliability.

……..

INTRODUCTION

Education is defined by Kpolovie, Iderima and Ololube (2014) as learning process where the skills, knowledge, values, habits and beliefs of a people are transmitted from one generation to the next through discussions, storytelling, training, teaching and/or research. In Nigeria, our secondary education is shared into levels: the Junior and Senior Secondary schools. The Universal Basic Education, which comprises the primary and junior secondary stages of education, was introduced in Nigeria in September 1988 (FRN, 2004). Thus, the Federal Government through the Nigerian Educational Research and Development Council (NERDC) established and initiated the 9-year Basic Education Curriculum (BEC) in schools by rearranging all existing primary and junior secondary school Curricula to achieve the key objectives of the UBE programme. The idea of the 9-year Basic Education Curriculum is centered on acquiring appropriate levels of literacy, numeracy, manipulative, communicative and life-skills in addition to principled, moral and civic values essential for laying a solid foundation for life time learning as a basis for scientific and philosophical thinking.

The Junior Secondary School Certificate Examination (JSSCE) is an external examination that supervised by a recognized examining body: The State Ministry of Education. It is an examination meant for year three (3) junior secondary school students who wish to carry on to senior secondary school in state public, private and technical schools in Nigeria (Igbokwe, 2015). In the educational system of Nigerian, the JSSCE is an exceptional tool. It is an essential prerequisite for admission into the senior secondary school: used to appraise the academic attainment of students upon their third year in the junior secondary school. The JSSCE is also a tool for employment.

However, this research study is based on English Language JSSCE questions. This subject is very important and occupies central position in academics. For instance, as an international language, English Language is an important means of communication; the subject must be credited in every external examination like the JSSCE.

English Language is very important as an international language and a medium of communication and instruction in schools. As well as other subjects like Mathematics, Social Studies and Basic Science and Technology, they are known as junior secondary school core or basic subjects. In Nigeria, it was introduced by the British colonial administration in the early 19th century. Nigeria as a country with diverse ethnic groups, religions and language, has the English language as an important means of interaction among the diverse ethnic groups, making it a lingua franca in the country. It is evident that English Language leads to studying of Art subjects. Today, English language is commonly used in Nigeria as the second language, it has succeeded in reducing the difficulties of multiple language in Nigeria (Danladi, 2013).

In Nigeria today, formal education is a creation of English language, without which there would have remained no possibility of anything more than the elementary formal education. English language is very useful in education where any person looking for admission into any tertiary institution is expected to have at least a credit in the subject before an admission can be given to him or her. English language is extensively used by all the sectors of the Nigerian economy: the government, in business, mass media.

However, students' performances over the years have been on a descending tendency despite the prominence of the English language. Reports revealed that the performance of students in written, reading and spoken English is not encouraging.

A test is used to assess students' performance in any subject including English Language. Onunkwo (2002) viewed a test as a tool meant for detecting some qualities, traits, characteristics, attributes etc. a person, an item or a thing possesses. A test may be given on paper, orally, on a computer or a restricted area that entails the person taking the test to perform skills physically. Administering a test may be informally e.g. by a parent in the home, or formally (in the school). In formal testing, marks or scores are often given; this may be inferred to a criterion or norm. Onunkwo further explained that the norm-referenced test permits comparison of students' performance, it determines how each student performs in comparison with the performance of other members of that class. The criterion-referenced test is also known as mastery, competency test, it only determines if the students can reach a certain performance level, but does not compare students' achievement.

A standardized test is a form of test given and scored in a consistent routine. It obliges all test takers to answer a selection of questions; it is a large-scale test administered to large populations of students, it is developed by experts, has uniform and standard procedure for administration and scoring of the test, covers a large content area (free online dictionary, 2015). Thus, the JSSCE is meant to be a standardized test. Tests are given for the benefits of the testees and the teachers. Data gotten from tests aid to advance the quality and usefulness of instructional activities. As identified by Orluwene (2012), the following serve as functions of tests:

- **Identifying what has been learned:** Tests are given after a lesson or unit to assess what the students have learned. The results can be analyzed to see where majority of the students are having problems with in their class especially when the tests are tied to lesson objectives. The feedback the test provides enables the teacher to periodically adjust and modify his methodology of teaching. Test results are used to foresee how a student will perform in future.
- **Identifying students' weaknesses and strengths:** Tests helps to identify these weaknesses and strengths when pretests are given at the commencement of units to discover what the students previously know and where the teacher needs to focus on. Students' ability levels are determined through tests.
- **Entry into a school programme:** Tests have been used over the years to judge a student centered on merit into a school programme. They may also be needed to take additional exams to be properly placed in classes, different programmes and courses.
- **Motivation:** The scores from the tests motivate hardworking students to maintain their standard, while the poor achievers would be energized to map out learning plans to aid them improve on their previous performance. Students therefore, would adopt good study habits for them to succeed in subsequent tests.
- **Placement/Certification:** Test results help to place students at a point or grade in a course dependent on their abilities, interests, and aptitudes. Information from the test after any educational programme show what the students have been able to achieve. So, certificates are given in accordance to students' performance (Orluwene, 2012).

Since English Language occupies the central position in academics, it is expected that every student performs very well and pass it in examinations including school-based and external examinations. Students' performance in an examination is determined by the students' ability levels and the psychometric properties of the test. Hence, to authenticate students' performance in a test, the test must possess some essential psychometric properties. Orluwene and Igwe

(2016) defined psychometric property of an examination as "the validity, reliability and item characteristics of the examination".

Reliability is the stability or consistency of measures while validity is concerned with determining if the examination measures what it is intended to measure (Kpolovie, 2010). Poor construction of items, ambiguous statements, unclear instruments, inappropriate terms, use of irrelevant technical terms, non-coverage of the content, difficulty of the instrument, incorrect arrangement of test items and poor arrangement of answer options can reduce the validity of an instrument. Nkwocha (2004) added that test duration, test civility and personal factors also affect test validity. It deals with the construction and authentication of instruments of measurement like questionnaires, tests, and personality assessments. Orluwene and Igwe (2016) defined psychometric property of an examination as "the validity, reliability and item characteristics of the examination. Some important psychometric properties include difficulty, discrimination indices and the distracters effectiveness. The faults in the psychometric properties of the Junior Secondary School Certificate Examination (JSSCE) questions can be identified by ascertaining whether the tests are reliable and valid.

Reliability is a very vital quality that scores on a test must possess for it to be accepted as a good one (Kpolovie, 2014). It is the total stability or consistency of a measure. For a testee's level of achievement to be measured with test scores, the scores are expected to be similar under repeated testing, alternate forms of the same test, two halves of a single test and individual item measurement of the test (Opara, 2016). Reliability index value range from -1 to +1. Kpolovie (2010) opines that for the validity to be significant statistically, the magnitude of its co-efficient should be high. The higher the reliability of a test, the higher the worth of the test. A test that is reliable may provide useful valid information, thus a test that is not reliable cannot possibly be valid (Davidshofer and Murphy, 2005).

The reliability of measuring instruments takes different methods, which include:

- **Internal Consistency Reliability:** It is also called total-item homogeneity consistency (Kpolovie, 2010). This measures the consistency of test results across items. The most known measure is Cronbach's alpha. Others are Kuder-Richardson 20 (KR20) and Kuder-Richardson 21 (KR21).
- **Split Half Reliability:** It is also called partial items consistency (Kpolovie, 2010). It treats the two splits of a measure as alternate forms. This can be done by adopting an odd-even split, in which the even-numbered items from one half of the test and the odd-numbered items form the other. A measure is the Spearman-Brown prediction formula.
- **Parallel Forms Reliability:** It measures the degree to which test scores are consistent when there is a difference in the approaches or instruments used. Both must contain test items that measure similar construct, skill, knowledge. Pearson Product Moment Correlation is used to correlate the instruments (Kpolovie, 2010).

Validity is the extent to which a test measures what it is supposed to measure (Elliot, Kratochwill, Cook and Travers, as cited in Orluwene, 2012). It is defined by Wolming and Wilkstrom (2010) as the point whereby the judgment teachers make about their students can be trusted depending on the worth of proof gathered. Opara (2016) views that the validity of an instrument must be evaluated to determine if there if there is enough backing for using the instrument in that specific manner. In assessing the validity of Junior Secondary School Certificate Examination (JSSCE), the following types of validity should be considered:

- **Face Validity:** This validity is not a very scientific type of validity. It estimates whether a test appears to measure a criterion, it does not promise that the test measures what it supposed to measure in that field. It is narrowly related to content validity. While content validity is hinged on a hypothetical basis for assuming if a test is measuring all domains of a certain criterion, face validity tells whether a test appears to be a good measure or not (Onunkwo, 2002).
- **Content Validity:** It is a type of validity that encompasses the logical check of the content of the test to decide whether behaviour field to be measured covers what would adequately represent a sample (Sidhu, 2005). It shows the extent an item represents the domain to be measured (Iweka, 2014). In content validity, the content of the test must match a content field related to the construct. A test's content validity is put together by careful choice of the items to be included. A test specification is drawn up, the items are selected to conform to it via an in-depth survey of the domain of the subject.

Other types of validity include construct, criterion-related, convergent, divergent reliability.

Item Analysis

This is a procedure that examines students' responses to individual test items to evaluate the importance or value of those items and the test itself. It is particularly valuable in improving items that might be sourced in subsequent tests, eliminating vague or misleading ones in a test administration. This analysis is also concerned in reviewing of test item content and statistics which describes testees' performance on the item (Orluwene, 2012). This psychometric analysis is the most operational way to improve reliability. It involves calculation of item difficulties indices and item discrimination indices, calculation of relationships between the items and the sum of items of the whole test. When items that are too difficult, too easy, have zero or negative discrimination are substituted with improved items, the reliability rises.

In carrying out item analysis, scores of tests are arranged, starting with the highest to the least scores. In practice, the criterion groups of 25% to 33% are taken while the middle papers are discarded (Hopkins and Antes, 1978). Using the classical approach, item analysis is often dependent on the P-value and item-total correlation (also known as point-biserial correlation coefficient (Iweka, 2014).

Item Difficulty: This shows the fraction of students in both the upper and lower ability group that got an item right (Ugodulunwa, 2008). It ranges from 0.0 (none of the students correctly responded to the item) to 1.0 (all that responded to the item correctly). The criteria by Sidhu as documented in Orluwene (2012), recommended a difficulty level with range from 0.4 to 0.9. An item that has a low difficulty value of less than 0.4, might have been miskeyed, too difficult in comparison to the general level of ability of the class, vague or not written clearly. Items with difficult index of 1.0 is too easy.

Item Discrimination: It describes how well an item discriminates between the upper and lower ability group of testees (Ugodunlunwa, 2008). The item-total correlation gives the p-value of the discrimination or differentiating power of the item. It will be positive if more students from the higher achievers responded to the item correctly more often than low achievers did, and negative if the opposite occurred. It ranges between -1 and +1. The discrimination is better when the value is closer to 1. Items with negative values should be reviewed.

Effectiveness of Distracters: The distracters must be distinctly not correct (not the best option), just as the correct option (the key) must be correct. They should seem reasonable to an

examinee that is not adequately conversant in the content area. Distracter index shows how the incorrect options distract the lower ability grouping from selecting the correct option (key). A good distracter should attract more students from the lower ability group than the upper ability group (Iweka, 2014). The distracter indices range from -1 to +1. A positive value shows that more students in the lower ability group choose it, while a negative value shows that more students in the higher ability group chose it. A zero index indicates that both groups were equally distracted (Orluwene, 2012).

This theory supposes that everyone has a true score attained if there are no errors in measurement. It shows that there is no perfection in the observed score, which is mainly due to influence of random errors in the measuring instrument (Iweka, 2014). The observed score is the real score an examinee gets on the test, the true score is the score that shows the exact amount of the trait an examinee possesses and demonstrates at every given time, the error score has to do with inconsistency, variance between the true score and the observed score (Kpolovie, 2010). This model of CTT is that, observed scores are composed of the true score and error score:

$X = T + E.$
where,
X is observed score
T is true score
E is error score.

Classical test theory is opposed to the latent trait theory that focuses on the item instead of its test-level focus. Classical Test Theory is for constructing, validating and interpreting norm-referenced tests; also, useful for the comparison of students' performance among themselves (Iweka, 2014). It focuses on the examination of item difficulty, discrimination and distracters. Thus, this theory is appropriate to this study on the psychometric properties of the tests.

It is obvious that students fail tests not just because they were not prepared for the tests, but failure can equally depend on the psychometric properties of the test items. Students would fail tests of high difficulty level above their mental level: questions that are ambiguous. Thus, the problem of this research study is therefore to identify the psychometric properties of English Language in the Junior Secondary School Certificate Examination from 2014 to 2016 academic sessions in Rivers State. The independent variables considered are psychometric properties like reliability (internal consistency), content validity, item analyses (difficulty, discrimination and distractor indices). The dependent variable is the JSSCE questions between 2014 and 2016. The researchers deemed it necessary to assess the psychometric properties of English Language in Junior Secondary School Certificate Examination in Rivers State.

The following research questions guided this study:

- What is the coefficient of internal consistency of English Language JSSCE questions from 2014 to 2016?
- What is the content validity of the objective test items of English Language JSSCE between 2014 and 2016?
- What are the difficulty indices of objective test items of English Language JSSCE questions from 2014 to 2016?

- What are the discrimination indices of objective test items of English Language JSSCE questions from 2014 to 2016?
- How effective are the distractive indices of the items in the English Language JSSCE objective test items from 2014 to 2016?

METHODOLOGY

Evaluation research design and descriptive survey research were adopted by the study. It deals with systematic setting of worthwhile goals for passing value judgments based on the set goals, to ascertain whether the objectives should be modified or improved on. The population of the study is all the 65,054 JSS 3 students of 2016/2017 academic session in junior secondary school 3 in the 280 public secondary schools and approved 397 private secondary schools in the 23 Local Government Areas of Rivers State. A sample of 1,068 junior secondary school students was randomly drawn from 16 selected Junior Secondary Schools in 8 selected Local Government Areas of Rivers State using a multi-stage sampling method. For this study, past objective English Language JSSCE questions from 2014 – 2016 were used; each having 60 objective test items. The Junior Secondary School curriculum and scheme of work were also used to generate data for the study. The instruments are valid instruments verified by experts, used over time in JSS 3 and were constructed by examining bodies as reliable instruments. The instruments (the English Language objective test questions for JSSCE in 2014, 2015 and 2016) were administered by the researcher and trained research assistants under the supervision of the researcher to the JSS 3 students in 2016/2017 academic session (an equivalent group to those that wrote the past examinations). The examination was given to the students at the beginning of the third term, so it served as "mock exam" in the preparation for their JSSCE, which they wrote immediately after. Each student answered the English Language objective questions for 2014, 2015 and 2016. Out of the sample of 1,068, only 1,035 JSS 3 students participated in all the examination, representing 97% of the sample. The data collected were analyzed using the following data analyses: percentages, Kuder-Richardson 20; formulae for item analyses (calculating difficulty, discrimination and distracters indices). In each examination, a sample of 558 students were used for item analysis.

RESULTS

Table 1: Kuder-Richardson20 Reliability of JSSCE English Language Objective Questions from 2014 – 2016.

JSSCE	k / k – 1	Σpq	VARIANCE	Σpq / VARIANCE	1 -Σpq / VARIANCE	INTERNAL CONSISTENCY
2014 English Language	1.017	12.929	119.115	0.109	0.891	0.907
2015 English Language	1.017	12.747	113.316	0.112	0.888	0.903
2016 English Language	1.017	12.577	79.870	0.157	0.843	0.857

In table 1, Kuder-Richardson 20 reliability indicated that:

- JSSCE 2014 English Language had a variance of 119.115, standard deviation of 10.914 and an internal consistency of 0.907;
- JSSCE 2015 English Language had a variance of 113.316, standard deviation of 10.645 and an internal consistency of 0.903;
- JSSCE 2016 English Language had a variance of 79.870, standard deviation of 8.937 and an internal consistency of 0.857;

Since all the tests had high internal consistency, all the tests are reliable. The table further reveals that JSSCE 2014 English Language had the highest co-efficient, thus having the highest reliability. JSSCE 2016 English Language had the least co-efficient of 0.857, thereby having the least reliability. Furthermore, JSSCE 2015 English Language had a co-efficient of 0.903.

Table 2: Weights assigned to weeks used to teach each topic in JSS Scheme of Work and topics in the objective questions JSSCE English Language from 2014 to 2016.

S/N	TOPICS	Weights Based on Number of Weeks Used to Teach Topic	Weights Based on Number of Questions from Each Topic	Weights Based on Number of Questions from Each Topic	Weights Based on Number of Questions from Each Topic
		Scheme of Work	2014 English Language	2015 English Language	2016 English Language
1.	Comprehension/Listening and speaking	25 (35%)	8 (13%)	8 (13%)	8 (13%)
2.	Intonation, stress and rhythm	3 (4%)	-	-	-
3.	Part of speech	8 (11%)	25 (42%)	28(47%)	24 (40%)
4.	Letter writing	2 (3%)	-	-	-
5.	Consonant sounds/Vowel Sounds/Diphthongs	17 (24%)	5 (8%)	3 (5%)	5 (8%)
6.	Literature:	2 (3%)	10 (17%)	10 (17%)	10 (17%)
7.	Questions and questions tags	4 (6%)	-	-	1 (2%)
8.	Identifying main ideas	2 (3%)	10 (17%)	9 (15%)	10 (17%)
9.	A speech	2 (3%)	-	-	-
10.	Commands	2 (3%)	2 (3%)	2 (3%)	2 (3%)
11.	Compound words	3 (5%)	-	-	-
		70 Weeks	60 Items	60 Items	60 Items

Table 2 indicated that the scheme of work of English for JSS 1 to 3 covered 11 topics. The weights assigned to each of the topics were based on the number of weeks spent in teaching the topics (70 weeks). It also shows that the 60 objective questions of JSSCE English Language 2014 to 2016 covered 11 topics. Weights assigned to each topic were based on the number of questions arising from each topic. The JSS 1- 3 Scheme of work for English Language had the greatest weight to be Comprehension, Listening and Speaking, followed by Consonants, Vowels

and Diphthongs. The least weight was Literature, Letter Writing. The objective questions for the 3 years being studied had Parts of speech as having the greatest weight, followed by Literature, Letter Writing. Unlike the scheme of work, the JSSCE English Language had Command as having the least weight.

Table 3: Difficulty indices of objective test items of JSSCE English Language

	ENGLISH LANGUAGE		
	2014	**2015**	**2016**
Good Items	45	41	36
% Good Items	75%	68%	60%
Difficult Item	14	18	24
% Difficult Items	23%	30%	40%
Easy Item	1	1	-
% Easy Item	2%	2%	-

In table 3, it was revealed that JSSCE 2014 English Language had 14 (23%) items that were difficult, 45 (75%) items adequate for JSS 3 students and 1 (2%) was too easy. JSSCE 2015 English Language had 18 (30%) difficult items, 41 (68%) adequate items and 1 (2%) very easy item. For JSSCE 2016 English Language, 24 (40%) items were difficult, 36 (60%) adequate.

Table 4: Discrimination Indices of Objective Test Items of JSSCE English Language

	ENGLISH LANGUAGE		
	2014	**2015**	**2016**
Good Items	49	45	38
% Good Items	82%	75%	63%
Marginal/Poor Items	11	13	20
% Marginal/Poor Items	18%	22%	33%
Bad Items	-	2	2
% Bad Items	-	3%	3%

In table 4, the JSSCE 2014 English Language had 6 poor items, 5 marginal items and 49 (82%) good items; 2015 English Language had 2 bad items, 6 poor items, 7 marginal items and 45 (75%) good items; 2016 English Language had 2 bad items, 12 poor items, 8 marginal items and 38 (63%) good items.

Table 5: Distracter indices of objective test items of JSSCE English Language

	ENGLISH LANGUAGE		
	2014	**2015**	**2016**
Effective Items	58	59	58
% Effective Items	97%	98%	97%
Ineffective Items	2	1	2
% Ineffective Items	3%	2%	3%

Table 5 showed that JSSCE 2014 English Language had 2 items that did not distract effectively, while 58 items were effective distracters. For 2015 English Language, an item was not an

effective distracter while 59 items distracted effectively. The JSSCE 2016 English Language had 2 items that did not distract effectively while 58 items distracted effectively.

DISCUSSION OF THE FINDINGS

The internal consistency reliability analysis of English Language JSSCE within the three years under review, using Kuder-Richardson 20 (K-R$_{20}$) indicated that from the coefficient of internal consistency of the 3 question papers (0.907, 0.903 and 0.857), there was consistency among the items that constitute each of the tests. This agrees with Kpolovie (2014), that he higher the reliability of a test, the higher the quality of the test.

There was consistency in the weights of the test in the three years under study, but they were not parallel with the scheme of work for JSS 1 to JSS 3. This implies that not all that were covered in the scheme of work were set in the Junior Secondary School Certificate Examination. The opinion is that before a student sits for the examination, he must have been exposed to all the topics covered in Junior Secondary School syllabus: the standard set for Junior Secondary School Certificate Examination.

Based on the categorizations of Sidhu as documented in Orluwene (2012), items with coefficients ranging between 0.4 and 0.9 are recommended, less than 0.4 are difficult items and greater than 0.9 are too easy. JSSCE 2014 English Language had 14 items that were difficult, hence they do not qualify for selection based on this criterion; 1 item was too easy to be included in the test. The remaining 45 items (75% of the 60 items) were adequate to be included in the test, which is good. The easy item (item 46) did not actually have the correct answer given, so a bonus mark was awarded all the examinees, making all of them to pass it, whether they ticked it or not. JSSCE 2015 English Language had 18 difficult items out of the 60 items, and thus do not qualify to be selected; 1 item was very easy to be included in the test. The remaining 41 items (68% of the 60 items) were adequate to be included in the test, which is good. The easy item (item 41) was also a bonus mark. In JSSCE 2016 English Language, 24 items out of the 60 items were difficult, hence they do not qualify for selection based on this criterion. The remaining 36 items (60% of the 60 items) were adequate to be included in the test, which is good. Of all the tests, JSSCE 2014 English Language had the least difficult items.

Also, based on the categorizations of Ebel and Frisbie as documented in Orluwene (2012), items with coefficients greater than 0.40 are very good; between 0.3 and 0.39 are fairly good; between 0.20 and 0.29 are marginal and needs some revision or eliminated; below 0.19 are poor and need major revision; items with a negative discrimination index are bad and should be eliminated. In JSSCE English Language 2014, 6 items were poor items, while 5 items were marginal items. These 11 items need to be reviewed and tried again. The remaining 49 items passed the discrimination indices criteria and represent 82% of the total number of questions constituting the test, which is good. JSSCE English Language 2015 had 2 items were bad items, 6 items were poor items, while 7 items were marginal items. The 2 bad items failed in the discrimination indices criteria, while the 13 poor and marginal items need to be reviewed and tried again. The remaining 45 items passed the discrimination indices criteria and represent 75% of the total number of questions constituting the test, which is good. In JSSCE English Language 2016, 2 items were bad items, 12 items were poor items, while 8 items were marginal items. The 2 bad items failed in the discrimination indices criteria, while the 20 poor and marginal items need to be reviewed and tried again. The remaining 38 items passed the discrimination indices

criteria and represent 63% of the total number of questions constituting the test, which is good. JSSCE 2014 English Language had more items that discriminated effectively.

Based on the recommendations of Mozaffer and Farhan (2012) for distracter indices, options should be selected by at least 5% (0.05) of the students. For 2014 JSSCE English Language, options B, C, D, E in item 26 and item 46 did not distract effectively. These options should be changed. The remaining options in the 58 items passed the distracter indices criteria, representing 97% of the 60 items. In 2015 JSSCE English Language, item 41 did not distract effectively. This option should be changed. The remaining options in the 59 items passed the distracter indices criteria, representing 98% of the 60 items. For 2016 JSSCE English Language, options A, B, C, E in item 1 and option B in item 22 did not distract effectively. These options should be changed. The remaining options in the 58 items passed the distracter indices criteria, representing 97% of the 60 items.

Implications of the Findings

The findings of the study imply that:

- The JSSCE English Language objective questions from 2014 to 2016 are not valid and reliable instruments; and
- The English Language course contents were not adequately covered in the JSSCE.

Recommendations

It was recommended that:

- Experts in measurement and evaluation should be used to carry out the rigorous process of item analyses by determining item difficulty, discrimination and distracter indices, not by hand-picking items from past questions.
- The wordings and clarity of the JSSCE test items should be revisited by psychometricians.
- Care should be taken to ensure that test options should be arranged in order so that the testees and the markers would not be confused.
- Questions that failed to meet the difficulty, discrimination and distracter indices should be eliminated.

CONCLUSION

Based on the findings of this study, it was concluded that the English Language Junior Secondary School Certificate Examination (JSSCE) from 2014, 2015 and 2016 lacked content validity. The co-efficient of internal consistency of each of the instruments was high among items that constitute each of the items. Based on the criteria set for item analyses, all the tests had good difficulty and discrimination indices. JSSCE 2014 English Language had the highest internal consistency, the best difficulty and discrimination indices, indicating that it was the best test among the three tests. All the tests passed the distracter indices criteria; The difficulty and discrimination indices observed could be attributed to unfamiliarity of the questions to the

students, test questions may have been too challenging relative to the ability of the students, some items were ambiguous, not clearly written.

REFERENCES

Danladi, S. S. (2013). Language policy: Nigeria and the role of English Language in the 21ˢᵗ century. *European Scientific Journal, 9*(17).

Davidshofer, K. R., & Murphy, C. O. (2005). *Psychological testing: principles and applications (6ᵗʰ ed.).* Upper Saddle River, N. J: Pearson/Prentice Hall.

Federal Republic of Nigeria (2004). *National Policy on Education (4ᵗʰ ed.).* Lagos: NERDC Press.

Igbokwe, C. O. (2015). Recent curriculum reforms at the basic education level in Nigeria aimed at catching them young to create change. *American Journal of Educational Research, 3*(1), 31 -37.

Iweka, F. (2014). *Comprehensive guide to test construction and Administration.* Omoku: Chifas.

Kpolovie, P. J. (2010). *Advanced research methods.* Owerri, Imo State: Springfield publishers.

Kpolovie, P. J. (2014). *Test, measurement and evaluation in Education (2ⁿᵈ ed.).* Owerri, Imo State: Springfield publishers.

Kpolovie, P. J., Iderima, C. E., & Ololube, N. P. (2014). Computer literacy and candidate performance on computer-based tests. In N. P. Ololube (Ed). *Advancing technology and educational development through blended learning in emerging economies,* (pp. 80-106)*.* Hershey, PA: Information Science Reference DOI: 10.4018/978-1-4666-4574-5.ch005.

Mozaffer, R. H., & Farhan, J. (2012). Analysis of one-best MCQs: the difficulty index, discrimination index and distracter efficiency. *Journal of Pakistan Medical Association*. (n.p).

Nkwocha, P. C. (2004). Measurement and evaluation in the field of Education. In Nwaogazie, J. I. Influence of content validity on secondary school students' academic achievement in Imo State, Nigeria. *JORIND, 12*(1), 182 – 190.

Onunkwo, G. I. N. (2002). Fundamentals of educational measurement and evaluation. In Iweka, F. *Comprehensive guide to test construction and Administration.* Omoku: Chifas.

Opara, I. M. (2016). *Test construction and measurement: concepts and applications.* Owerri: Career Publishers.

Orluwene, G. W. (2012). *Introduction to test theory and development process.* Port Harcourt: Chris-Ron integrated services.

Orluwene, G. W., & Igwe, B. N. (2016). Psychometric properties of WASSCE questions in Civic Education (2015) using Rivers State students. *International Journal of Psychology and Counselling, 8 –15*(1), 92 – 115.

Ugodulunwa, C. A. (2008). *Fundamentals of educational measurements and evaluation.* Jos: Fab Anieh.

Wolming, S., & Wilkstrom, C. (2010). The concept of validity in theory and Practice. *Assessment in Education: Principles, Policy and Practice, 17*, 117 – 132.

----

[i] Josephine Nkem Bosede Joe-Kinanee is of the Department of Educational Psychology, Guidance & Counselling, Faculty of Education, University of Port Harcourt, Nigeria.

[ii] Dr. (Mrs.) Opara, Ijeoma Margaret is a lecturer in the Department of Educational Psychology, Guidance & Counselling, Faculty of Education, University of Port Harcourt, Nigeria.