

ANALYSIS OF THE DISCRIMINATION INDEX AND ITEMS DIFFICULTY OF THE DELTA STATE JSCE MATHEMATICS MULTIPLE CHOICE ITEMS FROM 2007 TO 2009

Stephen T. Ogunkoru

ABSTRACT

This study assessed the quality or adequacy (i.e the difficulty level and discriminating power) of the Delta State Junior Secondary School Examination Test items in mathematics for three years, 2007, 2008 and 2009. This is in order to provide an empirical base for determining the standard of students' performance in the examination and give adequate information on the reliability of the JSCE to the stakeholder. The researcher used the ex-post-facto research method to extract data from Delta State Ministry of Education records, which have been obtained for the said three years. Simple random sampling technique was employed to select 320 students from four schools from four Local Government Area in Delta North Senatorial District; using multiple-stage random sampling technique. The results of the study showed that the JSCE Mathematics multiple choice questions had more difficult items and more items had low discrimination level. Based on the findings, recommendations were made.

The first three years of secondary education in Nigeria is known as the Junior Secondary School, which is normally rounded up with a uniform examination taken by those who attain that level. The State and Federal Ministries of Education conducts this examination called the Junior School Certificate Examination (JSCE). The Delta State Ministry of Education is solely responsible for the conduct of the JSCE in both public and private secondary schools in the state, this started in 1988.

Stanley (1981) believed that educational assessment includes all the processes and products, which describe the nature and extent of children's learning. Its degree of correspondence with the aims and objectives of learning and its relationship with the environments whose purpose is to make learning easier, he justifies assessment in education because of the need to evaluate the effectiveness of teaching, provide incentive and feed back information to children, certification, performance, obtain data for accountability and ensure the maintenance of educational standards.

The major measuring instrument of cognitive assessment is the test. A test can be regarded as an instrument made up of questions or tasks designed and presented to individuals to respond to independently and the results of which can be used to determine quantitative academic change in individuals and for the quantitative comparison of performance of different individuals or their level of achievement. Ughamadu, Onwuegbu and Osunde (1991) believed that, it is an attempt to ascertain how an individual will function in a set of actual situation and a systematic procedure for comparing the behaviour of two or more persons. A test consists of a group of items that are designed to be taken under specific condition. The outcome of which are used to evaluate the performance of the person giving the response where the response is taken as an index of how much the respondent has learned. Testing is based on the assumption that people vary in their behaviors in the extent to which they

make or choose the most appropriate response to any given situation and that, these differences are generalized to an extent.

Achievement tests however are designed to measure what a student has learnt to do or to indicate the degree of success in some past learning. Tests may be used for the purposes of examination, selection, diagnosis, guidance, research, evaluation and the test characteristics (reliability, difficulty and discrimination indices). Test may also be classified as maximum performance test as objectives or subjectives depending on the form of response, as written or oral, standardized or formal, verbal or non verbal and as well as culture fair or culture laden tests.

Tests construction in the psychometric sense is very complex and requires a high degree of personal knowledge. Owie (1997) asserted that to be good and effective, a teacher should be able to demonstrate confidence in his ability to critically evaluate the effectiveness of his teaching.

Mehrens and Lehmann (1993) also asserted that good test do not just happen, that it requires adequate and extensive planning so that the goals of instruction (objective), the teaching strategy to be employed, the textual materials and the evaluation procedures are related in some meaningful fashion, so to have a good test, one has to plan for it, for poor planning leads to the inclusion of ambiguous, poorly worded and grammatically incorrect items in the test and of test items that are either too easy or too difficult, not scorable to those that are measuring trivial details. Time, energy and proper planning are required to produce valid, reliable and objectively scorable test items.

In formal test construction as stated by Pyrozak (1983) project drafts of test items are usually administered to same examinees representative of those with whom the items ultimately are to be used. On the bases of the examinees response, items analysis are carried out to ascertain whether or not a particular test item contributes it's appropriate quota to the entire test and to estimate the difficulty level of each test item.

At the end of the items analysis the test items are listed according to their degree of difficulty (easy, medium, hard) and discrimination (good, fair, poor,). This distribution provides a quick overview of the test and can be used to identify items which are not performing well and which can perhaps be improved or discarded.

Whatever items format is used in a test, the items must be valid and reliable. This is only achieved as stated by Mehrens and Lehman (1993) such test items must be expressed in clear unambiguous language, the students are not given any clue to the correct answer and a table of specification has been prepared and followed.

It also follows that items must maintain some specific level of difficulty if it is to discriminate appropriately. According to Owie (1997). A test should be difficult enough such that the brighter students can effectively attack it while a good proportion of the weaker students should also be able to handle the problem adequately.

Statement of the Problem

The National Policy on Education (1998 revised) classified mathematics as one of the core subjects to be offered by all students from primary to senior secondary school level of education. Since the inception of the junior secondary school certificate examinations in Delta State in 1988, student's performance in mathematics has not been quite encouraging. Every year the number of students who write the resit examination in the subject are on the increase despite the fact that the subject is taught everyday in schools and as stated by Ojogwu (2001) that the teachers in Delta State

appear appreciably satisfied with their job and their task performance was relatively high. This is due to the fact that their conditions of service are fairly good, salaries and promotions are regular.

With the emphasis placed on the teaching and learning of mathematics as well as the usefulness of the subject and the good condition of service of Delta State teachers, it is expected that not fewer than 90% of the students should perform above average in the JSCE mathematics. However, this is not the case. Several factors have been advanced as reasons for poor academic performance in mathematics in secondary schools, generally. Among them are lack of interest in the subject, the nature of the subject, the subject itself, teacher related factors etc.

Could the nature of the questions asked in the JSCE Mathematics be responsible for students' poor performance and for the subject tagged "difficult"? Therefore, there is the need to investigate the level of validity of the items difficulty and discrimination indices of the Delta State JSCE mathematics multiple choice test items between 2007 and 2009.

Research Questions

The following research questions were raised.

1. What is the difficulty level of the JSCE mathematics between 2007 -2009?
2. Does the discrimination index of the multiple-choice items in the JSCE mathematics differ significantly between 2007-2009?
3. What is the reliability coefficient of the JSCE mathematics multiple choice test items between 2007-2009.

Research Design

The research design used for the study is the expost-facto research method because variables cannot be manipulated.

Population and Sample

The target population for this study consists of the students in JSS III attending public secondary schools in Delta North Senatorial District.

The sample for the study consists of 320 students selected from JSS III using multistage random sampling technique.

Firstly, 4 Local Government Areas, 4 schools were selected using simple random sampling procedure and from each school 20 students were selected through simple random sampling technique. Thus a total of 320 students constitute the research sample for the study.

Research Instrument

The research instruments used for this study were the JSCE mathematics multiple-choice question papers for the years 2007, 2008 and 2009.

The content validity of the instrument was determined by graduate mathematics teachers who are involved in the generation of the JSCE mathematics questions. This was done by comparing the test items with junior secondary school syllabus in mathematics.

Split-half methods of estimating reliability was used. A single test was administered and then split into two halves and their correlation determined.

Data were collected by the researcher from the Delta State Ministry of Education for the JSCE mathematics examination conducted by the Ministry of Education from 2007 to 2009.

The statistical analyses used included item difficulty, simple percentage and discrimination indices. These are calculated for three years (2007-2009).

Research Question One

What is the difficulty level of the JSCE mathematics multiple choice items between 2007 and 2009?

To answer this question the data collected were analyzed by finding the difficulty and discrimination indices of the test items for the three-year period.

**Table1:
Difficulty Index and Discrimination Index of the JSCE Mathematics Multiple Choice Questions for June 2007**

Difficulty Index			Discrimination Index		
Range	Freq	Percentage	Range	Freq	Percentage
0.00-0.40	42	70	0.0-0.50	48	80
0.41-0.70	16	27	0.51-1.00	10	17
0.71-1.00	2	03	Negative	2	03
Total items	60	100		60	100

From the total of 60 item, 42 items (70%) are within the range of 0.00- 40, 16 items (27%) are in the range 0.41-0.70 while only 2 items (3%) fall within the range 0.71-1.00

The table also revealed that 48 items (80%) fall within the range of discrimination index 0.00 – 0.50 while 10 items (17%) are in the range 0.51- 1.00. the table also show that 2 items (3%) were negative.

**Table 2:
Difficulty Index and Discrimination Index of the JSCE Mathematics Multiple Choice Questions for June 2008**

Difficulty Index			Discrimination Index		
Range	Freq	Percentage	Range	Freq	Percentage
0.00-0.40	48	63	0.0-0.50	51	85
0.41-0.70	18	30	0.51-1.00	08	13.3
0.71-1.00	4	07	Negative	01	01.7
Total items	60	100		60	100

A close look at the table shows that out of the 60 items 38 items (63%) were within the range of difficulty 0.00 -0.40, 38 items are in the range 0.41- 0.70, while 4 items (07%) fall within the range 0.71-1.00.

The table also shows that 51 items (85%) fall within the range of discrimination index 0.00 – 0.50 while 8 items (13.3%) are in the range 0.51 - 1.00 and I item 1.7% was found to have a negative discrimination.

Research Question Two

Does the discrimination index of the multiple-choice items in the JSCE mathematics differs significantly between 2007 and 2009.

**Table 3:
Difficulty Index And Discrimination Index of the JSCE Mathematics Multiple Choice Questions for June 2009.**

Difficulty Index			Discrimination Index		
Range	Freq	Percentage	Range	Freq	Percentage
0.00-0.40	38	65	0.0-0.50	59	98.3

Analysis of the Discrimination Index and Items Difficulty of the Delta State JSCE Mathematics Multiple Choice Items From 2007 To 2009

0.41-0.70	18	30	0.51-1.00	1	1.7
0.71-1.00	3	05	Negative	-	-
Total items	60	100		60	100

Table 3 reveals that out of the 60 items in the 2009 JSCE mathematics multiple choice items 39 items fall within the range of difficulty level of 0.00 to 0.4 representing 65%. 18 items representing 30% fall within the range 0.41 – 0.70 while 3 items, representing 5% fall within the range of difficulty 71-1.00.

The table also shows that in terms of discrimination 59 items fall within the range 0.00 – 0.5. This represents 98.3% of the item while no items had a negative discrimination.

Research Question Three

What is the reliability coefficient of the JSCE mathematics multiple choice questions between 2007-2009.

To answer this question the split-half method of estimating reliability was employed. The data collected was split into two halves and was tabulated and the pearson product moment method of estimating reliability was used to find the reliability indices for the test. Table 4 shows the results obtained.

**Table 4:
Reliability Coefficient of the JSC Mathematics Multiple Choice Question From 2007-2009.**

Year	Reliability
2007	0.77
2008	0.85
2009	0.83

From the table above we observe that the JSC mathematics multiple choice question for 2007 is 0.77, 2008 is 0.85 while 2009 is 0.83.

Discussion of Results

From the results in table .1, which is on the difficulty and discrimination indices of the JSCE examination, it was observed that in the 2007 multiple items 42 items were within the range of difficulty level 0.00 to 0.40. This accounted for 70% of the items, 16 items representing 27% were within the range difficulty level 0.41 to 0.70 while 2 items accounting for 3% of the total items were between the range 0.71 – 1.00.

Table 2 shows that in the 2008 JSCE mathematics multiple choice items, 38 items accounting for 63% were in the range of difficulty level 0.00 to 0.40 for range 0.41- 0.70 there were 18 items representing 30% while for the range 0.71 to 1.00, 4 items which accounted for 7% were found.

The results in table 3 showed that 39 items, accounting for 65% fell within the range of difficulty level of 0.00 – 0.04, 18 items representing 30% were within the range 0.41 – 0.70 and 3 items, representing 5% were within the range of 0.71 – 1.00.

Ughammadu (1991), on the interpretation of difficulty index of a test, noted that a test is termed “easy” if its items have high difficulty level, from tables 4.1 we can see that items 26, 27, 38, 39 42, 44, 53, 56, and 60 were difficult, while item 5 appears easy.

From table 2 items 8, 23, 27, 30, 31, 40, 42, 44 53, and 57 were difficult items while items 16 and 17 were easy.

In table 3 items 5, 6, 25, 23, 37, 38, 39, 40, 42, 43, 44, 48, 54, and 60 were difficult while items 9 and 24 were easy. Those items that fell within the range of 0.00 - 0.40 for the years 2007 – 2009 were regarded as very difficult. From the analysis above, it can therefore be said that the JSCE mathematics multiple choice questions from 2007 to 2009 were difficult.

Also from the table .1, it is observed that 48 items fall within the discrimination range of 0.00 – 0.50 while only 10 fall with in the range of 0.51-1, 000 which showed that more items had low discrimination index.

In table 2, 51 items fall within the range of 0.00 – 0.50 while 8 items fall within the range of 0.51 to 1.00 which also showed poor discrimination index.

In table 3 within the range of 0.900 – 0.50, there were 59 items while only I was in the range of 0.51 – 1.00 which once again showed that the items discriminated badly.

Mehrens and Lahmann (1993) posited that extremely difficulty or easy items will have low ability to discriminate, that they are needed, to adequately sample course content and objectives. They also claimed that an item may show low discrimination if the test measure many different content area and cognitive skills.

Tables 1 and 2 revealed that there were 1 item and 2 items respectively that had negative discrimination. A negative discrimination occurs when more students in the lower group answered an item correctly than those in the higher group.

According to Lord (1955), items with negative discrimination indices should be examined to determine why a negative was obtained, may be the items were miskeyed, but Ughamadu (1991) suggested that such items should be eliminated as they have discriminated in the wrong direction.

It can therefore be inferred that due to the difficult nature of the JSCE mathematics multiple choice items from 2007 to 2009 there was low item discrimination for the three years period.

The results on table 4 revealed that the reliability coefficient of the JSCE mathematics multiple choice items as 0.77, 0.85 and 0.83 for the year 2007, 2008 and 2009 respectively. This shows that the questions were reliable.

Summary

This study focused on the analysis of the discrimination index and items difficulty of the Delta State JSCE mathematics multiple choice test items from 2007 - 2009. It was designed to answer the following research questions.

1. What is the difficulty level of the JSCE in mathematics between 2007- 2009
2. Does the discrimination index of the multiple choice items in the JSCE mathematics differs significantly from 2007-2009.
3. What is the reliability coefficient of the JSCE mathematics multiple choice test items between 2007- 2009.

The result of the study revealed that:

1. The JSCE mathematics multiple choice questions from 2007to2009; had more difficult items
2. More items had low discrimination level
3. Three items, one from 2008 and two from 2007 discrimination were negative discrimination.
4. The JSCE mathematics multiple choice question from 2007 to 2009 had high reliability coefficients and so the questions were reliable.

Conclusions

From the findings of the study, the following conclusions were arrived at;

1. The JSCE mathematics multiple choice questions between 2007 and 2009 were difficult for the students.
2. The difficult nature of the questions could also be responsible for the low discrimination level.
3. The JSCE mathematics multiple choice questions between 2007 and 2009 were highly reliable.

Recommendations

Based on the findings and conclusions derived from this study, the researcher wish to make the following recommendations.

1. Those involved in developing questions for the Delta State JSCE should be well educated on the procedures for test construction by organizing seminars and workshops in test construction and validation for them. This also should be extended to secondary school teachers so as to assist them in drawing questions for their school examinations
2. The Delta State Ministry of Education should employ experts in measurement and evaluation who will be able to draw questions with appropriate levels of difficulty. This is to ensure better students performance in examinations
3. Students on their part should be well counseled on the need to be serious with their studies particularly in mathematics as it is the mother of all subjects.
4. It is also essential that, the test items are pretested on different but similar group of testers and detailed analysis carried out on each item to ensure the quality of each item.

References

- Federal Republic of Nigeria (1981). *National Policy on Education*, Lagos Federal Ministry of Information.
- Lord, E.M. (2001). *The relationship of the reliability of multiple choice test to the distribution of item difficulties*. Oxford Back Well Press Ltd,
- Mehrens, W.A. & Lehmann, F.J. (1993). *Measurement and evaluation in education and psychology* USA Holty Rhinehard and Winston Inc.
- Ojogwu, C. N. (2001). Relationship between teachers morale and their test performance in Secondary Schools in Delta State. *Journal of Nigerian Educational Research Association*. 15 (1).
- Owie, I. (1997). *Measurement and evaluation principles and procedures*. Mindex Publishing Ltd. Benin City.
- Pyrozak, F. (1993). Validity of the discrimination index as a measure of item quality *Journal of educational measurement* 10 (3).
- Stanley, J. C. (1981). *Assessment in schools*. Oxford Blackwell Press Ltd.
- Ughamadu, K. A. Onwuegbu, O.C. & Osunde, A. U. (1991). *Measurement and evaluation*, Benin City. Word of books.