Educational Research and Review Vol. 4 (11), pp. 585-587, November 2009
Available online at http://www.academicjournals.org/err
ISSN 1990-3839 © 2009 Academic Journals

Short Communication

Reliability and levels of difficulty of objective test items in a Mathematics achievement test: A study of ten senior secondary schools in five local government areas of Akure, Ondo State

S. O. Adebule

Faculty of Education, University of Ado-Ekiti, Nigeria. E-mail: doctorfemiadebule@yahoo.com.

Accepted 28 September, 2009

This study examined the reliability and difficult indices of Multiple Choice (MC) and True or False (TF) types of objective test items in a Mathematics Achievement Test (MAT). The instruments used were two variants- 50-items Mathematics achievement test based on the multiple choice and true or false test formats. A total of five hundred (500) students randomly selected from ten senior secondary schools in five Local Government Areas of Akure, Ondo State served as sample for the study. The results of the null hypotheses tested at p=0.05 showed that there was no significant difference between the reliability coefficients of MC and TF test items. Also, there was no significant difference between the difficult indices of MC and TF test items. Hence, it was recommended that a combination of both MC and TF types of tests should be used to evaluate students' achievement in schools.

Key words: Reliability indices and coefficients, difficulty indices, achievement tests, multiple choice and true or false tests.

INTRODUCTION

Teaching and testing go together in schools. The preparation and writing of questions in our schools are the sole responsibilities of the subject teachers. According to Adebule (1995), majority of the teachers do no give adequate attention and due consideration to the quality of the questions they set. Most of the question items are not reliable while some items are either too difficult or too simple.

In addition, majority of the teachers take on cognizance of the merits and demerits of the test formals they use. Most often, test formals that come to their minds are used without any consideration for its appropriateness for the purpose of the test.

Some of the psychometric properties that every measurement procedure could posses are the qualities of reliability validity and usability or practicality. Ali (1988) and Adebule (2004) concluded that a test designed to measure such characteristic traits as ability, aptitude and achievements should focus attention on how much an individual has accomplished on a course of instruction or training on the above characteristics.

There are multifarious purposes for which a test is given to students. Test items are indispensable tools in the evaluation of students' achievement at school. According to Nwaobia (1990), item analysis (difficult and discriminating indices) is concerned with ascertaining the worth of the test items. A good test should be able to differentiate the brilliant students from the dull ones. This can only be realizable when carefully constructed tests are set, administered, marked and scored. Tests that are too difficult or too simple rarely make effective evaluation possible.

The National Policy on Education (1998) emphasizes continuous assessment of the progress of the students on a continual basis. This has to be systematic, comprehensive, and cumulative and guidance oriented (Alonge 1989; Adebule, 2004). Also in consideration of the increase in student's population in schools today, objective test items which include Multiple Choice (MC) and True or False (TF) become very useful and expedient assess-ment techniques and efficacious instruments to evaluate a large sample of students on a wide expanse of course contents and process objectives. Teachers would find it more convenient to assess students with MC and

Appendix 1. Difficult indices of the MC and the TF items.

S/N	MC	TF	S/N	MC	TF
1	0.63	0.76	26	0.24	0.61
2	0.33	0.56	27	0.45	0.43
3	0.63	0.87	28	0.33	0.63
4	0.61	0.58	29	0.35	0.65
5	0.41	0.57	30	0.44	0.44
6	0.35	0.57	31	0.43	0.57
7	0.52	0.69	32	0.46	0.59
8	0.28	0.69	33	0.35	0.59
9	0.39	0.43	34	0.33	0.52
10	0/35	0.54	35	0.20	0.54
11	0.35	0.67	36	0.56	0.33
12	0.67	0.48	37	0.2	0.43
13	0.57	0.41	38	0.22	0.48
14	0.44	0.56	39	0.35	0.41
15	0.41	0.61	40	0.33	0.37
16	0.65	0.82	41	0.41	0.39
17	0.48	0.57	42	0.17	0.56
18	0.50	0.61	43	0.2	0.52
19	0.39	0.54	44	0.22	0.67
20	0.41	0.67	45	0.37	0.67
21	0.48	0.57	46	0.24	0.48
22	0.35	0.61	47	0.31	0.48
23	0.13	0.48	48	0.19	0.33
24	0.43	0.54	49	0.28	0.41
25	0.48	0.61	50	0.5	0.57

TF tests than essay test.

To this end, this study examined the reliability and levels of difficulty of MC and TF test items in a Mathematics achievement test so as to be able to make appropriate conclusion and recommendations on the suitability of the test format to evaluate students' achievement on which decisions are based (Appendix1).

Statement of the problem

The major problems of this study include:

To examine the coefficients of reliability of MC and T F test items in a Mathematics achievement test.
To determine the difficult indices of MC and TF test items in a mathematics achievement test.
To compare statistically, the equivalence or otherwise of the reliability and indices of MC and T F test formats.

Research hypotheses

Two hypotheses were generated from the above problems and tested for significance are 0.05 alpha levels.

Table 1. Research design.

Group	Item format
X	Multiple choice (MC)
Y	True False (T F)

Ho1: There is no significant difference between the reliability coefficients of MC and TF test items in a Mathematics achievement test
Ho2: There is no significant difference between the difficult indices of MC and TF test items in a Mathematics achievement test

RESEARCH DESIGN

This is a descriptive research of the survey type, in accordance with the format shown below (Table 1).

Population and sample

The population for the study consisted of all senior secondary school three students who were preparing for the Senior School Certificate Examination (SSCE) in Mathematics. A total of 500 students selected through stratified random sampling technique from ten schools in Akure North, Akure South, Owo, Akoko North and Akoko South Local Government Areas of Ondo State served as sample for the study.

Instrument

The instruments used were two variants; MC-MAT and TF-MAT that contained 5-0 items each based on the current SSCE mathematics syllabus. All questions were to be answered and there was no penalty for guessing.

The MC and TF test formats were certified to experts in mathematics and psychometric and thus found to be suitable for administration. Also the pre test carried out showed that the difficult indices of MC and TF test items ranged between 0.30 and 0.70 while the reliability coefficients were 0.83 and 0.79 for MC and TF test, respectively

The students sat for the test without been told of the discrepancy in the question format. After one hour, the scripts were collected and marked according to the keys certified to be the correct responses.

ANALYSIS OF DATA AND RESULTS

To test hypothesis one, a split half reliability coefficient was computed for each of the MC and TF formats using Pearson Product Moment Correlation analysis. Also Spearman- Brown prophecy formula was used to adjust the full- range reliability coefficients. Later equivalent Fishers Z-Transformation for the two Pearson's r's were computed. The Z equivalent values got were also subjected to student t- test statistics which was finally used to ascertain if any significant difference exists between the computed coefficients and the Table values. The result is as shown in Table 2.

Table 2. Differences between the reliability coefficients of MC and T F test formats.

Item format	N	R	Z	df	t_{cal}	$t_{critical}$
MC	50	0.35	0.3654	98	0.524	1.980
T F	50	0.25	0.2554			

P > 0.05 (Results not significant).

Table 3. Differences between the difficult indices of MC and TF test items in a Mathematics achievement test.

Format	N	X	SD	df	t_{cal}	$t_{critical}$
MC	50	0.39	0.13	48	1.60	1.980
TF	50	0.43	0.12			

P > 0.05 (results not significant).

The reliability coefficients for MC and TF were 0.35 and 0.25, respectively. Also was found to be 0.3564 and 0.2554 for both MC and T F. The t –calculated was 0.524 with the critical t value =1.980. Since p > 0.05, it suggests that the null hypothesis is not rejected. Thus there is no significant difference between the reliability coefficients of a-TF and a-MC test formats.

To test hypothesis two, the number of testes with the right or wrong responses to each of the MC and TF items were determined. Next the difficult indices were calculated. Also the mean and standard deviations were found and were compared statistically using parametric t-test

An examination of Table 3 showed that MC and T F have means of 0.39 and 0.43, respectively. This meant that TF items were much easier for the students than the MC test items. The t calculated valve was 1.60 while the t critical valve was 1.980. Since 1.60 < 1.980 the null hypothesis was not rejected. Thus there was no significant difference between the difficult indices of Mc and T F test.

The result of hypothesis one supports Alonge (1988) and Adebule (2004) who found out that there was no significant difference between the reliability coefficients of a 3- alternative MC and 4- alternative MC tests. The relevance of the finding lies in the popularity of MC item coupled with technicalities of constructing very good MC decoys, since it is obviously more convenient to construct a three variable option than four. The implication is that a 3-atternative MC would perform comparable functions with a 4- alternative MC or a 5- alternative MC.

Also the finding of hypothesis two was in line with Hopkins and Stanley (1981) that strongly believed that well constructed TF tests compare favourably with MC test and that TF tests have a legitimate use in achievement tests. Alonge (1988) and Adebule (2004) found that TF test item is psychometrically comparable to a 3-Alternative MC and that a MC format is as dependable as a TF format. Both tests have 0.5 and 0.2 probabilities of guessing items correctly. But TF test format is not as popular as MC format among test users despite the fact that TF format takes a shorter time to answer as well as a shorter space. This study has therefore reveals its suitability, particularly for classroom tests where MC tests might have been used.

Conclusion and Recommendations

Since the finding of the study indicated that MC and TF, reliability coefficients of MC and TF are not significantly different, a classroom teacher or any person involved in the testing process could use any of the two formats to achieve desirable results.

However, both MC and TF formats have their usefulness and shortcomings. So a combination of the two formats should be judiciously used by test users to arrive at desired goals. A variety of test forms can help keep students' interest at a high level in examination situations.

Since test designed for specific purposes utilize items with difficult values or indices that come closest to the desired selection ratio, test items should be easy and simple and so that tests could have a peaceful and less tense psychological environment. Teachers could be given information as regard this and other basic principles of test construction and administration.

The curriculum development and continuous assessment units of the Ministry of Education and Examination bodies like WAEC, NECO, JAMB, NTI etc could use a combination of MC and TF test formats to measure achievement of students.

REFERENCES

Adebule SO (1995). A comparative Analysis of Difficult and discriminating indices of objective test items in a Mathematics Achievement test. An unpublished M.Ed Thesis, Ondo State University, Ado Ekiti

Adebule SO (2004) Relationship between Difficulty and Discriminating indices of Multiple Choice and True False test items in a Mathematics Achievement test, J. Res. Dev. 3(7): 26-30

Ali A (1988) Educational Measurement and Evaluation Awka Meks Unique Publishers.

Alonge MF (1988). A Comparative Analysis of the Reliability and Validity of Multiple Choice and True False Test item in a Mathematics Achievement test. J. Ondo State Math. Assoc. 1(1): 37-42.

Alonge MF (1989). Measurement and Evaluation in Education and Psychology, Ado Ekiti, Adedayo Printing (Nig) Ltd.

Hopkins KD, Stanley JC (1981) Educational and Psychological Measurement and Evaluation. Eaglewood cliffs N.J, Prentice- Hall Inc National Policy on Education (1998)

Nwaobia EM (1990) Monograph on Introduction to Advanced Measurement and Evaluation University of Nigeria, Nsukka.