



**COMPARATIVE EFFECTIVENESS OF CONVENTIONAL TESTING,
LIBERAL MARKING AND CONFIDENCE SCORING ON
PSYCHOMETRIC PROPERTIES OF CHEMISTRY MULTIPLE-CHOICE
TEST ITEMS IN OSUN STATE SECONDARY SCHOOL**

Adebule, Samuel Olufemiⁱ

Faculty of Education, Ekiti State University, Ado-Ekiti, Nigeria

Awodele, Babatunde Adedeji

Faculty of Education, Ekiti State University, Ado-Ekiti, Nigeria

Abstract:

This study compared the effectiveness of Conventional Testing (CT), Liberal Marking (LM) and Confidence Sharing (CS) of scoring objective test in Chemistry. It examined which of the three scoring methods is more effective, reliable and valid. These were with a view to enhancing the reliability and validity of chemistry objective test for standard assessment. The population for the study comprised of senior secondary school three (SSSIII) students in Osun State. The sample consists of 150 SSSIII chemistry students in their intact classes from three randomly selected schools and in three randomly selected local government areas of the state. The classes were randomly assigned to the three scoring methods. The instruments used for this study were Chemistry Multiple-choice Test. The 40-item Chemistry multiple-choice test was administered on the students in each school. Three scoring methods used to score the test items. Data collected were analysed using Kuder-Richardson (kp-21) formula and fisher Z-test. Result obtained revealed that significant different existed in the reliability coefficient of CT, LM and CS ($f_c=1.444195$; $p<0.05$) df (2,147). There is no significant difference existed among CT, LM and CS in the validity of the test score. CS performed better than LM at observed mean different of 2.2.0000 $P<0.05$. While LM is better than CT at mean observed different of 0.78451. It was therefore, concluded that CS could be used to authentically assess Chemistry students' performance and able to identify students with genuine learning difficulties

ⁱ Correspondence: doctorolufemiadebule@yahoo.com

Keywords: conventional testing, liberal marking, confidence sharing, test effectiveness, psychometric properties

1. Introduction

Teaching activity may not be completely assessed. One major instrument for such assessment is test, which Omirin (1999) in Awodele (2013) called systematic method of gathering data for the purpose of making intra and inter comparisons between individuals within class or in a school system. Today multiple-choice test is the most common and widely used assessment tool for the measurement of knowledge, ability and complex outcomes. A multiple-choice item usually consists of a stem which presents a problem situation and several alternatives which provide possible solutions to the problem. The alternatives include the correct answer and several possible answers called distracters.

In multiple-choice test the examinees are prone to blind guessing, which enables testees to be credited with undeserved scores, where an academically poor student scores more marks than knowledge he has in the subject. Therefore making it difficult to discriminate between the bright students and poor student. To preserve the advantages of objective tests in general and that of multiple choice test in particular, to sustain the continued usefulness of multiple choice test, a number of scoring procedures have been developed. Such as liberal marking, logical-choice weight, eliminating testing, confidence scoring, probability testing, partial order, complete order, permutational multiple-choice test is few.

A conventional multiple choice test is one of the most widely used assessment methods. When faced with a question in a conventional multiple choice test, a candidate must evaluate each option and choose the appropriate one. The candidate's test score is usually calculated by tallying the number of correct responses in the test, and is used as a measurement of the candidate's knowledge of the materials and contents covered by the test. However, the number of correctly answered questions may be composed of two numbers, the number of questions where the candidate actually knows the answer, and the number of questions where the candidate correctly guesses the answer. Sometimes, the candidate knows only part of the answer, or is uncertain of the answer. But this partial knowledge is not captured or taken into consideration in the conventional multiple choice test method.

The limitations of the conventional multiple choice test method have been noted and alternative methods for administering multiple choice test that increase the complexity

of responding and scoring have been introduced. There is a variety of multiple choice test methods nowadays to discourage scoring by guessing and give examinees an opportunity to reflect the partial knowledge actually possessed by candidates. Liberal multiple choice test allows candidates to select more than one answer to a question if examinees are uncertain of the correct answer.

The term liberal is used to denote the extra dimension of choices. A method which turns out to be a variation of the liberal multiple choice test is elimination testing, which requires candidates to select the answers which they believe are wrong; rather than selecting those they believe are or may be right.

Today, there is a variety of multiple choice test methods. Comparison of alternative multiple choice test methods with respect to conventional multiple choice test methods choice Bradbard Parker & Stone (2004) found that partial wondering produced higher reliabilities than traditional method, and produced the highest validity coefficient Afolabi (1990) revealed that confidence marking offered advantages over traditional form in term of measurement accuracy. Eliminating testing is equivalent to liberal multiple choice, both liberal and elimination methods are better than.

Conventional on Bradbard, (2004) et al liberal multiple choice was found to reward partial knowledge more generously and punish misinformed examinees more severely than conventional testing found that in elimination procedure. The chance of guessing is reduced and partial knowledge is measured. To sum up, the current finding general not help us identify which multiple choice method is the best to use. A comprehensive evaluation on the typical multiple choice methods, which are conventional or traditional, liberal and confidence marking has not been investigated on a single study in chemistry multiple choice test.

2. Statement of Problem

The current scoring methods in use by the secondary school teachers in measuring students' achievement in chemistry multiple choice test is mainly conventional testing which may not be guess free. Confidence and liberal scoring method has been used to measure the true performance of student in some subjects. There is therefore, the need to find out which of these methods of scoring are more effective in chemistry multiple choice testing and their effectiveness, hence this study compare which of conventional testing, liberal and confidence scoring procedure is the best to score multiple choice test.

3. Purpose of the Study

The broad objectives of this study is to compare conventional testing, liberal marking and confidence scoring on psychometric properties of multiple choice test in chemistry achievement test. The objectives of the study are:

1. Compare the reliability of conventional testing liberal marking and confidence scoring of multiple-choice test.
2. Determine the validity of conventional testing, liberal marking and confidence scoring of multiple choice-test

4. Research Hypothesis

From the purpose of study, the follow non hypotheses were postulated:

1. There is no significant difference in the reliability of conventional testy, liberal marking and confidence scoring of multiple-choice test.
2. There is no significant difference in the validity of conventional testing, liberal marking and confidence scoring of multiple-choice test.

5. Significance of the Study

In conventional method of scoring objectively, tests the testes are prove to greater propensity to do blind guessing. Blind guessing enables the testes to be credited with underserved scores where an academically poor test-wise student would score higher points than the knowledge he has in the subject.

To minimize guessing in multiple-choice test experts have introduced different procedures of scoring objectives test which have been narrowed to liberal marking, conventional testing and confidence scoring. The study is an attempt to compare the typical multiple choice methods and find out a better way of assigning marks to student response in multiple-choice test.

6. Scope of the Study

This study was delimited to secondary schools in Osun state. One school each will be selected in each three senatorial discrete in the state. Chemistry, a qualitative subject to be used and the participants will be senior secondary school three (S.S.3) students. The concepts used are Hydrocarbons and Alkanol, carbohydrate and alkanoate.

7. Population and Sample

The population of the study consisted of all secondary school III (SSSIII) students in Osun State. Three secondary school were selected, one from each senatorial discrete in the state by purposive script to represent the population in which the researcher is interested and for the purpose of convenience to reach out of students. Firmly, from each school 50 students of senior secondary school three classes were selected randomly. All the schools used have been accredited by West African Examinations Council (WAEC) of having minimum requirement for offering Chemistry, in terms of personal and facilities. A simple of 150 students were sampled by purpose and proportional sampling techniques.

8. Research Instrument

The research instrument was 40 items, 4 options multiple-choice test in chemistry. Items were both adopted and adopted from past West African Certificate Examinations. The items covered such consents chemistry they are Hydrocarbons, Alkanol, carbohydrate, Alkanoate, since the items had been validated by West African Examinations Council (WAEC) and standardized, they were administered on the one hundred

9 Data Collection

The 40 items multiple choice test was administered on the students with the assistance of the subjects' teacher in each school. The items were present to the students as mid-term test, answer were indicated on the question sheets. The mid-time hint was liberal the total answer scripts collected were 150. The response, in Ife-Ijesa secretarial district were scored based on conventional method, the students from Osun central secretarial district were scored using liberal method and students from Osun-west secretarial district were scored based on confidence scoring method.

In confidence scoring candidates have to assign a confidence level to their best answer to each question. It starts from absolute confident (AC) to fairly candidate (FC) and not confidence (NC) for instance. *"A mole of neon is (A) diatomic (B) Ionic (C) positively charged (D) Triatomic (D) monatomic (AC) (FC) (NC)"*.

For each question, candidates will be awarded +5 for picking Absolute confidence (AC) and right answers +3 for fairly confident and right answer and +1 not

confident and right answer. Candidate are awarded -2 for panicking (AC) but wrong answer -1 for (FC) but wrong answer and zero for (NC) and wrong answer

10. Liberal Marking

In liberal marking, the candidate's scores depend on the degree of nearest to the correct option. According to Kolawole (2006) in a five option items a candidate with correct option obtained 1 mark, second best option 0.75. third best option 0.50, fourth best option 0.25 and fifth best option zero. The grading of the options should be done by experts in the subject concern.

Cronbach's coefficient alpha is used to estimate the quiz reliability of each testing method. To identify which multiple choice method would bet candidates achieve better performance on the three multiple choice method analysis of variance was used (Anova).

To know which of the method is the best in term of performance Scheffe-multiple was used.

H_0^1 - There is no significant difference among the reliability of conventional testing (CT) liberal marking (LM) and confidence scoring (CS) of multiple choice test.

VAR001	SUM OF SQUARE	DF	MEAN SQUARE	F	P
Between the Groups	123.368	2	61.684	1.4441.95	0.00
Within the Group	0.628	147	0.004		
Total	123.996	149			

$P > 0.05$ $df = 2, 147$

Table 1: ANOVA

At probability level of 0.05 and degree of freedom 2,147. The calculated of value $f_c = 1.44195$ $P < 0.05$, $df (2,147)$. This indicates that a significant difference existed among the three scoring methods (CS, LM, and CT). The multiple comparisons Scheffe was also carried out on the group means for the significant comparison of the mean values of the three groups. The result is presented in Table 2:

Adebule Samuel Olufemi, Awodele Babatunde Adedeji –
 COMPARATIVE EFFECTIVENESS OF CONVENTIONAL TESTING, LIBERAL MARKING AND
 CONFIDENCE SCORING ON PSYCHOMETRIC PROPERTIES OF CHEMISTRY MULTIPLE-CHOICE TEST
 ITEMS IN OSUN STATE SECONDARY SCHOOL

					95% confidence interval	
(1)VAR002	(J) VAR 002	Mean difference (1-J)	Std Error	Sig	Lower bound	Upper bound
1	L.M	-78451	0.1307	0.00	0.8168	-0.7522
	C.S	-2.200	0.1314	0.00	2.2325	-2.1675
2	C.T	0.78451	0.1307	0.00	0.7522	0.8168
	C.S	-1.41549	0.1301	0.00	-1.4477	-1.3833
3	C.T	2.20000	0.1314	0.00	2.1675	2.2325
	L.M	1.41509	0.1301	0.00	1.3833	1.4477

Table 2: Multiple-comparisons Scheffe

From the table 2, the observed difference of -78451 was significant at 0.05 levels ($P < 0.05$). This shows that there was a significant different in the performance of the CT and LM. LM is better them CT. also, between CT and CS, The observed difference of -2.2000 was significant at 0.05 level of significant and CS produce better performance than CT.

Comparing LM and CT observed difference was 0.78451 at significant level of 0.05, ($P < 0.05$) LM is better that CT. In the vain, LM and CS have observed different of -1.41549 at ($P < 0.05$) but CS performed better that LM.

To compare CS and CT, the observed difference was 2.2000 which is significant at $P < 0.05$ and C.S performed better than CT. Also CS also CS also better than LM at observe mean difference of 1.41509.

In conclusion CS performed better than LM at observed mean different of 2.2000 $P < 0.05$. While LM is better than CT and means observed difference of 0.78451 CS thereby, produce a better reliability than LM and CT.

VAR 00001	SS	DF	MS	F	P
Between group	0.005	2	0.003	0.34	0.71
Within group	1.096	147	0.007		
Total	1.111	149			

$P > 0.05$ $df, = 2, 147$

Table 3: ANOVA

At probability level of 0.05 and degree of freedom of value; $f_c = 0.34$ $P < 0.71$, $df(2,147)$, $P > 0.05$. This indicates that no significant difference existed in validity of CT, LM and CS method of scoring objective test. Therefore, the null hypothesis is hereby accepted.

11. Conclusion and Recommendations

The confidence scoring method is more adequate to capture students' genitive status in multiple choice test. Also, it reflects the degree of knowledge students have on the items. Finally, it is relatively less laborious and safe time unlike convention testing and liberal marking.

From the findings, the following recommendations were made:

1. The confidence scoring method should be encouraged and used in schools as it has been found to be effective in reducing contribution of random guessing of the testes on multiple choice tests.
2. Confidence scoring method considerable reduces the 'craze' for a do or die affair to pass examination at all cost, hence should be used in all schools.
3. Public examination councils such as West African Examination Council, the National Examination Council (NECO), Joint Admission and Matriculation Board(JAMB) should apply confidence scoring methods in setting and scoring multiple choice tests.
4. Classroom teacher should be encouraged to develop skills on how to conduct and score multiple choice tests using confidence scoring methods. This can be achieved through organizing seminars and training on testing and scoring techniques.

References

1. Afolabi E.R.J (1990). Towards an accurate assessment of students cognitive capacity. An evaluation of three scoring procedures to objective testing. Unpublished manuscript, faculty of education, Obafemi Awolowo University, Ile-Ife.
2. Alonge F.M (2001), African journal of educational studies in mathematics and science Vol.1 pp 1-4.
3. Omirin M.S (1999) compares of the Psychometric properties of three multiple choice test, using confidence scoring procedure. An unpublished Phd. thesis, faculty of education, EkitiState University, Ado-Ekiti.

4. Awodele B.A (2013) journal of educational and social research Comparative Effectiveness of logical-choice weight and confidence scoring methods on reliability and validity of Chemistry multiple-choice test items in Nigerian Secondary Schools. Vol.3 pp. 387-39.
5. Bradbard D.A. Parker D.F and Stone G.L (2004). An alternative multiple choice scoring procedure in a macroeconomics course. Division scheme journal of innovative education Vol.2 pp11-15
6. Kolawole (2006) Statistical methods. Bolabay academic publishing consultants, Ikeja Lagos.

Views and opinions expressed in this research article are the views and opinions of the author(s), Open Access Publishing Group and European Journal of Physical Education and Sport Science shall not be responsible or answerable for any loss, damage or liability etc. caused in relation to/arising out of the use and provenience of the content. The author(s) retain(s) the copyrights of the published research article and consent(s) on being fully responsible for any conflict of interests regarding copyright violations, provenience and improper or inaccurate use of third part content. This research article is meeting the Open Access Publishing requirements and can be freely accessed, shared, modified, distributed and used in educational, commercial and non-commercial purposes under a [Creative Commons Attribution 4.0 International License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).