ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.2, No.12, 2012

Determining the Effect of Guessing on Test Scores

William J. Ubulom[1]* Clifford M. Amini[2]

1. Faculty of Technical and Science Education, Rivers State University of Science and Technology, P. M. B. 5080, Port Harcourt, Nigeria
2. REPRIBAL, National Open University of Nigeria, 14/15 Ahmadu Bello Way, Victoria Island, Lagos, Nigeria

*E-mail of the correspondence author: will.ubulom@yahoo.com

Abstract

The purpose of this study was to determine the effect of guessing on test scores. The population used for this study consists of twenty (20) pupils of a primary school randomly selected from a total of one hundred (100) primary six pupils in Calabar, Nigeria. Ten (10) multiple choice Chemistry test items were constructed and pilot tested on five (5) primary school pupils of the same grade and age. The result of the study fails to agree with the Perfect Guttman Scale rather it tends to agree with the opinion of Lord (1977). The tests were constructed and applied in the classical test theory pattern and the pupils do not possess the ability demanded by the test items. It is suggested that examiners should give appropriate and clear instructions to the examinees to avoid guessing. In order to assign scores to the responses of the examinees, it is recommended that examiners should adopt the correction of guessing formul**a.**

Key Words: Academic Achievement, Classical Test Theory, Electrolysis, Item Difficulty, Item Discrimination, Test score, True Test Theory

1. Introduction

Over the years, experts in psychometric have used teacher made and standardized test to measure students' abilities and academic achievement in specific and general subject areas. The practice is so common that there tend to be an agreement that test scores are absolute indicators of students' attainment of cognitive gains and hence their academic achievement. There thus tend to be a forgetting of the implications of other factors that may contribute positively or negatively to the score of an examinee in a test or examination. One of such influencing factors is guessing. Guessing is the process of trying to give an answer to a question or make a judgment about something without being sure of the facts (Hornby, 2001). For example, when the pupils are required to choose their answers from among the answer choices supplied by the examiner, there is the tendency that they can get some of the answers correctly or wrongly by guessing. This is an indication that guessing, according to Thorndike, Cunningham, Thorndike & Hagen (1991), is a serious problem on objective tests when there are only two answer choices, when the test is very short, when the items on the test are too difficult for the examinees, when the test is highly speeded and when the test items are poorly constructed.

The influence of guessing on test scores is fundamental that some studies seek to determine the effect of guessing on test scores. For instance, Waters (1969) studied the effect of different scoring formulas on the number of items omitted on a test. The scoring formulas used were rights only, rights - ¼ wrong, rights-wrongs, rights – 2X wrongs and rights – 4X wrongs. He then discovered that the number of omits did increase as the penalty for wrong answers increased, but that the examinees with the more severe scoring penalties did not omit as many items as they should have if they were to achieve the highest score. Waters also discovered that the average number of items right was about the same for all scoring patterns.

Basically, every test must contain instructions on how the testees will respond to the test items. For the Water's study, it seems that neither the instructions not to guess, which attracts penalty for wrong answers nor the instructions to answer every question whether examinees are sure of the answers or not make them respond the same way on an objective test. Hence, the probability of the examinees taking a risk appears to be dependent on their personality based on the instructions on the test.

Guessing on tests has been recognized as the most serious problem in test items that have only two answer-choices or tests that are highly speedy. Despite that a number of suggestions have been made for more complex scoring patterns of objective tests, particularly that of the multiple choice items, it is ideal that most of them increase scoring time and to some extent, the methods appear to produce little result in precision of the measurement for each test item.

The ability of a student is often being estimated using the number of correct scores to the items in the test score (Sotaridona, Pornel & Vallego, 2003). One concern of using a test score as ability estimate is that the estimate is less sensitive to the characteristics of the items. Item characteristics such as an item difficulty and item discrimination were seem to be ignored when estimating an examinee's ability or proficiency. It is appropriate to consider the item difficulty in the scoring process. This is because some items should have more weight than the other items.

One of the basic assumptions in testing is that test items are able to separate high ability students from low ability ones. For a good test with good psychometric properties, this is supposed to be so. Such test will have items constructed in such a way that the distracters are able to attract low ability students easily. The foregoing presupposes that if all the students in a testing situation do not possess the ability or theta (θ) demanded by the test items, they should all score zero. The mean and standard deviation scores of the students will be zero. According to the True Score Theory,

$$X_o = X_{mf} \pm X_c \tag{1}$$

Where: X_o = Observed score
X_{mf} = True score of an individual
X_c = Error score

The theory suggests that an examinee's observed score (X_c) is made up of his true score plus the error in the score. These errors assume the form of course, systematic and random errors. The error score due to guessing takes the form of random since no particular pattern is followed.

The purpose of this study was to determine the effect of guessing on test scores of primary school pupils in Calabar, Nigeria. The results of this study will be useful to test experts, as it will highlight guessing as one of the contributory factors to errors in scores obtained from tests.

2. Method

A perfect Guttman's Scale developed in 1944 was employed in this study. This is because although Guttman's approach was originally devised to determine whether a set of attitude statements is un-dimensional. The technique has been used in many different kinds of tests (Gregory, 2006).

The population used for this study consists of twenty (20) pupils of a primary school randomly selected from a total of one hundred (100) primary six pupils in the school. The study was conducted in Calabar, Nigeria. Ten (10) multiple choice Chemistry test items were constructed and pilot tested on five (5) students of the grade and age. The content area selected was Electrolysis, which is considered very high for primary school pupils. All the pupils in the pilot group scored zero. The item by person matrix of the pilot study is given on Table 1.

Table 1 reveals that all the pupils who participated in the test scored zero. This shows that none of the pupils in the pilot group possess the ability demanded by the test items. The same test items were administered on the twenty (20) sampled primary school pupils. However, in order to guide the examiners to adequately score the responses of the examinees, there is the need for them to adopt a correction for guessing formula, which is:

$$X = R - \frac{W}{n - 1} \tag{2}$$

Where: X = Correct score

R = number of items answered correctly
number of items answered incorrectly
N = number of options

3. Results

Table 2 contains the scores of primary school pupils that participated in the test. The table shows that pupils on serial numbers 1, 2, 7, 9, 10, 11, 12, 13, 14, 18, 19 and 20 scored zero. The table also reveals that pupils on serial numbers 4, 5, 8, 15 and 16 had 1 test item correct, those on serial numbers 6 and 17 had 2 test items correctly and a pupil on serial number 3 had 3 test items correctly. From the foregoing, it could be noted that the mean and standard deviation scores for the 20 primary school pupils that responded to the 10 multiple choice test items in electrolysis, a topic in senior secondary school chemistry are 0.55 and 2.5 respectively. This implies that these pupils do not possess the ability demanded by the test items hence the test items they were able to answer correctly were based on guessing.

Table 3 is the display of the test scores of primary school pupils based on item by person matrix. The table reveals that 1 out of 20 primary school pupils who responded to the test answered test items 3, 5, 9 and 10 correctly, 2 out 20 of them got test items 1, 2, 4 and 6 correctly and none of them had test items 7 and 8 correctly. The foregoing is an indication that these primary school pupils do not possess the ability demanded by the test items rather the test items they were able to answer correctly was based on guessing.

4. Discussion

The result as shown on Table 4 fails to agree with the Perfect Guttman Scale. Guttman (1944) was concerned with the consistency of response pattern and the role it can play in determining index of any characteristics under measurement. A Perfect Guttman scale will assume a triangular pattern due to students' responses to items according to the amount of characteristics or traits or ability possessed by them. For example, most testing practitioners make use of the classical test theory, whether they are familiar with or not because their basic tools, according to Warm (1978:15) are:

1. The P-value, which is the proportion of examinees selecting an item alternative (also called item difficulty)
2. D-value, which is the bi-serial correction between the item alternative and the test (also called item discrimination).
3. Mean of examinees (number right) scores
4. Standard deviation of examinees scores
5. Skewness and Kewtosis of examinees scores
6. Reliability of the test using KR-20

Lord (1977) however observed that the classical test theory does not provide an appropriate framework for dealing with the task of assigning at the end of the testing, rather a numerical score (or interval estimate) somehow represent the ability or overall level of performance of the individual tested.

The result of the study tends to agree with the opinion of Lord above. Since the tests were constructed and applied in the classical theory pattern and knowing that the pupils do not possess the ability demanded by the items, students 3, 4, 5, 6, 8, 15, 16 and *17 who scored 3, 1, 1, 2, 1, 1, 1, and 2 respectively did so by guessing. Thus, the true total score of the pupils, which was expected to be zero, is shown to be 12. The total score of 12 by the 20 pupils constitute the total error score of the pupils who took part in the study. This implies that guessing is a serious factor which affects examinees test scores. If appropriate instructions are given to the examinees, the [problem of guessing will be eradicated.

5. Conclusion

The study was conducted on primary school pupils using 10-item multiple choice test in chemistry, a subject not familiar to the examinees. While it was expected that students do not possess the ability demanded by the items, some scored some items correctly even though these scores do not fall into the regular consistency of responses as propounded by Guttman (1944) it shows that guessing was the main factor responsible for the error in the test scores and should be checked by all examiners.

It is advocated that examiners should give appropriate and clear instructions to the examinees to avoid guessing, stressing that this might attract penalty on the part of the pupils. Also, in order to assign scores to the responses of the examinees, it is important that examiners should adopt the correction of guessing formula. This will assist in discouraging guessing.

References

Gregory, R. J. (2006). *Psychological testing: History, principles and applications.* New Delhi: Pearson Education Inc.

Guttman, L. (1944). Response pattern and test score consistency. *Psychometrika,* 11, 28-43

Hornby, A. S. (2001). *Oxford advanced learner's dictionary for current English.* London: Oxford University Press.

Lord, F. M. (1977). A study of item bias, using item characteristic curve theory. In Y. H. Poortingal (ed.), *Basic problems in cross cultural psychology*. Amsterdam: Sweets & Zeitlingera

Sotaridona, L. S., Pornel, J. B., & Vallejo, A. (2003). Some applications of item response theory to test*ing. The Philippine Statistician*, 53(1-4), 81 - 92.

Thorndike, R. M., Cunningham, G. K., Thorndike, R. L., & Hagen, E. P. (1991). *Measurement and evaluation in psychology and education*. New York: Macmillan Publishing Company.

Warm, T. A. (1978). (1978). *Primer of item response theory*. Springfield: National Technical Information Service U. S. Department of Commerce.

Waters, L. K. (1967). Effect of perceived scoring formula on some aspects of test performance. *Educational and Psychological Measurement.* 27, 1005-1010.

Dr. William James Ubulom was born at Okoroboile Town in Andoni Local Government Area of Rivers State, Nigeria on 5th March, 1961. He became a member of the Association of Business Educators of Nigeria in 2002, a member of the Nigeria Institute of Management in 2001, Institute of Management Consultants in 2001 and a member of the Nigeria Association of Educational Researchers and Evaluators in 2008. He attended the Rivers State University of Science and Technology, Port Harcourt where he obtained his B.Sc. Degree in Business Education in 1990. He also attended the University of Port Harcourt, Nigeria where he obtained his M.Ed. Degree in Educational Measurement and Evaluation in 1997. Dr. Ubulom further attended the University of Nigeria, Nsukka where he obtained his Ph.D. Degree in Science Education (Measurement and Evaluation Option) in 2006.

Dr. Clifford M. Amini was born at Egendem Town in Andoni Local Government Area of Rivers State, Nigeria on 14th July, 1960. He became a member of the Science Teachers Association of Nigeria (STAN) in 1997, Nigerian Association of Educational Psychology in 2000 and also a member of the Nigeria Association of Educational Researchers and Evaluators in 2008. He attended the University of Calabar, Nigeria where he obtained his B.Sc. (Ed.) Degree in Chemistry in 1989. He also attended the University of Port Harcourt, Nigeria where he obtained his M.Ed. Degree in Educational Measurement and Evaluation in 1997. Dr. Amini further attended the University of Calabar, Nigeria where he obtained his Ph.D. Degree in Test and Measurement in 2008.

Table 1: Item by Person Matrix for Pilot Group

	1	2	3	4	5	Total
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
Total	0	0	0	0	0	0

Table 2: Pupils Score in Multiple choice Tests

S/N of Pupil	Total Score
1	0
2	0
3	3
4	1
5	1
6	2
7	0
8	1
9	0
10	0
11	0
12	0
13	0
14	0
15	1
16	1
17	2
18	0
19	0
20	0

Mean (X) = 0.55 SD = 2.5

ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.2, No.12, 2012

Table 3: Item by Person Matrix of Pupils Test Scores

	1	2	3	4	5	6	7	8	9	10	Total
1	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	1	0	0	1	1	3
4	1	0	0	0	0	0	0	0	0	0	1
5	0	0	0	0	1	0	0	0	0	0	1
6	1	1	0	0	0	0	0	0	0	0	2
7	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	1	0	0	0	0	0	0	1
9	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	1	0	0	0	0	0	0	1
16	0	0	1	0	0	0	0	0	0	0	1
17	0	1	0	0	0	1	0	0	0	0	2
18	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0
Total	2	2	1	2	1	2	0	0	1	1	12

Table 4: Perfect Guttman Response Pattern

	1	2	3	4	5	Total
1	1*	0**	0	0	0	1
2	1	1	0	0	0	2
3	1	1	1	0	0	3
4	1	1	1	1	0	4
5	1	1	1	1	1	5
Total	5	4	3	2	1	

*Correct Response = 1 ** Incorrect Response = 0