

Journal of Education

Vol. 1, No. 2,

January, 1992

**Published by the Faculty of Education,
University of Uyo, P. M. B. 1017, Uyo.**



AN ASSESSMENT OF THE PREDICTIVE POWER OF TEACHER-MADE ACHIEVEMENT TESTS IN SENIOR SECONDARY SCHOOL CHEMISTRY

Eme E. Ansa
University of Uyo

ABSTRACT

Students are often exposed to different types of examinations ranging from teacher-made tests to externally set examinations. The purpose of this research is to ascertain the predictive ability of teacher-made tests on externally constructed instruments. Results from this study have shown that most teacher-made tests do not predict student performances on externally made instruments since the results of such tests do not correlate significantly with an externally constructed and validated test. Based on this finding, there is therefore, the need to expose practising teachers to appropriate methods of test construction, administration and scoring.

With the introduction of continuous assessment, there has been an increasing need for classroom teachers to prepare and administer tests in order to obtain certain vital information on what have been achieved during teaching-learning processes. That means testing can be used by teachers as a teaching instrument as well as an instrument for the assessment of student performances. These purposes are essential in the teaching of chemistry as well as other subjects. When used as a teaching instrument, the outcome of a testing exercise provides a quick feedback on what learning has taken place in the class. As an aid to teaching, testing detects problems with teaching strategies adopted and also detects weaknesses and strengths of the students taught.

For ascertaining the gains of instructional processes, achievement tests are administered by teachers. These, according to Stanley and

Hopkins (1972) are tests that measure the extent to which a person has acquired certain information, mastered certain skills, usually, as a result of specific instruction. These tests are in most cases not scrutinized for validity or reliability which are some of the most essential qualities of a good test. This use of invalid and unreliable tests runs contrary to the views of Ohuche and Akeju (1977) who are of the opinion that, the choice of any assessment procedure greatly depends on whether or not it is valid, reliable, objective, usable and discriminatory. In the face of Ebel's (1979), assertion that, most teacher-made achievement tests are not valid, the use of invalid and unreliable tests raises a lot of concern.

The validity of a test has often been defined as the extent to which the test measures what it is supposed to measure. However, according to Ghiselli, et al (1981); the most recent definition of validity is that it refers to the appropriateness of inferences made from test scores or other forms of assessment. Thus the validity of a test is the accuracy of specific conclusions made from its scores. There is also the criterion validity of a test. It is known as an empirical or statistical validity which describes in an objective and quantitative fashion the degree of relationship between predictor scores and the criterion scores (Ghiselli, et al, 1981). To establish criterion validity, correlational coefficients are commonly used to describe the degree of predictive validity observed. The coefficient indicating the degree of relationship between predictor and criterion scores are termed validity coefficient statistic (Okoh, 1983) The higher the coefficient, the greater is the predictive power of the test assured, Ghiselli et al, (1981).

This study is principally aimed at ascertaining the degree to which students' performances at a teacher-made chemistry test could predict their performance on an externally set and validated chemistry test.

RESEARCH QUESTIONS

The following questions were formulated for the study:

- (i) Is there any significant relationship between students scores on an externally validated chemistry test and teacher-made Senior Secondary Two (SS2) final examination scores in Chemistry?
- (ii) Is there any significant difference between the number of schools with correlated scores and the schools whose scores did not correlate in the externally validated chemistry test and teacher-made SS2 final examinations scores?
- (iii) Is there any significant difference between the scores of students

from schools with correlated scores on the externally validated chemistry test and the teacher-made SS2 final examination scores and students from schools with uncorrelated test scores on the external and teacher-made tests scores?

MAJOR ASSUMPTIONS

For this research to be conducted, the following assumptions were made:

- (a) All schools considered in this study adopted the same core curriculum.
- (b) All the students have the aptitude for the study of chemistry, with the understanding that they willingly chose chemistry as one of the optional subjects in their Senior Secondary School programme.
- (c) The students opting for chemistry in each of the sampled schools made up a fairly homogenous ability group.

RESEARCH SAMPLE

The target population for the study consisted of chemistry students who had completed their SS2 chemistry programme in secondary schools in Akwa Ibom State. A sample of eighty (80) students from three (3) schools selected through stratified random sampling technique were used for the pilot study.

For the main research, chemistry students from twelve schools selected through stratified random sampling were used. A total of 300 students participated in the study.

CONSTRUCTION OF INSTRUMENT

With the use of the SSS chemistry core curriculum, the researcher constructed a Chemistry Achievement Test (CAT). In order to collect teacher-made test scores in chemistry, students academic records were used to extract their SS2 final examination scores in Chemistry.

An eighty-item Chemistry Achievement Test (CAT) was constructed from a pool of one hundred and ten multiple choice items constructed by the researcher. To assess achievement, based on the six levels of Bloom's taxonomy of educational objectives for the cognitive domain; the 80 multiple choice items were compiled in the ratio of 30:21:11:7:6:5 for knowledge, comprehension, application, analysis, synthesis and evaluation respectively.

PILOT TEST

The CAT was administered with the school teachers serving as invigilators. The students responded to the test under strict examination conditions. During the test, the students were expected to complete the 80 items within 90 minutes. One point was allotted to each question thus the highest score of 80 points was expected for the unvalidated instrument.

VALIDITY AND RELIABILITY OF INSTRUMENT

The eighty item CAT was scrutinized for content validity by a Chemistry Education specialist, a Measurement and Evaluation lecturer and two practising secondary school chemistry teachers. On ascertaining the content validity of the instrument, a pilot study was conducted using eighty students from three schools in Akwa Ibom State.

On the basis of the pilot study results, forty multiple choice items were selected to make up the final validated Chemistry Achievement Test (VCAT). This test was composed of questions with discriminatory indices (D) of 0.4 and above; difficulty indices (p) between 0.30 and 0.73; also a few items with the discriminatory indices between 0.27 and 0.37 were accepted with corrections.

The VCAT was assembled in the ratio of 18:10:5:3:3:1 from each of the six levels (in ascending order) of the cognitive domain as proposed by Bloom, et al, (1956).

The VCAT had a reliability coefficient of 0.78 using the test retest method and an internal consistency of 0.78 using the Kuder-Richardson 21' formula.

ADMINISTRATION OF VCAT

With the aid of resident teachers in the participating schools, the VCAT was administered under favourable examination conditions. A duration of 50 minutes was allowed for the 40 items. One point was allotted to each question and a maximum of 40 points was expected on the test.

PROCEDURE FOR DATA ANALYSIS

(1) In order to test the correlation between the students' VCAT scores and their scores in the SS2 final examination in chemistry, the Spearman Rank Order Correlation formula was used.

(2) To further ascertain the significance of the correlation, a t-test formula for establishing relationships was employed. The resultant values were compared with the table t-value at 95% confidence level ($\alpha = 0.05$) at various degrees of freedom depending on the number of students in each school used for the study.

(3) The z-score analysis for significance of proportions at $\alpha = 0.05$ was used to ascertain the significance of the proportions of schools with correlated scores.

RESULTS

The validated instrument was administered to three hundred (300) SS2 chemistry students in twelve schools from six local government areas in Akwa Ibom State. On the whole, 147 students were males while 153 students were females. The students' scores ranged between 6 and 38 out of the forty questions on the VCAT. The mean score was 18.21 with a standard deviation of 5.94.

TABLE I
The Relationship Between the VCAT Scores and SS2 Final Examination Scores in Chemistry for the Sampled Schools

| Schools | N | D ² | r _s | df | t |
|---------|----|----------------|----------------|----|---------|
| A | 17 | 198.50 | 0.757 | 15 | 4.483* |
| B | 20 | 368.50 | 0.709 | 18 | 4.270* |
| C | 28 | 1421.50 | 0.599 | 26 | 3.960* |
| D | 50 | 11355.25 | 0.455 | 48 | 3.540* |
| E | 8 | 78.50 | 0.066 | 6 | 0.395 |
| F | 45 | 9776.50 | 0.360 | 43 | 2.498* |
| G | 52 | 16085.00 | 0.314 | 50 | 2.340* |
| H | 10 | 201.00 | 0.220 | 8 | 0.638 |
| I | 4 | 0.50 | 0.950 | 2 | 19.490* |
| J | 19 | 911.00 | 0.201 | 17 | 0.846 |
| K | 14 | 324.50 | 0.287 | 12 | 1.040 |
| L | 10 | 109.00 | 0.339 | 8 | 1.020 |

* Significant at $\alpha = 0.05$

Table I shows great variations in the number of subjects used for the study. It is worthy to mention that, all the SS2 students who offered chemistry in each of the participating schools were used for the study. Thus, some schools had as low as 4 or 8 students while some others had 45 students and above.

To establish a criterion validity of the test, the students scores for VCAT was correlated against the SS2 final examination score in chemistry for each of the studied schools. This facilitated the verification of the research question which sort to find out if there is any significant relationship between students scores on an externally validated chemistry test and teacher-made SS2 final examination scores in chemistry.

From the correlational analysis using Spearman Rank Order Correlation formula and the student t-test for establishing relationships, teacher-made test scores of seven of the twelve schools correlated with their scores from VCAT. The correlation coefficients of 0.757; 0.709; 0.599; 0.455; 0.360; 0.314 and 0.950 with the corresponding t-test results of 4.483; 4.270; 3.960; 3.540; 2.498; 2.340 and 19.490 were obtained for schools A,B,C,D,F,G and I respectively. It is worthy to note that most of these schools are situated in the urban areas except for school I which incidentally has an "outrageous" correlational coefficient of 0.95. This exceptionally high coefficient must have been caused by the very small size which allowed the teacher a closer interaction with her students.

The seven schools with correlated results accounted for only 58% of the total number of schools studied. On applying the z-transformation; a 0.55 result was obtained. This result showed that, the number of schools with correlated results was insignificant when viewed against the table z-score of 1.96 at $\alpha = 0.05$. This further substantiates Ebel's (1979) assertion that most teacher made achievement tests are not valid. The low z-score (0.55) thus indicated that there is no significant difference between the number of schools with correlated scores and schools with uncorrelated scores in the externally validated chemistry test and teacher-made SS2 examination scores.

The result from the z-score analysis raised some doubts about the validity of the VCAT. This led the researcher to investigate, if, there was any existing differences in the VCAT scores of the students from schools with correlated scores and the students from schools whose scores had no significant relationships – this is the third research question. Table II shows the finding which attest to the credibility of the VCAT.

Jegade, 1986). To further buttress this fact, some investigators, Elgood (1967) and Ekpo (1984) also discovered that some biology teachers usually omit certain topics in the syllabus because of their unfamiliarity with those topics.

The proportion (0.58) of schools with significantly correlated scores is not appreciable as shown by the z-score analysis. This is an indication that teacher made achievement tests in some schools do not predict the possible performance of their students on externally constructed and validated tests. This situation thus gives an idea as to why students in most schools do not have a corresponding performance level in external examinations like the Senior Secondary Certificate Examination (SSCE) and the Joint Matriculations Examinations (JME)

Most of the schools whose teacher-made tests scores did not correlate with their student scores on the VCAT were located in the rural areas. These schools also had a relative small class size. Although the small class size is ideal for closer student-teacher relationship, the researcher suspects the adverse effect of class size on the teachers' scores. This suspicion results from the fact that, a t-test value of 0.19 was obtained in trying out if there were any significant difference in the VCAT scores of students from schools whose test scores correlated significantly and those whose test scores did not correlate significantly. The low result (0.19) shows a no significant difference in the scores of the two sets of schools. Such a result therefore indicates that the VCAT had tested both groups without any discrimination or disparity. The possible discrepancy may lie in the teacher-made test scores, hence the researcher suspects that there may have been some degree of bias and a resultant ambiguity in the scores from the teacher-made achievement tests (final SS2 examination in Chemistry), especially in those schools whose test scores did not correlate significantly with their VCAT scores.

CONCLUSION AND RECOMMENDATIONS

Findings from this research have shown that, most teacher-made achievements are not valid, thus these tests can not predict the performances of students on externally set achievement tests. In order to improve this situation, there is the need to organise workshops and seminars on techniques of test construction. To reduce subjective scoring which results in ambiguous test scores, the use of objective type tests (multiple choice items) instead of essay questions should be encouraged.

REFERENCES

- Ebel, R. L. (1979). **Essentials of Educational Measurement**, (3rd ed.) New Jersey, Eaglewood Cliffs: Prentice Hall Inc.
- Ekpo, O. E. (1984) "The Nigerian Biology Curriculum: Perspective on Course Articulation" Published Ph.D. Thesis, Cornell University, Ithaca, New York.
- Elgood, J. (1967) "Biology Teaching in West Africa", **Journal of Biological Education**, Vol.1, No. 4 Dec.
- Ghiselli, E. E. et al, (1981) **Measurement Theory for the Behavioural Science**, San Francisco: W.H. Freeman and Company.
- Ohuche, R. O. and Akeju, S. A. (1977) **Testing and Evaluation in Education**, Lagos: African Educational Resources
- Okoh, N. (1983) "The Problems of Measurement in Psychology and Education" In N.Okoh (ed) **Professional Education: A Book of Readings**, Benin City: Ethiope Publishing Co.
- Senior Secondary School (SSS) Chemistry Core Curriculum (1980). Comparative Education Study and Adaptation Centre (CESAC) and the Federal Ministry of Education Science and Technology, Lagos: University of Lagos.
- Stanley, J. C. and Hopkins, K. D. (1972) **Educational and Psychological Measurement and Evaluation**, (5th ed.) New Jersey: Prentice Hall Inc.