

Dr. Casmir N. Ebuoh

INFLUENCE OF SCORING BY SECTION, INDEPENDENT SCORER AND CONVENTIONAL PATTERNS ON SCORER'S ATTITUDE IN BIOLOGY ESSAY TESTS.

BY**CASMIR N. EBUOH (Ph.D).**

Department of Science and Computer Education, Enugu State University of Science and Technology, Agbani.

Abstract

The scoring of essay tests had be criticized for not being reliable and that the level of reliability were lower than expected. Literature revealed that there is scorer's negative attitude in scoring essay test. A survey study was adopted 42 Biology teachers from the 23 secondary schools in Enugu Education zone were used for the study. The scoring of the attitude scale was done for 15 minutes by the 42 scorers respectively. The instrument was developed by the researcher. The research questions were answered by calculating the mean scores and standard deviations. The hypothesis was tested using two way analysis of variance. The results revealed that use of independent scorers (UISP) and scoring by section (SBSP) had significant influence on scorer's attitude. Use of independent scorers Contrary, influence of conventional pattern (CPSAI) was not significant.

Background of the study

One of the claimed factors that is responsible for the unreliability in scoring essay tests is the attitude of the scorers. The attitude is the psychological (attitudinal) dispositions which include fatigue, anger, hatred these influence the ability of the scored to distinguish the scorer reliability. As a test scorer reads through a large pile of answer scripts, there are systematic changes resulting from the attitudinal factor (Ezeoke, 1985).

The scoring of essay test had been criticized for not being reliable. There was evidence too to show that the level of unreliability in scoring essay tests appears more likely to be more in the internal than in the external examinations.

Table 1:
Biology Essay Test.

S/N	Script Number	Examiners Number	Examiners		Scores	Deviation (Highest Lowest)
1.	525260/313	5024/870	Assistant	Original	15	
		5024/130	Team Leader	Vetted	17	-2
2.	525260/017	5024/873	Assistant	Original	33	
		5024/030	Team Leader	Vetted	35	-2
3.	525260/056	5024/892	Assistant	Original	19	
		5024/291	Team Leader	Vetted	28	-9
4.	525260/001	5024/820	Assistant	Original	37	
		5024/120	Team Leader	Vetted	39	-2
5.	525260/182	5024/874	Assistant	Original	35	
		5024/380	Team Leader	Vetted	32	+3
6.	525260/073	5024/890	Assistant	Original	16	
		5024/160	Team Leader	Vetted	26	-10
7.	4290708/57	45042/570	Assistant	Original	28	
		5042/131	Team Leader	Vetted	19	-9
8.	4290708/31	5042/723	Assistant	Original	33	
		5042/13	Team Leader	Vetted	20	+13
9.	4290708/88	5042/70	Assistant	Original	33	
		5042/314	Team Leader	Vetted	27	+6
10.	4290708/236	5042/900	Assistant	Original	17	
		5042/60	Team Leader	Vetted	17	0
11.	429072/001	5042/1238	Assistant	Original	36	
		5042/48	Team Leader	Vetted	31	+5
12.	429072/010	5042/1236	Assistant	Original	14	
		5042/148	Team Leader	Vetted	19	-5
13.	429072/111	5042/394	Assistant	Original	26	
		5042/94	Team Leader	Vetted	15	+17
14.	429072/91	5042/1260	Assistant	Original	39	
		5042/230	Team Leader	Vetted	36	+3
15.	429072/187	5042/1336	Assistant	Original	27	
		5042/066	Team Leader	Vetted	17	+10
16.	429072/143	5042/1214	Assistant	Original	16	
		5042/151	Team Leader	Vetted	20	-4
17.	4290703/137	5042/187	Assistant	Original	18	
		5042/65	Team Leader	Vetted	18	0
18.	4290703/119	5042/195	Assistant	Original	12	
		5042/75	Team Leader	Vetted	20	-8
19.	4290703/218	5042/1411	Assistant	Original	33	
		5042/80	Team Leader	Vetted	25	+8
20.	4290703/317	5042/1314	Assistant	Original	30	
		5042/92	Team Leader	Vetted	27	+3

Source: West African Examination Council (2001).

Table 1 shows that out of the 20 randomly selected vetted scripts of assistant examiners by Team leader, 9 was under marked, and 9 was over marked while only 2 scripts have no deviation. Further, analysis of the relationship between the assistant examiner's scores and team leader's scores using spearman's rank order statistic indicated a coefficient of -0.10. This very low negative relationship is a pointer to the problem of different scores in scoring student's scripts even with the same marking guide.

Table 2. Scores given to student by five scores in mathematics essay test.

Students	Scores					Deviation (highest-Lowest scores)
	A	B	C	D	E	
1	0	1	17	13	8	17
2	1	1	18	11	7	17
3	1	6	13	5	2	12
4	11	2	10	8	5	9
5	2	10	10	3	0	10
6	5	14	24	3	1	23
7	0	9	15	3	3	15
8	2	9	19	13	3	17
9	1	1	9	15	21	20
10	1	3	20	14	6	19

Source: Cox (1989)

Variations indicated in table 2 showed that the conventional pattern which was used in the scoring appeared to be unreliable.

The low level of the scorer reliability had been attributed to a number of factors. It was claimed that the low level of scorer's attitude is caused by the use of inappropriate scoring patterns in scoring essay tests (Cox 1989 and Ebuoh 2004).

It appeared that in Nigeria, scorers are more conversant with the conventional method (scoring all items at a time in a script before picking another one) of scoring essay tests than with other methods. A question that arises is: is the conventional scoring pattern better than the other patterns in achieving higher scorer's positive attitude. It appears that there is no empirical study so far that compared relative effectiveness of the various patterns of scoring Biology essay tests in terms of which one engenders higher positive attitude than others. Moreover, it is not certain which scoring pattern is more associated with higher positive attitude of the scorer in scoring of Biology essay tests.

Statement of the problem

The scoring of essay tests had been criticized for not bring reliable. The low level of scorer's reliability appears to be more likely in the internal than external examinations as shown in the tables one and two above. The two level of the scorer's negative attitude had been attributed to a number of factors. It was claimed that the low level of scorer reliability had been blamed on negative attitude of scorers and inappropriate scoring patterns in scoring essay tests.

Purpose of the study

The purpose of this study was to investigate the influence of the use of different scoring patterns on scorer's attitude.

Scope of the study

Scoring of essay test is practical technique that cut across all levels of Nigerian Education system: Primary, junior/senior secondary school and tertiary institutions. The choice of the senior secondary school level is necessitated by the fact that Biology is taught only in senior secondary schools.

Consequently, the study is designed to find out the influence of the three scoring patterns (scoring by section pattern, SBSP; use of independent scorers patern, UISP and conventional pattern of scoring all items, CPSAI) on scorer's attitude in biology essay tests in Enugu Education zone of Enugu State.

Research question one

What is the mean scores of the scores in Biology essay when the three different scoring patterns are employed?

Research Hypothesis one

There is no significant difference in mean scores awarded by scorers who scored Biology essay test using the three different scoring patterns (USBSP, UISP, and CPSAI).

Literature review

Literature on scorers' attitude and scoring patterns in scoring Biology essay tests were reviewed.

Attitude refers to either mental readiness or implicit predispositions that exert some general and consistent influence on a fairly large class of evaluation responses. Attitude as mental and neural state of readiness organized through experience exerting a direct and dynamic influence upon the individual's response to all objects and situations with which it is related (Ugwu, 1997). For the researcher, attitude is the feeling of favourableness towards some "psychological actions such as teaching, learning, interpretation and scoring of students' tests. Attitude helps scorers to come out with

desired scoring pattern. It is a quality that arouses predispositions that hold and maintain attention during scoring.

Characteristics that are associated with score's attitude

A number of characteristics that are associated with the attitude of scorers are discussed here.

Gagne (1997) identified attitude as highly individual actions and examined six characteristics which are associated with the attitude namely:

- i. Attitude is predisposition to respond, that is, readiness to behave rather than actual behaviour towards an objective.
- ii. Attitude is amenable to change, but that which is strongly held, requires substantial effort.
- iii. Attitude produces consistency among its' various manifestations in an individual's behaviour towards object.
- iv. Attitude has intensity and a directional quality WHICH connotes preference regarding outcome involving the object, evaluations of object or positive affections for the object.
- v. Attitudes result from experience. Some attitudes are acquired through experiences that have profound affective component. More than this different forms of learning are transmitted through the process of imitation, modeling and identification within the peer group.
- vi. Attitude remains latent until signs or objects evoke it to influence the behaviour of the individual towards the stimulus or object. Based on the foregoing, attitude can be seen as having three components namely: Behaviour (action tendency), cognition and effect (emotion).

The behaviour (action tendency) concerns overt actions, action tendencies, verbal statements concerning behaviour (as seen in number (I) and (II) above.

The cognition involves perceptual responses, beliefs about or factual knowledge of object and verbal statements of cognition as in numbers (iii) and (iv) above.

The effect is the feeling of like/dislike, love/hate, sympathetic nervous system responses and verbal statement of affect as in numbers (v) and (vi) above (Zimbardo and Ebbeson, 1977). These three components are related and exert mutual influence on a scorer's attitude in scoring Biology essay tests.

Factors influencing scorer's attitude

The factors that influence the scorer's attitude are discussed in this section of the work.

In studying the scorer's attitude, care should be taken to identify these undesirable or negative attitude areas. A scorer's attitude has to be guided so that the scoring objective might be directed towards eliminating undesired scoring patterns.

Ugwu (1997) identified negative attitude as major factors. 'Horrocks and Shoonover (1968) discovered that:

- (i) Perception of scoring of essay test
- (ii) Anxiety towards scoring essay tests
- (iii) Value/importance attached to scoring
- (iv) Self-concept
- (v) Employment of scores of essay tests and
- (vi) Positive Attitude motivates the scoring of essay tests

These are the important factors in measuring the attitude of scorers. Both old and modern examiners are concerned with the problem of developing scorer's attitude. However, for both the old and modern scorers, psychological knowledge about attitude, experience, interest and appreciation formation is of importance if for no other reason than development (good or bad) of a scorer's interest, appreciation, experience and attitude will take place during the course of scoring whether the scorer wishes or not.

Every classroom teacher faces the problem of how to improve the unreliability of external and internal examinations which are reflections of the poor attitude of the scorers towards the scoring of essay tests. The attitudes of scorers have led to the declining popularity of scoring essay tests over the year (Piper & Hough 1982). Teachers seemed to have expressed deep concern over this unfortunate trend of our time and unless something is done about it, position of scoring essay tests in the school may be jeopardized.

Consequently, stones (1982) highlighted a number of major roles which attitude plays in scoring of essay tests namely: attitude influences devotion to duties, fairness, firmness, honesty, endurance and discipline.

In summary, the review showed that any pattern of scoring which facilitates meaningful scoring is bound to enhance the scorer's attitude. Evidence in literature has consistently shown that a number of characteristic factors are associated with scorer's attitude.

Scoring patterns;

Scoring patterns in this study mean the various methods that are employed by scorers to obtain the quantitative performance of learners. Various scoring patterns have been reported in literature:

They are;

- i. Scoring by section pattern (Ukeje, 1984 and Ezeudu, 1997).
- ii. Ranking all scripts before scoring all items (Ezeoke, 1986).
- iii. Scoring an item across board (Harbor peters, 1997).

Dr. Casmir N. Ebuoh

- iv. Use of independent scores (Ezeoke, 1986, Harbor peters 1999).
- v. Another reported pattern of scoring essay test is where the scoring involves dividing the task of scoring into sessions (Lovegroove, 1984).
- vi. Rearrangement of the scripts according to the order of quality, (Horrocks and Shoonover, 1998).
- vii. Conventional pattern of scoring all items in a script before picking up another script (Maduabum, 1984 and Ezeoke, 1986).

However, among these scoring patterns reported above, patterned scoring in form of scoring by section, use of independent scorers pattern and conventional pattern of scoring all items in a script before picking up another script appeared to be most popular in scoring essay tests.

In view of the above, the essay test scoring patterns employed in this study are:

- i. use of independent scorers pattern (UISP)
- ii. Scoring by section pattern (SBSP).
- iii. Conventional pattern of scoring all items (CPSAI).

Use of Independent Scorers Pattern (UISP)

In this pattern, use two or more independent scorers to score a student's script and use the average as a final score. In a situation where this is employed, the scores should not be recorded on the test booklet without finding the average of the scores but should be written on a separate sheet (Ezeoke, 1986).

This method has the advantage of checking possible biases arising from rater-ratee interaction. Again, it may rectify errors caused by oversight on the part of the rater. This advantage holds if the idea of independent scoring is maintained by ensuring that two scorers do not sit close enough to discuss their scores as they may influence each others, thereby defeating the idea of independent scoring (Harbor-Peters, 1999 and Quereshi, 1974).

Scoring By Section Pattern (SBSP)

This pattern is where one scorer specialized in scoring a section (part) of the test. The scorer scores his section and passes on the script to the next scorer. Evidence abound in support of this method. For instance Lovegroove in Ukeje (1984) advised that in an examination in which crucial decision may be taken, two or more scorers should be allowed to score a section of the script independently. It is argued that it increases the speed of scoring and gives opportunity of inter-item comparison rather than complete test comparison. Another major critique in favour of the method is that it increases the risk of error due to the possibilities of failure to hear correctly (Ezeoke, 1986).

Conventional Pattern of scoring all items (CPSAI)

The conventional pattern presupposes that the teacher/scorer scores all items in a script before attempting to score any other script. The method has been strongly criticized by scholars. One of its major criticism is that teacher/ scorer biases are made manifest in the way they score the scripts. It would be most difficult to be consistent when handling different items in a script one after another. That is, to give the same score to similar answers to a particular question in all the scripts that are scored (Ezeoke 1986).

Another disadvantage of this pattern is that it causes "halo effect." This means that if the answer to a question is a very good one. It influence the scoring of the next answer/items. Likewise, if the first item to be scored is a poor item, the subsequent items could be scored within the same context.

In view of the above the avoidance of the halo effect can be achieved according to Powell and Lobster (1974) if each response is judged on its merit without regard to other success or failure. The scorer must guard against allowing his scoring to be influenced by any general impression the scorer formed of the subject/candidate's ability. There is a natural tendency to over-estimate the ability of a bright, self-confident, talkative child. An examiner has to be very careful to avoid scoring the responses of such a child too leniently. Scoring must not be tempered with any conviction that the subject could have been answering correctly. The task is to score the response which has actually been given.

Research methodology**Design of the study**

This study is a survey research design. This design is a study that has to do with gathering of facts, rather than manipulation of variable, to provide important and useful information for educational decision-making. Therefore, this study is aimed at gathering information on the scorer's attitude in Biology essay tests without manipulating the subjects.

Area of study

The study was designed to cover all the schools in Enugu Education zone of Enugu State. The researcher adopted the educational administrative structure in which Enugu State is divided into six education zones. These are Awgu, Enugu, Nkanu, Oboloafor and Udi zones. The choice of the zone is because of logistical convenience and the researcher sees the zone as a thickly populated zone in terms of biology teachers among all the six zones in Enugu State.

Population of the study

The population for this study comprises all the 42 senior secondary school biology teachers in all the 23 secondary schools in Enugu Education zone.

Sample and sampling techniques

In consideration of the fact that only senior secondary school biology teachers were used for the study and because the number of biology teachers is not too large, the researcher used all the biology teachers for the study. The use of all the 42 biology teachers further helped the researcher in avoiding sampling errors.

Instrument for data collection

The researcher constructed a written attitude scale on scoring biology essay test (ASISBET). The attitude scale in scoring Biology test (ASISBET) was 20 item attitude scale developed by the researcher. The ASISBET was developed to measure the attitude attributes of biology teachers. It was 4–point modified likert-type response scale. The respondents were expected to indicate their degree of agreement or disagreement on a number of positive and negative attitude statement about scoring Biology essay tests.

Validation of the instrument

The ASISBET were face validated by five experts drawn from the sub-department of science education (two Biology specialist and three measurement and evaluation specialists) of the university of Nigeria, Nsukka. Their criticisms and vetting helped in modifying and/or replacing some items. The items measured objective in the affective domains. This is because the attitude mainly measures the effective domain of behaviour.

Reliability of the instrument

In order to determine the reliability coefficient of ASISBET, the instruments was administered to 20 Biology tests in Udi education zone of Enugu State and the scores generated from the respondents used for the trial test were applied to Cronbach Alpha formula. Using the Cronbach Alpha formula, an internal consistency of 0.80 was obtained for ASISBET. The Cronbach alpha formula for establishing the reliability was considered appropriate for items that are not dichotomously scored (non-dichotomous). Test re-test method (stability measure) of ASISBET was conducted and was found to be 0.85. The assumption of the method is that if the same test is given twice to the same respondents, the scores obtained were to maintain the same or slightly different position in the group (Ebuoh, 2005).

Administration of the instrument

The researcher administered the ASISBET to the 42 senior secondary school Biology teachers in Enugu Education Zone of Enugu State using three trained assistant

Dr. Casmir N. Ebuoh

researchers. Each assistant researcher was to cover Enugu North, Enugu East and Isiuor Local Government Areas respectively. This arrangement made it possible for one hundred percent of the return of the questionnaire.

Method of data analysis

The research question was analysed using the mean and standard deviation while the two ways analysis of variance (ANOVA) was used in testing the null hypothesis. The ANOVA is more useful when it is employed to determine the nature and scope of variance existing within and between three or more comparative sample means investigation.

Results

Research Question one

What are the mean scores of the scorers in Biology essay test when the three different scoring patterns are employed?

The table below shows responses of respondents on their attitude towards scoring essay test in Biology.

Table Three: Means and standard deviation of scorers test attitude scores (by three scoring patterns) in Biology essay test.

	Group I			Group II			Group III		
	Use of independent scorers pattern			Scoring by section pattern			Conventional method		
	Mean (X)	SD	N	Mean (X)	SD	N	Mean (X)	SD	N
Post Attitude Scores	2.86	0.99	14	2.69	0.92	14	2.50	1.06	14

Data in table three shows that the mean attitude scores for group I, group II and group III were 2.86, 2.69 and 1.50 respectively. This meant that respondents who scored using independent scorers pattern had the highest mean attitude score of 2.86 followed by those who scored by section pattern. The group who score with the conventional method of scoring all items had least mean (2.50) attitude score.

Hypothesis one:

There is no significant different in the mean scores awarded by respondents who scored Biology essay test using the three different scoring patterns (SBSP, UISP and CPSAI).

Table Four: Analysis of Variance (ANOVA) of scorers' post attitude scores using the three scoring patterns (SAIAB, RASBSAI and CMSAI).

Source of Variation	DF	Sum of squares (SS)	Mean square (MS)	F calculated	F-Critical
Group	140	64.25	36	10.27	3.34
Column Group rater	3	7.16	2.39	1.64	
Error	25	24.59	1.64		
Total	42	116.00			

Note: *Significance at $P < 0.05$.

Data in table four provided data for testing hypothesis one of this study. The data showed that the calculated F-value of 10.27 was higher than the critical ratio of 3.34 for 3 and 14 degrees of freedom at 0.5 level of significance. A significance F-ratio suggests the rejection of the null hypothesis. Consequently, the number one null hypothesis of this study was rejected. This implies that scoring patterns has a significant effect on the attitude of scorers towards the scoring of Biology essay tests.

When a Scheffe post-hoc pairwise multiple comparison test was conducted on mean reliability scores of the three treatment groups, it was observed that significant difference existed between groups I and II; and group I and group III. This implied that group I showed more attitude in scoring Biology essay test than both group II and group III while the level of attitude shown between group II and group III was not significantly different. This means that use of independent scorers pattern influenced scorer's attitude towards scoring Biology essay test than the other two patterns of scoring under study.

Effects of the three scoring patterns on scorer's attitude in scoring Biology essay test

Evidence obtained in this study regarding the level of attitude in scoring Biology essay test as two (UISP and SBSP) two scoring patterns significantly demonstrated more attitude in scoring Biology essay test than the CPSAI. It was revealed that when a post-hoc comparison of the two groups were made, the use of independent scorers pattern group showed significantly more positive attitude in scoring Biology essay test than the scoring by section pattern.

Dr. Casmir N. Ebuoh

Use of independent scorers and scoring by section pattern involved scorers' active participation in the scoring of Biology essay test. These were strategies that required the scorers to identify and name the interrelationship between and among the main idea (Amer, 1994). The findings of this study in respect of scorer's attitude in scoring Biology essay test is in agreement with the observation made by Masling (1994) that the effect of interpersonal and emotional interaction on test scoring are two factors which can likely engender scorers' attitude in scoring biology essay tests.

Attitude according to (Osisioma, 1995) if true scores of pupils are to be appropriately obtained in subject areas, there is the need to ensure that there is a positive attitude among the scorers. The ranking of all scripts before scoring all items and scoring an item across board provided the scorers the opportunities of active participation in identifying the interrelationship between and among main ideas. This had accounted for the more positive attitude towards scoring of Biology essay test under this study.

Use of independent scorers and scoring by section patterns had significant positive effect on scorer's attitude in scoring Biology essay test. Again, use of independent scorer was found to be more outstandingly efficacious than scoring by section pattern on scorer's attitude in scoring Biology essay test. The effect of the conventional method of scoring all items at a time was not significant on scorers in scoring Biology essay test.

Based on the findings of the study the following are the recommendations.

1. Since the use of independent scorers and scoring by section patterns were found efficacious in engendering scorer's attitude in scoring Biology essay test and since the techniques are not yet popular in our school system, they should be incorporated in the curriculum for teacher training institutions.
2. On acquiring the necessary skills, the teachers should be encouraged to employ these techniques more in scoring Biology essay test.
3. The opinion of the researcher is that should be given opportunity to use the scoring patterns (use of independent scorers, scoring by section) because they engender positive attitude.

References

- Ali, A. (1988) Inferential Statistics techniques. In S. O. Olaitan & G.I.Nwoke (eds) **Practical research methods in Education**. Onitsha Summer education publishers.
- Amer, D.P. (1994) The effects of knowledge-map and underlying training on the reading comprehension of science test. **English for specific purpose** 13(1), 35-45.
- Clift, J.C. & Imrie, B.W. (1981). **Assessing students and appraising teaching**. London John Wiley and sons.
- Cox, R. (1999) Reliability and validity of Education: London: Evans Brothers.
- Ebuoh, C.N. (2004) **Educational Measurement and evaluation for effective teaching and learning**. Enugu: Sky Printing press
- Ferguson, A.C. & Yoshua, T. (1989) **Statistical analysis in psychology and education**. New York: Mc Graw Hill.
- Gagne, R.M. (1997) **Condition of learning**. (3rd ed) New York: Holt Reinhart and Winston.
- Masling, J (1999) The effects of warm and cold interaction on the administration and scoring of an intelligence test. **Journal consult in psychology**. (23, 336-341).
- Nworgu, B.G. (2006) **Introduction to Educational Measurement and Evaluation: theory and Practice** (2nd ed) Nsukka-Nigeria: Hallman publisher.
- Okafor, G.A. (2000). Effects of note-taking patterns on students' academic achievement, interest and retention in Geography. **Unpublished Ph.D. thesis** Nsukka: University of Nigeria.
- Osioma, U.I.N. (1995) Effects of mode of concept mapping and gender on students acknowledgment in and attitude towards integrated science. **Unpublished Ph.D. Thesis**. Nsukka: University of Nigeria.
- Piper M.R. & Hough, L.W. (1982) The relationship between attitude towards science and science achievement. **Journal of Research in Science teaching** 19 (1). 33-38.

Dr. Casmir N. Ebuoh

Stones, E.(1982). **An Instruction to Educational Psychology**. London: Methen Co. Limited.

Ugwu, A.B.C. (1997). **Developmental psychological and education** Enugu: Fred-Ogah Publisher.

Zimbardo, P.O & Ebbessan, C. (1977) **Influencing attitude and changing behaviour**. Philippnes: Addion Wesley Publications.