Article

Measurement and evaluation in institutions of learning in Nigeria

Wahab Akanmu Aboyade
Obafemi Awolowo University, Ile-Ife

Modupeola Aduke Aboyade
Federal Polytechnic, Ede,

Abstract

The paper examines the importance of quality control and the perils of examination inadequacies in tertiary institutions in Nigeria especially in Library and Information Science Schools. It discusses various steps that are germane to good examination. The necessary steps discussed and examined exhaustively include: items generation, administration, scoring, reporting, interpretation and test security. In essence, the paper gives prominence to the construction, administration, scoring, reporting and interpretation of teacher made essay and multiple choice exams system. As part of far reaching recipes for examination inadequacies, the paper recommends that all stakeholders should be alive to their responsibilities at ensuring proper, effective and efficient examination system in our tertiary institutions in Nigeria. When this is done, it is believed that, the technological advancement of Nigeria would become realisable.

Key words: examination, measurement, evaluation, test, Nigeria.

Introduction

Examination, which is a common feature in almost all human enterprises, is also an integral part of the teaching process. It is through examination that library and information science students are tested and evaluated to find out the quality of knowledge they have acquired within a specific period. This process of logical decision making is so important that, the library educator as an examiner should be familiar with the basic principles and techniques of instrument of assessment and the technical rudiments underlying such so as to bring out from learners, learning outcome that is qualitative and quantitative.

Library and Information Science training schools sprang up in the mid-70s in Nigeria due to the dire need to grow indigenous library manpower to cater for the rising needs of the emerging libraries and information centres in Nigeria. The few universities that offer library science education cannot cater for the demands of the manpower requirements in libraries hence, the justification for establishment of library and information science departments in polytechnics in Nigeria. Librarianship as a course deals majorly

with information dissemination and information is power. Thus, a measuring instrument of good quality in terms of validity, reliability and usability is imperative when its outcome is to express students' performance in quantitative and qualitative terms. It is believed that examination inadequacies pose serious challenges on the academic performance of students who are at the receiving end.

Examination, according to Adeyemi and Akindele (2002), is the process which comes after a period of learning, and it is an organised assessment of an individual's performance on the basis of his or her institutional procedural exposure. To Nsude (1998), examinations are tools intended for the evaluation of the progress made by an individual in the course of acquiring skills or knowledge over a period of time. Such examination could be in form of short quiz, multiple choice tests, or even essay type that requires elaborate answers from candidates. Meaning that, examination has to do with the passing of value judgment on an individual based on his/her performance in a set of questions, statements or series of tasks given. Examination is believed to be confined to testing intellectual skills with the intention of assessing how much of a desired trait, skill or knowledge the individual possesses. It could come either as internal, external, oral, written or both (Benard, 1988).

Examples of internal examinations include continuous assessment, terminal, semester, annual or promotion examinations while those conducted by West African Examinations Council (WAEC), National Examinations Council (NECO) are external. Although examination is not the only instrument for assessing students' knowledge in tertiary institutions, it has however emerged as the major established yardstick and the most practical way of assessment. In the context of this paper, test and examination will be used interchangeably to mean the same thing. A test or exam is a standard measurement and evaluation technique used to determine the academic standing of a student. Whether test or exam, it should pass through certain litmus test and cover all the course contents.

Measurement and evaluation

Measurement may be regarded as a process of assigning dimensions or figures to an attribute. In measurement, standard instrument such as rulers, scales, thermometers may be used to determine how long, tall or short, heavy, hot or cold the phenomenon is. Similarly, a teacher may use a test to measure the achievement of a student. Measurement is just a collection of information relative to some established rule or standard. A raw score of 50% in an examination is a measure of achievement of a student.

The term evaluation on its own is a process of providing information for a meaningful decision making. In other words, evaluation enables decision makers to make good scientific judgement and

decisions about a given situation or phenomenon. In short, evaluation, in behavioural term, refers to judgement and classification of objects, situations, people or condition according to defined or specified criteria. A raw score of 60% in an examination is a measurement, however, comments like excellent or credit is an evaluation.

Examination as a measuring instrument performs the following functions

- To determine the extent to which students have benefitted from a course of instruction.
- It helps in predicting students' future performance.
- It assists in diagnosing students' areas of strength and weaknesses.
- It also serves as feedback to teachers in monitoring the effectiveness of their methods of instruction.
- It is used as feedback instrument to curriculum development through improvement of learning theory.
- Used for selection, certification, placement and of course as a quality control device.

Examinations could be classified according to the attributes they measure, the purpose and criteria used. However, when classified, the criterion used and the purpose for which the exam is meant must be specified.

- **Mental ability test/exam**: Mental ability exam accesses and measures an individual general mental ability. It measures the intelligence level of a person as well as certain attributes which comprise individual's cognitive skills for obtaining information on knowledge. These skills are perception, conception, memory, language, reasoning and creativity.

- **Aptitude test/exam**: It is a test/exam which measures what individuals are capable of learning to do. It could also be regarded as the potential for success in an area after a period of training. Aptitude test does not test present ability, rather that of future ability. It measures ability which is in an undeveloped form which may eventually flourish after a period of training.

- **Achievement test/exam**: This measures a person's performance after going through a course of instruction in a specified subject or field. It is often based on a syllabus which an individual has gone through in the course content in training. A person is thus examined on the content of the course covered. They may be teacher-made or standardised exams. It could also be essay or objective. Examples include teacher-made exam, WASCE, NECO exams, NABTEB exams, UTME etc.

- **Personality test/exam**: Personality refers to the totality of an

individual. The exam measures the more enduring, more lasting and more stable characteristics of an individual. Tests of perseverance, sociability, anxiety among others, are examples of personality tests/exams.

All these tests/exams could be administered formally or informally**.** A formal exam may be standardised and designed to be administered according to a standard set of rules and circumstances. It has time limits as well as sets of directions to be followed strictly. The informal exam does not have a set of standard directions to be strictly followed. It has a great deal of flexibility in administration and is usually constructed by teachers with no psychometric properties such as validity and reliability.

Tips on construction of examination items

Constructing tests/exams involves series of steps which include the following:

1. **Test/exam objective**: Objectives should be clearly defined. Any statement in this context would be a clear statement of the behaviour expected of students. The task of the teacher is to try to arrange a situation so that the learners have the opportunity to behave in the ways specified by the objectives. Objectives should be formulated to a level of specificity such that it is possible to readily infer some learning activities appropriate for helping students achieve each objective and also to devise means of assessing the achievement.

2. **Test/exam blueprint or specification**: Test/exam blue print is a two-dimensional table which shows the objectives in one direction and the content in another. A test blue print should have both behavioural and cognitive objectives. Since the teacher might have successfully covered the content and knows the objectives, the only major decision to be made is the number of items to be included in the test. This is based on the type of test format to be used either essay or objective and the time available for the test.

3. **Generation of test/exam items**: The examiner should construct as many test/exam items as possible depending on the types (multiple choice/objective, essay, fill in, matching etc). For instance, answers in a multiple choice test/exam should be plausible, meaning that, the alternatives should have equal chances of being correct.

4. **Item difficulty**: A good instrument should be moderately difficult for the group intended. In other words, a good test/exam should neither be soft nor hard. To determine the

difficulty level of a test item, the examiner needs to select a sample for which representative would be possible. That is, the examiner should find the percentage (%) of students that get the item right and those that get it wrong.

5. **Discrimination index**: A test/exam item which does not discriminate is not good for administration. A good instrument, apart from being valid and reliable, should be able to make distinction between poor and good learners. It should show the slight differences between learners' attainment and achievement that will make it possible to distinguish between poor, average and brilliant learners.

6. **Distractor analysis**: A distractor/distracter is an alternative that looks very similar like the correct answer. This concerns alternative answers to a stem especially in objective test/exam. Aside the correct answer, there should be a near similar one to distract an ill prepared examinee from the correct one.

7. **Validity or accuracy of assessment**: Any good instrument must be valid. That is, it should measure what it is supposed to measure. Validity refers to the accuracy with which a test measures what it is supposed to measure. For instance, a ruler is a

valid instrument for measuring length. If it is used to measure weight, then it is not valid.

8. **Reliability/Precision**: Reliability, on the other hand is a test/exam psychometric property that refers to the consistency to which a test measures what it is supposed to measure. Together with validity, reliability is a fundamental property which test constructors hope to achieve. It refers to the extent to which the test is dependable, stable and consistent, when given to different people and or administered in different occasions.

Hints on test/exam administration

1. **Test/exam venue**: The test/exam venue should be well prepared. It is necessary to provide maximum conducive conditions to help the examinees put in their best. There should be good seating arrangement. The seating furniture should be enough and well spaced. The spacing is necessary to prevent spying and cheating. The venue must be well ventilated and lighted.

2. **Test/exam materials**: Long before the examination date, adequate preparation should be put in place in terms of adequacy of the question papers, answer sheets, writing materials, etc. There should be more than enough so as to forestall any shortage.

3. **Training of test/exam proctors**: Proctors are assistants that help the chief examiner in the administration of the test/exam. They are also known as invigilators. The ratio of one proctor to 30 examinees is advisable for proper test administration. Proctors should be well trained on the administration of the examination. The training should involve how to distribute and retrieve answer scripts from the examinees. Furthermore, proctors should be trained on the timing of the examinees as well as maintenance of security at the test/exam venue.

4. **Test/exam administration proper**: Examinees should be kept informed in advance of the actual testing date, days and time of the examination. On the day of the examination, the proctors should call the examinees in and get them seated. After that, rapport should be established with the testees, meaning that, they should be put at ease.

Attributes of a good test/exam

1. All the examinees must be exposed to the same conditions.
2. The test/exam must consist same questions, that is, items must be the same.
3. There must be uniform conditions of administration of the test/exam.
4. There must also be a uniform scoring procedure. In objective exam, there must be one and only one correct answer.
5. It is necessary for the test/exam to have psychometric properties such as validity and reliability.
6. There must be some measures of standard in the test, that is, difficulty and discrimination indices must be established. It must also contain distracter analysis.

Teacher made examinations could be essay or objective types.

Essay test/exam

Essay exam requires an examinee to structure answer in a sentence, paragraph or short composition. It requires the testee to demonstrate ability to: recall factual, conceptual and procedural knowledge; organise this knowledge and; interpret the information critically in a logical and integrated manner. Essay question requires the testee to give a written explanation or analysis on a subject matter. It also measures higher level of thinking such as analysis, synthesis, evaluation and creativity. It is a type of examination that requires freedom of expression and response. According to Clay (2001), essay exam is appropriate for small group and when the test is not to be reused. It is also to reward and encourage the development of writing skills in students.

Benefits of essay test/exam

1. Ease of construction. Constructing essay examination is less labourious and tedious.

2. Examinees have the freedom of choice and response. That is, an examinee responds in his/her own language and ideas depending on his/her understanding of the question.

3. Essay exam involves concept definition, application of concept, integration of ideas and problem solving procedures.

4. It is also less time consuming in terms of construction. In some cases, the questions can be written on board.

5. Essay exam can be utilised to measure the mastery of language. In addition, it can be used to improve language and expression skills. This is so because, examinees are required to put down their ideas in their own language.

Demerits

1. It has a limited content validity. This is so because, it may be difficult to cover the whole syllabus in an exam.

2. Scoring is subjective.

3. It may be difficult and laborious to establish discrimination and difficulty indices.

4. It consumes a lot of time on the part of the examinee.

5. Scoring of essay questions can be time consuming.

Suggestions for setting essay test/exam items

1. Make the instruction for each type of question simple and brief. Indicate legibly the number of questions to be answered, the time allocated and so on.

2. Avoid the use of long, complex and ambiguous statements. Such statements will present some of the students with a time consuming problem of digging out the elements that are important in answering the item.

3. Test items must be drafted far ahead of time so as to minimise errors. Always trial-test your question items before setting them for examinees.

4. Test items should be clearly worded, and must not be presented ambiguously. It should be prepared to standard.

5. Test should not be worded in a manner that the response could easily be deduced from the item itself. It should rather be intellectually challenging to the learners.

6. Test items should be constructed on what the testee should be able to know or be able to make meaning out of. It should be within the limit of the course content and the scheme of work.

7. The items should be carefully reconsidered by a superior officer (Head of Department, external moderator, a colleague or an expert) to ascertain its qualities (quality control).

8. Adequate instructions should be given. Examinees should be aware

of the length required and the duration.

9.   Essay test/exam should consist of several short stems that require short answers rather than one, long and holistic question that requires extended response items.

Hints on scoring essay test/exam

Quite unlike the objective test which appears easy and simple to score, essay questions take a considerable time and effort in that the examiner has to be consistent across the examinee responses. Technically, test construction includes preparation of a suitable and dependable marking scheme. Marking scheme according to Kpala, Onocha and Oyedeji (1999), is a sketchy compilation of all points that are essential to earning the possible maximum score from a question paper. Scoring of essay exam could be **analytic** or **global/holistic**. In the analytic scoring, each answer is compared to an ideal answer and points are assigned for the inclusion of necessary elements. The elements or points expected from the examinee should have been exhaustively included in the marking scheme to be prepared by the examiner. The global or holistic approach involves reading the entire answer and then assigns a score after comparing it with an ideal answer in the marking scheme. Global scoring is suitable for longer essays and it is based on overall impression and not on points

1.   Marking scheme or scoring guide should be prepared in advance, and

should be explicit so as to be fair in scoring and to minimise examiner's biases. Marks should be assigned to points on the basis of their importance.

2.   A particular question should be scored at once for all the candidates. This improves consistency.

3.   To minimise halo effect, essay questions should be scored anonymously. That is, the identity of the examinee should not be disclosed to the examiner. Examinees should be instructed not to put down their names at the back of their answer scripts.

4.   Periodically, the examiner should check to make sure that the same criteria have been applied to all examinees.

5.   The examiner should cross through every blank page with a red pen.

6.   There should be a tick for each scoring point and a cross for a missing point on every   page to indicate where marking took place.

7.   If papers are to be returned to the examinees, make sure to write comments and   correct errors where necessary

Objective test/exam

Objective test/exam is a form of teacher made test which does not give room for subjective inferences and judgement. It is a form of test that has very clear and unambiguous scoring criteria. Objective test could be classified

into true/false, completion, matching and finally, multiple choice systems. Clay (2001) opines that, objective examination is appropriate when the population of testees involved is very large and there is the need for the test to be reused, eg., UTME. It is also suited when reliable scores are necessary. It is also free from biases of the examiner. Furthermore, there is impartiality, fairness in evaluation and it is also free from possible test scoring influences.

Benefits of multiple choice objective test/exam

1. It is suitable for machine scoring.
2. Through item analysis, the strength and weaknesses of stems can be provided and determined.
3. The scoring is not subjective. It is objectively scored.
4. It is easy to score. Even a layman can score it as long as there is a stencil or key.
5. A large percentage of the course content can be tested because, examinees do not require much time to write the answer.
6. It is suitable for computer based test (CBT).
7. It can check discrimination ability of students.

Suggestions for writing multiple choice objective test/exam

1. Always construct more items than will eventually be used in the examination. This gives the examiner the opportunity to discard items that are not very good, that is, items that are not valid and reliable.
2. Develop a test specification or test blue print.
3. All the choices provided should look plausible.
4. The correct choice for the stem should not be at the same place in all, eg., putting the right answer in option (a) or (d) alone.
5. Avoid extraneous cues (indicators). Check carefully for information a student can use to get an item.
7. Avoid the use of negative statement whenever possible and never use double negatives.
8. The language of construction of the test items should be at the level of the examinees.
9. Avoid giving clues to one item in the statement of another.

Scoring of multiple choice test

Scoring of objective test could be tackled from two perspectives, that is, manual and machine scoring techniques. The manual scoring could be stencil/template or inspection/ordinary scoring. In hand scoring, the examiner scores with the aid of the answer protocol, i.e., answer keys. It is a form of comparing physically, the answer protocols and the examinee's script. In stencil scoring, the correct answers have to be punched. The answer sheet is specifically designed to match the stencil.

The stencil is then aligned properly on the answer sheet to avoid error of marking. It is much faster than hand scoring.

The machine scoring involves specially designed answer sheets that are fed into the computer. The specially designed answer sheets have sensors in them that make them sensitive to the computer system. The marking is done by the computer.

Interpretation of result

After a successful administration of test using the above guidelines, the examiner can be sure that the test will provide information that is useful about the knowledge level of testees. Raw scores generated in an examination remain mere data until processed to become information. Interpretation of result can take the form of criterion reference or set standard system.

In criterion reference system, the performance of an individual is compared to the overall performance of the group. Depending on the examiner or the examination body, an individual's score may be marched against the group average. Thus, a score below the class or group average may be classified as poor performance while the one above the group average may be termed good performance.

In a set standard system, an individual score is compared against a set standard score by the institution usually 40% or above. Thus, a score below 40% standard is regarded as failure or poor performance while a score above may be given some grades such as ordinary pass, credit pass, excellent, etc. Some institutions may adopt alphabetical grading system such as A for distinction, B for excellent, C for credit, P for ordinary pass and F for failure. Most higher institutions in Nigeria adopt cumulative grade point average system to arrive at an evaluation of students.

Test/exam security

Test/exam security has remained an important issue in the final production, administration, scoring and reporting of examination. This is necessary in that, the time, energy and resources, both human and material, expended in having a good examination may be a waste if adequate and appropriate security is not maintained at every stage of development of a good examination. Test security may be taken to mean all the steps, processes and procedures adopted to prevent unauthorised and negative interference with test items, materials, administration, scoring and test reporting. Examination insecurity affects the examinees, the examination integrity and the testing institution from the time the first item of the exam is written to the time the results are reported.

Test insecurity or test abuse occurs from two different perspectives. The first scenario is when an examinee has a contact with life questions or items prior to the time of examination. This phenomenon is rampant in both internal

and external examination exercises and is often referred to as exam leakages. The second source of examination insecurity involves a deliberate and conscious manipulation of examination scores. This type of test insecurity or abuse is more damaging with grave consequences on the examinee, the testing institution or body, and the society in general.

Sources of exam insecurity

Examination insecurity or exam leakages could be traced to every individual connected with the test. These include test developers or writers, test administrators including test proctors, test scorers, test reporters, security officers, exam supervisors, typists, printers, among others. If test developers are careless with the test items, examinees may have access to the test items in advance of testing time. Also markers, scorers and reporters could also erode the credibility of an examination if they become unscrupulous by altering the test scores and grades. In addition, typists and test proctors who do not have sound moral character, and who do not realise the grave consequences of examination insecurity often make themselves easy source of examination abuse. In addition, the examinees themselves and their agents constitute nefarious agents of examination abuse. They go extra mile to mount pressure on personnel connected with examination to have an unauthorised access to life questions.

Consequences of examination insecurity

Examination insecurity or abuse has far reaching and profound consequences on the exam, the examinees, the examining body/institution, the society and the nation at large. Test insecurity affects the psychometric properties of the actual test. In terms of validity, the examination becomes an inappropriate and bad measure of the ability it is meant to evaluate. This is so because, students who do not possess the ability but who come in contact with the question in advance end up obtaining spuriously high scores. Also, the reliability of the test is impaired in that, the recorded performance at different occasions for different individuals remain inconsistent thus rendering the whole examination useless.

Test insecurity leads to wrong decision making. A number of personal, institutional, corporate and national decisions are made on the basis of examination results. On the part of the examinees, an individual with spurious scores may be misguided by him/herself, a teacher, parents or a career counsellor into a future course of study that is in dissonance with his/her ability. This may eventually lead to career failure and consequently deflated ego and self-concept. Furthermore, during recruitment exercises, wrong prospective employees may be employed thus leading to a square peg in a round hole. For instance, a medical doctor who is a product of test insecurity may turn out to be a butcher rather than life saver if eventually

employed on the basis of spurious grades maliciously obtained. In addition, during admission exercise into higher institutions of learning, students with spurious scores and grades may be wrongfully admitted for a course thus denying honest and hard working students of admission.

Test insecurity also erodes the integrity and credibility of the certification issued by the testing body and the testing institution. Once this happens, the public would eventually lose faith and confidence in the products and certificates issued by the discredited institution.

Recipe for good test/examination security

Conducting good examination and eventually good test/exam security is a joint responsibility of all stakeholders. Examination security is an issue that concerns all those connected with examination from items writing stage to the stage of reporting test scores as well as the general public. Some measures that could enhance test security are as follows:

1. Examination materials should be kept in secured places.
2. There should be adequate remuneration and compensation for all the personnel connected with examination from the stage of items generation to the stage of test reporting.
3. Adequate training should be given to all those connected with examination. In particular, they should be educated on the consequences of test abuse.
4. There should be greater penalty for test abusers irrespective of their status

(students, teachers, supervisors, exam scorers, recorders etc).

Conclusion and recommendations

In conclusion, for best practices, global acceptability, quality control and assurance to be achieved in teaching, learning and research activities in our higher institutions in Nigeria, it is imperative to focus on the right steps towards achieving excellence in measurement and evaluation of library and information students in particular and all the students in general. Examination has become a unique instrument for gauging the success of students in any academic pursuit. As such, good examination involves item generation, item administration, item analysis, item scoring, item reporting and test security. It is necessary to ensure quality control and assurance at every stage of assessment and evaluation of students' academic performance. When this is done, it is envisaged that, the much talked about technological development of Nigeria and vision 20:2020 would soon become realisable. This guide is therefore, a way of providing support to quality testing in our nation's library training schools in particular and tertiary institutions including universities in Nigeria. To this end it is hereby recommended as follows:

- All library educators, lecturers, course advisers and examining bodies/officials should be kept informed about their expected role

and the consequences of their actions.

- Team effort on the part of all academics to tackle examination inadequacies at all levels of our educational system is essential.
- All examiners should undergo requisite training in measurement and evaluation so as to always take cognisance of examination ethics.
- Apathy on the part of examiners must be highly discouraged and all stakeholders in education sectors must be kept alive to their responsibilities.
- All the stages involved in the conduct of good examination such as items generation, administration, scoring and security should pass through a litmus test so as to have appropriate and objective assessment and evaluation of our students.
- Institutions of higher learning must learn to review their examination policies and methods from time to time so as to achieve the best approaches of assessment andlearning outcomes.

References

Adelabu, I.T. (2012). *Course unit system.* Being a paper presented at the 2012 orientation programme organized for the fresh students at the Federal Polytechnic, Ede.

Adesokan, A. (1994). *Fundamental principles of evaluating students'*
achievement for prospective and practising teachers at all levels of educational system. Ibadan: Dove Communications.

Adeyemi, J. & Akindele, I. (2002). Examination malpractices in Nigeria educational system: causes, effects and the way out. *Nigerian Journal of Clinical and Counselling Psychology,* 14(1), 15-18

Alkin, M.C. (2001). Product for improving educational evaluation. *Comment*, 2(3), 1-4

Anastasi, A. (1996). *Psychological testing.* New York: Macmillan.

Awofala, A. O. A & Babajide, V. F. T. (2013). Examining attitude towards continuous assessment practices among Nigerian preservice STM teachers. *Journal of Education and Practice*, 4(3), 37-49.

Brown, F.G. (1983). *Principles of educational and psychological testing.* New York: Holt, Rinehart and Winston.

Clay, B. (2001). *Is this a trick question? a short guide to writing effective test questions.* Retrieved from http://www.k-state.edu/ksde/alp/resources/Handout-Module6.pdf.

Egbegbulem, C.E. (2012). *Constructing tests for continuous assessment.* Ibadan: Macmillan Nigeria.

Federal Government of Nigeria (2004). *National Policy on Education.* Lagos: Federal Government Printing Press.

Nsude, I. (1998). *Principles and practice of education (Education I)*. Enugu: Hillys Press.

Ogunsanya, M. (ed.) (2000). *Foundation courses in education*. Ibadan: Vantage Publishers.

Okoji, B. I. (2014). *Ethics in examination administration and human behaviour*. Being a paper presented at the School of Business Studies Annual Seminar held on Monday, 1st December, 2014 at Akogun Lawal Hall, Federal Polytechnic, Ede.

Oladunni, M. O. (2010). Students' perception towards examination and examination ethics in College of Education, Ikere-Ekiti, Nigeria. *Journal of Sociology, Psychology and Anthropology in Practice: Int'l Perspective*, 2(1-3), 47.

Author biography

Wahab Akanmu Aboyade PhD, is a Principal Librarian with the Hezekiah Oluwasanmi Library, Obafemi Awolowo University, Ile-Ife, Nigeria. His areas of interest are: human resources management, ICT, Library administration etc. He has published in both local and international journals. He can be reached at abowakan@yahoo.com:

Modupeola Aduke Aboyade is a Senior a lecturer and head of Department of Library and Information Science, Federal Polytechnic, Ede. Osun State. She has published in both local and international journals and has also participated and read papers at both local and international conferences. She can be contacted at abowakan@gmail.com: