TOWARDS IMPROVING THE QUALITY OF CLASSROOM ASSESSMENT TOOLS IN NIGERIA

ILOGHO, FEGHA EVELYN, Ph.D
Email: feghaevelyn@gmail.com

and

OSAKUADE JOSEPH OLUWATAYO Ph.D
Department of Guidance and Counselling,
Faculty of Education Adekunle Ajasin University,
Akungba Akoko, Ondo State, Nigeria
Email: osakuade_tayo@yahoo.co.uk

Abstract
Assessment is an integral part of education. The challenges being faced by students in terms of their learning outcomes such as interest, attitudes and achievement could be as a result of poor quality assessment tools being used by teachers.If the quality of assessment tools is not well ascertained by teachers, it would have negative implications on decisions made onstudents. Wrong decisions could mar the future of the learners as right decisions can improve students' learning in the classroom. Assessment can play a vital role in improving education if used well and is of good quality. Hence, this paper examines the meaning of assessment, functions of assessment, assessment tools, towards enhancing the quality of assessment tools in classroom. It was concluded that if the quality of assessment tools is ascertained, it will go a long way in taking accurate decisions on learners. The way forward to further enhance the quality of assessment tools were recommended.

Keywords: Assessment, assessment tools, classroom assessment, quality, psychometric properties

Introduction

Assessment is said to be an integral part of teaching learning process. Every passing day in the education sector is to find ways of improving teaching and learning. Assessment can be a vital tool in the efforts towards improving education if it can be implemented well. Assessmentrefers to the set of procedures for obtaining information about a person or thing. It is a means of gathering information about a subject

ability or behaviour for the purpose of making informed decisions about the subject and future instruction (Elliot et al, 2000). Assessment is the act of measuring learning outcomes and other attributes so as to make decisions about an individual. Okpala and Onocha (1994) defined assessment as the process of fashioning data into interpretable form using data collection methods. Some decisions that can be taken based on the scores obtained from assessment tools according to Alonge in Osakuade and Aina (2020) are: instructional decisions by classroom teachers, curriculum decisions by the school authorities and Ministry of Education, selection decagons by employers or educational institutions, placement or classifications decisions by examination bodies like WAEC, NECO, and NABTEB, personal decisions made by friends and administrative decisions made by the school administrators.

There are many assessment tools that can be used for collecting datain education. Examples of assessment tools used in education are: questionnaire, tests, rating scales, checklists, interview, observations, opinnionaire and so on (Bandele, 2004)

Functions of Assessment

Assessment serves several functions in education. Some of the functions according to Ogunniyi in Ukwuije (2009) are:

1. To determine the relative effectiveness of the programme in terms of students behavioural output
2. To make reliable decisions about educational planning
3. To provide a just basis for determining at what level of education the possessor of a certificate should enter a career
4. To understand and improve students' learning
5. For systematic data collection, analysis and interpretation
6. Determine whether learning meets expectations and standards
7. To document, explain and improve performance.

There are different forms of assessment available to the teacher in the school setting. These include:

1. The closed book assessment method: This is the method in which the student is required to answer a number of questions which are not known to him. The student is expected to recall what has been taught before, put it down on his/ her own to complete the questions given

over a stipulated period of time. Eg. SSCE/NECO Examinations, HND and final degree examinations.

2. The open book assessment method: as the name implies, this is where students are allowed to use textbooks, lecture notes or handouts during testing.
3. The pre-published assessment method: This is the method where students have been informed before testing where questions will be asked or where students are given area of concentration in advance before the date of testing.
4. Open time (take home) method: this is when the examinees are allowed to take questions home to prepare their answers at their own pace and submit at a specific time given by the examiner.
5. Oral assessment method: this is a face-to-face situation where students are asked questions in form of interview by an examiner.
6. Practical assessment: this is a practical demonstration of what is learnt by student in a course is applied, and this could be done in the laboratory or field work.

   Assessment is one of the vital components of teaching learning process. Assessment provides a means for teachers to obtain the level of achievement of students as well as providing feedback to parents, guardians, students and education stakeholders. Assessment can be formative and summative in nature.

Assessment involves collection of a lot of data on the pupils. These data are not only on the cognitive domain but also cover the affective and psychomotor domains. In the old system of assessment, only tests were made use of. With tests only, cognitive domain was assessed. Continuous assessment makes use of instruments that apply to all the three domains. Tests, class assignments, homework, quizzes, interview and projects are the instruments used to assess cognitive domain. Interview, questionnaire, project, checklist, observation, anecdotal records and social metric technique are used to assess affective domain. Project, checklist and observation are used to assess psychomotor domain (Gbore, Osakuade & Ilogho, 2020).

1. **Test**s are in form of written exams applied within or at the end of the term or year to assess progress
2. **Class assignment or homework** is given at the end of a lesson or a group of lessons. **Quizzes** are given to assess the readiness of the

learners. This is why they are given in impromptu that is unannounced) basis.

3. **Anecdotal records** are records of importance of incidental behaviours. Sometimes, the learner exhibits certain important but not so common behaviours e.g returning lost items, accepting responsibility for an action that was committed when nobody was there, extreme quietness in the class, unruly behaviour in the class, exhibiting high level of leadership roles. Records of these types of behaviours can be kept and later used for assessment purposes. These records are known as anecdotal records. They are therefore records of incidental behaviours.
4. **Checklist** is particularly useful if behaviour can be broken down into specific tasks, then it will be possible to check which tasks are present and which tasks are absent.
5. **Observation** is very useful especially as it gives a reliable set of information provided the observer is not biased. Observation generally requires a schedule which specifies in detail what guidelines to follow in rating what is being observed. In this method the researcher enters into the field, becomes a part of the population under investigation. For instance if someone wants to use the method of participant observation to examine and report a case of masquerade groups in oral literature, the observer may be expected to join the group in order to obtain the information.
6. **Interview:** This is a systematic way of obtaining detailed information about a person's opinion, beliefs, experiences and so forth, through a one-on-one interaction between the researcher and the respondent. In-depth interview involves the interviewer engaging the respondents in a personal or person-to-person interaction with the client in order to obtain information concerning the variables under study. Information can be obtained through a questionnaire. It is unstructured, that is, the respondent is given freestyle on the topic of interest. There are no restrictions as regards the structure and timing of responses. It is important to note that interviews can be recorded with either a video or an audio recorder but care must be exercised to obtain permission from the interviewee before adopting either. Also, the interviewer should transcribe such information immediately after the interview to reduce loss of information. To obtain much information from the respondents, it is necessary that the interviewer uses open ended questions and

encourage the interviewee to respond in a detailed manner to the raised questions. With this, the interviewee freely expresses his opinions, experiences and beliefs concerning the subject. The interviewer is also supposed to provide a plan of action, an interview protocol and guide, if possible and when needed train interview assistants, collect, analyze and interpret the data. The interviewer must be a good listener, must be patient, must be open-minded, must be able to transcribe the information obtained from respondents or interviewees and must be able to note non-verbal cues in the interviewee. Hence, it could be in the following form; structured, semi-structured and unstructured.

7. **Questionnaire:** A questionnaire is an assessment instrument that is mostly used by individuals/counsellors to rate themselves on a specific list of attributes. It could also be used by the individuals to respond to questions about themselves regarding such matters as their feelings, attitudes, interest and so on. A questionnaire can be *open-ended* (i.e open response) or *closed-ended* (i.e structured response). In an open-ended type of questionnaire, the options are not given; instead the respondents are free to provide their response using any format that suits them while closed-ended questionnaires, options are provided to the questions or statements (items). That is, the respondents are limited in their responses to their items. This makes the analysis of the questionnaires easy and gives a proper direction as regards the trait being measured.

Towards enhancing the Quality of Assessment Tools

Quality according to Adeyemo (2011) means the worth of something. According to UNESCO as cited in Ilogho (2022) defines quality as a combination of input (resources put in place), process (teachers methodology, classroom environment) and output variables (students' performance or results). Quality of education is defined as the systematic management, supervision, monitoring, assessment, evaluation and strategic review of the resources, input, teaching-learning process and output continually, in order to achieve academic standard that meet the set goals and expectation of the stakeholders in the classrooms and education sector at large.

Quality assessment according to Afolabi (2001) is vital in determining the level of education and improvement in teaching learning

process. A good assessment tools according to Osakuade and Aina (2020) must be valid, reliable, objective, usable and practicable, comprehensive and precise, easy to administer, economical, easy in scoring, easily available and so on.

Validity of Assessment Tools

Validity is simply the extent to which a test measures what it is supposed to measure. The degree of validity is the single most important aspect of a test. Hassan (1984) defined validity as the extent to which an instrument or test is capable of achieving certain aims. To further explain, to determine how accurate or valid a test is, it is necessary to gather the appropriate evidence needed for the various aim of testing and this lead us to the types of validity.

Significance of Validity of Assessment Tools

Anikwezein Osakuade and Aina (2020) is of the view that it is very important for teachers to ascertain the validity of assessment tools because of the following reasons:

i. For accurate and reliable prediction of students' future success
ii. Valid test can serve as a reliable reference point for the promotion of learners to higher class
iii. Judgments based on the results from valid tests cannot be contradicted
iv. For the objective grading of learners
v. To confirm if desired changes have taken place from the teachers' instruction
vi. To provide a reliable basis for the comparison of teachers' efforts

Types of Validity

i. **Content validity.** The extent to which the items in the test adequately covers the contents or domain of behaviour it intends to measure. In evaluating the content validity of a test, experts in the field are simply expected to decide whether or not the items covers the content of instructions or subject matter of instruction and the behaviour that should be included in the test. Such experts are to decide whether the instrument covers the domain or behaviour to be measured and are evenly distributed under the topic in the subject area.

ii. **Construct validity.** This is the extent to which a test measures a specific trait or construct. It is the determination of the extent to which scores on the test are related to measures of behaviour in situations where that construct is thought to be an important variable. Construct validity is the degree to which the test score can be accounted for by certain explanatory construct in a psychology theory. A construct is an ability, aptitude trait, attributes or characteristics that are hypothesized to explain some aspect of behaviour such as mechanical ability, intelligence, anxiety or extraversion. Construct validity determines the extent a test is consistent with a given theory or hypothesis under consideration.

Criterion-related validity is divided into;

- Predictive validity
- Concurrent validity

Criterion-related validity is the extent to which scores obtained with an evaluation instrument are in line with current criterion measured (concurrent validity) or predict future criterion measured (predictive validity).

a. Predictive Validity: A test that makes exact prediction concerning future performance for which it is designed is said to be predictive validity.

Steps for establishing predictive validity are;

- To administer the test (predictor)
- Wait until the performance predicted by test matured
- Correlate the test (predictor) score and the actual performance the test is designed to measure.

The higher the correlation coefficient the more effective is the test in predicting the criterion measured.

Concurrent validity. This is when the teacher is comparing or relating performance to some other current measure of performance. Scores obtained from instruments which the teacher wants to use are correlated with those obtained from another instrument. For example scores from mathematics multiple choice tests could be concurrently validated with chemistry practical tests. A high correlation coefficient (of about 0.70) is a

mark of good concurrent validity and the newly developed multiple choice tests could be guided to be suitable for a practical test.

According to Gbore et al (2022) the following are the factors affecting the validity of assessment instrument:Length of test, Language content, Cultural bias of the test items, Identifiable patterns of answers, Clues within the test items, Vague instruction, Poor test administration and scoring, Spread of scoring and Factors affecting students response

Reliability of Assessment Tools

Reliability simply means the degree of consistency of assessment tool. That is, a tool is reliable if it measures whatever it intends to measure. For example, mathematics teacher administered a test of mathematics on his students on two separate occasions. The responses were scored and item analyzed. The teacher observed that some students who obtained high scores on the first administration had lower scores on the second administration. A few students obtain the same scores on both occasions. This illustration shows the test scores were not consistent.

Reliability refers to the degree of consistency between two sets of scores obtained with the same instrument or equivalent forms of instruments.Correlation coefficient is what is used in determining the degree of reliability. A correlation coefficient of + *1.00* is a perfect positive relationship,-*1.00* is a perfect negative relationship while one of *0. 00* indicates no relationship. The nearer the correlation is to + *1.00*, the more reliable the results.

Reliability of test scores can be affected by variation in conditions of administration, item contamination and faulty marking procedure.

Significance of the Assessment Tools

Anikwezein Osakuade and Aina (2020) also justifies the need for teachers to ascertain the reliability of assessment tools because of the following reasons:

i. Assessment tool with established reliability index makes testing meaningful and dependable
ii. The reliability of Assessment tool attests to the teachers' honesty in evaluation
iii. Only reliable test produces test scores that can usually discriminate learners' abilities

iv. It is tests with proven reliability that can perform the important function of motivating student learning
v. Teachers can depend on reliable test outcomes as basis for realizing prediction of learners' future attainments
vi. Only reliable teacher-made tests can effectively prepare students for more serious external examinations.

There are four methods that are used in estimating the reliability of assessment instruments (Ogunmakin, 2003). They include:

1. **Test- retest:** This involves giving out a test to the same set of students on two different occasions. The two set of scores obtained are used in computing a correlation coefficient which is interpreted as an estimate of reliability.

   The same test is administered twice, with a time interval of two weeks between administrating; the reliability obtained is a measure of consistency of the test scores. If the reliability coefficient is high, then student with high in the first one will tend to have high score in the second one.
2. **Equivalent Form:** This is concerned with two tests which are equivalent or parallel, similar and measuring the same thing. Two tests which are equivalent/ parallel are administered and the scores are computed to find the correlation coefficient between these two sets of scores, the correlation coefficient obtained is the coefficient of reliability.
3. **Split Half Reliability:** This requires the administration of a test to a set of students once and dividing the responses into two halves A and B.

For example a test of 20 items

A= 1-10 B =11-20 or

A= Odd numbers, B= Even numbers

The correlation coefficient obtained between A and B is the reliability coefficient. We use Spearman Brown formula in estimating the reliability coefficient.

$r_h = \frac{2\,X Reliability of \frac{1}{2} test}{1 + Reliability of \frac{1}{2} test}$ or $r = \frac{2r\square}{1+r\square} - 2$

Where r is estimated reliability of the whole test and r_h is the reliability of half of the test. For example if the correlation coefficient between the two halves of a given test is 0.70 the reliability of the test

$$= \frac{2\,(0.70)}{1+0.70} = \frac{1.4}{1.7} = 0.82$$

4. Kuder - Richardson Reliability

This is the fourth method in estimating the reliability of a test. It is a single administration of test. It is based on the agreement or consistency of all items in the test. The underlying assumption is that whatever the test tends to measure, the items must exhibit high degree of agreement among themselves, then they are consistent.

KR-20 $r = \frac{K}{K-1}\left[1 - \frac{\sum ppq}{\partial^2}\right]$

KR-21 $r = \frac{K}{K-1}\left[1 - \frac{M(K-M)}{KS^2}\right]$

Where r = reliability coefficient

K = number of items in the test

p = proportion of people who got the item right

q = proportion of people who got the item wrong.

S = Standard deviation of the test scores

$\sum$ = summation sign

$\bar{x}$ = means of the entire test

According to Alonge (2003) the following factors affect the reliability of assessment instruments: Length of test, Spread of scores, Speed of the test, Objectivity of scoring and Test difficulty

Assessment Tool Usability and Practicality:

This implies a good Assessment tool should be useable in the sense that it must have simple and clear instruction, time required for administration should be stated, ease of scoring, ease of interpretation, such test should have equivalent form and should discriminate adequately among pupils/testees. The cost of producing test should not be outrageous both in terms of time and money. There must a good balance of efficiency and cost. A good test should have a good impact on teaching and learning process.

Conclusion

It is imperative to use quality assessment tools in the classroom setting to be able to take valid and accurate decisions on the learners. Assessment tools should be valid and reliable before use in the classroom. The principal developer and user of classroom assessment toolsare the teachers. To maximaze the benefits of assessment, the teachers should be properly taught on how to construct and validate assessment tools. Besides, they must know and when to use different assessment tools in the classroom in order to yield the desired goal of improving learning outcomes in the school system.

The way forward

Assessment involves a lot of tools which if properly constructed and validated would lead to improvement of teaching and learning. To further ensure the quality of assessment tools in classroom, the following are the ways forward:

1. Teachers should be involved in the development and validation of assessment tools for their students
2. All assessment instruments should possess all the psychometric properties
3. Seminars, conferences and workshop should be organized regularly for teachers to enable them ensure the quality of assessment tools
4. Teachers should be ready to introduce new innovations to their instructions while teaching in the classroom.
5. School records must be properly kept for continuity
6. Teachers should use assessment results to improve quality of their classroom assessments in agreement with their learning objectives
7. Assessment results should be used by teachers to help students to learn and improve their learning, not to separate those who fail and pass.

References

Adeyemo, E.M. (2011). Attaining standard through quality in assessment practices and record keeping in Nigerian secondary schools. *Nigerian Journal of Educational Research and Evaluation 10(3),* 192-199.

Afolabi, E.R. I. (2011). Modern techniques of students assessment and record keeping in Nigeria primary schools being paper presented in a seminar at OAU, Ile-Ife, 29th -30th October, 2011

Areekkulnyil, S (2021). Issues and concerns in classroom assessment practices. *Edutracks*, 20(8) 20-23.

Gbore, L.O, Osakuade, J.O. &Ilogho, F.E. (2022). Module writing on Edu 402: Test and Measurement. Institute of Education, Adekunle Ajasin University, Akungba Akoko, Ondo State.

Ilogu, B.N. (2022). Assessing the quality of English language achievement items of senior secondary II Students in Lagos State using factor analysis in Creent issues in Education: Trends, views and Analysis. Published by Esthorn Graphics Prints 44-49.

Ilogho, F.E. (2022). Effect of Assessment for learning strategy on Students achievement and interest in senior secondary schools in OndoState.*EastAfrican scholars journal of education, Humanities and Literature 5(5)*133-138. Kenya.

Osakuade, J. O. & Aina, B. B. (2020). Reliability and Standard error of, measurement: Implications for teachers' decisions on students' academic achievement. *International Journal of Educational Research and Development, 1*(1), 11-20

Ukwuije, R. P. I. (2009). Test and measurement for teachers. Chandik Printing press, Port-Harcourt, Nigeria

Wagbara E.S & Ubuluom (2022) Utilization of information and communication technology in educational assessment practices in River State owned universities.*Nigerian Journal of Educational Research and Evaluation 21(1),* 240-254.