

1-1-2011

Assessment Practices in English Language at the Nigerian Secondary School Level: A Psycholinguistic Issue

Oluwole Akinbode

Anthony Dairo

Follow this and additional works at: <https://scholars.fhsu.edu/alj>



Part of the [Educational Leadership Commons](#), [Higher Education Commons](#), and the [Teacher Education and Professional Development Commons](#)

Recommended Citation

Akinbode, Oluwole and Dairo, Anthony (2011) "Assessment Practices in English Language at the Nigerian Secondary School Level: A Psycholinguistic Issue," *Academic Leadership: The Online Journal*: Vol. 9: Iss. 1, Article 11.

DOI: 10.58809/HETC4009

Available at: <https://scholars.fhsu.edu/alj/vol9/iss1/11>

This Article is brought to you for free and open access by the Peer-Reviewed Journals at FHSU Scholars Repository. It has been accepted for inclusion in Academic Leadership: The Online Journal by an authorized editor of FHSU Scholars Repository. For more information, please contact ScholarsRepository@fhsu.edu.

Academic Leadership Journal

Introduction

The major purposes to which the results of assessment could be put were diagnosis, evaluation, guidance, grading prediction, and selection. In Nigeria, it is unfortunate that the results of many examinations have been used almost exclusively for grading and promotion exercises and inevitably there has been a neglect of diagnosis, guidance and evaluation although considerable informal use of assessment for these purposes are made by teachers in the classroom.

The Concept of Assessment

Assessment is the process of organizing test data into interpretable forms on a number of factors (Belts, 2001).

In order to obtain data, a wide variety of instruments, such as tests, questionnaires and observations are used.

In all Nigerian secondary schools the most reliable method of measuring students performance is assessment and the process of assessing students' performance is done by using a variety of techniques among which testing is one.

The Concept of Testing

Belts (2001) defines a test as an instrument or system procedure for measuring a sample of behaviour. An examination is a more formal term used for practical or oral test to appraise students' progress, ability or knowledge at the end of term, or session for school based assessment.

The aim of testing may be to classify, diagnose, compared with established norms on standards or simply to check on understanding and learning. Testing is the most widely used technique for assessing students (Macintosh and Hale, 1976). Tests constructed by teachers help to appraise all the things they want their students to remember, understand and think about. Also, this will enable a decision to be taken regarding those who may need individual attention.

The Concept of Evaluation

Evaluation is the systematic process of collecting analyzing and interpreting information to determine the extent to which students are achieving instructional objectives. It also refers to value judgement about a learner's level of performance, using different assessment instruments (Macintosh and Hale, 1976). It enables the teacher to determine how much students have learnt in any given subject at the end of the term or year. It also helps the students and their parents to know how well they are progressing; hence, they will be able to facilitate their children's work at home.

Usefulness of Assessment

1. Tests may be used to determine the selection of elements and the way in which they are to be tested.
2. Tests may also indicate how well certain material has been taught and assimilated by the students.
3. Tests can also be conducted to eliminate some applicants and select a few outstanding candidates from a large group of students from different institutions of learning. This type of test is referred to as selection test.
4. Tests can also be used as the determining factor in placing a new student in a class working at an appropriate level.
5. Language aptitude tests are also developed to identify students who will learn English with some facility.
6. Tests can also be used by the teacher to help a particular student with serious problems in some area of skill.
7. Tests can also be used to identify underachievers in English language class.
8. Tests can also be devised to establish the level of proficiency which a student has reached.
9. Achievement tests are administered to groups of students from different institutions who are presenting themselves as candidates for a common terminal examination at the end of a certain period of study such as the Senior Secondary Certificate Examinations conducted by the West African Examinations Board (WAEC) and National Examination Council (NECO) in Nigeria.
10. Achievement tests are usually based on a public course of study, however sketchy and students expect to be tested on what they are supposed to have been learning. Such tests have considerable influence on methods of teaching since teachers try to ensure that their students are adequately trained in the particular areas which will be tested.

Purpose of Testing

Tests may be designed to indicate to both the teacher and student areas of strength and areas of weakness. The results of such tests will show what sections of the work should be re-taught or restudied and where further practice is essential; thus, giving the teacher a clear indication whether the students are ready to move on to new work.

These tests may be called diagnostic tests. Diagnostic tests are reduced in usefulness if they are not corrected thoroughly, returned promptly to the students, discussed in the class and rewritten where necessary by the students. Tests can also be used to determine whether the students can go to the next level or not. Tests of this nature are referred to as placement tests.

The Concept of a Language Test

A test can be referred to as a sample of behaviour while a language test can be referred to as a sample of language behaviour. In a typical language test, the student is presented with a set (i.e.

sample) of items to be answered or tasks to perform. It is impossible to present the student with all the items or tasks that we can think of (Williams, 1990).

In defining what a language test is, one needs to clarify two very important concepts; performance and competence in the language.

There is a clear distinction between performance and competence. Performance, according to Williams (1990) is elicited and observed while competence can only be inferred as an underlying ability. Competence has to do with whether a specific language ability is present; while performance has to do with the strength of that ability.

There are several ways in which the distinction between performance and competence can be exemplified. These have important implications for psychological procedures in language testing. It is possible for somebody to have an idea or a knowledge of a particular thing – he is about to remember a name of thing; the name does not just come out at the precise moment when it is needed. The memory of that person has actually failed at that particular point in time. He may remember the name later. This is natural with every human being. Humans memory cannot be trusted at all times.

There are at least two implications for the psychology of language testing from the above example:

1. Reduction of anxiety/Non-linguistic Variables: Care should be taken to remove from the testing situation anxiety or any such non-linguistic variables that will interfere with performance to the extent that they may mask the potential competence of the individual(s) being tested (Williams, 1990).

In any testing situation requiring the individual or the group to speak, a real difficulty lies in distinguishing between testing oracy as such and testing ability to be interviewed. According to Williams, (1990), testees may be nervous not because of fluency in English (or any language for that matter) but because of personality factors, or sheer unfamiliarity with the testing situation.

There are yet other extraneous variables to be taken into account. Those being tested may have to compete with outside noises in a typical classroom setting. The examiners themselves may be tired. Examiner's fatigue has been known to have an adverse effect on interaction in oral tests and consequently on test administration and scoring procedures.

2. Adequate Sampling at Receptive and Productive Levels: If competence is to be correctly inferred, there is need for proper sampling of the different language areas, receptive and productive. This is necessary because receptive and productive skills do not develop at the same rate among second language learners.

A student may be poor at reading aloud and good in answering listening comprehension questions. By the same token, he may be good at reading comprehension and poor in written composition. A comprehensive estimate of his language ability can only be arrived at from an adequate sampling based on scheme such as the following:

Linguistic Items	Communicative tasks
Receptive	

Productive

In word or sentence contexts, a student may be tested for his ability to recognize or produce items of phonology, grammar and vocabulary. Beyond the sentence level, he may be given tasks for assessing his receptive and communicative skills (Williams, 1990).

Qualities of a good Language test

In determining qualities of a good language test, two basic concepts which apply generally to evaluation in teaching must be considered. These are the concept of validity and reliability.

Validity

A language test, like any other test, is valid when it measures what it sets to measure. For example, for a test of reading comprehension to be valid, the questions must be tied to the passage, otherwise, they remain questions of general knowledge or other subject matter, and questions of reading comprehension as such.

Types of Validity

At least five types of validity can be distinguished.

1. Face Validity: The test looks as if it measures what it claims to measure. It is accepted by all as fair and credible.
2. Content validity: The test is an adequate sample of the content of the syllabus. One may reason that a School Certificate Examination has content validity if it reflects syllabus content in composition, comprehension, summary, lexis and structure.
3. Predictive validity: If language tests at one educational level predict performance at another level, or in courses and jobs where language ability is essential, we can then say that they have predictive validity.
4. Concurrent Validity: This is similar to predictive validity in the sense that it refers to performance in another situation as a criterion. The criterion is a test that has been validated. A teacher may want to replace an elaborate rating scale for assessing written composition a simpler one. He obtains similar results from both scales, and so establishes concurrent or criterion validity for the simpler one (Williams, 1990).
5. Construct Validity: The test is validated against a theory. It is connected on the basis of what the tester believes to be the nature of ability to be tested. For example, he may have a theory or construct of what reading is, of the factors that make up reading ability. So he embodies his theory in his reading ability test. He tests hypothesis derived from his theory. If these are supported by the results, his reading ability test can be said to have construct validity.

Reliability

Reliability refers to consistency or stability of measurement. There are two major types of reliability:

1. Scoring procedure/rater reliability
2. Internal consistency/item reliability

Two kinds of reliability in connection with scoring procedure can be distinguished:

(a) Intra-rater reliability: A teacher's mark for the same composition exercise may vary in accordance with his physical condition, emotional state, the time of day, or any other circumstances that might affect his judgement. He may also give a different mark if he were to assess the same composition after an interval of a month or a term (Williams, 1990).

(b) Inter-rater reliability: If two teachers mark the same composition exercise, the likelihood is that their marks would not tally closely unless they had coordinated their assessment on the basis of some marking scheme.

2. Internal consistency/item reliability: Another technical consideration in testing is reliability of test items. It is possible to have 100% agreement among raters/markers and yet not be able to rely on the goodness of the test in terms of results, because the test items themselves are not reliable. There are two points to be considered here:

(a) Some items may be too easy or too difficult, and

(b) Some items may not disseminate sufficiently between more able and less able students. (Davies, 1986, Williams, 1990).

Types of Tests/Testing

Objective testing and subjective testing: Objective testing is the type of testing in which the tester or rater can achieve maximum rater reliability while the problem of item reliability remains. In subjective testing, there is definitely a problem of rater reliability. Each of the two approaches to testing has its relative advantages.

The terms objective and subjective refer to marking or scoring. A multiple-choice test can be marked with 100% reliability by the teacher, no matter how many different times or under what different conditions he does the marking. The same result can be achieved by any number of markers. Reliability in marking is guaranteed once the correct choices are agreed upon.

One does not think in terms of 100% reliability in marking essay tests. Such tests are subjective in the sense that the marking must vary from one person to another. Subjective marking can be done globally by general impression: or piecemeal, with reference to a marking scheme. No matter how detailed a marking scheme is, the element of subjectivity cannot be ruled out (Williams, 1990).

Selection of items for objective test necessarily involves subjective judgment, whether the items are selected by one tester or a group of testers. One cannot guarantee 100% item reliability for ease/difficulty or discriminative value of the items. The following are examples of objective questioning:

1. The man ... the award is my uncle
- (a) which (b) whom (c) who (d) when

2. She forgot her dictionary at the hostel, so Paul..... her his own.

(a) loans (b) borrowed (c) lent (d) sponsored

Subjective testing is appropriate for oral production and written composition because these are problem areas where irrelevant personal biases are likely to creep in. For example, continuous writing may be assessed on the basis of whether the examiner agrees (or not) with the idea expressed by the candidate. Handwriting may cause the examiner to add or deduct marks, depending on whether it is neat or not. There is what is known as the 'halo-effect', an error that occurs when a rater's general impression of a person affects his rating of that person on individual traits.

Williams (1990) recommends the following procedures if subjective rating is to be relied on:

1. The rating should be confined to characteristics that can be directly observed e.g. the pronunciation of segmental and supra-segmental features in speech; appropriate vocabulary and variety of structures in continuous writing.
2. Points on rating scale should be precisely defined. Descriptions of each characteristic should be given. This is very important when one is dealing with terms such as "clarity" and 'fluency' of expression in speech or writing.
3. The number of points on a rating scale or schedule should be exactly designated. Fewer points on the scale are needed in assessing performance traits that permit only crude or broad judgements.
4. The pooled judgement of two or more raters will provide a more reliable estimate of performance, cancelling out the personal biases of individual raters. But this is not possible in classroom circumstances.
5. In tests of speaking or writing, more than one topic should be given in order to obtain a more comprehensive sample of the student's language ability.

Subjective marking by impression no doubt, takes features such as errors into consideration. It goes beyond the neutral or negative assessment; giving credit to relevance of content, coherence of thought, clarity and appropriateness of expression (Williams, 1990, Valette, 2002).

Criterion-Referenced and Norm-Referenced Tests

The purpose of a criterion-referenced test is to the ability of an individual without reference to the performance of others. The purpose of a criterion-referenced test is to discriminate between individual performances, to spread them out so that they fit some sort of curve. School Certificate Examinations results are reported in this way; a decision is made about the percentage of distinctions, credits, passes, and failures. For example, a candidate who scored a p8 with a mark of 34% in a given year would have found himself with an F9 the following year with the pass mark raised to 35% for English Language. The P8 and F9 are norm-referenced grades.

The implications which criterion-referenced and norm referenced tests have for English language teaching have to be based on mastery learning, and mastery testing.

Mastery learning is concerned with the barest essentials of the syllabus. These are skills which must be mastered by each student if he/she is to proceed to the next level of instruction.

Mastery testing on the other hand, enables the teacher to determine whether the essential skills have been mastered.

Discrete-Point and Integrative Testing

Discrete-point testing is quite literally, testing one item at a time. It may be a unit of pronunciation, or grammar or vocabulary. The word discrete means 'separate'. In discrete-point test, items of language are dealt with in isolation, as follows:

1. Listen to these words (e.g. 'pail/bail') and say whether they are the same or different (sound discrimination).
2. Complete this sentence using the correct form of the verb in brackets (Grammar (Tense)). I ... the letter last week (write).

A discrete-point test is referred to as that because of isolated bits of test content, because of the manner of response to the test items, and the way in which the responses are scored. One correct choice is required in a sentence-completion multiple choice, or some different format.

Integrative testing does not focus on isolated bits of language. It seeks to measure ability to make appropriate responses by harmonizing different language skills and a knowledge of different aspects of the language. It involves listening and reading comprehension tests in which inferences are to be drawn, deductions to be made, and the main point is to be grasped by weaving together details of vocabulary and grammar within and between paragraphs (Williams, 1990, Akinbode, 2004).

There are two kinds of reading comprehension questions which are particularly suitable for integrative testing. They can be categorized as follows:

1. Question for testing ability to discover relationships of thought in sentences, paragraphs, and the entire discourse;
2. Projective questions requiring the reader to associate or integrate data from the text with his own experiences, real or vicarious (i.e. experienced through or watching or reading about other people).

Conclusion

Testing is a very important aspect of teaching. It is also a very useful tool through which an effective teacher can assess the extent of curriculum content coverage, and the reliability of his instructional methods. This shows that there is a connection between teaching and testing and this connection must never be allowed to cut.

Recommendations

Teachers should ensure that they take continuous assessment as a very serious issue. Questions set should be clear, unambiguous and be based on the curriculum content covered by the teacher. All

benefits of continuous assessment should be explored especially during the time of making decisions on educational matters. Teachers should also be encouraged to attend seminars and workshops on Continuous Assessment. Universities and Colleges of Education should also train their students effectively on Continuous Assessment.

References

Akinbode, J. O. (2004). *"Language and Literature methods"*, Nigeria, Abeokuta, Goal educational Publishers.

Akinbode, J. O. (2006). "Repositioning English Language Education in Nigeria", Nigeria, Ikot-Ekpene, *International Journal of Research in Education*, 3 (1), 66-71.

Allen, J. P. B. and Davies, A. (1977). *Testing and Experimental Methods'. The Educating Course in Applied Linguistics*, London, Oxford University press vol. 4.

Belts, V. J. (2001). *'Tests for Schools and Universities'* London University Press Limited.

David, H. (2001). *'Testing English as a Second Language'*: New York, McGraw-Hill.

Davies, A. (1968). *'Language Testing Symposium'*, London. Oxford University Press.

Macintosh, H. G. and Hale, D. E. (1976). *'Assessment and Secondary School Teacher'*, London Routledge and Kegan Paul.

Valette, R. (2002). *'Modern Language Testing'*. A Handbook, New York, Harcourt Brace.

Williams, D. (1990). *'English Language Teaching: An Integrated Approach'*, Nigeria, Ibadan Spectrum Books Ltd. Pp. 196-226.

VN:R_U [1.9.11_1134]