academicJournals

Vol. 8(8), pp. 431-439, 23 April, 2013
DOI: 10.5897/ERR12.219
ISSN 1996-0816 © 2013 Academic Journals
http://www.academicjournals.org/ERR

Educational Research and Reviews

Full Length Research Paper

Measuring teachers' competencies in constructing classroom-based tests in Nigerian secondary schools: Need for a test construction skill inventory

Ngozi N. Agu*, Chika Onyekuba and Aloysius C. Anyichie

Department of Educational Foundations, Nnamdi Azikiwe University, Awka, Nigeria.

Accepted 22 February, 2013

Classroom-based achievement tests have been extensively used in Nigerian secondary schools especially after the introduction of continuous assessment in 1985. These achievement tests have been criticized by over the years for lack of proper psychometric properties of a test. These views bother on teachers' possession or non possession of competencies in test construction skills. This study developed and validated a Test Construction Skill Inventory (TCSI) for assessing the secondary school teachers' competencies in constructing classroom-based tests. Factor analysis was done on the 30-item instrument developed by the researchers. 25 items were found to be factorially valid. The TCSI was also found to be reliable with a coefficient of 0.73 and the secondary school teachers found almost all the 25 items important skills for quality classroom-based test construction. The TCSI was, therefore, recommended as an important measure for determining the secondary school teachers' test construction skill in Anambra State, Nigeria.

Key words: Classroom-based achievement tests, continuous assessment, teachers' test construction skills, Nigerian secondary schools.

INTRODUCTION

The issue of validity and reliability of classroom-based achievement tests in Nigerian secondary schools have engaged the attention of researchers (Baker, 2003; Dosumu, 2002; Alele-Williams, 2002). These researchers observed that most of these classroom-based tests in Nigeria lack validity and reliability because teachers seem to lack test construction skills and thus cannot construct good achievement tests. Most tests used for continuous assessments and end of term examinations in the secondary schools contain ambiguous and mis-leading questions which may be the reason why some of the students fail these tests (Adeola and Fajonyimi, 1999). The implication of this is that most teachers lack competencies in test construction, and may be using poorly constructed tests to measure students' achievements in various school subjects. When students' achievement levels are not properly measured and interpreted, the teachers and school administrators will not be able to provide educational opportunities and support each individual student needs.

Testing provides feedback on which educational decisions are made. These decisions may be the ones that require information about the success of learning programmes or about students who have reached particular levels of skill and knowledge (Izard, 2005). Whatever type of information needed, educational decisions depend upon valid and reliable measures to inform those who make the decisions (Izard, 2005).

*Corresponding author. E-mail: ngozi-uj@yahoo.com.

Achievement tests are relevant for measuring important aspects of a subject, accurately reflect the emphasis placed on important aspects of instruction and also measure appropriate level of student's knowledge in a school subject (Frey, 2007). These tests can be either classroom-based, which are produced by an individual classroom teacher for his/her class (teacher-made tests), or standardized, which are developed by test experts for more global purposes. The most instructionally-relevant achievement tests, however, are classroom-based and if carefully constructed, provide teachers with accurate and useful information about the knowledge retained by their students in particular school subjects (Childs, 1989).

Quality classroom-based assessment means adherence to standard procedures for test construction. Every classroom teacher is expected to possess and apply requisite skills in construction of good items for class assessments. A good test item must be both valid and reliable. A test is valid if it is suitable for the intended purpose. On the other hand, a test is reliable if it measures what it is supposed to measure consistently under all conditions (NTI Manual, 2006). Teachers today, perhaps more than ever before, have a need to be knowledgeable consumers of test information, constructors of assessments and protocols, and even teachers about testing (Rudner and Schafer, 2002).

Most achievement tests used in Nigerian secondary schools for continuous assessments are teacher-constructed tests. Teachers, therefore, need to apply some acceptable degree of test construction skills in order to be able to develop valid and reliable tests that will yield accurate feedback of students' achievement. Test construction skills include the competencies needed for developing quality tests based on stipulated principles of test construction (Ali, 1999). These competencies are outlined by Ujah (2001) as: objectivity, communicative, item validation skills and skills for applying appropriate strategies for ascertaining the reliability of test instruments. To Silker (2003), skill in test construction enables a teacher to construct tests with precision, appropriateness of language-use, objectivity and good grading scales. Teachers need not be experts in educational measurement and evaluation to construct valid and reliable tests, but there are basic test construction skills which every teacher ought to possess to construct quality tests. These skills help teachers to: structure items to elicit clear and concise answers from students; construct tests that will be appropriate for learners of different ages, abilities, and gender; set tests so that students finish within time and do not grow scared of tests (Ali, 1993).

Lack of test construction skills by teachers might result in false assessment of students' achievements. Some researchers (Esomonu, 2002; Paulson, 2003) see this incompetency in test construction by teachers as a major cause of malpractice in school examinations by both teachers and students in Nigerian secondary schools. A number of studies have suggested that faulty test items affect students' comprehension and ability to provide accurate answers to the items, the inference drawn about what a student knows and understands may be compromised (Koksal, 2004; Leighton and Gokiert, 2005). Jatto (1996) observed that classroom teachers generally write poor items and that a typical classroom teacher in a secondary school cannot construct good multiple choice test items.

Competencies for test construction

The quality of a test given by a teacher is closely linked with its ability to provide the kind of information needed regarding students' performances. A well written test allows the teacher to accurately and consistently measure students' mastery of specific contents taught in class. Results of such tests allow teachers to measure to some degree, how effective their instruction has been. Conversely, poorly designed test items can lead to inaccurate measurements of learning and provide false information regarding student performance as well as instructional effectiveness (Education Up Close, 2005). Any characteristics of a test item which distracts the examinee from the major point or focus reduces the effectiveness of that item (Frey, 2007). Any item answered correctly or incorrectly because of extraneous factors in the item results in misleading feedback to both examinee and examiner (Frey, 2007). Koksal (2004) outlined factors that are inherent in poorly designed tests which if well handled will lead to quality classroom-based tests. These are:

(i) Non specification of the target audience, what skill or area of ability the test intended to measure, how much time allocated for each test item, and what points the test-takers would get for each correct response.
(ii) Separate sections not clearly stated.
(iii) Test items having more than one possible answer because they were not conceptualized.
(iv) Not stating time allocated for each task on the papers. Only the total time available to perform all the tasks was given.
(v) Non consideration of level of students in test construction.
(vi) Unclear instructions.
(vii) Tasks students are expected to perform not being in concert with the tasks they are told to do during the classroom instructions.
(viii) Lack of representativeness of items concerning what the teacher intends to test.
(ix) Collage-production of some test items.

Let us take a look at the example of a poorly constructed test item given by a classroom teacher to her pupils as follows:

Teacher 'A' taught her pupils in second grade a lesson on 'magnet' and asked them, on the following day, to write a six letter worded object which picks things. She expected almost the whole class to return the word –'magnet' as their response. To her chagrin, the answer given by more than 50% of the class was 'mother' (Daily Bread, 2011).

The teacher must have wondered what actually went wrong. Was it that she did not teach well or that the pupils did not understand what was taught? The problem stems from neither the teaching nor the pupils' learning but from the way the test item was constructed. The task the pupils were required to perform in the test was not perfectly clear thus giving room for more than one possible correct response.

To construct good test items, therefore, classroom teachers should possess the competencies in the following areas as outlined by Chidolue (1999, p. 36):

determining the purpose of each testing exercise; stating specific, measurable educational objectives; making good content outline; preparing test plan which will guide item construction; choosing appropriate test item formats; constructing clear, precise and unambiguous items; constructing items that focus the attention of a group of students, often with widely varying background experiences, on a single idea; constructing items with appropriate difficulty and discriminative indices; developing marking guide suited for the test; performing item analysis of their test items; developing tests that are economical in time and money; giving clear directions on how the test should be administered and taken; reviewing the test in order to correct any errors made during item construction.

Adding to the outline of competencies, Koksal (2004) noted that in test construction, it is essential that the teacher asks the following questions: Is the task perfectly clear? Is there more than one possible correct answer? Can test-takers arrive at the correct response without having the skill supposedly being tested? Do test-takers have enough time to perform the task(s)?

Teacher-made tests and Nigerian secondary school students' academic achievements

Students who perform well at various levels of teachers' classroom-based tests are expected to equally perform well in the standardized tests like West African Examination Council (WAEC) and National Examination Council (NECO) taken at the end of secondary education in Nigeria. However, this appears not to be so. Researchers (Ejiogu, 1988; Itedjere, 1992) observed that the performance of secondary school students in external examinations in Nigeria have steadily been on the decline. Statistics on students' performance in 2009 revealed a 75% failure rate in WAEC and similar proportion in NECO. What is worrisome about these results is that students, often times, perform better in the various classroom-based achievement tests constructed and administered by their teachers. It would seem, therefore, that some of these teacher-made achievement tests are invalid and unreliable and thus, fail to provide accurate assessment of students' knowledge and understanding of the various subject areas (Esomonu, 2002; Ujah, 2001). Some teachers construct poor tests while some continue to duplicate test items because they seem to lack test construction skills (Onyechere, 2000).

Test construction has been found to be a major source of anxiety among many teachers in Nigerian schools, especially, less experienced ones (Ebinye, 2001). This anxiety stems majorly from lack of test construction skill by these teachers. This is surprising as the least qualified teacher in Nigerian secondary schools today have the Nigerian Certificate in Education (NCE) qualification. This NCE qualification is obtained after three years of study in a College of Education where teachers are trained in different education courses, including educational testing and assessment. A number of possible reasons could be deduced for this deficiency. It may be either that the teachers' pre-service training did not prepare them adequately for test construction due to little emphasis on assessment during their professional development (Stiggins, 2000) or that most of the teachers failed to acquire test construction skills needed for quality test item generations while in training. This implies the need to assess the test construction skills of teachers in the secondary schools. Surprisingly, there seems to be paucity of instruments developed for such assessment in Nigeria (Ujah, 2001). This study is set, therefore, to develop and validate an instrument for assessing the test construction skills of secondary school teachers in Anambra state, Nigeria. Specifically, the study was set to do the following:

1. Develop a Test Construction Skill Inventory (TCSI) for measuring test construction competencies of secondary school teachers in Nigeria.
2. Determine the factorial validity of the items in the TCSI.
3. Determine the internal consistency reliability of the items in the TCSI.
4. Determine the perceived importance and appropriateness of the items.
5. Determine the influence of factors such as gender and years of teaching experience on the TCSI.

Research questions that guided the study

1. Are the items in the Test Construction Skill Inventory (TCSI) factorially valid?
2. Are the items in the TCSI internally consistent?
3. Do the teachers agree with the items in the TCSI as

important?

Hypotheses

1. The difference in the mean rating of male and female teachers on the TCSI is not significant.
2. The mean ratings of experienced teachers on the TCSI do not differ significantly from that of less experienced teachers.

METHOD

The study was an instrumentation research geared purely towards the development and validation of an instrument for assessing test construction skills of secondary school teachers in Nigeria. 543 secondary school teachers in Onitsha education zone, Anambra state, Nigeria were used for the study. Proportionate stratified random sampling was used to ensure that teachers from different local government areas that make up Onitsha education zone were selected in the same proportion as they existed in the population. To do this, 40% of teachers in each of the three local government areas in Onitsha education zone were sampled (292 from Onitsha North LGA; 107 from Onitsha South LGA; and 268 from Ogbaru LGA). The sample comprised of 120 male and 423 female teachers. Teachers sampled for the study with less than ten years teaching experience were taken as 'less experienced' (n=200) while those that have been in the teaching profession from ten years upwards were taken as 'experienced teachers' (n=323). Cronbach alpha was used to analyze data for research questions 1 and 2 while mean and standard deviation statistical procedures were used to analyze data for research question 3. Hypotheses 1 and 2 were tested with z-test of significance of difference between two population means. The z-test was used because of the large sample.

RESULTS

The instrument – Test Construction Skill Inventory (TCSI) – was developed from review of literature and survey instruments. The instrument is a 30-item inventory structured on a 4-point scale with the following options: strongly agree (SA), agree (A), disagree (D), and strongly disagree (SD). The TCSI is presented in Table 1.

After the preliminary validation of the instrument, it was administered on the research sample. The data collected for the study was subjected to factor analysis. The results in Table 2 revealed that out of the five factors extracted during the analysis, only four factors (factors 1, 2, 4, and 5) have items sufficiently loaded on them, that is, up to four items. Only one item loaded on factor 3. Items 3, 22, 24 and 28 were not loaded on any factor and item 30 was loaded on two factors.

Factor 3 was dropped as the number of items for accepting a factor is four and the minimum loading of any item should be 0.35 (Baker, 2003). Factors 1, 2, 4, and 5 were accepted as factorially valid. Items 3, 22, 24 and 28 that were not loaded on any factor and item 30 which was loaded on two factors which were taken to be factorially impure were dropped. The four factors and the associated 25 items considered to be factorially valid are shown in Table 3 with their factor loadings. Factor 1 is made up of five items, factor 2, eight items, factor 4 has five items and factor 5 has seven items. All the items had factor loadings and communality above 0.35. This indicates that the twenty five items are factorially valid to be used in measuring the test construction skills of the teachers.

The surviving 25 items were subjected to a test of internal consistency using the Cronbach alpha. A summary of this analysis is shown in Table 4. The Cronbach coefficient alpha of the test construction skill inventory is 0.73. This coefficient is taken to be high and adequate for the instrument.

Analysis in Table 5 shows that the mean ratings of items 21 and 22 (X=1.64 and 1.48 respectively) were below the mean of weights (2.50) which is taken as the criterion for taking decision on any item. All the other items were rated above 2.50. This indicates that the teachers agreed that 23 of the items in the TCSI are important in constructing classroom-based tests. Two of the items (21 & 22) were not seen as important.

Data analysis presented in Table 6 shows that there was no significant difference between male and female teachers' mean ratings on the TCSI with z- value of 0.78 ($p \leq .05$). There was, however, a significant difference in the mean ratings of teachers of different teaching experiences on the TCSI. This is evident in the test of the hypothesis which yielded a z-value of 7.22 ($p \leq .05$).

DISCUSSION

A 30-item instrument for measuring teachers' test construction skill was constructed and the items were subjected to construct validation using the varimax-rotated factor analysis. Results in Table 2 show that the principal component extracted only five constructs, out of which only four factors had items substantially loaded on them. This indicates that the four factors (language use, content coverage, item organization and test guidance) are valid for measurement of teachers' test construction skill. Silker (2003) also found that the constructs of adequate language mechanics, organization of items and content coverage are indispensable constructs in test construction. It was also found that out of 30 items that were factor analyzed, only 25 items met the requirements for accepting an item as valid. The 25 items obtained up to 0.35 factor loading each which, to Baker (2003), is acceptable as valid. The other five items which were either loaded on more than one factor or were not loaded on any factor at all were dropped.

The 25 items that survived factor analysis were subjected to a test of internal consistency. The reliability coefficient showed an alpha of 0.73 which indicates that the instrument has high inter-item consistency and is

Table 1. Test construction skill inventory.

Items	SA	A	D	SD
A teacher takes the following steps in constructing tests for his/her class				
1. Outline the content covered for the term before setting test from them.				
2. Prepare a test blueprint as a guide in the test construction.				
3. Consult previous tests and adapt questions from them.				
4. Consult standard text books in the subject for guide.				
5. Organize test items in a logical manner.				
6. Give clear instructions to guide the test takers.				
7. Write test so that both high and low achievers can understand.				
8. Subject test items to item analysis.				
9. Keep a resource bank of questions that can be referred to when setting tests.				
10. Set tests with due regard to the time available for testing.				
11. Add enough test items to cover all the requisite levels of cognitive domain.				
12. Ascribe scores for each test item.				
13. Ensure that the items are measuring the determined objectives.				
14. Set essay items that elicit creative and imaginative answers from the students.				
15. Prepare a marking guide while constructing the test.				
16. Consider the age of learners during item writing.				
17. Avoid gender stereotypes in the test items.				
18. Add sufficient items to cover the appropriate instructional units.				
19. Submit items for vetting to the Head of Department or the principal.				
20. Submit tests meant for promotional examinations for expert editing on time.				
21. Avoid the use of clues in multiple choice questions.				
22. Number diagrams in tests clearly.				
23. Avoid the use of interlocking items.				
24. Avoid items that measure opinions.				
25. Review draft of the test at least two times in two days before administering.				
26. Limit essay tests to high level objectives.				
27. Avoid overlapping alternatives in writing objective tests.				
28. Use appropriate numbering and lettering formats in writing tests.				
29. Avoid too long questions or phrases in item writing.				
30. Set test items that elicits information on one thing at a time.				

therefore taken to be satisfactorily dependable. The results presented in table 5 on teachers' rating of the importance of the items on the TCSI, showed that only two items (21 and 22) scored below the acceptable mean of 2.50. The rest of the items had mean scores ranging from 2.56 to 3.86. This indicates that, in the teachers' opinions, all the statements in items 1 to 20, 23 to 25 are important in constructing tests. Baker (2003) found that teachers are aware of a variety of test construction skills. Since the teachers agreed with most items in the instrument, the instrument can be reliably used to assess which of the skills the teachers' apply when constructing test.

On the influence of gender, it was found that there was no significant difference in the mean ratings of male and female teachers ($p \geq .05$). This confirms that the TCSI is stable across gender and could be appropriately used to assess the test construction skills of both male and female teachers. Silker (2003) observed that all teachers irrespective of gender should develop some valid and reliable tests. It was also observed that there was a significant difference in the mean ratings of more experienced and less experienced teachers. This difference observed is an indication that the TCSI is sensitive to years of experience. Dosumu (2002) observed that the more experienced a teacher is, the more he begins to understand and appreciate some important test construction skills. The implication of this is that the TCSI could be administered on the teachers bearing in mind their years of experience as an independent variable.

Table 2. Factorial validation of items.

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
1	.47773	-.09001	.05883	.14730	-.02452
2	.38762	-.09324	.06184	-.25972	-.16391
3	.00314	-.18502	.27143	-.10267	-.08328
4	19995	-.33926	.24578	.12520	.47528
5	.58104	-.13848	.18947	.12849	-.11488
6	.37110	-.06427	.14099	.06436	.15248
7	.27896	.36745	.12327	-.17773	-.10949
8	.46956	-.04282	.16879	-.15513	-.03021
9	.60455	.12051	.04938	-.03648	-.11009
10	.35377	-.18209	-.21079	.28420	.21713
11	.50169	-.12204	-.5829	.30293	.02857
12	.26800	-.4967	.00844	.46144	.07765
13	-.00261	.12610	.13921	.10010	.57458
14	-.11179	.49349	-.34645	-.45270	.01533
15	-.02152	.51687	-.01027	-.04650	-.34733
16	.00140	-.15108	.07014	.48839	-.26692
17	.30346	-.20066	.21645	.36388	.04797
18	.16276	-.15878	-.12120	.56300	.11527
19	.07051	.41684	-.19324	.15768	.13311
20	.10164	-.20337	.06309	.29617	.38796
21	.07492	.05600	-.06268	.16364	.57911
22	.25079	.087808	.027080	.16153	.23118
23	.05677	.00409	.17344	-.62172	.20785
24	.30262	.12341	.1176	.24118	.31007
25	.17665	.17344	.18253	.13990	.481570
26	.63059	.50299	-.14961	.11900	.13854
27	.18335	.40238	.026481	.18311	.036722
28	.07990	.25341	.22801	.000138	.05491
29	.17510	.14199	.62521	.31442	.21531
30	.79901	.04531	.12886	.03335	.64101

Table 3. Item loadings for the surviving 25 items on the valid factors.

	Factor	Item	Item loading	Communality
1	Language use	6	.37110	.64970
		7	.36745	.58954
		17	.36388	.66308
		12	.46144	.54308
		29	.62521	.67238
2	Content coverage	1	.47773	.59853
		8	.46956	.58243
		11	.50169	.59927
		13	.57458	.72273
		14	.49349	.58189
		15	.51687	.63025
		18	.56300	.51488
		26	.63259	.62068

Table 3. Contd.

4	Item organization	5	.58104	.58452
		10	.35397	.48132
		21	.57911	.76605
		23	.62122	.60992
		27	.40238	.60996
5	Test guidance	2	.38762	.49531
		4	.47528	.65003
		9	.60455	.60673
		16	.48839	.62731
		19	.41684	.52384
		20	.38796	,53921
		25	.49157	.61670

Table 4. Summary of Cronbach alpha measure of internal consistency of the TCSI.

Statistics	Mean	Variance Minimum	Standard Dev Maximum	Range	Max/Min.	Variance
Item means	3.091	1.455	4.007	2.552	2.754	.5201
Item varian.	2.641	2.160	13.4110	13.195	62.093	14.992
Inter-item Covariance	.5082	.3593	3.8260	4.1853	10.6482	.0653
Inter-Correlation	.680	.2071	.4403	.6474	2.1262	.0107

Reliability co-efficient for 25 items = .73.

Table 5. Teachers' mean ratings on the importance of items TCSI.

Item A teacher takes the following steps in constructing tests for his/her class	Mean	SD	Rmks
1. Outline the content covered for the term before setting test from them.	3.37	0.69	Agreed
2. Prepare a test blueprint as a guide in the test construction.	3.63	0.88	Agreed
3. Consult standard text books in the subject for guide.	3.45	0.79	Agreed
4. Organize test items in a logical manner.	3.39	0.72	Agreed
5. Give clear instructions to guide the test takers.	3.40	0.72	Agreed
6. Write test so that both high and low achievers can understand.	3.50	0.73	Agreed
7. Subject test items to item analysis.	3.83	0.99	Agreed
8. Keep a resource bank of questions that can be referred to when setting tests.	2.97	1.05	Agreed
9. Set tests with due regard to the time available for testing.	3.76	0.47	Agreed
10. Add enough test items to cover all the requisite levels of cognitive domain.	3.77	0.46	Agreed
11. Ascribe scores for each test item.	3.27	0.75	Agreed
12. Ensure that the items are measuring the determined objectives.	3.39	0.73	Agreed
13. Set essay items that elicit creative and imaginative answers from the students.	2.75	1.06	Agreed
14. Prepare a marking guide while constructing the test.	3.86	1.07	Agreed
15. Consider the age of learners during item writing.	2.94	0.95	Agreed
16. Avoid gender stereotypes in the test items.	3.60	0.96	Agreed
17. Add sufficient items to cover the appropriate instructional units.	3.41	0.98	Agreed
18. Submit items for vetting to the Head of Department or the principal.	3.73	3.00	Agreed
19. Submit tests meant for promotional examinations for expert editing on time.	3.42	0.51	Agreed

Table 5. Contd.

20. Number diagrams in tests clearly.	3.31	0.71	Agreed
21. Avoid the use of interlocking items.	1.64	0.88	Disagree
22. Avoid items that measure opinion.	1.48	0.77	Disagree
23. Limit essay tests to high level objectives.	3.77	0.78	Agreed
24. Avoid overlapping alternatives in writing objective tests.	2.56	0.78	Agreed
25. Avoid too long questions or phrases in item writing.	3.49	0.64	Agreed

Table 6. z-test analysis of teachers' rating of the TCSI based on gender and teaching experience.

	N	Mean	SD	z-cal.	z-crit.	DF	Decision
Teachers' gender							
Male	120	3.83	0.99				
Female	423	3.76	0.47	0.78	1.96	541	NS*
Teachers' teaching exp.							
0-10 years	200	3.83	2.96				
11 years and above	323	3.41	0.98	7.22	1.96	541	S*

Silker (2003) made a similar observation and concluded that years of experience was a significant factor that affects the validity of teacher-made test.

Conclusion

The findings of this study show that a 25-item inventory, which is factorially valid and reliable, has been developed for assessing the test construction skills of teachers. This implies that the Anambra State Education Commission can use this instrument to assess the extent to which teachers possess the skills to develop valid and reliable tests. They may also use the instrument in recruiting and evaluating staff for the test development unit of the commission. The findings of this study also have some implications for in-service training. The instrument can be used to determine the skills of test construction which teachers lack and as such form the basis for designing and approving in-service training on test construction skills. The instrument can also be used to identify teachers with poor test construction skills so that in-service training will be organized for them. Teachers can also use the TCSI for self evaluation and as guide in constructing tests. The TCSI may guide them to develop valid and reliable tests that will give dependable estimate of students' performances.

REFERENCES

Adeola OA, Fajonyomi AA (1999). Adult learners' performance on standardized and non-standardized test as a function of sex and location. Nigerian J. Counsel. Dev. 4(1):32-36.

Alele-Williams G (2002). Measurement and evaluation in mathematics: The way forward. J. Basic Sci. 1(1):1-7.

Ali AA (1999). Basic research skills in education. Enugu: Orient Printing and Publishing.

Baker JO (2003). Testing in modern classrooms. London: George Allen and Urwin Ltd.

Chidolue M (1999). Introduction to tests and measurements. Unpublished mimeograph on measurement and evaluation. Awka, Nnamdi Azikiwe University.

Childs RA (1989). Constructing classroom achievement tests. ERIC Digest. ED 315426.

Daily Bread. RBC Ministries (2011). Our daily bread: For personal and family devotions. Michigan, USA: Discovery House Publishers.

Dosumu CT (2002). Issues in teacher-made tests. Ibadan: Olatunji and Sons Publishers.

Ebinye PO (2001). Problems of testing under the continuous assessment programme. J. Qual. Educ. 4(1):12-19.

Education Up Close (2005). Writing effective tests: A guide for teachers. Retrieved on October 17, 2007 from http://www.glencoe.com/sec/teachingtoday/educationup.phtml/print/40.

Ejiogu AM (1988). Landmark in educational development in Nigeria. Lagos: Joja Educational Research and Publications.

Esomonu NP (2002). Assessment of non-cognitive behaviours in secondary schools for national development. In: Ebenebe RC & Akudolu LR (Eds.). Education for National Development and Integration. Awka: John Best Publishing pp.88-93.

Frey BB (2007). An introduction to quality test construction. Retrieved on October 19, 2007 from http://www.specislconnection.ku.edu/cgi bin/cgiwrap/speccom/mainphp?cat=assessment&section=main&section=qualitytest/main.

Itedjere PO (1992). Obstacles to full professionalism of teaching in Nigeria. In V. F. Peretomode (Ed.). Introduction to the Teaching Profession in Nigeria. Owerri: Totan Publication.

Izard J (2005). Overview of test construction. In K. N. Ross. Quantitative Research Methods in Educational Planning. Paris: UNESCO International Institute for Educational Planning.

Jatto YA (1996). The continuous assessment system in Nigeria: The problems and challenges ahead. In G. A. Badmus & P. I. Odor (Eds.). Challenges of Managing Educational Assessment in Nigeria. Book of Readings on Educational Assessment. Kaduna: Atman Ltd.

Koksal D (2004). Assessing teachers' testing skills in Elt and enhancing

their professional development through distance learning on the net. ISSN 1302-6488. Turkish Onlines J. Distance Educ. 5(1).
Leighton JP, Gokiert RJ (2005). The cognitive effect of test item features. Retrieved on October 17, 2007 from http://www.education.ualberta.ca/educ/psych/crame/files/NCMEfinal LeighandGokiert.pdf.
NTI (National Teachers' Institute) (2006). Manual for re-training of primary school teachers: School-based assessment. Kaduna: National Teachers Institute.
Onyechere I (2000). New face of examination malpractice among Nigerian youths. The Guardian Newspaper July 16.
Paulson TU (2003). The study of teaching. New York: Holt, Rinehart and Winston Inc.
Rudner LM, Schafer WD (Eds.). (2002). What teachers need to know about assessment. Washington, DC: National Education Association.
Silker RT (2003). Teachers and tests. London: Basil Blackwell.
Ujah EU (2001). Development and validation of an introductory technology achievement test. Unpublished M.Ed. thesis, University of Nigeria, Nsukka.