

ASSEREN JOURNAL OF EDUCATION (AJOE) 2017

EDITORIAL BOARD

Editor-in-Chief

Prof. Christiana A. Ugodulunwa
Department of Educational Foundations,
University of Jos

Editors

Dr. S. I. Dadughun
University of Jos

Dr. A. Y. Mustapha
University of Jos

Dr. S. G. Wakjissa
University of Jos

Dr G. C. Imo
University of Jos

Consulting Editors

Prof. P. O. Awotunde
University of Jos

Prof. Alice E. Asim
University of Calabar

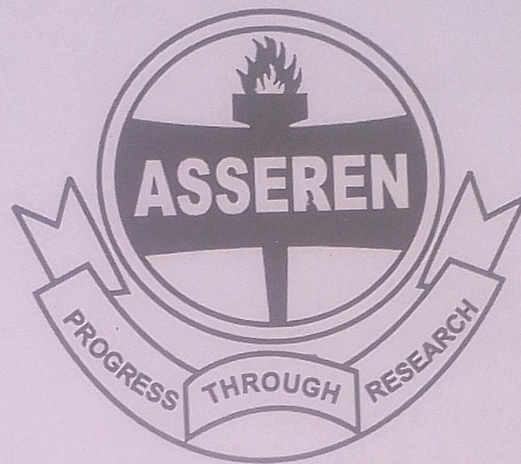
Prof. A. Hamman-Tukur
University of Maiduguri

Prof. E. D. Ozoji
University of Jos

Prof. C. M. Anikweze
Nasarawa State University, Keffi

Prof. T.O. Afemikhe
University of Benin, Benin- City

ASSEREN JOURNAL OF EDUCATION (AJOE)



A JOURNAL OF THE ASSOCIATION OF
EDUCATIONAL RESEARCHERS AND EVALUATORS
OF NIGERIA (ASSEREN)

ISSN 2504 - 9968

Vol. 2, No. 1, July 2017

LIST OF CONTRIBUTORS

- | | |
|----------------------|--|
| Adams O.U. Onuka | Institute of Education, University of Ibadan, Ibadan, Nigeria |
| Justina N. Dimson | Department of Life Science Education, Faculty of Education Imo State University, Owerri |
| Idowu O. Eluwa | Department of Science Education, College of Education, Michael Okpara University of Agriculture, Umudike |
| Janet A. Kebbi | Niger Delta University, Wilberforce Island, Bayelsa State, Nigeria |
| Patrick U. Osadebe | Department of Guidance and Counselling, Delta State University Abraka |
| Margaret E. N. Orubu | Delta State University Secondary School, Abraka |
| Petrolina N. Ukah | Department of Psychology/Guidance and Counseling Alvan Ikoku Federal College of Education, Owerri |
| Angela O. Dara | Dept. of Psychology/Guidance and Counseling AlvanIkoku Federal College of Education, Owerri |
| Opeyemi Ogunwale | Kwara State Teaching Service Commission, Ilorin, Kwara State |
| Dorcas S. Daramola | Department of Social Sciences Education, Faculty of Education, University of Ilorin |
| Adekunle T. Olutola | Department of Educational Foundations, Faculty of Education, Federal University Dutsin-Ma, Katsina State |
| Henry O. Owolabi | Department of Adult and Primary Education, Faculty of Education, University of Ilorin, Nigeria |
| Cobbinah Andrews | Department of Educational Foundations, University of Cape Coast, Ghana |
| Chidera E. Amaechi | Department of Science Education, Faculty of Education, Michael Okpara University of Agriculture, Umudike |
| Benedict I. Onuoha | Department of Science Education, Faculty of Education, Michael Okpara University of Agriculture, Umudike |
| A. O. Ovute | Department of Science Education, Faculty of Education, Michael |

	Okpara University of Agriculture, Umudike
Celestine C. Jiwueze	Department of Physical Science Education, Faculty of Education, Imo State University, Owerri
Catherine N. Ugaseh	Federal University of Agriculture, Makurdi
M. O. Aduloju	Federal University of Agriculture, Makurdi
S. O. Emaikwu	Federal University of Agriculture, Makurdi
P.M.C. Ogomaka	Department of Physical Science, Imo State University, Owerri
T. A. Nwanebo	Department of Physical Science, Imo State University, Owerri
M. C. Nosike	Department of Physical Science, Imo State University, Owerri
Eugene C. Unamba	Department of Primary Education Studies, Alvan Ikoku Federal College of Education
C N. Onuekwusi	Department of Educational Psychology, Alvan Ikoku Federal College of Education
Joy N. Onyeka	Department of Educational Psychology, Alvan Ikoku Federal College of Education
M. U. Orheruata	Department of Educational Evaluation and Counselling Psychology University of Benin, Benin City.
O. K. Omorogiuwa	Department of Educational Evaluation and Counselling Psychology University of Benin, Benin City.
A. U. Osunde	Department of Educational Evaluation and Counselling Psychology University of Benin, Benin City.
Ifeoma R. Ezechukwu	Department of Physical Science Education (Measurement and Evaluation), Faculty of Education, Imo State University, Owerri
Casmir N. Ebuoh	Department of Science and Computed Education, Enugu State University of Science and Technology, Enugu State, Nigeria
Gabriel Okafor	Department Of Science Education, Ebonyi State University, Abakiliki.

- Emma C. Ekwonye
Department of Physical Science Education Imo State
University, Owerri
- Samuel A. Okoroafor
Department of Physical Science Education Imo State
University, Owerri
- A.C. Akukwe
Department of Computer Education, Alvan Ikoku
Federal College of Education, Owerri
- Idris U. S. Bauchi
Department of Social Sciences, Faculty of Education,
University of Ilorin, Kwara State
- Ezeamagu M. Ugochukwu
Department of Social Sciences, Faculty of Education,
University of Ilorin, Kwara State
- Aderinto M. Olalere
Department of Social Sciences, Faculty of Education,
University of Ilorin, Kwara State
- Haliru Shehu
Department of Curriculum Studies, Sokoto State
University, Sokoto
- Terhemba G. Atsua
International Center for Educational Evaluation,
Institute of Education, University of Ibadan
- Umaru Abdullahi
International Center for Educational Evaluation,
Institute of Education, University of Ibadan
- Lawal O. Omonike
Department of Education, Yobe State University,
Damaturu, Yobe State

**Procedures for Improving Content Validity of Teacher-Made Achievement Test:
A Way Forward to Ensuring Quality Research and Evaluation**

**Idris U. S. Bauchi¹, Ezeamagu M. Ugochukwu², Aderinto M. Olalere³
Haliru Shehu⁴**

^{1,2,3}*Department of Social Sciences, Faculty of Education,
University of Ilorin, Kwara State*

⁴*Department of Curriculum Studies, Sokoto State
University, Sokoto*

Abstract

The significance of validity as a psychometric property of test in research and evaluation can never be overemphasized. Invalid test instruments rather lead to invalid interpretations and conclusions in research and evaluation. Based on this predicament, this paper sampled simple procedures to follow by teachers and researchers in constructing test items that will represent objectives of instruction and have substantial coverage of contents of subjects/courses. It provided detailed steps to be followed in constructing table of specification as a vital component for test construction. The paper also discussed some guidelines on writing objective and essay items. It also showcased some disadvantages of haphazard selection and writing test items for assessment and data generation. It recommended among others that teachers and researchers should strictly follow the procedures explained in setting their test items for both formative and summative assessments in schools.

Key words: table of specifications, test blue print, content validity, teacher-made test

Introduction

Over the years, great number of researches have been on board, and numerous tests have been constructed for data generation. Much attention has been given to an outstanding psychometric property of these tests which is "test validity". Without any iota of doubt, invalid research and test instruments rather yield invalid results, interpretation and conclusions. This consequently affects the quality of research and assessment in education.

Validity is defined as what test measures and how well it does so (Anastasi & Urbina, 1997). It is also seen as the degree to which evidence and theory support the interpretation of test scores (AERA, APA & NDCE, 1999). Berlin and Westein (2008) viewed validity as the extent to which a test measures what it is purported to measure. In a more elaborate and comprehensive way, it is defined as an overall evaluative judgment of the degree to which empirical evidences and theoretical rationales support the adequacy of and appropriateness of interpretations based on test scores and other modes of assessment (Messic, 1999).

Content validity is defined as the type of validity that addresses how well items developed provide adequate sample of all items that might measure the construct of interest (Kinberlin & Westein, 2008). In line with this definition is the view of Idris and

Alfa (2012) who defined content validity as the extent to which a test measures an intended content area. This simply indicates that content validity is all about how topics and sub-topics are represented in a test instrument for the purpose of formative or summative evaluation. It is expected that learners should have mastery to certain level of all the contents taught during instruction within which test items are selected in order to assess the level of the achievement of instructional objectives. By the time instructional objectives are achieved on daily basis, there is every tendency that the overall educational goals would be achieved with time.

As a type of validity, construct validity is seen as the degree to which a test measures the intended construct (Gay, Mills & Airasian, 2009). Kinberlin and Westein (2008) defined construct validity as judgement based on accumulated evidences from several studies using specific measuring instrument. Content validity if established helps in the determination of construct validity of any given test instrument.

Establishing Content Validity of a Test

As stated previously, content validity is concerned with content coverage of a test. The most widely method used in obtaining content valid test is through the construction of Table of Specification (TOS) otherwise known as Test Blueprint (TB). It is a two way chart which maps instructional objectives with course or subject contents (Kolawole, 2010). TOS is also seen as a table that assists teachers and test developers to align objectives, instruction and assessment (Alade & Omoruyi, 2014). Table of specification as a guide to establishing content validity have been in existence for a very long time in education. TOS have several advantages in education. It helps detect which topics are being stressed in a test, helps in preparing test that reflects what is learnt, ensures no omission of important topics in a test, helps in producing valid and robust test items and ensures that only instructional aims and objectives are assessed (Akem & Agbe, 2003). It is important to note that judgement about students' understanding and learning can be meaningful only if tests are aligned to the contents taught (Alade & Omoruyi, 2014). The question remains that are our teachers and researchers make use of it in developing test items? Are they conversant with the procedures involved in the construction of the table? By and large, validity, as significant as it is can never be compromised. Neglecting the steps to achieve it can serve as nuisance to quality instrumentation, assessment and interpretation in research. It is therefore paramount for both teachers and test developers to be conversant with the initial stages involved in bursting the validity of test instruments.

Construction of Table of Specifications Using Bloom's Taxonomy

Bloom's taxonomy is the most widely used taxonomy of cognitive objectives. It has six levels which include knowledge (know.), comprehension (comp.), application (appli.), analysis (analy.), syntheses (synth.) and evaluation (eval.). This taxonomy is widely accepted because it exhibits substantial coverage of cognitive domain which is the focus of most of the achievement tests. The second domain is the psychomotor domain that is assessed by the performance test instruments. The third domain is the affective domain usually assessed by personality inventories such as perception inventory, self-esteem scale, motivation scale and many more.

Now, let's proceed to the main steps to be followed in the construction of table of specifications using blooms taxonomy:

1. Write the purpose or aim of the test, for instance, why testing? What to be tested? Is it students' aptitude, achievement, performance.... etc?
2. Outline the learning objectives, the topics treated and the periods taken per topics, in other words, the scheme of work and the lesson plans should be revisited.
3. Decide on which format of the test will suit to measure such objectives since evaluation depends on the objectives of learning. Is it objective, essay or short answer format that will fit the purpose and subjects to be tested?
4. Write the topics and the periods or weeks taken in the classroom instruction to get the total number of periods, weeks or days spent in teaching and learning activities
5. Get the proportion of the topics by dividing the topic periods by the total number of periods taken.
6. Decide the total number of questions needed and multiply by the percentage of each topic to get the number of question(s) per topic
7. Depending on the domain tested, for example cognitive domain, you can use Bloom's taxonomy which contains six levels of cognitive objectives which are knowledge, comprehension, application, analysis, synthesis and evaluation; to develop your table.
8. Write the topics across the six levels of objectives and distribute the number of items based on the previous calculations to each level until they tally with the total number of questions predetermined.

Following these steps, below is the example of how the test blueprint can be created :-

Step 1. The purpose of the test was to measure the academic achievement of students in EDU321 (Test and measurement)

Step 2. All the learning objectives are within the cognitive domain, and seven topics were taught in 14 lesson periods.

Step 3. The format of the test was Multiple choice format to enable wide coverage of the topics.

Step 4. The topics covered are as follows:-

Table 1: Scheme of Work for Fourteen Periods

TOPIC	PERIOD
i. Meaning of test and measurement	1
ii. Types of test	2
iii. Classification of test	3

Table 1 consists of treated topics and their corresponding periods. Most of the topics were taught in a period (45 minutes) with only one topic having three periods (135 minutes)

Step 1. One can get the percentage of items to be derived from each topic as follows:-

Table 2: Percentage of Each Topic Out of the Treated Topics

1. Meaning of test and measurement	$= 1/14 \times 100 = 7.14\%$
2. Types of test	$= 2/14 \times 100 = 14.29\%$
3. Classification of test	$= 3/14 \times 100 = 21.42\%$
4. Test validity	$= 2/14 \times 100 = 14.29\%$
5. Test reliability	$= 2/14 \times 100 = 14.29\%$
6. Scoring systems	$= 2/14 \times 100 = 14.29\%$
7. Measures of central tendency	$= 2/14 \times 100 = 14.29\%$
Total	= 100 %.

Table 2 contains the proportion of each topic compared to all the topics treated. It is extracted by dividing the period(s) spent teaching the topic by the total number of the topics and multiplying it by 100. At this juncture, number of questions can be arbitrarily decided by the school authority or management based on the level of the learners capacity. For example, Lets assume 60 questions was agreed upon to be set for the test by a school authority, then actual number of items per topic can be obtained by multiplying the proportion by 60 as follows in the table below :-

Table 3: Extraction of Actual Number of Items from Percentages

1. Meaning of test and measurement	$= 7.14\% \times 60 = 4.28$
2. Types of test	$= 14.28\% \times 60 = 8.57$
3. Classification of test	$= 21.42\% \times 60 = 12.85$
4. Test validity	$= 14.29\% \times 60 = 8.57$
5. Test reliability	$= 14.29\% \times 60 = 8.57$
6. Scoring systems	$= 14.29\% \times 60 = 8.57$
7. Measures of central tendency	$= 14.29\% \times 60 = 8.57$
Total	= 59.98 app. 60 items

Table 3 shows how the desired number of questions are obtained objectively from the percentages. In this example, 60 questions were filtered from the topics accordingly by dividing the percentage of the topic by 100 and multiplying it by 60 (the target number of questions). It also indicates that one can approximate the decimal points as to get whole numbers where necessary. And, in case of incomplete number of questions due to decimal points or fractions, one can add an item to a topic that has significant decimal point as to make them complete as designed. By so doing, the objectivity of the exercise is not compromised.

Table 4: Test Blue Print

CONTENTS	KNOW 20%	COMP 27%	APPLI. 30%	ANALY. 16%	SYNTH. 5%	EVA 2%	TOTAL
1. Meaning of test and measurement	1	2	1	-	-	-	4
2. Types of test	2	1	2	2	1	-	8
3. Classification of test	2	4	4	1	-	1	12
4. Test validity	2	2	3	1	1	-	9
5. Test reliability	2	2	3	2	-	-	9
6. Scoring systems	2	2	2	2	1	-	9
7. Measures of central tendency	1	3	3	2	-	-	9
Total	12	16	18	10	3	1	60

Table 4 depicts objective distribution of questions across the Blooms learning objectives as derived from the calculated percentages maintaining the actual number of questions (60) set. It will be easier for one to start writing items by directly sampling the allotted number of questions or items of each topic. Using TOS at the stage of writing questions would enable teachers and other test developers guarantee that the resulting test instrument has proper balance of topic coverage and desired range of cognitive skills (Alade & Omoruyi, 2014). In addition, following the guidelines for test construction as outlined by Kolawale (2010) and Alade & Omoruyi (2014) will go along way in helping test developer construct good and valid test items.

Considerations Necessary for Writing Test Items

For a teacher to write outstanding test items capable of enhancing effective assessment of the learners, here are some considerations for both essay and objective test to be kept abreast as outlined by Kolawale (2010), Alade and Omoruyi (2014) :

A. Considerations For Writing Good Essay Questions

The following are some of the considerations while constructing essay questions:-

1. There is need to give enough time and thought in the preparation until one is satisfied that ;
 - i. all items are measuring what they are set for, for instance, items set to measure knowledge on types of pollution should not be measuring effects of pollution.
 - ii. there is simple wording of items i.e. simple and clear to the understanding of the learners.

- iii. all items are answerable by the students regardless of their achievement levels, that is to say, the items should be moderately difficult which can be determined through item analysis.
2. Test questions should be written to elicit the type of behavior to be measured.
3. Essay format items should delimit the area covered by the question and direct the student on precise pattern of response expected.
4. Indicate the marks or points value for each question and also the time limit. This will help test takers to judiciously utilize the time given for a test
5. Do not write so many essay questions and mandate test takers to answer all within a limited time, rather ask few and rich questions to curtail wrong responses from the test takers.

B. Considerations for writing Objective Questions

The following are some of the issues to be considered while writing any type of objective test:

1. Develop a table of specification for the test. This is vital as it involves matching contents with the instructional objectives to improve the content validity of such test..
2. Test for important facts or knowledge, i.e. the test should assess one central idea
3. Consider students' age, ability levels and purpose of the test in the course of construction
4. Check for the difficulty, discrimination and distractor indices of the items which can be done through item analysis after construction of items
5. Clearly write the items and avoid poor arrangement and ambiguous wordings, it must be written with right font size, font type, colour, and free of grammatical errors
6. Reading difficulty and vocabulary level should be simple appropriately.
7. Ensure that there is only one correct or best answer in a question and should appear at different positions. For instance, do not maintain A for the key answer through out the test, rather use the A,B,C,D,E interchangeably.
9. Minimize writing negative questions as that usually confuses test takers
10. Never give hint(s) to the key answer, for example do not bold, italicize, increase font size of the key answer to avoid clues to the right answer. For example:
 - grammatical clues like a, an, the, are, is, etc
 - length, case, boldness of key answers must be considered
 - mentioning key answer in other items of the test with the intention of asking another question in the test.

Possible Consequences of Haphazard Item Selection for Evaluation

It is not and never a good practice for a teacher or researcher to develop a habit of selecting or writing test items haphazardly without following due procedures. Below are some of the possible consequences to be encountered by teachers and other test developers with the habit of indiscriminate selection and writing of test items :-

To the teachers and researchers:-

- i. It will instill poor evaluation skills among teachers and researchers
- ii. It will Make teachers and researchers violate principles of testing as there is every tendency for selecting items beyond the contents taught in during classroom instruction
- iii. It is a sign of lack of integrity among teachers and researchers particularly if they

have the knowledge of constructing TOS

iv. It will subject teachers and researchers to be having invalid assessment consistently in schools.

To the learners:-

i. There will be partial assessment of what students learned, thereby abandoning important areas of subjects or courses..

ii. Item bias will be high because some items selected may favour some learners that are lucky to have read the selected portion, or may favour males or females

iii. It can indirectly disrupt cordial teacher-student's relationship as students may lose confidence in their teachers when they detected such lapses from them.

To the process:-

i. It can erode objectivity of the assessment

ii. It can produce invalid and unreliable test instruments

iii. It may lead to wrong interpretation and conclusion

iv. It violates the standard of test and measurement

iv. It violates ethics of test and measurement

Conclusion

From the discourse on table of specification so far in this paper, it is clear that it enhances valid instrumentation in research and evaluation. It can also be agreed that the use of TOS remains one of the best preliminary procedures in establishing content validity of any teacher-made achievement. By the time much attention is given to the use of TOS, there would be tremendous improvement in the quality of research and evaluation in various institutions of learning, not only in Nigeria but the world in general.

Recommendations

The following recommendations were made based on the content of this paper:-

1. Teachers and researchers should learn how to construct table of specification and also use it when setting questions for tests and examinations.

2. They should strictly follow the guidelines for writing test items as highlighted in the paper

3. Procedures for setting questions should be monitored by the respective authorities to ensure quality assessment and data generation

4. Seminars should be constantly organized by the institutions of learning to keep exposing teachers and researchers to the best and up-to-date practices in evaluation.

References

Akem, J. A., & Agbe, N. N. (2003). *Rudiments of measurement and evaluation in Education and psychology*. Markurdi: The Return Press.

Alade, O. M., & Omoruyi, I. V. (2014). Table of specification and its relevance in educational development assessment. *European Journal Of Educational Developmental Psychology*, 2(10), 1-17.

American Educational Research Association, A Psychological Association, and National Council of Measurement in Education (1999). *Standards for educational and psychological testing*. Washington DC: AERA

Anastasi, A., & Urbina, S. (1997). *Psychological testing*. Seventh Edition. Upper saddle river: Prentice Hall

Idris, U.S.B., & Alfa, A.M. (2012). *Test and measurement*. Unpublished EDU321 courseware, Federal University of Technology Minna.

Gay, L. R., Mills, G. E., & Airasian, P. (2009). *Educational research: Competencies for analysis and applications*. Singapore: Pearson Education International

Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measuring instruments used in reseach. Retrieved 10th May, 2016 from <http://www.ashpfoundation.org/fundamentalKimberlinarticles>

Kolawole, E.B. (2010). *Principles of test construction and administration*. Lagos: Bolabay Publications.