

**AN ASSESSMENT OF SECONDARY SCHOOL CHEMISTRY TEACHERS' SKILL
OF CONSTRUCTING MULTIPLE CHOICE TEST IN BOSSO AND CHANCHAGA
LOCAL GOVERNMENT AREAS OF NIGER STATE.**

BY

ADEGBOYE, Abdulmalik Olawale

2014/1/51422BE

**A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE AWARD OF THE DEGREE OF BACHELOR OF
TECHNOLOGY (B. TECH) IN CHEMISTRY EDUCATION**

NOVEMBER, 2019.

TABLE OF CONTENTS

Title page	i
Declaration	ii
Certification	iii
Dedication	iv
Acknowledgement	v
Table of contents	vi
List of Tables	x
Abstract	xi

CHAPTER ONE

1.0 INTRODUCTION	1
1.1 Background of Study	1
1.2 Statement of Problem	5
1.3 Aim and Objectives of the Study	5
1.4 Research Question	6
1.5 Significance of Study	6
1.6 Scope and Delimitation of the Study	6
1.7 Basic Assumption of the Study	7
1.8 Operational Definition of key Terms	7

CHAPTER TWO

LITERATURE REVIEW

2.0	Introduction	9
2.1	Brief History of Test	9
2.2	Definitions of Test	10
2.2.1	Types of Test	11
2.2.2	Intelligence Test	11
2.2.3	Aptitude Test	12
2.2.4	Achievement Test	12
2.2.4.1	Essay Test Item	13
2.2.4.2	Objective Test Item	13
2.2.5	Characteristics of a good test	14
2.3	Multiple Choice Test	14
2.3.1	Multiple Choice Test Construction	15
2.4	Item Analysis	20
2.4.1	Item Difficulty Index	20
2.4.2	Item Discrimination Index	21
2.4.3	Distracter Analysis Index	22
2.5	Test Item Validity	23
2.5.1	Types of Test Item Validity	23

2.6	Reliability	25
2.6.1	Method of Test Reliability	25
2.7	History of Science Education in Nigeria	26
2.8	Concept of Chemistry Education	27
2.9	Teacher competencies in test construction	31

CHAPTER THREE

RESEARCH METHODOLOGY

3.0	Introduction	33
3.1	Research Design	33
3.2	Population of the Study	33
3.3	Sample and Sampling Techniques	33
3.4	Research Instrument	34
3.5	Validity of the Instrument	34
3.6	Reliability of the Instrument	34
3.7	Method of data collection	34
3.8	Method of Data Analysis	35

CHAPTER FOUR

4.0 DATA PRESENTATION AND ANALYSIS

4.1	Introduction	36
4.2	Research Question One:	36
4.3	Research Question Two:	39
4.4	Summary of Findings	40

4.5	Discussion of Results	40
-----	-----------------------	----

CHAPTER FIVE

5.0 SUMMARY, CONCLUSION AND RECOMMENDATION

5.1	Introduction	41
-----	--------------	----

5.2	Summary	41
-----	---------	----

5.3	Conclusions	41
-----	-------------	----

5.4	Recommendations	42
-----	-----------------	----

5.5	Suggestions for Further Study	42
-----	-------------------------------	----

REFERENCES		43
-------------------	--	-----------

APPENDICES		48
-------------------	--	-----------

List of Tables

Table 2.1: List of Topics and number of Periods taken.

Table 2.2: Treated Topics Percentage.

Table 2.3: Number of test items extracted from percentage of treated topics.

Table 2.4: Test Blueprint or Table of specification (TOS).

Table 4.1: Shows frequency and percentage of chemistry teachers level in skill of constructing multiple choice test.

Table 4.2: Shows frequency and percentage of chemistry teachers gender in skill of constructing multiple choice test.

Abstract

The study investigated the assessed Secondary School teachers' skill of constructing multiple choice test in Bosso and Chanchaga Local Government of Niger State. The aim of the study was to evaluate the chemistry teachers' skill of constructing multiple test by attempting to find out the level and gender difference of chemistry teachers' skill in constructing multiple choice test. Two research question were formulated to guide this study, the first research question is; what is the level of chemistry teachers' skill of constructing multiple choice test? and the second research question is; is there any gender difference on chemistry teachers' level of skill in multiple choice test construction. The instrument for data collection was the questionnaire for assessing of secondary school chemistry teachers' skill of constructing multiple choice test (SCMCT) consisting 21 items with four point likert scale to indicate perception of the respondents. Cronbach alpa was used to calculated the reliability of the instruments which generate 0.75 and SPSS was used to analyze the items. Based on the outcomes of the analysis, chemistry teachers in Bosso and Chanchaga Local Government of Niger are moderately skilled and there is significant gender difference on chemistry teachers' level of skill in multiple choice test construction. It was recommended that Niger State Ministry of Education and other educational bodies should often organize seminar/workshop for chemistry teachers on the skill of constructing multiple choice test to bring about achievements of educational goals. Also, chemistry teacher should constantly practice the skill of constructing multiple choice test learnt in seminar/workshop to enrich the efficiency of their skill.

CHAPTER ONE

INTRODUCTION

1.1 Background of Study

Educational accountability has a great impact on assessment procedure in schools which rotate across the cognitive, psychomotor and affective domains of students' learning experiences in various subject. Test construction and quality is an essential tool required by any teacher in order to achieve educational goals. The importance of test in school settings is to gather information about the learning progress of the students in which the curricula remains the pathway of educational growth, personal inspiration and institutional excellence. Crooker and Algina (2008) described test to be a standard procedure for obtaining a sample of behavior from a specified domain.

Tamakloe, Amedahe and Atta (1996) cited by Quansah (2019) described a test as a device or procedure for measuring a sample of an individual's behaviour in a specific learned activity or discipline, therefore it should be made clear that testing is a key component in educational assessment, these tests are normally conducted after a period of instructional objective achievement. Test is usually a small scale task carried out to determine a candidate level of performance. Assessment therefore encompasses the utilization of students' empirical data on learning for further improvement of programs and enhancement of students' learning (Allen and Yen, 2002). The National Policy on Education (2004) laid a special weight on the importance of continuous assessment by stating that "Educational Assessment and Evaluation will be liberalized by basing them in whole or in part on continuous assessment of the progress of the individual". This insinuate that assessment should be carried out at every segment

during the teaching and learning process as this will disclose the progress of the students during the procedure.

In a broad term, a test can either be; discrete or integrative. Discrete test is expected to test one item or scheme at a particular time e.g reading-skill test, Integrative test combines various items or skills in a single test based on the aim and objective of the test e.g achievement test. Forms of tests can be identified as; multiple-choice tests, matching tests, true-false tests, short-answer tests, problem test, essay tests. Multiple-choice tests

In view of the role that test play in making educational decisions and evaluation for students, management and research, it is important to point out that school management must make judicious and conscious effort to ameliorate the validity and the reliability of the test in order to get an unbiased information that describe the individual's true behaviour, which the test developer and educational stakeholders seeks to qualitate.

The increase in desire for school effectiveness and improvement at all levels of education in Nigeria, demands that classroom tests should be valid and reliable measure of students' real knowledge and skills, the quality of a test in terms of validity and reliability is affected by a number of factors. Some of these factors may be as a result of teacher's characteristics in terms of experience, level of education, knowledge of test construction, professionalism and gender. Validity is a tool that requires to be fully understood by teachers and test developer in any effort to improve quality of a test, it is also the most important instrument of an assessment.

According to Rudner and Schafer (2002), test validity refers to the degree with which the inferences based on test scores are meaningful, useful and appropriate. This definition implies that validity is expressed in degree from low, moderate to high. Validity requires that an instrument is reliable but an instrument can be reliable without being valid. On the other hand,

a test is reliable if it measures what it is supposed to measure consistently under all conditions (NTI Manual, 2006).

Most educational researchers identify different types of validity which are; content validity, construct validity, criterion-related validity and face validity. However, content validity which indicate the comprehensiveness of the instrument in covering the content area that has been treated during instruction, is considered to be of most importance for classroom teachers.

Content validity requires judgment of experts to determine if the item is representative of knowledge and skills that are supposed to be measured. This involves consistency between curriculum content, test objectives and content of the test. The degree of content validity depends on the test's coverage of necessary objectives and content as well as adequate sampling of important curriculum content. This definition implies that for an instrument to be valid, it must measure what it purports to measure and nothing else for the aim of formative and summative evaluation. On the other hand, predictive validity is the extent to which a student's current performance on a test (Predictor) estimates the student's later performance on a criterion measure. Although face validity is not a type of validity in a technical sense, it is the degree to which an instrument appears to measure what it measures, it is usually confused with content validity by teachers (Ugodulunwa & Wakjissa, 2016).

Test reliability refers to the degree of accuracy with which a test measures what it is designed to measure or how dependably or consistently a test measures a characteristic. A test maybe measuring something other than what is expected to measure, but if it does that consistently such test is not considered to be valid but reliable, this means that a test maybe reliable but not valid. Test developers have the responsibility of reporting the reliability estimates that are relevant for a particular test, these reliability estimates include; test-retest reliability; which is

determined by administering the same instrument to the same set of students under certain condition, split-half reliability; involves single administration of instrument to obtain two set of scores which are then correlated by Spearman-rank order correlation method. other methods include; Cronbach's Alpha method which is applied to instrument that are polynumerously scored and Kuder-Richardson method that is applicable to instrument that measures internal consistency of dichotomously scored items (Ugodulunwa & Wakjissa, et al., 2016).

Considering the importance of validity in classroom testing, teachers need to possess necessary knowledge of validity and how to gather validity-related evidence for their tests and other forms of assessment to enable them draw valid inferences and take relevant decisions based on students' assessment scores. Unfortunately, it has been observed (Dosumu, 2002; Adeola & Fajonyimi, 1999; Imo, 2012; Agu, Onyekuba & Anyichie, 2013) that teachers lack test construction skills. Okpala (2002) stated that few teachers in Nigeria are aware that good teaching is characterized by assessments that motivate and engage students in ways that are consistent with their philosophies of teaching and learning and with theories of development, learning and motivation. Lack of test construction skills by teachers might result in false assessment of students' achievements.

Some researchers (Esomonu, 2002; Paulson, 2003) noted 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. Silker (2003) stated that teachers need not be experts in educational measurement and evaluation to construct valid and reliable tests, but basic test construction skills can be employ to construct quality tests. These skills will help teachers to structure instrument to deduce clear

and concise answers from learners, construct appropriate tests for learners, set tests that works with the time frame and to conquer test anxiety.

1.2 Statement of Problem

One of the major reasons of student's poor performance in Chemistry test is attributed to the teacher's lack of test construction skills. (Esomonu, 2002; Paulson, 2003 et al.) stated that the incompetency in teachers' test construction skills is found to be the cause of exam malpractices in Nigerian senior secondary schools, due to the fact that the students are unable to interpret and comprehend the question tabled before them, therefore if continuous assessment instrument is to succeed, teacher must enhance their test construction skills. Thus, this study examined the teacher's competence on test construction procedure in Bosso LGA selected senior secondary schools. Research have shown how the teacher's incompetency in test construction procedure affects students' academic performance in most private and public senior secondary school in Nigeria.

Senior secondary school students require specialized assessment instrument for evaluation and decision making, which educators should adopt to meet these needs to improve educational standard.

1.3 Aim and Objectives of the Study

The aim of this study is to assess the teachers' skill of constructing multiple choice test by attempting to achieve the following objectives.

1. To evaluate the level of chemistry teachers' skill of constructing multiple choice test.
2. To determine gender differences on chemistry teachers' skill of constructing multiple choice test.

1.4 Research Question

1. What is the level of chemistry teachers' skill of constructing multiple choice test?
2. Is there any gender difference on chemistry teachers' level of skill in multiple choice test construction?

1.5 Significance of Study

This research is conducted with the objectives that the data and information collected will be of huge valence and help to practicing teachers, measurement and evaluation lecturers in Colleges of Education, Faculty of Education in the universities, policy makers, curriculum developers, ministries of education, examination bodies, etc. Science educators especially chemistry teacher should endeavor to use the item analysis tools in the method, techniques and procedure of test construction to enable a judicious and standard evaluation of test result. This study therefore is to inspire the minds of chemistry teachers towards the use of recommended guidelines in test construction of multiple choice.

This study will be of great help to examination bodies as it provides them with a method to access students' performance for evaluation purpose and also to encourage them to inspect test item critically.

It will aid measurement and evaluation specialist in various institutions to instill proper ideas of test construction unto their students as they as seen as potential teachers in effective evaluation of student's achievement.

1.6 Scope and Delimitation of the Study

The scope of this research is focused on chemistry teachers in secondary school in Bosso and Chanchaga Local Government of Niger State. With a questionnaire that will elicit information from the sample population.

1.7 Basic Assumption of the Study

The following postulate were made the course of this research:

1. That in Bosso and Chanchaga Local Government there are experienced and qualified chemistry teachers.
2. That experienced and qualified chemistry teachers are more skilled in multiple choice test construction.
3. That male and female chemistry teachers are equally skilled in construction of multiple choice test.
4. That chemistry teachers in Bosso and Chanchaga Local government operate the same syllabus and school calendar year.

1.8 Operational Definition of the key Terms

Test construction: is the set of activities involved in developing and evaluating psychological function in education.

Multiple Choice test: it is a form of an objective assessment in which students are asked to select only correct answers from a random option.

Professionally Qualified Teachers: Chemistry teacher holding a teaching qualification (NCE, B.Ed., B.Sc. Ed, P.G.D.E.) with chemistry as area of specialization.

Professionally Unqualified Teachers: Chemistry teachers without teaching qualifications (HND, B.Sc, B.A. B.Eng.)

Senior Secondary Schools (SSS): These are schools with students undergoing studies in the second tier of the 3 – 3 secondary school programme.

CHAPTER TWO

REVIEW OF RELATED LITERATURE

2.0 Introduction

Review of related literature involves the process of identification and analysis of documents that contains information related to the research topic that is being investigated. Therefore, this chapter present similar articles carried out by other researcher on multiple choice test construction and teachers' skill on test construction. Literature reviewed was based on two broad frameworks, these are Conceptual framework and Empirical studies.

Conceptual Framework

2.1 Brief History of Test

In 1904, the French government invited a psychologist named Alfred Binet to investigate and decide which students were mostly likely to experience challenges in school. Alfred Binet and his assistant Theodore Simon, developed questions that focused on content that had not been taught in school such as attention, memory, and problem-solving skills. Using these questions, Binet realized that some students were able to answer advanced questions than other students, while other students of the same age bracket were able to answer questions with the same as younger students. The result of this observation leads to Alfred Binet recommendation of mental age (Kamin, 1974) cited by Brink (2011). Alfred Binet and his assistant Theodore Simon first intelligence test is referred to today as the Binet-Simon Intelligent Scale Test. Group test was developed during the World War I when soldiers' intelligence is needed to be measured in order to assign them different tasks and operations. Important contributions were also made by

Karl Pearson, Edward L. Thorndike in achievement test and in 1864 George Fisher developed the first standardized objectives test.

West African Examination Council (WAEC) was established in Nigeria in 1952 to contribute to standard of education in Anglophonic countries of West Africa (Ghana, Nigeria, Sierra Leone, Liberia, and Gambia) by conducting a public examination and award a certificate equivalent to international examination bodies. Other examination bodies in Nigeria includes; National Examinations Council (NECO) which was established in 1999 to conduct Senior Secondary Certificate Examination (SSCE) and Junior Secondary Certificate Examination (JSCE) Joint Admission and Matriculation board (JAMB) established in 1976 which conducts entrance examination into universities, The National Business and Technical Education Board (NABTEB) established in 1992 is responsible for organizing examinations and certification of business and technical or vocational examinations.

2.2 Definitions of Test

Test is an essential instrument in educational system, which is useful in defining the course objective by the teacher. In educational terms, a process of measuring a person's ability or knowledge in a given area is referred to as test (Brown 2004). Simply put, a test is considered to be a kind of measurement device typically used to find out something about a person. Test refers to formal situation(s) deliberately created by a tester (teacher) to make the test-taker (student) respond to stimulus (stimuli) from which desired information could be elicited' (Abodurin, 1999). Test is an experience that the teacher creates to serve as basis for grading a learner in order to group them according to a laid down standard by an institution. Joshua (2005) defined test as an instrument used systematically to measure a sample of behaviour. Umoinyang and Nenty (2003) as cited in Ovat and Ofem, (2017) defined test as an instrument

used systematically to elicit response from respondent in order to determine the presence of a particular ability.

In summary, a test is instrument of evaluation which systematic measure the test respondent's ability, knowledge, and performance before or after a learning period has taken place in order to ascertain a value judgment. Therefore, the successfulness of the teaching and learning can be seen in the test's results.

2.2.1 Types of Test

Measures of maximum performance and Measures of typical performance are the two major categories of test classification base on nature of measurement. The measure of maximum performance which determines student's performance by motivation attain a satisfactory result consist of; Aptitude test, Achievement test and Intelligence test. Meanwhile, measures of typical performance are used to measure student's behavioral aspect of assessment which deals with social adjustment (NOUN 2006).

2.2.2 Intelligence Test

Alfred Binet is known as father of intelligence testing along with Theodore Simon in 1905. According to Binet and Simon, intelligence is, "judgement, good sense, initiative, the ability to comprehend and to reason well and to adapt one's self circumstances". Thorndike defined intelligence as the power of good response from the point of view of truth or fact and Charles Spearman stated that "intelligence is the analytic and synthetic ability of mind (Priyansh, 2018). An Intelligence test consist of wide variety of tests so as to sample several aspects of cognitive function. Some people believe that Intelligence can be expressed only in speech and writing and therefore cannot be tested.

2.2.3 Aptitude Test

Aptitude test is next to intelligence test which is also considered to predict student's outcomes during a learning process, both intelligence and aptitude test provide sufficient information about the potential of a student. Van Dusen defined aptitude test as the "measure of the probable rate of learning which results in interest and satisfaction and is relatively specific and narrow". According to Traxler "Aptitude is a condition, a quality or a set of qualities which is indicative of the probable extent to which an individual may be able to acquire, under suitable training, some knowledge, understanding, or skill" (Disha, 2019). Therefore, aptitude test is relatively based on indication of one's potentialities. An example is the Common Entrance Examination into Vocational Schools and even Secondary Schools.

2.2.4 Achievement Test

Achievement means one's learning attainments, accomplishments, proficiencies, etc. It is directly related to the pupil's growth and development in educational situations. Achievement test is defined by NM Downie as "The type of ability test that describes what a person has learned to do". According to Thronk and Hagen achievement test is "A systematic procedure for determining the amount a student has learned through instructions" (Amanda 2014). Achievement tests are designed to measure the effects of a specific programme of instruction or training which the learners attained usually by their effort. Generally, they represent a terminal evaluation of the learner's status on the completion of a course of study or training. That is, it is used to determine how much the learner has learned from specified content via systemic and controlled instructions. End of term examinations and classroom tests are mostly achievement tests. Achievement test can be broadly classified by format of test item; objective test item and essay test item. By mode of response; Oral test, Written test and Practical test. By purpose of testing; Placement test, Diagnostic test, Formative test,

Summative test. By Degree of Rigour Employed in Preparation and Scope of Applicability; Teacher made test and Standardized test (NOUN 2006).

2.2.4.1 Essay Test Item

An essay test is one which contains items (questions) that requires the student to compose a free response, usually lengthy ones. Essay test is appropriate for measuring outcomes that involve high-order skills, such as application and synthesis. It should be noted that the scoring of objectives tests is generally subjective. Therefore, in essay test there is no only one answer; but there can be a number of alternative ways that would correctly answer the question involved (Ridwan 2017). They also have poor psychometric quality or measurement qualities although popular among classroom teachers especially those who are deficient in the skill required for item construction. The classification of essay test item is based on the degree of freedom namely the Restricted response type and the Extended response type.

2.2.4.2 Objective Test Item

Simply put, an objective test are those test items in which a random options are provide with only one correct answer available among the options. Objective tests, in which students are not required to compose responses but rather to select from among a number of given alternatives, are more valid and reliable than essay items (Ridwan et al., 2017). The two major classification of objective test item include the free-response type that requires the test-taker to provide a word or phrase as an answer such as short-answer type and completion type while the fixed-response type which provides the testee with options as an answer can be further divided into; True or False test item, Matching test item and Multiple Choice test item.

2.2.5 Characteristics of a good test

The construction of an item for the purpose of a classroom test cannot be underestimated, there is lower chance for an inexperienced teacher to construct a good item. However, there are

numbers of characteristics which are desired for all tests developed by experts in educational measurement and evaluation. The characteristics include; Validity that is the extent to which test measure what is intended to measure, Reliability which is refer to as consistency of test, Test objectivity is the situation whereby the tester's opinion, bias, or individual judgement does not affect the scoring system, Discrimination in test is concerned with the construction in a way that will detect or measure small differences in achievement, Comprehensiveness of a test focused on including sufficient number of lesson objectives to provide a valid measure of the student's achievement (Ridwan et al., 2017).

2.3 Multiple Choice Test

Multiple choice tests have been found as an effective way of identifying the misconceptions of the student's measurement and evaluations by researchers rather than oral or essay test. Multiple choice tests are tests consisting of questions with a series of alternative answers from which learners select the one which they believe to be correct and do not require any additional input or adjudication during scoring. A multiple choice test item consists of two components: A problem (stem) – that may be in the form of a question/problem or an incomplete statement, at the beginning of each item and it should contain qualities such as succinctness, positive wordings and clear statement of question. A list of options (alternatives) – that contains one correct option (the answer) and a number of incorrect options (distractors) with a minimum of four options that are; similar in length, correct grammar, one correct answer, mutually exclusive alternatives, and absence of “All of the above” and “None of the Above” (Lurdes, José, Cristina, & Ana 2017).

2.3.1 Multiple Choice Test Construction

The role of test in educational system is inestimable which is crucial for the teachers to know how to design a good test which require adequate and extensive planning for effective

evaluation. Inadequate planning of test item usually leads to the test being poorly conceived, poorly worded ambiguous and grammatically incorrect. Furthermore, the test may contain items that are either not scorable or have more than one correct answer. It may be either too easy or too difficult and measures content out of range. Therefore, writing items that are valid, reliable and objectively scorable requires time, energy and adequate planning. To ensure successful planning, the teacher can follow these steps in order to create a good test item (Marjorie, 2014).

- Determining the instructional objectives.
- Content survey of the subject matter.
- Preparing test blueprint.
- Test item writing.

Determining the instructional objectives: this is achieved by producing the test items from stated instructional objectives that constitute the intended learning outcomes of instruction in terms of subject matter and student behavior. Instructional objectives should be critically reviewed when creating a test item because it encloses what the students are expected to possess by the end of the classroom instruction. The Bloom's taxonomy of educational objectives is a useful guide in classifying the instructional objectives stated for the assessment of behavior in the cognitive domain into knowledge, comprehension, application, analysis, synthesis and evaluation. Educational instructional objectives are the soul of a standard test item construction (Ridwan et al., 2017).

Content survey of the subject matter: the next step in test item construction is outlining and assigning weight which depends on the importance and emphasis given to the content (topics) to be covered in the test, the this is important because the subject matter is the vehicle through which behavioral objectives are taught and achieved. The importance of content survey is the

means of determining the level of mastering of the instructional objectives (Ridwan et al., 2017).

Preparing test blueprint:

The test blueprint is also known as the Table of specification, is a two-way table of behaviors (objectives) and content. It provides a method for relating the content to the objectives and makes it possible to classify each item in terms of objective and content. A completed table shows the number of items needed to obtain a balanced measure of instructional objectives and the subject matter emphasized in the instruction. It describes the relative emphasis of the various subject area and the several types of behavior. Without the table there is a tendency to exaggerate the importance of one behavior to the neglect of others (Marjorie, 2014).

A well planned table of specification enhances content validity of that test for which it is planned. The two dimensions (content and objectives) are put together in a table by listing the objectives across the top of the table (horizontally), a useful guide in defining objectives in the cognitive domain is Bloom's taxonomy of educational Objectives which is classified into Knowledge (define, identify, list, label, select, name, recognize, reproduce), Comprehension (classify, derive, estimate, expand, calculate, convert, describe, give example), Application (apply, compute, show, construct, draw, demonstrate, use, prepare), Analysis (separate, determine, deduce, discriminate, estimate, make inference, outline), Synthesis (explain, summarize, categorize, develop, combine, compile, compose, formulate, plan), Evaluation (compare and contrast, evaluate, judge, justify, contrast, decide, criticize).

The taxonomy categorized the cognitive skills from the simple to the complex and the content down the table (vertically) to provide the complete framework for the development of the test items (Marjorie, 2014).

Step 1: Outline the instructional objectives, topics and number of periods per topics using the scheme of work and lesson plan as guide.

Table 2.1: List of Topics and number of Periods taken

Topics	Periods
Definition of Hydrocarbon and Organic compounds	1
Classification of Hydrocarbon	1
Homologous Series	1
Functional group	1
Alkanes	1
The Basics of Organic Nomenclature: Naming Alkanes	2
Physical and Chemical properties of Alkanes	2
Total number of periods	9

The table above comprises of treated topics in organic chemistry taught in 45 minutes per periods with three periods (135 minutes) per week for three weeks.

Step 2: Derive the percentage of each topic treated

Table 2.2: Treated Topics Percentage

Topics	Percentage
Definition of Hydrocarbon and Organic compounds	$1/9 \times 100 = 11.1\%$
Classification of Hydrocarbon	$1/9 \times 100 = 11.1\%$
Homologous Series	$1/9 \times 100 = 11.1\%$
Functional group	$1/9 \times 100 = 11.1\%$
Alkanes	$1/9 \times 100 = 11.1\%$
The Basics of Organic Nomenclature: Naming Alkanes	$2/9 \times 100 = 22.2\%$
Physical and Chemical properties of Alkanes	$2/9 \times 100 = 22.2\%$
Total	100%

Table 2.2 above consist the percentage of treated topics calculated by dividing number of period(s) used on each topic by the total number of periods multiplied by 100.

Step 3: establish the number of test items per treated topics from a proposed number of items using the treated topics percentage

Table 2.3: Number of test items extracted from percentage of treated topics

Topics	Number of items
Definition of Hydrocarbon and Organic compounds	11.1% X 30=3.33
Classification of Hydrocarbon	11.1% X 30=3.33
Homologous Series	11.1% X 30=3.33
Functional group	11.1% X 30=3.33
Alkanes	11.1% X 30=3.33
The Basics of Organic Nomenclature: Naming Alkanes	22.2% X 30=6.66
Physical and Chemical properties of Alkanes	22.2% X 30=6.66
Total	30 items

Table 2.3 justify how number of items are acquired from percentage of treated topics for 30 proposed test items by multiplying the percentage of the treated topics by the number of proposed test items.

Step 4: construction of the test blueprint or test of specification using the Bloom's taxonomy (knowledge [know], comprehension [comp], application [app], analysis [analy], synthesis [syn], evaluation [eva]) and Contents (Topics treated).

Table 2.4: Test Blueprint or Table of specification (TOS)

Content	Know. 30%	Comp. 17%	App. 20%	Analy. 13%	Syn. 10%	Eva. 10%	Total 100%
Definition of Hydro carbon and Organic Compound	2	1	-	-	-	-	3
Classification of hydrocarbon	1	1	1	-	-	-	3
Homologous Series	1	-	-	1	1	-	3
Functional Group	1	1	1	-	-	-	3
Alkanes	1	1	1	-	-	1	4
The Basics of Organic Nomenclature: Naming Alkanes	2	-	2	1	1	1	7
Physical and Chemical Properties of Alkanes	1	1	1	2	1	1	7
Total	9	5	6	4	3	3	30

Table 2.4 is the test blueprint or table of specification that represent the distribution of test items across the cognitive domains using Bloom's taxonomy of learning objectives. Test blueprints enables test developer, examination bodies, teachers etc. achieve a desired content coverage of test item in cognitive skills to assure validity and reliability (Alade & Omoruyi, 2014)

Item writing: the final task in constructing the test item is item writing, Item writing is the preparation of assessment tasks which can reveal the knowledge and skill of students when their responses to these tasks are inspected (Izard 2015). To be able to write good items the

teacher must have a thorough mastery of the subject matter, understand psychologically and educationally the students enough to be able to adjust the complexity and difficulty of her items to a levels that is appropriate for them, must be familiar with the various forms and types of tests to select that which best serves her purpose, ensure communication in order that the item communicates the desired task as simple as possible to avoid misinterpretations.

2.4 Item Analysis

Item analysis involve the uses of statistical tools that can provide pertinent information to improve the accuracy and quality of an objective test item by eliminating ambiguous and false item during test administration (Ary, 2002).

Lange, Lehmann, and Mehrens (1967) cited by Siri and Freddano (2011) point out that analysis of test item enables test developer to realize extremely difficult and easy items, items that are able to differentiate between students who have learned the content and those who have not, or questions that have distractors to detect their level of understanding of the particular item. In item analysis, the test developer is concerned with item difficulty, item discrimination and distractor analysis.

2.4.1 Item Difficulty Index

Item difficulty in analysis of a test item indicates the degree to which an item is difficult for the test taker in relation to their cognitive ability (Boopathiraj & Chellamani, 2013). It is estimated by the percentage of test taker that got the item right after it has been administered. Larger percentage indicate the students who have learned the content and lower percentage indicate students who have not achieve learning of the content effectively which is measured by the item. However, an item is said to be too difficult when the index is 0.3, too easy when the index is above 0.7 and for maximizing validity and reliability the optimal difficulty index is 0.5 which is usually preferred in summative assessment such as those handle by West

African Examination Council. High difficulty index occur when the content of such item was not taught, if the item lacks direction or when the item contains grammatical errors (Odukoya, Adekeye, Igbino, & Afolabi, 2017). The difficulty index (p) of an item is given by the formula;

$$p = \frac{R}{N}$$

Where:

P = Item difficulty index

R = Total number of correct responses in upper and lower groups

N = Total number of test taker in upper and lower groups

Note:

The best 25% or 27% highest total score constitute the upper group while poorest 25% or 27% lowest total score form the lower group.

2.4.2 Item Discrimination Index

Item discrimination index is also known as discrimination power of an item; it is an indication of how well an item discriminates between high and low scores of test taker (Sabri, 2013). It determines the extent to which the given item discriminates among examinees in the function or ability measured by the item. Item discriminations of $D = 0.50$ or higher are considered excellent, $D = 0$ means the item has no discriminatory power, while $D = 1.00$ means the item has perfect discrimination power (Odukoya et al., 2017). It is obtained using the formula below.

$$D = \frac{R_u - R_l}{T}$$

Where:

D = Discrimination index

R_u = Total number of test taker who choose correct responses in the upper group

R_l = Total number of test taker who choose correct responses in the lower group

T = Total number of students in upper and lower groups

2.4.3 Distractor Analysis Index

A distracter analysis allows you to examine how many test taker in the upper group and lower groups selected each option on a multiple choice item (*The role of testing and assessment in Education*, n.d.). The distraction power of a distractor in a test item is its ability to differentiate between those who do not know and those who know what the item is measuring. That is, a good distracter attracts more test taker from the lower group than the upper group.

According to Instructional Assessment Resources (Instructional Assessment Resources, 2011), student performance in an exam is very much influenced by the quality of the given distractors (Brink et al, 2011).

In order for a distractor to be acceptable it should attract at least one candidate. If no one selects a distractor it is important to revise the option and attempt to make the distractor a more plausible choice (Educational Development Center, 2017). The formula is giving below.

$$D_o = \frac{R_l - R_u}{T}$$

Where:

D_o = Distractor power

R_l = Total number of test taker who choose distractor in the lower group

R_u = Total number of test taker who choose distractor in the upper group

T = Total number of test taker in both groups

2.5 Test Item Validity

Cronbach posited test item validity as the process in which a test developer collects evidences to support the test scores' interpretations. He recommended that when a test developer planned to conduct a test validity, clear identification of the intended outcome is required (Rufina, 2016). Therefore, a test item validity simply refers to the degree in which the item is measuring what it intended to measure.

2.5.1 Types of Test Item Validity

In this research, three (3) types of Test Item Validity which are; Face validity, Content validity, Construct validity, and Criterion validity are discussed below

Face validity:

Face validity is also known as logical validity, it the simplest and weakest form of validity (Stephanie, 2015), a test item has face validity when the contents of the test item simply 'looks' relevant to the test developer, it is a quick method of establishing item validity after it construction by assessing the item whether it is relevant, reasonable, unambiguous and clear (Oluwatayo, 2012).

Content validity:

Content validity is the extent to which a measure "covers" the construct of interest, it involves the process of comparing the test item with the instructional objectives, it is also called Rational validity, Curricular validity, Internal validity and Intrinsic validity. Content validity refers to the degree or extent to which a test consists items representing the behaviors that the test developer wants to measure. The extent to which the items of a test are true representative of the whole content and the objectives of the teaching is called the content validity of the test (Disha et al., 2019).

Criterion validity:

This is concerned with the effectiveness of a test item in predicting a test taker's future performance in another given assessment. Criterion validity (or criterion-related validity) measures how well one measure predicts an outcome another measure. A test has this type of validity if it is useful for predicting performance or behaviours in another situation (past, present or future) (Stephanie et al., 2015). There are two types of criterion validity namely; Predictive validity and Concurrent validity.

- **Predictive validity:** a test item is predictively valid when it theoretically predicts a future outcome by correlating the predictor variable and the criterion variable by the means of correlation coefficient using either Spearman Rank Correlation Pearson Product-Moment Correlation (No title, 2019). An example is a situation whereby a formative assessment (classwork) and a summative assessment (terminal examination).
- **Concurrent validity:** the term 'concurrent' means 'simultaneous' or 'existing'. Thus concurrent validity indicates the process of validating a new test by correlating its scores with some existing criterion obtained shortly before or shortly after the new test is administered. Predictive validity tells us the relationship between measures over a long period of time, Concurrent validity is concern in estimating present status between the measures obtained concurrently (Disha et al, 2019).
- **Postdictive validity:**
For this type of validity, the criterion is in the past. That is, the criterion (e.g., another test) was administered in the past. It is a form of criterion-referenced validity that is determined by the degree to which the scores on a given test are related to the scores on another, already established test or criterion administered at a previous point in time (Hamed, 2016).

Construct validity:

Educationally, Construct means psychological quality that exist in relation of some aspect of behavior such as creativity and intelligence. Therefore, construct validity is the degree in which a test item measures a psychological quality or trait that the test developer intends to measure (Ogomaka & Caroline, 2017).

2.6 Reliability

Reliability is an essential quality of any test item. It provides the consistency that makes validity possible (Linn and Gronlund, 2000). A test item is said to have a high reliability if it produces similar results under consistent conditions. If a test-taker has a certain skill level and she or he is able to demonstrate the same level when retested, then the skill level is reliable (Erwin & Najib, 2011). Reliability of a test item can be obtained by Test-retest method, Split-half method and Internal Consistency method (Cronbach's Alpha)

2.6.1 Method of Test Reliability

The method of measuring the reliability of a test which include; Test-retest, Split-half and Internal Consistency (Cronbach's Alpha)

Test-retest method:

Measuring reliability of a test item using test-retest method is determined by administering the test twice to the same group of learners with a given time interval between the two administrations (Webpage, 2019). The two test are correlated and the test is considered highly reliable when there is no significant difference between the two tests.

Split-half method:

Split-half reliability is a measure of consistency between two halves of a test that is obtained after a single administration. In split-half reliability, a test for a single content area is split into two parts and then both parts are administered to one group of students at the same time. The scores from both parts of the test are correlated using spearman-rank order correlation method (Mcleod, 2019).

Internal Consistency (Cronbach's Alpha):

Internal Consistency Reliability test determine how all factors on the test relate to all other factors, internal consistency of a test is measured from a single test administration. One of the common measure of internal consistency is Cronbach's alpha which allows the test developer to assess the test reliability and determine the score variance and covariance. Cronbach's alpha increases as the inter-correlations among test factors increase, since inter-correlations among test factors are maximized when all factors measure the same behavior that are poly-numerously scored such as questionnaire (Rosaroso, 2016).

2.7 History of Science Education in Nigeria

Teaching of science education in Nigeria can be traced to the era of Christian missionaries who introduced western education. It started with the establishment of Church Missionary Society (CMS) grammar school in Lagos, 1859 (Brown, 2015). The foundation of science education was created and infused into school curriculum with arithmetic, algebra, geometry and physiology as the main subject. Before the year 1931, only very few students in secondary schools tried science at external examinations conducted by Oxford and Cambridge examination board and those students that tried it, failed (Ogunleye, 1999). Brown (2015), noted that before the year 1932, there were no post-secondary institution for the teaching and

learning of science education in Nigeria after the specialized institution, the Medical School attached to CMS Theological College in Abeokuta close down. Between the period of 1883 - 1930 indicate the beginning of Colonial government involvement in the development of science in secondary education and the education ordinance of 1980 also marked another phase in the development of science teaching in schools. In the year 1932, Yaba college was established which was a major development in science curriculum and was later upgraded to Yaba College of technology, to run courses in engineering, medicine, science, agriculture, survey and teacher training college to filled vacancy in relevant government development. Another major development in the history of science education in Nigeria was found in the year 1948 when University College Ibadan was established as College University of London following the report of Elliot commission Higher education set up in 1943, (Ogunleye, 1999). Chemistry, Biology, Physics and laboratory work were offered after the introduction of Higher School Certificate (HSC) in 1951. In 1952, an examination board was set up with its headquarters in Accra, Ghana and the board later became the west African examination council (WAEC), which received the curriculum of school subjects including science, with its first examination in 1995 (Brown, 2015). On 30 November, 1957, The Science Teacher Association of Nigeria (STAN) was established to review the science curriculum of WAEC and HSC in May 1968. The Federal Colleges of Arts, Science and Technology at Ibadan in 1950; Zaria in 1952 and Enugu in 1954 were established with the aim of promoting the teaching of science.

2.8 Concept of Chemistry Education

Chemistry is an essential and a popular subject among science senior secondary school students in Nigeria due to its application across many fields. Nigeria needs a functional chemistry education in schools that will boost the development of the nation through industrialization,

therefore chemistry education has been identified to be one of the major cornerstone for the development of our National economy (Emendu, 2014).

The three basic science subject contains in the National Science Curriculum are; Biology, Chemistry and Physics which are strictly offered in Science Senior Secondary School in Nigeria education (Ihuarulam, 2008).

Among these basic science subjects, chemistry stand out and occupies a unique position due to its requirements in study of many science courses, such as medicine, biochemistry, pharmacy, agricultural science, laboratory technology, geology, etc. (Ihuarulam et al., 2008). Chemistry education plays an important role in global sustainable economic development by enhancing research and quality of teaching to ensure students are provided with standard knowledge of chemistry which is the nucleus of science (Okieimen, 2007). It contributes to the major development of food (fertilizers and insecticides), clothing (textile fibers), housing (cement, concrete, steel, bricks), Medicine (drugs), Transportation (fuel, alloy materials) which are vital needs that affect human one or the other (Shamsuddin, Arome, Aminu, Isah, & Adamu, 2018). Chemistry is one of the naturally and well established means through which the nation's abundant natural and mineral resources can be harnessed into useful ventures for the overall economic and socio-political wellbeing of its citizenry (Uwague, 2008). Okoro (2013) noted that Chemistry is a systematic investigation of nature with a view to understanding and applying them to serve human needs.

Emendu et al., (2014) defined Chemistry Education as the acquisition of knowledge or ideals relevant to chemistry which is concerned with the impartment of knowledge on properties, components, transformations and interactions of matter. Therefore, Chemistry education is the methodical process of attaining ideas and exploration about natural entity for the benefit of man. According to Emmanuel (2013), Chemistry education is the vehicle through which

chemical knowledge and skill reach the people who are in need of capacities and potentials for development.

The function of chemistry education in national development in Nigeria will be directly related to the teachers who contribute greatly in the teaching and learning process and will help them to inculcate in their students the vast knowledge in chemistry, the positive value includes the increase in employment rate and the economy.

State of Chemistry Education in Nigeria

According to several research, it has been reported that secondary school students perceived chemistry as an extremely difficult science subject, hence increasing the ‘phobia of chemistry’. It was equally observed that the ‘phobia of chemistry’ is attributed to several factors in chemistry education such as; lack of qualified and dedicated teachers, quality of teaching, lack of practical skills, poor funding and mismanagement, and defective curriculum. These factors are discussed below;

Lack of qualified and dedicated teachers:

Educational quality depends largely on the teachers because of the vital roles they play during teaching and learning process (Muhammad, 2013). Achieving the goals of chemistry education requires qualified and highly scientifically literate teachers. Meanwhile, if teachers are unmotivated, passionless, uncommitted, uninspired, the negative impact of such attitudes to the development of the nation is immeasurable. If they are ignorant in their discipline and impart wrong knowledge, they become not only useless but dangerous (Musa 2005). Qualified and dedicated Teachers remains the most important resource in the school that predicts student achievement by providing standard test score and test quality (Dahor, Dahar & Faize, 2011).

Quality of Teaching:

Quality teaching is defined as the ability of the teacher to transform written instructions into forms that are pedagogically accepted and adaptive to the students' abilities and educational background (Okafor, 2007) cited by (Omorogbe & Ewansiha, 2013). Quality teaching is effective teaching and effective teaching is the process whereby students learn and achieve scientific knowledge and scientific goals (Omoifo, 2012). Lack of teaching quality has been one of the major problem of chemistry education in Nigeria senior secondary schools (Omorogbe & Ewansiha, 2013). Ayodele (2006), outlined the use of inappropriate teaching methodology as a major factor hindering students understanding and achievement in science, which does not require theoretical and lecture approaches because it makes the lesson un-interactive and difficult to grasp some scientific principles, concepts and skills rather it should involve more practical aspect.

Lack of practical skills:

Chemistry teaching is more effective when the students are involved in laboratory practical with adequate laboratory equipment. Lack of practical knowledge led to the failure of students in various external examination which can be traced to the teacher's lack of practical skills (Efe, 2008). Chemistry practical skills are science process skills which are the foundation for both enquiry and development of intellectuals and individual skills needed for learning science concepts (Efe et al., 2008). Therefore, shortage of laboratories equipment and lack of teacher's practical skills could have serious implication on the quality of chemistry education (Adeyemi, 1998; Ige, 2000).

Poor funding and mismanagement:

Chemistry education in Nigeria senior secondary school has no doubt been faced with lack of funding and mismanagement, most schools especially the tertiary institution gain grants by government specific project and international agencies (Emmanuel et al., 2013). The poor

funds provided by the government and other educational agencies are mismanaged in the process of providing a functional laboratories and equipments thereby decreasing national education through chemistry education (Ibrahim, Adamu, Ibrahim, Ismaila & Abubakar, 2018).

Poor curriculum:

Curriculum means the planned experiences given to learners by the teachers under the guidance of the school. The chemistry education curriculum of senior secondary school is found out to be overloaded and focused on theoretical teaching which is supposed to laid more emphasis on practical class. Practical exercises are not well structured in the curriculum, practical are not taught in schools until examination bodies send practical specimens thereby subjecting the students to unwarranted assimilation (Ibrahim, Adamu, Ibrahim, Ismaila & Abubakar et al., 2018). In order to effectively implement chemistry education curriculum in senior secondary school, provision of laboratory facilities should be by a combined effort of the government, Parent Teachers' Association and other education stakeholders, chemistry teacher should be involved the decision-making and planning of the curriculum, state government should employ adequate number and qualified chemistry teachers to handle secondary education chemistry curriculum (Iheanyi, 2015).

Empirical Studies

2.9 Teacher competencies in test construction

Teacher's knowledge or competency is highly need in school assessment in order to ascertain unbiased evaluation. Test quality is concerned with its power to produce the information needed regarding students' performance, hence adhering strictly to standard procedures for test construction (Rhoda, 2017). UNESCO specify standards requirement for assessing knowledge or skills that teacher should possess in order to perform well in their evaluation effort (Ololube,

2008). Chemistry teachers are expected to acquire relevant skills in item construction for classroom assessment in order to develop an item that is valid and reliable for accurate feedback of students' achievement. Agu, Onyekuba and Anyichie et al., (2013) noted that test construction skills consist of the knowledge needed to process a quality item based on outlined principles of test construction.

In a study carried out by Agu, Onyekuba and Anyichie et al., (2013) on teachers' competencies in constructing classroom based tests in Nigeria and the need for a test construction skill inventory, the researcher constructed 30 items measuring teachers' test construction skills validated by varimax-rotated factor analysis (maximizes the sum of the variance of the squared loadings, where 'loadings' means correlations between variables and factors) eliminating 5 items and the remaining 25 items were subjected to internal consistency to obtain reliability coefficient of 0.73. The study found no significant difference between male and female teacher and significant difference was detected between experienced teacher and inexperienced teacher.

In another research study investigated by Ololube et al., (2008) on competencies of professional and non-professional teachers in Nigeria. He reported that professional teachers tend to construct more valid and reliable evaluation instrument more than non-professional teachers. The author also stated that professional teachers have the tendency to apply various evaluation techniques correctly than the non-professional teachers. The researcher concludes that professional teacher kept assessment records more accurately than non-professional teachers (Rufina, Abubakar & Stephen et al 2015).

CHAPTER THREE

RESEARCH METHODOLOGY

3.0 Introduction

This chapter presents the description of a specific methodology the researcher used in assessing secondary school chemistry teachers' skill of constructing multiple choice test in Bosso and Chanchaga Local Government of Niger State. This would enable the researcher to answer the research questions formulated to guide this study. This chapter was subdivided into the following subheadings: Research design, Population of the study, Sampling and Sampling technique, Research instrument, Validity of the instrument, Reliability of the instrument, Method of data collection and Method of data analysis.

3.1 Research Design

The research design used for this study is a descriptive research of the survey type, which required the researcher to design questionnaire used to gather relevant information on the assessment of secondary school chemistry teachers' skill of constructing multiple choice test.

3.2 Population of the Study

The population of the study include the entire public senior secondary school teachers in Bosso and Chanchaga local government, Niger State. According to the Niger State Secondary Education Board, there are thirty-five (35) teachers in Bosso and Chanchaga Local Government of Niger State.

3.3 Sample and Sampling Techniques

Chemistry teachers in Bosso local government of Niger State were selected using total sampling which is a type of purposive sampling technique. Therefore, all thirty-five (35) chemistry teachers in secondary schools of Bosso and Chanchaga Local Government of Niger State were selected for the purpose of this study.

3.4 Research Instrument

The research instrument was a questionnaire designed by the researcher which was used for data collection on teachers' skill of constructing multiple choice test (SCMCT). The research instrument is divided into two section, Section A: consist of the respondent name of school and gender and Section B: contained twenty (21) items which assess the respondent skill of constructing multiple choice test whereby four rating scales point namely: Strongly Agree (SA) = 4, Agree (A) = 3, Disagree (D) = 2 and Strongly Disagree (SD) = 1 were instructed to represent their perception on each items.

3.5 Validity of the Instrument

To ensure validity of the instrument used in this research, the researcher's supervisor and a lecturer both from Science Education Department, Federal University of Technology Minna validated the instrument. Their advice and suggestions were used to modify the items in the instrument and also prepared the final copy, the instrument is certified to have content validity.

3.6 Reliability of the Instrument

The reliability of the instrument was determined by using Cronbach alpha. This procedure involves the situation where by the researcher analyzes each items to determine the degree of internal consistency. The reliability coefficient of this study was estimated at 0.75

3.7 Method of data collection

The researcher visited and distributed copies of questionnaire to the samples in each secondary school, Bosso and Chanchaga respectively. The researcher collected back the copies of the questionnaire after it had been completed and collated for analysis.

3.8 Method of Data Analysis

The data obtained from the administered questionnaire were subjected to descriptive statistics. Statistical package for social science (SPSS) were used to analyze the data collected, using t-test, frequency, percentage and mean analysis to analyze each items.

CHAPTER FOUR

4.0 DATA PRESENTATION, ANALYSIS AND DISCUSSION

4.1 Introduction

Data presentation and analysis are important aspect of any research work serving as the deciding point of any research effort and giving meaning and shape of the raw data collected. In the course of this research work, data were collected from Secondary Schools in Minna Niger State of Nigeria through administering 35 copies of questionnaires to the sampled respondents. It implies that the research questions formulated of this study was based on the information gathered from the questionnaire administered. This was done with the aid of tables using t-test, mean, percentages and frequency analysis of each items.

4.2 Research Question One: What is the level of chemistry teachers' skill of constructing multiple choice test?

Table 4.1: Shows frequency and percentage of chemistry teachers level in skill of constructing multiple choice test.

S/N	ITEMS	SA(%)	A(%)	D(%)	SD(%)	MEAN	RMK
1.	A table of specifications relates test item to instructional objectives.	8(22.9)	13(37.1)	12(34.3)	2(5.7)	2.77	Agreed
2.	The table of specification is a two-way table of behaviors (objectives) and content.	5(14.3)	10(28.6)	19(54.3)	1(2.9)	2.54	Agreed
3.	A multiple choice item can be reliable without being valid.	4(11.4)	14(40.0)	15(42.9)	2(5.7)	2.57	Agreed

4.	Inadequate planning of test item usually leads to the test being poorly conceived, poorly worded, ambiguous and grammatically incorrect.	5(14.3)	14(40.0)	14(40.0)	2(5.7)	2.63	Agreed
5.	To be able to write good test items the teacher must have a thorough mastery of the subject matter.	4(11.4)	17(48.6)	12(34.3)	2(5.7)	2.66	Agreed
6.	Multiple choice tests are tests consisting of questions with a series of alternative answers.	5(14.3)	14(40.0)	16(45.7)	-	2.69	Agreed
7.	Test items are subjected to item analysis.	11(31.4)	9(25.7)	15(42.9)	-	2.89	Agreed
8.	The best method for assessing students is through the use of multiple choice test.	7(20.0)	10(28.6)	16(45.7)	2(5.7)	2.63	Agreed
9.	A list of options (alternatives) contains one correct option (the answer) and a number of incorrect distractors.	4(11.4)	17(48.6)	12(34.3)	2(5.7)	2.66	Agreed
10.	Multiple choice test consists of two components; a stem and alternatives.	8(22.9)	15(42.9)	10(28.6)	2(5.7)	2.83	Agreed
11.	Test blueprints is also known as Table of Specification	7(20.0)	16(45.7)	10(28.6)	2(5.7)	2.80	Agreed
12.	Instructional objectives should be critically reviewed when constructing a test item.	7(20.0)	17(48.6)	9(25.7)	2(5.7)	2.83	Agreed

13.	Test reliability refers to the degree of accuracy with which a test measures what it is designed to measure.	14(40.0)	14(40.0)	6(17.1)	1(2.9)	3.17	Agreed
14.	The final task in multiple choice test construction procedure is the item writing.	6(17.1)	19(54.3)	9(25.7)	1(2.9)	2.86	Agreed
15.	Item analysis involve the uses of statistical tools to provide pertinent information to improve the accuracy and quality of a test item.	9(25.7)	12(34.3)	13(37.1)	1(2.9)	2.83	Agreed
16.	Without the table of specification there is a tendency to exaggerate the importance of one test item to the neglect of others.	4(11.4)	15(42.9)	13(37.1)	3(8.6)	2.57	Agreed
17.	Face validity is the weakest and simplest form of validity.	5(14.3)	15(42.9)	14(40.0)	1(2.9)	2.69	Agreed
18.	A well planned test blueprint enhance content validity.	15(42.9)	12(34.3)	8(22.9)	-	3.20	Agreed
19.	Test blueprints enable test developer achieves a desired content coverage of test item to assure validity and reliability.	2(5.7)	18(51.4)	15(42.9)	-	2.63	Agreed
20.	The Bloom's taxonomy of educational objectives are useful guide in classifying the instructional objectives when constructing test item.	7(20.0)	18(51.4)	7(20.0)	3(8.6)	2.83	Agreed

21.	Determining the instructional objectives is the first measure to be considered in multiple choice test construction procedure.	14(40.0)	14(40.0)	6(17.1)	1(2.9)	3.17	Agreed
GRAND MEAN						2.78	

Table 4.1 indicates that the frequency, percentage and mean of items, the frequency and percentage of items that respondents strongly agree and agree (positive) is 42 (52.5%) and the frequency and percentage of items that the respondents strongly disagree and disagree (negative) is 40 (48.7%). Also the grand mean response of the item is 2.78 which is higher the decision rule (2.5). Therefore, chemistry teachers are moderately skilled in constructing multiple choice test, this shows an improvement on the level of chemistry teachers skill in constructing multiple choice test.

4.2 Research Question Two: Is there any gender difference on chemistry teachers' level of skill in multiple choice test construction?

Therefore, to answer this question a null hypothesis was formulated which stated as follows

HO₁: There is no significant gender difference on chemistry teachers' level of skill in multiple choice test construction.

To test the formulated hypothesis an independent t-test was used and the findings is presented in the Table 4.2.

Table 4.2: Shows t-test result of gender difference on chemistry teachers' level of skill in constructing multiple choice test

Gender	N	df	Mean	Std.Dev	t	P
Male	19	33	2.67	0.824	17.056	0.00
Female	16		2.93	0.803		

Table 4.2 using t-test reveals that there is significant gender difference on chemistry teachers' level of skill constructing multiple choice since the t-cal (t) is greater than p-value (Sig.). Therefore, the null hypothesis is rejected.

4.3 Summary of Findings

1. Chemistry teacher are moderately skilled in constructing multiple choice test.
2. There is significant gender difference on chemistry teachers' level of skill in multiple choice test construction.

4.4 Discussion of Results

The aim and objective of this study is to assess the chemistry teachers' skill of constructing multiple choice test in secondary school of Bosso and Chanchaga Local Government of Niger State. Two research question were generated and answered by analyzing the t-test, mean, frequency and percentage of the respondents' items.

The first research question was formulated to find out the level of chemistry teachers' skill of constructing multiple choice. Table 4.1 shows that chemistry teachers have moderate level of skill in constructing multiple choice test as 52.5 % agreed that chemistry teachers are skilled in constructing multiple choice test.

The second research question was generated to find out the gender difference on chemistry teachers' level of skill in constructing multiple choice test. Table 4.2 There is significant gender difference on chemistry teachers' level of skill in multiple choice test construction.

CHAPTER FIVE

5.0 SUMMARY, CONCLUSION AND RECOMMENDATION

5.1 Introduction

This chapter entails the summary of the research work, major findings of the study, conclusion, recommendation and suggestion for further study.

5.2 Summary

The aim and objective of this research was to assess the secondary school chemistry teacher skill of constructing multiple choice test in Bosso and Chanchaga Local Government of Niger State. Two research questions was generated to guide the researcher, the first research question seek to find out the level of secondary school chemistry teachers' skill of constructing multiple choice test, the second research question sought to uncover the gender difference of chemistry teachers' skill of constructing multiple choice test.

To attain these purposes, a questionnaire tagged SCMCT was administered to thirty-five chemistry teachers total sampled population from public secondary school in Bosso and Chanchaga Local Government of Niger State. Each items from the respondents were analyzed using t-test, frequency, percentage and mean table.

The analysis of the data collected revealed that chemistry teachers are moderately skilled in constructing multiple choice test and there is significant gender difference on chemistry teachers' level of skill in multiple choice test construction.

5.3 Conclusions

It was concluded base on the result of analysis that majority of chemistry teachers in Bosso and Chanchaga Local Government of Niger State are skilled in constructing multiple choice test and There is significant gender difference on chemistry teachers' level of skill in multiple

choice test construction. This could be as a result of exposure to test constructing skills seminar and workshop.

5.4 Recommendations

Based on the findings in this study, the following recommendations are made:

1. Niger State Ministry of Education and other educational bodies should often organize seminar/workshop for science teachers on the skill of constructing multiple choice test to bring about achievement of educational goals.
2. Chemistry teachers should get themselves acquainted with relevant textbooks on multiple choice test construction to guide them in planning and administering of appropriate tools for better evaluation of students' learning outcomes.
3. The use of multiple choice items should be encouraged to reduce the effect of inappropriate scoring in the use of essay type tests.
4. Chemistry teachers should consistently practice the test construction procedure learnt in the organized workshop and seminar to enhance the effectiveness of their skills

5.5 Suggestions for Further Study

The researcher makes the following suggestions for further study:

1. This study should be replicated using larger population like Chemistry teachers of Niger State.
2. The study can be further conducted by comparing the skill of multiple choice test of art or commercial teachers.

REFERENCES

- Abodunrin, (1999). *Evaluating Teaching/Learning Effectiveness: A basic course in Educational, Test, Measurement and Evaluation*. Cardinal Information Technology.
- Achievement of Students at Secondary Stage in Punjab (Pakistan). *European Journal of Social Sciences*.19:1
- Adeola, A. O., & Fajonyomi, A. A. (1999). Adult learners' performance on standardized and nonstandardized tests as a function of sex and location. *Nigeria Journal of Counselling Development*, 4(1), 32-36.
- Adeyemi, T.O., 1998. *School and Teacher Variables Associated with the Performance of Students in the Senior Secondary Certificate Examinations in Ondo State Nigeria*. Ph.D. Thesis, University of Hull, Kingston Upon Hull, England, UK.
- Agu, N. N., Onyekuba, C., & Anyichie, A. C. (2013). Measuring teachers' competencies in constructing classroom-based tests in Nigerian secondary schools:Need for a test construction skill inventory. Retrieved from http://www.academicjournals.org/article/article1379700583_Agu%20et%20al.pdf
- Alade, O. M., & Omoruyi, I. V., (2014). Table of Specification and its Relevance in Educational Development Assessment. *European Journal of Educational and Development Psychology*. Vol.2, No.1, pp.1-17
- Allen, M. J., & Yen, W. M. (2002). Introduction to measurement theory. *Illinois: Waveland Press*.
- Amanda J. (2014). *Achievement test- Definition, objectives, functions, characteristics*. Retrieved from <http://www.nsgmed.com/achievement-test-definition-objectives-functions-characteristics>
- Ary, D., Jacobs, L.C., Razavieh, A. (2002). Introduction to Research in Education. *6th edn*. Wadsworth, California Press
- Ayodele O. O., (2006). Building A Sustainable Science Curriculum in Nigeria; Accommodating Local Adaptation, Leveraging Technology and Enhancing Ares of Improvement for Quality Assurance. *Journal of Science Teachers Association of Nigeria* 1(7)
- Boopathiraj, C., & Chellamani, K., (2013). Analysis of Test Items on Difficulty Level and Discrimination Index in the Test for Research in Education. *International Journal of Social Science & Interdisciplinary Research*, ISSN 2277 3630, Vol.2 (2).
- Brink, Carole Sanger, "A Historical Perspective of Testing and Assessment Including the Impact of Summative and Formative Assessment on Student Achievement" (2011). *Dissertations*. 409. <https://aquila.usm.edu/dissertations/409>

- Brown, D. H., (2004). *Language Assessment: Principles and Classroom Practices*. Longman, San Francisco State University.
- Brown, M. G.,(2015). Science and Technology Education in Nigeria: A Historical Perspective. *SCSR Journal of Education Research (SCSR-JER)*. Vol. 1, No. 1, pp. 33-41.
- Crocker, L., & Algina, J. (2008). Introduction to classical and modern test theory. *Ohio: Cengage Learning Press*.
- Dahor, M. A Dahar R. A Dahar R. T. and Faize F. A, (2011) Impact of Teacher Quality on Academic
- Disha, M. (2019). *Aptitude Tests: Definition, Measurement and Uses*. Retrieved from <http://www.yourarticlelibrary.com/statistics-2/aptitude-tests-definition-measurement-and-uses>
- Disha, M. (2019). *Validity of a Test: 6 Types*. Retrieved from <http://www.yourarticlelibrary.com/statistics-2/validity-of-a-test-6-types-statistics/92597>
- Dosumu, C. T. (2002). Issues in teacher-made tests. *Ibadan: Olatunji and Sons Publishers*.
- Educational Development Center, (2017). *Scantron Guides: Item Analysis*
- Efe O. M (2008). The effect of practical skills on students' achievement in practical chemistry amongst some selected secondary schools in Zaria metropolis.
- Emendu, N. B., (2014). The Role of Chemistry Education in National Development. *The International Journal of Engineering and Science (IJES)* Vol. 3, pp. 12-17, ISSN (e): 2319 – 1813 ISSN (p): 2319 – 1805
- Emmanuel, B. (2013). *The place of Nigeria Certificate in Education Chemistry Teachers in UBE Basic Science programme. STAN 54th Annual conference proceedings. 177-181.*
- Erwin, A.,& Mohamed, Najib, A, (2011). The Validity and Reliability of Assessment for Learning (AfL). *Education Journal*. Vol. 4, No. 2, pp. 64-68. doi: 10.11648/j.edu.20150402.13
- 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.
- Federal Government of Nigeria (2004). National Policy on Education. *Lagos: Nigeria Educational Research and Development Council Press*.
- Frank, Q., Isaac, A., & Francis, A., (2019). Teachers' Test Construction Skills in Senior High Schools in Ghana: Document Analysis. DOI: 10.21449/ijate.481164

- Hamed, T., (2015). Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research. *International Journal of Academic Research in Management (IJARM)*. Vol. 5, No. 3, 2016, Page: 28-36, ISSN: 2296-1747
- Ibrahim Muhammad Shamsuddin, Adamu Tairu Arome, Ibrahim Aminu, Ismaila Itopah Isah, Abubakar Mahmud Adamu. Solving the Problems of Chemistry Education in Nigeria: A Panacea for National Development. *American Journal of Heterocyclic Chemistry*. Vol. 3, No. 4, 2017, pp. 42-46. DOI: 10.11648/j.ajhc.20170304.12
- Ige, T.A., 2000. The school science curriculum as an effective agent for training in environmental management. *University of Ado-Ekiti, Nigeria. J. Edu. Foundations Manage., 1: 190-191.*
- Iheanyi O. I., (2015). Secondary Education Chemistry Curriculum Implementation in Nigeria: Contending Issues and Innovative Approaches for the Future. *IJSAR Journal of Life and Applied Sciences (IJSAR-JLAS) ISSN: 2408-7610 Vol. 3, Issue 1, pp. 24-33.*
- Ihuarulam, A. I., (2008). Chemistry teachers understanding and implementation of the senior secondary chemistry curriculum. *Nigerian journal of Science and Educational Research 4(1) 151 – 158*
- Instructional Assessment Resources, (2011): *Item analysis. Instructional Assessment Resources*. Retrieved from <http://www.utexas.edu/academic/ctl/assessment/iar/students/report/itemanalysis.php>
- Izard, J., (2015). Overview of Test Construction. *UNESCO International Institute for Educational Planning*.
- Joshua, M. T. (2005) *Fundamentals of test and Measurement in Education*. Calabar: University of Calabar Press
- Kamin, L. (1974). The science and politics of IQ. *Potomac, MD: Erlbaum*.
- Linn, R. L., & Gronlund, N.E., (2000). *Measurement and assessment in teaching*. 8th ed. New Jersey: Prentice Hall.
- Lurdes, B., José, A., Cristina, T., & Ana, P., (2017). *Moodle and Multiple-Choice Tests*. Institute of Accounting and Administration of Oporto – ISCAP Polytechnic Institute of Oporto – IPP Oporto – Portugal.
- Marjorie, Arandela., (2014). *Educational Testing and Evaluation*. India, Kerela: Research Signpost Press.
- Mcleod, S., (2019). *What is Reliability?* Retrieved from <https://www.simplypsychology.org/reliability.html>
- Muhammad, B. B., (2013). Qualitative Chemistry Education: The Role of the Teacher. *IOSR Journal of Applied Chemistry (IOSR-JAC) e-ISSN: 2278-5736. Vol. 4, Issue 5, pp. 10-14.*

- Musa, H. (2005). *Obstacles to Qualitative Science Education for National Advancement in Colleges of Education and Strategies for Improvement*. A paper presented at the 4th national conference of federal college of education, Katsina.
- National Open University, (2006). *Measurement and Evaluation*. Lagos, Nigeria: The Regent Press.
- NTI (National Teachers' Institute) (2006). Manual for re-training of primary school teachers: School-based assessment. *Kaduna: National Teachers Institute*.
- Odukoya, J. A., Adekeye, O., Igbino, A. O., & Afolabi A., (2017). Item analysis of university-wide multiple choice objective examinations: the experience of a Nigerian private university. *Qual Quant*; DOI 10.1007/s11135-017-0499-2
- Ogomaka, P. M., & Caroline, I. O., (2017). Criterion – Referenced Measurement for Educational Evaluation and Selection. *Nigerian journal of Educational Research and Evaluation, a publication of The Association of Educational Researchers and Evaluators of Nigeria (ASSEREN), Vol. 16, Page: 160-168, ISSN 0795-3607*
- Ogunleye, A.O. (1999): *Science Education in Nigeria*. Lagos: Sunshine Publishers.
- Okieimen, F. E., (2007): The role of chemistry in national development. *ICCON News Nov. 2007. Edition Vol. 2, pp 7*.
- Okoro S. U. C., (2013). *Attaining the MDGs through Effective STAN Education Delivery. STAN 54th Annual Conference proceedings. 108-118*.
- Okpala, E. O., (2002). Effective Implementation of the Continuous assessment policy in primary schools: A keynote presented at the train-the-trainers workshop on Continuous Assessment: Lagos.
- Ololube, N. P., (2008). Evaluation Competencies of Professional and Non Professional Teachers in Nigeria, *Studies in Educational Evaluation (SEE)*, 34(1), 44-51.
- Oluwatayo, J. 2012. Validity and reliability issues in educational research. *Journal of Educational and Social Research* 2, 391-400.
- Omoifo, C. N., (2012). Dance of the Limits, Reversing the Trends in Science Education in Nigeria, *Inaugural Lecture University of Benin, Benin City*.
- Omorogbe, E., & Ewansiha, J. C., (2013). The Challenge of Effective Science Teaching in Nigerian Secondary Schools. *Academic Journal of Interdisciplinary Studies (AJIS) e-ISSN: 2281-4612. Vol.2, No. 7, Doi:10.5901/ajis.2013. v2n7p181*
- Ovat, S., & Ofem, U. S. (2017). Teachers Variables and Application of Test Blue Prints in Learners Assessment in Secondary Schools in Cross River State. *International Journal*

of *Scientific Research in Education*, 10(1), 112-118. Retrieved from <http://www.ijrsre.com>

Paulson TU (2003). *The study of teaching*. New York: Holt, Rinehart and Winston Inc.

Priyansh, T. (2019). *Intelligence Test in school*. Retrieved from www.yourarticlelibrary.com/education/intelligence-test-in-school-education/90042

Rhoda E. Camble Abubakar Hamman-Tukur, (2017). Teachers' Characteristics and Knowledge of Test Construction in Day Secondary Schools in Maiduguri Metropolis, Borno State, Nigeria. *Research Journal of Education* Vol. 5 No. 4, ISSN 2347-8225

Ridwan, M. O., (2017). *Educational Evaluation and Testing*. African Virtual University Press

Rudner, L. M., & Schafer, W. D. (2002). What teachers need to know about assessment. *Washington, DC: National Education Association*.

Rufina, C. R., (2016). Using Reliability Measures in Test Validation. *European Scientific Journal* June 2015 edition vol.11, No.18 ISSN: 1857 – 7881 (Print) e – ISSN 1857- 7431

Rufina, S. H., Abubakar, H., Stephen S. H., (2015). Assessing Teacher Competence in Test Construction and Content Validity of Teacher Made Examination Questions in Commerce in Borno State, Nigeria. *Education* 2015, 5(5): 123-128 DOI: 10.5923/j.edu.20150505.01

Silker RT (2003). *Teachers and tests*. London: Basil Blackwell Press.

Siri, A., & Freddano, M., (2011). *The use of item analysis for the improvement of objective examinations*. Paper presented at International Conference on Education and Educational Psychology (ICEEPSY 2011), 29 (2011) 188 – 197, DOI:10.1016/j.sbspro.2011.11.224

Stephanie, A. (2015). *Face Validity: Definition and Example*. Retrieved from <https://www.statisticshowto.datasciencecentral.com/face-validity/>

Stephanie, A. (2015). *Criterion Validity: Definition and Example*. Retrieved from <https://www.statisticshowto.datasciencecentral.com/criterion-validity/>

Ugodulunwa, C. A., & Wakjissa, G. S., (2016). What Teachers Know About Validity of Classroom Tests: Evidence from a University in Nigeria. *IOSR Journal of Research & Method in Education (IOSR-JRME)* e-ISSN: 2320–7388, p-ISSN: 2320–737X Volume 6, Issue 3 Ver. I (May. - Jun. 2016), PP 14-19

Uwague, A. & Ojebah, C. K. (2008). The sorry state of chemistry –The nucleus of science and technology in Nigeria polytechnic education today. *Ozoro journal of general studies*. Vol. 1, pp 363-370.

APPENDIX

FEDERAL UNIVERSITY OF TECHNOLOGY, MINNA

DEPARTMENT OF SCIENCE EDUCATION

QUESTIONNAIRE FOR ASSESSING OF SECONDARY SCHOOL CHEMISTRY TEACHERS' SKILL OF CONSTRUCTING MULTIPLE CHOICE TEST IN BOSSO LOCAL GOVERNMENT OF NIGER STATE. (QSCMCT)

Dear Respondent,

I am a student in the School of Science and Technology Education, Federal University of Technology, Minna. This research instrument is constructed to collect information by the means of assessing the secondary school chemistry teachers' skill of constructing multiple choice test.

Kindly tick the column that best describe your perception on each item with sincerity and faithfulness. Your response will be treated strictly and confidential for the purpose of this research, this is not a test and your name is not required.

Adegboye, Olawale Abdulmalik

INSTRUCTION: use the following 4 rating scale keys to represent your perception on each item by ticking the options that best suits your level of agreement provided in the column. The keys are:

Strongly Agree - SA

Agree - A

Disagree - D

Strongly Disagree - SD

Section A:

Name of School:

Teacher's Gender:

Male Female

Section B: Chemistry Teachers' Skill of Constructing Multiple Choice Test.

S/No	Items	SA	A	D	SD
1.	A table of specifications relates test item to instructional objectives.				
2.	The Table of specification is a two-way table of behaviors (objectives) and content.				
3.	A multiple choice item can be reliable without being valid.				
4.	Inadequate planning of test item usually leads to the test being poorly conceived, poorly worded, ambiguous and grammatically incorrect.				
5.	To be able to write good test items the teacher must have a thorough mastery of the subject matter.				
6.	Multiple choice tests are tests consisting of questions with a series of alternative answers.				
7.	Test items are subjected to item analysis				
8.	The best method for assessing students is through the use of multiple choice test.				
9.	A list of options (alternatives) contains one correct option (the answer) and a number of incorrect distractors				
10.	Multiple choice test consist of two components; a stem and alternatives.				
11.	Test blueprints is also known as Table of specification.				
12.	Instructional objectives should be critically reviewed when constructing a test item.				
13.	Test reliability refers to the degree of accuracy with which a test measures what it is designed to measure.				
14.	The final task in multiple choice test construction procedure is the item writing.				
15.	Item analysis involve the uses of statistical tools to provide pertinent information to improve the accuracy and quality of a test item.				
16.	Without the table of specification there is a tendency to exaggerate the importance of one test item to the neglect of others.				
17.	Face validity is the weakest and simplest form of validity.				
18.	A well planned test blueprint enhance content validity.				
19.	Test blueprints enables test developer achieve a desired content coverage of test item to assure validity and reliability.				
20.	The Bloom's taxonomy of educational objectives is a useful guide in classifying the instructional objectives when constructing test item.				
21.	Determining the instructional objectives is the first measure to be considered in multiple choice test construction procedure.				