



ASSESSING JUNIOR SECONDARY SCHOOL TEACHER'S COMPETENCY ON ITEM CONSTRUCTION AT BAUCHI STATE CENTRAL SENATORIAL ZONE

BABAYO MUHAMMAD OTHMAN

Department of psychometrics, AD Rufa'i College of Education, Legal & General Studies Misau, Bauchi State

ABSTRACT

The study examined junior secondary school teacher's competency on items construction at Bauchi state central senatorial zone, using survey design. Participants were 155 teachers (Male=123, Female=32) selected using stratified random sampling technique from 25 out of 124 public secondary schools within the Zone. Research instrument was a self-constructed questionnaire, 'Teacher's Competency Scale (TCS)'; the instrument consisted of 2 sections A and B. Section A solicited for Demographic information of respondents while section B posed questions related to item construction at various degree. The response format was on 4 point Likert scale weighted as, strongly agree (SA=4pts) Agree (A=3pts), Strongly Disagree (SD=2pts) and Disagree (DA=1pts). A measure of internal consistency using Cronbach alpha was determined and reliability coefficient of 0.84 was obtained which showed high degree of reliability of the instrument. Data was analyzed using frequency count, simple percentages and chi-square cross tabulation. Results showed that majority of the teachers that write items in negative form didn't capitalized or underline the items, similarly, the finding also revealed that part of the respondent didn't arrange items in order of difficulty while constructing test item. Qualifications of Junior Secondary Teachers do not significantly influence their Competency in Item construction. The outcome also indicated that gender difference was statistically not significant; hence possession of technical skill on item construction is uniform among the participant. The researcher recommends that Teacher should consider

arranging item in order of difficulty, especially from simple to complex. Doing this will motivate students in developing confidence while answering questions. An in-house training or workshop on item construction should be conducted, especially to those that lack education qualification as well as those having lowest teaching qualification.

Keywords: *teacher's competency, item construction, Junior Secondary schools*

INTRODUCTION

The aim of JSS level of education according to the National policy on Education (FRN 2004) is to prepare the children for useful living within the society and education. This level is to raise a generation of people who can think for themselves, respect the dignity of labour and provide literacy and factual knowledge, these goals of FRN (2004) can be achieved largely based on the competency of the teacher to tailor the learning activities to the learner. The teacher need to be knowledgeable, versatile, and should be able to employ variety of techniques in carrying out teaching and assessment of the students. The issue of assessment by teachers has been a thing of concern in the area of measurements using a valid and reliable instrument. This issue has engaged the attention of researchers, according to Agu, Onyekuba and Anyichie (2013). Recent research indicated that most of these classroom-based tests in Nigeria lack validity and reliability, because teachers seem to lack test construction skills and thus cannot construct good achievement tests.

Lack of test construction skills by teachers might result in false assessment of students' achievements. Some researchers (Esomonu, 2002; Paulson, 2003) see this incompetency in test construction by teachers as a major cause of malpractice in school examinations by both teachers and students in Nigerian secondary schools.

Most tests used for continuous assessments and end of term examinations in the secondary schools contain ambiguous and mis-leading questions which may be the reason why some of the students fail these tests (Adeola

and Fajonyimi, 1999). The root cause of this was that most teachers lack competencies in test construction (Othman, 2022), leading them to using poorly constructed tests in measuring students' achievements in various school subjects. When students' achievement levels are not properly measured and interpreted, the teachers and school administrators will not be able to provide educational opportunities and support that will fit each individual student needs. (Agu, et al, 2013)

Nenty and Fetogang (2014) found out that teachers use past examinations to prepare classroom and school test. This is an indication that the teachers do not have understanding of how to develop a reliable and valid test. Nenty and Lusweti (2014) are of the opinion that teachers should be assessment literate i.e. having good and practical knowledge of testing procedure for valid result.

Quality classroom-based assessment means adherence to standard procedures for test construction and analysis. Every classroom teacher is expected to possess and apply requisite competences in construction good items, for class assessments. A good test item must be both valid and reliable. A test is valid if it is suitable for the intended purpose. On the other hand, a test is reliable if it measures what it is supposed to measure consistently under all conditions (Sidhu, 2012). Teachers today, perhaps more than ever before, have a need to be knowledgeable consumers of test information, constructors of assessments and protocols, and even teachers about testing (Kolawole,2006)

In the same vein, competency on the part of the teacher demands ability to effectively develop and analysed test items. It has been observed by Ohuche and Akeju in Nworgu (2003) that the qualities of the items that make up a test determine the quality of the test as a whole. Assessment of these essential qualities of the items in a test constitutes item analysis. Anastasi and Urbina (2003) identified item analysis as qualitative and quantitative situation whereby the content and form of a test are addressed. Quantitative item analysis is concerned with the statistical properties of the items. For the effective assessment, this is the bedrock of effective teaching, demands that teachers should be very familiar with the item development.

LITERATURE REVIEW

Assessing students learning is something that every teacher has to do, usually quite frequently too. This is why a typical teacher can spend more than one third of his class time engaged in one form or another type of measurement activity (Stiggins, 1994). However, despite the amount of time teachers spend assessing students learning, it is a task that most of them dislike and that only a few do well. Research indicated that if experience teachers will guide the new intake (in-experience) the problems will be resolve (Othman, 2022) It is also believed that teachers with more training in assessment use more appropriate assessment practices than do teachers with less training (Bielher & Snowman, 1997)

Assessment in education is best described as activities to determine the importance size, or value of teaching and learning (Ojerinde 2014) while Anikweze, (2013) opines that assessment to include teacher, peer and self-assessment process. Stiggins and Chappins in Nenty and Lusweti (2014) emphasized on teachers having deep knowledge of why they assess and hence what decisions are to be made, who will make the decision, and what kind of information will be helpful. Teachers should find ways through which they can enhance the learning aspects as well as how to assess the quality of learning and teaching exhibited.

Classroom teachers must have a good mastery of achievement targets that their students are to achieve and such targets must be completely and clearly defined. It must be mastered by the teachers for him to be able to communicate them clearly to the students. This called competency in teaching. As noted by Akuto, Aduloju, Odeh (2012) competency in teaching involves a number of principles which include among others, aligning the three major components of instructions namely learning objectives, assessments and instructional activities. In other to accomplish these tasks, competency on the part of the teachers in the area of item construction and utilization of teaching materials becomes critical.

Test construction has been found to be a major source of anxiety among many teachers in Nigerian schools, especially, less experienced ones (Ebinye, 2001). This anxiety stems majorly from lack of test construction skill by these teachers. This is surprising as the least qualified teacher in

Nigerian secondary schools today have the Nigerian Certificate in Education (NCE) qualification. This NCE qualification is obtained after three years of study in a College of Education where teachers are trained in different education courses, including educational testing and assessment. A number of possible reasons could be deduced for this deficiency. It may be either that the teachers' pre-service training did not prepare them adequately for test construction due to little emphasis on assessment during their professional development (Stiggins, 2000) or that most of the teachers failed to acquire test construction skills needed for quality test item generations while in training (Othman, 2022). This implies the need to assess the test construction skills of teachers in the secondary schools. Surprisingly, there seems to be shortfall of instruments developed for such assessment in Nigeria (Ujah, 2001).

Item construction is the ability of a teacher to write items that cut across all the content specified. While item analysis lies at the heart of evaluating the quality of tests developed. Item difficulty and discrimination are two statistics in an item analysis according to Ariyo and Lemut (2015). Item difficulty is the mean item score and item discrimination is the correlation between the item score and test scores. The authors are of the opinion that the statistics allow instructors to identify problematic items such as those that are too easy or too difficult for students and item that are unrelated to the overall score. This process will help eliminating poorly functioning items. This can be done by having the knowledge of measurement and evaluation. This cut across all the domains in educational objectives. Test construction has been found to be a major source of anxiety among many teachers in Nigerian schools, especially, less experienced ones (Ebinye cited in Agu et al, 2013). This is why they attributed anxiety majorly to lack of test and measurement knowledge by these teachers. This is surprising as the least qualified teacher in Nigerian secondary schools today have the Nigerian Certificate in Education (NCE) qualification.

In item construction, certain procedures are followed as highlighted by Enyi (2009) which include; Determine the purpose of testing, Building a table of specification, Selecting appropriate item types, Preparing relevant test items, Assembling the test, Administering the test, Appraising the test,

and Using the results. To effectively do these, it calls for competency on the part of the teacher. Test blue print is one and most important aspect of this development. It takes competency on the part of the teacher to employ the use of a table of specification. There is a need for teachers to acquire adequate knowledge of test development and continuous assessment practices to enable them make necessary improvement in teaching and evaluation, this opinion was stressed by Okpala and Onocha (1985), Ipaye (1987) and Nwosu in Aduloju (2010).

Most achievement tests used in Nigerian secondary schools for continuous assessments are teacher-constructed tests (Teacher-made test). Teachers, therefore, need to apply some acceptable degree of test construction skills in order to be able to develop valid and reliable tests that will yield accurate feedback of students' achievement if we want to achieve the goals of FRN (2004).

The hindrances to the teacher made-test from close observation and analysis is the construction of a blue print. Kolawole (2006) argued that, in an ideal situation, a teacher who constructs items without a blue print is just like a brick layer who is building a house without any plan. Teachers especially science teachers have been found wanting in the area of basic contents determination and this call for concern if the nation is to move forward in science education

Qualitative item analysis involves consideration of content validity and the evaluation of items in terms of effective item-writing procedures (Nworgu, 2003). Competency of the teachers in doing this, will serve to improve the test by revising or discarding ineffective items as well as providing diagnostic information in what examinees know and what they do not know. Strengthening this view, Anastasi in Nworgu (2003) opined that it involves essentially the "systematic examination of the test content to determine whether it covers a representative sample of the behaviour domain to be measured. What is implied here is that, test items are analyzed to ascertain the adequate reflection of the major aspects of behavior domain under consideration in relation to content and intellectual skills. This view is shared by Nworgu (2003,) Enyi (2009), and Akem (2011). A number of studies have suggested that faulty test items

affect students' comprehension and ability to provide accurate answers to the items, the inference drawn about what a student knows and understands may be compromised (Koksal, 2004; Leighton and Gokiart, 2005). Similarly, Agu et al (2013) observed that classroom teachers generally write poor items and that a typical classroom teacher in a secondary school cannot construct good multiple choice test items. On this note, Akem (2008) submitted that the teachers are not well equipped with the skills to handle evaluation of all abilities of students due to the late introduction of the course to the students. The author argued that the duration of the course which is only one semester is full of other activities like strike. Since decision is made based on the result of testing and the instrument used for this purpose, it is important that the teachers who developed this instrument should be assessed whether they are competent in the item construction and analysis. This study therefore sought to assess the teacher's competence in this area.

STATEMENT OF PROBLEM

The development of good items plays a very important role in the selection, placement and measurement of student's achievement in school and classes. The greatest challenge faced by teachers has been the competency in development of reliable and valid items according to Ariyo and Lemut (2015). Students who perform well at various levels of teachers' classroom-based tests are expected to equally perform well in the standardized tests like Junior Secondary School Certificate Examination (JSSCE), West African Examination Council (WAEC) and National Examination Council (NECO) This has been a different issues entirely and a great concern because the same student that failed this standard examination have been doing well in the internal examination (Agu et al, 2013).. It would seem, therefore, that some of these teacher-made achievement tests are invalid and unreliable and thus, fail to provide accurate assessment of students' knowledge and understanding of the various subject areas.

Quality in test development is critical factor in Test, Measurement and Evaluation. The stage in writing a good test are many and varied which

include items construction, items moderation, item trial testing, items analysis, item calibration, construction of equivalent items and items banking (Alonge 2000). For the purpose of this study, only item construction and analysis were considered. Ariyo and Lemut (2015) in their study revealed that teachers (test developers) response to training needs assessment confirms their need in the area of test construction and modern test theory. This shows that teachers are not competent in developing a good item for student's evaluation and when these students eventually presented for external examination which the items pass through the stages in test development, their performance become poor. Researchers (Kolawole 2006, Njoku 2006, Ugwu 2008) argued that teacher's incompetence in item construction, analysis and evaluation of students at each stage of learning contribute to the poor performance of students, there is need for teachers in the area of item construction and analysis to be competent.

One of the most used tools which are to assess students' achievement in classroom is a test instruments, these tools highly fall into select (objectives) or supply (essay) item format (Susana, 2014), both of them are often fraught with reliability and validity problems as the process for constructing such tests is often not followed or misunderstood, and this leads to significant measurement error into the measurement process. Inaccurate data-based inference as a result of poor measurements will lead to bad decision making, with this there is a need of competency from the side of lectures on test item development. It has already been pointed out that poorly designed tests could make the students loose interest in a particular subject (Osadebe, 2001). When the test instrument does not possess the necessary characteristics it ought to, this means that the test is not valid and reliable, and the effort to achieve the instructional objectives will be in futility. This perhaps attract the researcher's interest in conducting the research on Assessing Junior Secondary School Teacher's Competency in item Construction at Bauchi State Central Senatorial Zone

OBJECTIVES

The objective of the study is thus,

To find out the level of Teacher's Competency in item Construction

RESEARCH QUESTIONS

The study aims at finding answers to the following questions:

What are the common errors Teachers' made while Construction items

2- Will gender difference in the students' evaluation of teaching quality be significant?

Research Hypotheses

The following hypotheses were tested at 0.05 level of significance:

HO₁- Qualification of Junior Secondary Teachers do not significantly influence their

Competency in Item construction

HO₂- There will be no significant gender difference on Competency in Item construction

Method

The study adopted a survey research design on reconnoitering the existing problem of Teacher's Competency in item Construction in Bauchi State. The target population encompasses of only 6 local government areas in Bauchi Central Senatorial Zone, the study is restricted to public junior secondary school in the zone,

The researcher developed an instrument tittle Teachers Competency scale (TCS) which was used for the collection of data, the instrument consisted of 2 sections A and B. Section A solicited for Demographic information on respondents while section B posed questions related to item construction at various degree. The response format was on 4 point Likert scale weighted as, strongly agree (SA=4pts) Agree (A=3pts), Strongly Disagree (SD=2pts) and Disagree (DA=1pts). A measure of internal consistency using Cronbach alpha was determined and reliability coefficient of 0.835 was obtained which showed high degree of reliability of the instrument.

RESULTS

Table one: Demographic information

Variables	Categories	Frequencies	%
Gender	Male	133	85.8
	Female	022	14.2
	Total	155	100.0
Teaching	1-9 years	18	11.6

Experience	10-18 years	50	32.3
	Above 20 years	87	56.1
	Total	155	100.0
Qualification	Degree with education	58	37.4
	Degree without Educ.	51	32.9
	NCE	40	25.8
	Other	06	03.9
	Total	155	100

Table 1.0 shows summary of the Demographic information of the sample variable, from the table 85,8% of them are Males, while 14,2% are female, their teaching experience ranges from 18 Teachers (11,6%) are having less working experience ranging from 1-9 years, 32,3% which made up to 50 teaching staff have experience ranging from 10-18 years. Similarly, 87 academic staff (56,1%) has experience above 20 years. On qualification, 37.4% of the entire teaching staff possesses Degree with education, while 51 among them possesses Degree without education making up to 32.9% similarly, 25.8% of the population are having NCE Certificate made up of 40 teaching staff and lastly, 6 staff among them making 04% possess other qualification

Research Question one: What are the common errors Teachers' made while constructing test items

Table 2: Summary of Descriptive Statistics on the common errors Teachers made while constructing test items?

S/ N	Statements	OPTION			OPTION			OPTION			OPTION		
		OPTION	FREQUENCY	PERCENTAGE	OPTION	FREQUENCY	PERCENTAGE	OPTION	FREQUENCY	PERCENTAGE	OPTION	FREQUENCY	PERCENTAGE
	Writing of the stem	SA			A			D			SD		
1	I am familiar with the characterist		28	70.0		8	20.0		4	10.0		00	00

	ics of test items												
2	I write more of the items in the positive term		9	22.5		11	27.5		14	35.0		6	15.0
3	I underline/capitalized the negative items		6	15.0		5	12.5		18	45.0		11	27.5
4	I do review and editing of the items		28	70.0		10	25.0		2	5.0		00	00
5	I arrange the items in order of difficulty		19	47.5		4	10.0		17	42.5		00	00
6	I write each item to measure a single objective		14	35.0		10	25.0		13	32.5		3	7.5
7	I ensure that the stem of the item clearly formulate a problem		23	57.5		17	42.5		00	00		00	00

The table 2 above depicts respondents view on common errors that need proper consideration while constructing test items. Understanding the characteristic of item is the foundation, on this regard, 28 out of 40

respondents making up to 70% are familiar, in that only 30% disagree with the statements. A 50-50 respondent opinion was received on writing more items in positive than negative term. 83.5% of those that write items in negative form didn't capitalize or underline the items. 95% of the respondents agree that they do review and edit item before taking decision. 40% of the respondent didn't set their item to measure single objective, whereas 60% totally disagree with them. Lastly, 100% of them agree that the stem of their item clearly formulate a problem

Research Question Two: Will gender difference in competency on item construction be significant?

Table 3: Summary of Descriptive Statistics and analysis of chi-square cross tabulation on competency on item construction among gender?

Statements	Agree		Disagree		Chi-square (cross tabulation)	Decision
	Male	Female	Male	Female		
I ensure that the stem of the item clearly formulate a problem	70	14	53	08	.483	Accepted, $p > 0.05$
I do review and editing of the items	105	12	28	10	.505	Accepted, $p > 0.05$
I write each item to measure single objective	91	16	42	06	.001	Rejected, $p < 0.05$
I am familiar with the characteristics of test items	115	14	18	08	.166	Accepted, $p > 0.05$

Table 2 above, revealed respondents view on ensuring that the stem of the question clearly formulate a problem while constructing items, 40% (14) of female respondents disagree with the statements, whereas 53% (70) male respondents agree with the statements. The p-value obtained using chi-square cross tabulation was 0.483, the result was statistically not significant ($p > 0.05$). This indicated that there is no gender difference

regarding the statements. Regarding the second statements, 12 out of 105 male respondents didn't review and editing of their items before final construction, on the other hand 12 out of 22 female participants do review and editing of their items (agree) with the statements. The p-value of 0.505 was obtained based on chi-square cross tabulation, the result was statistically not significant ($p > 0.05$). This indicated that there is no gender difference regarding the statements.

The third item is on ascertaining whether the respondents write each item to measure single objective,? 91 out of 133 male respondents agree with the statements whereas only 6 out of 22 female disagreed. The chi-square cross tabulation result obtained was 0.001($p < 0.05$) The result was statistically significant between the two categories. Hence, gender disparity do existed on writing each item to measure single objective while constructing test item.

Finally, majority of the participants agreed with the statements "I am familiar with the characteristics of test items|" in which 115/18 out of 133/22 unanimously agree with the statements. The p-value obtained was 0.166 ($p > 0.05$), this indicated that The result was statistically not significant, hence , the said hypothesis was retained

HO₁- Qualification of Junior Secondary Teachers do not significantly influence their Competency in Item construction

Table 4: Summary of Descriptive statistics and Chi- Square Cross Tabulation on Analysis Junior Secondary Teachers Competency on Item construction

Statements	Determinant	Category of Respondent	Frequency of Responses		p-value	Decision
			A	DA		
I possess technical skills required for constructing test items	Qualification	Degree with education	22	15	.015	Rejected
		Degree without Educ	08	21		
		NCE	61	22		

		Others	00	06		
Total			94	61		

Table 3 above, describes the descriptive statistics on the frequency count of respondents. The summary for the frequency count of the respondents indicated that 94 out 155 (61%) agree with the statements that they possess technical skills required for constructing test items whereas 61 out of 155 (39%). disagree. Similarly, the analyses for chi-square (cross tabulation) result indicate that the calculated p-value obtained was of 0.015 with a degree of freedom 2. This detected that the p-value 0.015 is approximately (0.02) less than the *alpha* value of 0.05. Thus, the null hypothesis which states that, Qualification of Junior Secondary Teachers do not significantly influence their Competency in Item construction is rejected. Hence qualifications play a significant role in competency on item construction.

The chart below further illustrates the differences.

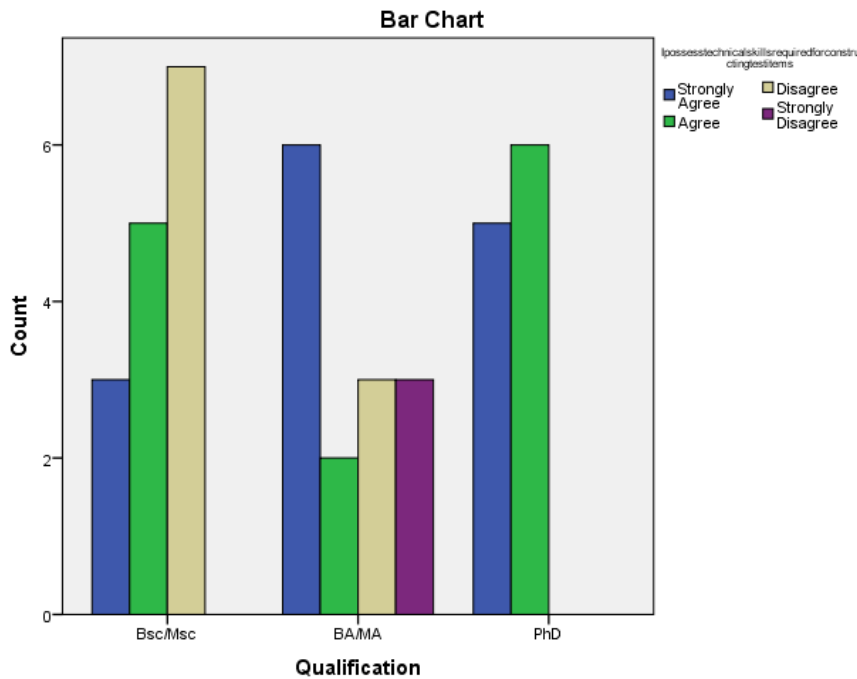


Figure 1: Respondents' talent on items construction in relation to Qualification

HO₂- There will be no significant gender difference on Competency in Item construction

Table 5: summary of descriptive statistics on gender difference on Competency in Item construction

Statements	Agree		Disagree		Chi-square	Decision
	Male	Female	Male	Female		
I ensure that the stem of the item clearly formulate a problem	70	14	53	08	.483	Accepted, p> 0.05
I possess technical skills required for constructing test items	93	16	40	06	.550	Accepted, p> 0.05
I do review and editing of the items	105	12	28	10	.505	Accepted, p> 0.05
I write each item to measure single objective	91	16	42	06	.001	Rejected, p< 0.05
I arrange the variable alphabetically or numerically	97	13	36	09	.011	Rejected, p< 0.05
I always discriminate among options that vary in degree	99	12	34	10	.000	Rejected, p< 0.05
I am familiar with the characteristics of test items	115	14	18	08	.166	Accepted, p> 0.05

Source: result analysis 2022

The table 5 above portrays the respondents view up on considering the statements that says, "I write more of the items in the positive term" while constructing test item. 60% (80) male respondents out of 133 agreed with the statements, whereas 68% (15) female respondents out of 22 disagree. The chi-square cross tabulation result on gender difference between these two groups was statistically significant, the p-value obtained is less than, $p < 0.05$. hence gender difference existed

-On ensuring that the stem of the item clearly formulate a problem while constructing items, 40% of female respondents disagree with the statements, whereas 53% (70) male respondents agree with the statements. The p-value obtained using chi-square cross tabulation was 0.483, the result was statistically not significant ($p > 0.05$). This indicated that there is no gender difference regarding the statements

- I possess technical skills required for constructing test items, is the view of the majority of the respondents. In that 68% (90) of male respondents and 73% of female respondents agreed simultaneously. Regarding the analysis of chi-square cross tabulation, the p-value obtained was 0.550 which greater than the predicted ($p > 0.05$). The result indicated that gender difference was statistically not significant; hence possession of technical skill on item construction is uniform among the participant

-12 out of 105 male respondents didn't review and editing of their items before final construction, on the other hand 12 out of 22 female participants do review and editing of their items (agree) with the statements. The p-value of 0.505 was obtained based on chi-square cross tabulation, the result was statistically not significant ($p > 0.05$). This indicated that there is no gender difference regarding the statements.

-on ascertaining weather the respondents write each item to measure single objective,? 91 out of 133 male respondents agree with the statements whereas only 6 out of 22 female disagreed. The chi-square cross tabulation result obtained was 0.001($p < 0.05$) The result was statistically significant between the two categories. Hence, gender disparity do existed on writing each item to measure single objective while constructing test item.

-99 out of 133 of the male respondents agree that they always discriminate among options that vary in degree while constructing items, whereas 10 from the 22 female disagreed. The result from the analysis for chi-square cross tabulation was 0.005, the result for the two categories was statistically significant considering that ($p < 0.05$). Conclusively the said hypothesis was rejected.

-finally, majority of the participants agreed with the statements "I am familiar with the characteristics of test items|" in which 115/18 out of 133/22 unanimously agree with the statements. The p-value obtained was 0.166 ($p > 0.05$), this indicated that The result was statistically not significant, hence, the said hypothesis was retained

DISCUSSION OF FINDINGS

The focus in this study was on Junior Secondary School Teachers competency on item developments at Bauchi central senatorial zone, as well as finding out the degree of their competency in relation to qualification, as well as gender.

Research question one is on what are the common errors Teachers' made while Construction items, this research finding revealed that. 83.5% of those that write items in negative form didn't capitalized or underline the items. The finding was in accord with that of Marmah and Impraim (2013), similarly, it was observed that 76% of the respondent didn't arrange items in order of difficulty while constructing test item, as confirmed by Zimmaro, D. M. (2004),

The second research question is on determining if gender difference in competency on item construction is significant?, three among the four results obtained indicated that the p-value obtained using chi-square cross tabulation was statistically not significant ($p > 0.05$). This indicated that there is no gender difference regarding the statements, hence the hypothesis was retained. This contradict that of Derri et al (2012), his finding by comparison revealed that female teachers presented higher knowledge of assessment techniques than male teachers.

Table 3 portray the analyses for chi-square (cross tabulation), result indicated that the calculated p-value obtained was of 0.015 with a degree

of freedom 2. This detected that the p-value is approximately (0.02) less than the *alpha* value of 0.05. Thus, **the null hypothesis which states that, Qualification of Junior Secondary Teachers do not significantly influence their Competency in Item construction is rejected. Hence qualifications play a significant role in competency on item construction.** This is in accordance with that of (Othman .B.M, 2018) that investigated the cognitive demand on Teacher-Made mathematics test with respect to qualification, experience and gender.

The second hypothesis is trying ascertain if there will be no significant gender difference on Competency in Item construction, in answering this, six among the seven test conducted revealed that the result was statistically not significant ($p > 0.05$). The result indicated that gender difference was statistically not significant; hence possession of technical skill on item construction is uniform among the participant, therefore the Null hypothesis was retained. This outcome contradicts that of (Othman, 2018).

Conclusion

Based on the preceding discussion, the following conclusions were drawn: Those that write items in negative form didn't capitalized or underline the items, but while constructing item, the stem of their item clearly formulate a problem.

Majority of Teachers respondent didn't arrange items in order of difficulty while constructing test item,

The null hypothesis which states that, Qualification of Junior Secondary Teachers do not significantly influence their Competency in Item construction were rejected. Hence qualifications play a significant role in competency on item construction

The second hypothesis is trying ascertain if there will be no significant gender difference on Competency in Item construction, The result indicated that gender difference was statistically not significant; hence possession of technical skill on item construction is uniform among the participant, therefore the Null hypothesis was retained

RECOMMENDATION

Based on the findings of this study, the following recommendations are put forward for consideration.

Teacher should consider arranging item in order of difficulty, especially from simple to complex. Doing this will motivate the students in developing confidence while answering questions

Before display of final print-out of questions paper, teacher should ensure that they underline or capitalized all item written in negative form.

An in-house training or workshop on item construction should be conducted, especially to those that lack education qualification as well as those having lowest teaching qualification.

The current status queue should be maintained through providing equal opportunity or treatments on training and workshop gender wise

Reference

- Adeola OA, Fajonyomi AA (1999). Adult learners' performance on standardized and non standardized test as a function of sex and location. *Nigerian J. Counsel. Dev.* 4(1):32-36.
- Aduloju, M.O. (2010). Teachers Competency in Evaluating Senior Secondary School Students Cognitive Achievement in Chemistry in North Central, Nigeria. Unpublished Ph.D Thesis, University of Ado-Ekiti.
- Agu, N. N; Onyekuba, C & Anyichie, C.A. (2013). Measuring Teachers Competencies in constructing classroom-based tests in Nigeria Secondary School Need for a test construction Skill inventory: vol.8 (8) pp.431-439. doi: 10.5897/ERR 12.219
- Akem, J.A. (2011). *Research and Statistics for Higher Education*. Makurdi: Destiny Ventures.
- Anastasi, A. & Urbina, S. (2004). *Psychological testing, (7ed.)*
- Anikweze, C.M (2013). *Assessment Practical's and Challenges of Innovation in Nigeria Education*. Being a paper Presented at an institutional seminar organized by Joint Admission and Matriculation Board (JAMB), Abuja on 3rd September, 2013 at the Jamb auditorium
- Ariyo, A.O. & Lemut, T.I (2015). Ensuring Quality in the Test Development process through Innovations in Item Calibration: A comparison of Classical Test Theory and Item Response Theory Eras in Jamb, *Nigeria. Journal of Education Assessment in Africa, 10, 50-68*
- Ebinye P .O (2001). Problems of testing under the continuous assessment programme. *J. Qual. Educ.* 4(1):12-19.
- Enyi, S.G. (2009). *Introduction to Measurement and Evaluation in Universities, Polytechnics and Colleges of Education. A Professionals Approach*. Enugu: Linco Press Ltd.
- Esomonu N.P (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 Republic of Nigeria (2004). *National Policy on Education (4th ed)*, Federal Ministry of Information, Abuja: Printing Division, Abuja
- Kolawole, E.B (2006). Evaluation of the Perception of the Cognitive levels of mathematics test items among the prospective and practicing mathematics teachers in Nigeria. *Journal of Research in Education 3 (1), 30-35*

- Marmah, A.A and Impraim, K.A (2013) University Lecturer's Competence In the Construction Of Multiple Choice Test Items. The International Journal Of Humanities & Social Studies Vol 1 Issue 4,
- Nenty, H. J. & Fetogang, E.B (2014) Influence of Syllabus and Past Examination on Assessment Behaviour of Botswana Secondary school Teachers. *African journal of theory and practice of Educational Assessment (AJTPEA) 1-18*
- Njoku, Z. N. (2006). Level of Chemistry Practical Skills Acquired by senior Secondary school Students. *Nigeria journal of professional Teachers*
- NTI (National Teachers' Institute) (2006). Manual for re-training of primary school teachers: School-based assessment. Kaduna: National Teachers Institute.
- Nworgu, B.G. (2003). *Educational Measurement and Evaluation. Theory and practice*. Nsukka: University Trust Publishers.
- Ojerinde, D. (2014). New approaches on university placement Examinations: Nigerian Experience. Being a paper presented at the Abant meeting 2014. Turkey.
- Othman B.M (2022). Investigating college lecturer's competency in the construction of multiple choice items (mci) Being a paper presented at the international conference on Education, Bayero University, Kano.
- Othman B.M (2018) Investigating the cognitive demand in Teacher-Made Mathematics test in Misau local government , Bauchi state. Unpublished theses, Bayero university Kano, Nigeria
- Sidhu, k.S (2012), New Approaches to Measurement and Evaluation. Sterling Publishers, New Delhi
- Ujah EU (2001). Development and validation of an introductory technology achievement test. Unpublished M.Ed. thesis, University of Nigeria, Nsukka.
- Urbina, S. (2004). *Essentials of Psychological Testing*. Hoboken, New Jersey: John Wiley and Sons, Inc.
- Zimmaro, D. M. (2004), Writing Good Multiple-Choice Exams. Measurement and Evaluation Center. University of Texas at Austin.

Appendix

Reliability Statistics	
Cronbach's Alpha	N of Items
.084	31

Chi-Square Tests						
I ensure that a stem	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	.828 ^a	1	.363	.483	.252	
Continuity Correction ^b	.457	1	.499			
Likelihood Ratio	.849	1	.357	.483	.252	
Fisher's Exact Test				.483	.252	
Linear-by-Linear Association	.823 ^c	1	.364	.483	.252	.127
N of Valid Cases	155					

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 8.94.
b. Computed only for a 2x2 table
c. The standardized statistic is -.907.

Chi-Square Tests						
I write more item in positive term	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	19.560 ^a	3	.000	.000		
Likelihood Ratio	25.536	3	.000	.000		
Fisher's Exact Test	22.886			.000		
Linear-by-Linear Association	.910 ^b	1	.340	.354	.202	.059
N of Valid Cases	155					
a. 2 cells (25.0%) have expected count less than 5. The minimum expected count is 2.98.						
b. The standardized statistic is -.954.						

Chi-Square Tests						
I posses technical skill	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	2.226 ^a	3	.527	.550		
Likelihood Ratio	3.757	3	.289	.376		
Fisher's Exact Test	1.859			.644		
Linear-by-Linear Association	.315 ^b	1	.575	.625	.335	.085
N of Valid Cases	155					
a. 1 cells (12.5%) have expected count less than 5. The minimum expected count is 1.56.						
b. The standardized statistic is -.561.						

Chi-Square Tests						
I do review	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	1.581 ^a	2	.454	.505		
Likelihood Ratio	2.846	2	.241	.355		
Fisher's Exact Test	1.044			.623		
Linear-by-Linear Association	.749 ^b	1	.387	.446	.259	.115
N of Valid Cases	155					
a. 1 cells (16.7%) have expected count less than 5. The minimum expected count is 1.28.						
b. The standardized statistic is -.865.						

Chi-Square Tests						
I write each item to measure single objective	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	18.177 ^a	3	.000	.001		
Likelihood Ratio	23.459	3	.000	.000		
Fisher's Exact Test	18.409			.000		
Linear-by-Linear Association	1.494 ^b	1	.222	.238	.136	.045
N of Valid Cases	155					
a. 1 cells (12.5%) have expected count less than 5. The minimum expected count is 1.56.						
b. The standardized statistic is 1.222.						

Chi-Square Tests						
I always discriminate	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	20.454 ^a	2	.000	.000		
Likelihood Ratio	26.979	2	.000	.000		
Fisher's Exact Test	23.803			.000		
Linear-by-Linear Association	4.446 ^b	1	.035	.039	.024	.013
N of Valid Cases	155					
a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 6.25.						
b. The standardized statistic is -2.108.						

Chi-Square Tests						
I am familiar with the characteristic of item	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	3.726 ^a	2	.155	.166		
Likelihood Ratio	6.200	2	.045	.061		
Fisher's Exact Test	3.909			.153		
Linear-by-Linear Association	1.357 ^b	1	.244	.256	.160	.073
N of Valid Cases	155					
a. 2 cells (33.3%) have expected count less than 5. The minimum expected count is 2.55.						
b. The standardized statistic is -1.165.						