Development And Validation Of Mathematics Test (MT) For Senior Secondary Schools

Otuoku Karl

Psychology, Guidance and Counseling, Institution: Isaac Jasper Boro College of Education, Sagbama. Bayelsa State Nigeria

DOI: 10.29322/IJSRP.12.07.2022.p12741
http://dx.doi.org/10.29322/IJSRP.12.07.2022.p12741

Paper Received Date: 21st June 2022
Paper Acceptance Date: 6th July 2022
Paper Publication Date: 12th July 2022

Abstract- The purpose of this study was to construct and validate a mathematics test. Five research questions were employed in conducting the study. Test items were constructed and sent to two content specialists who jointly rated the relevance of the MT items. Based on the recommendation of these resource persons, nine to ten items were randomly sampled on each domain specifications and built into the form 1 MT. In a situation where items assessed two or more behaviors, other items constructed to measure the behavior were eliminated. This practice helped to cut down form 1 MT from 100 items to 54. Form 1 MT was face-validated in terms of clarity of words and plausibility of distracter by specialist in educational measurement and evaluation, and mathematics teachers. This MT was also tried tested and the difficulty and distracter indices of the items determined. The form II MT (i.e. 54 items) was administered to 324 students and their response were used in determine the proportions of students who showed mastery. The findings of this study show that the MT II is a valid and reliable instrument. The major implication of these findings is that MT should be used in both formative and summative evaluation of student's performance in mathematics tests.

Index Terms- Development, Mathematics, Schools, Test, Validation

I. INTRODUCTION

Mathematics is believed to be the most exact, useful and important branch of human knowledge (Jayanthi, 2014). This is because in our homes, farm, workplaces, industry, in war, peace time, in our leisure etc, we make use of mathematics. No meaningful thing can be done without some mental or physical competition. Due to all-round importance of mathematics in our lives, Opara and Magnus-Arewa (2017) cited the work of Karl Fredrick (1777-1885), a German mathematician who stated that "mathematics is the queen of the science and arithmetic is the Queen of mathematics". Mathematics is a very important component of human activities and survival; it is useful in science and technological activities as well as commerce, economics, and educations (Rafina, Abubarkar & Stephen, 2015). According to Darling (2002) the acquisition and development of mathematical concepts and skills during early years forms the foundation of future mathematics performance and achievement.

A study done by Ololube (2008) affirmed that high challenging and accessible mathematics education in secondary school is a vital foundation for future mathematics learning. The author also added that mathematics help students make sense of their world outside of school and also help them construct solid foundation for success in school. Furthermore, a UNESCO (1990) report shows that teachers efficacy towards mathematics is a powerful predictor of how and whether he or she will act with students. Additionally, teachers who have high self efficacy in mathematics tend to persist in failure situation, take more risk with the curriculum, use new teaching approaches to get better gains in students achievement and have more motivated students (UNESCO, 1990). These qualities of a mathematics teacher give him or her ability to provide students with appropriate mathematics experience and attitude towards mathematics. Studies that have been done to assess teachers efficacy towards mathematics shows that it is affected by number of factors such as teacher training, teaching experience, type of schools and scores in mathematics in college, model and school subjects (Akhter & Bahoo, 2015). Lastly, good working conditions and high motivation of teachers in private schools than public schools make teachers in private schools to be more committed to their work in class.

For instance, in Nigeria, despite the wide importance of mathematics in school curricula and its applicability in many fields, many students have not found their feet in mathematics as a result of their perennial failure in the subject (Ukeje, 2005). The performance in maths at the secondary school has been very lo as compared to other compulsory subjects such as English language as affirmed by Clement & Sarama (2007) who asserts that mathematics is the most dreaded examination paper at O'level. For instance, in many schools in Nigeria, more than half of the students score grade "E" which represents total failure (Rafina et al. 2015). The Ololube (2008) posits that due to these poor results in mathematics; many students have developed a phobia for the subject to the extent that even when test

and examination questions are very simple, student fails**.** The situation is not different in early mathematics education in lower secondary schools which forms the basis for later mathematics learning and achievement. For instance, a survey conducted by Unodiaku (2009) revealed that majority of students in lower schools do not possess the mathematics skills and learning passion required for high grade performance in mathematics.

One probable reason for students dislike in mathematics is the inability of most teachers to expose students to adequate mathematics skills through their method of teaching. For instance, before one can learn how to know how to add numbers with 3 digits, such as 256+128, it is important to know how to add number with just one digit or two such as 1+5 or 24+50. After instructions on these skills, there is need for teachers to determine the level of achievement of their students in these skills. To be able to do this effectively, teachers require valid and reliable test in mathematics skills, but these test (in mathematics) are either scarce or absence in most public and private secondary school in Nigeria. The type of test used by most teachers is teachers-made test which are constructed by class-room teachers and most often lack psychometric properties (Nasreen, Ahmad & Sabiha, 2019). Thus, decisions taken based on scores obtained from the administration of such instruments are inaccurate and misleading.

According to Odual (2013) some instructor without teaching ability, possess rather vague idea about the requirements for effective evaluation and learning achievement. Thus, it is necessary that students and specialist in education construct and validate tests to use in secondary schools to help improve teaching and learning**.** Therefore, the main purpose of this study was to find out how to prepare a standardize mathematics test items at the senior secondary school level. Specifically, the study objectives are as follows:

1. Determine the suitability of (MT) items
2. Find out the content validity of the MT
3. What is the cut –off score for the MT
4. What is the reliability coefficients of the various MT
5. How is the difficulty and distracter indices of the MT established

To get the desired study, the research questions is analyzed

1. What is the suitability of the MT items as rated by the content expert?
2. How is the content validity of MT established?
3. How is the cut-off score of the MT determine?
4. How is the reliability coefficient of the MT established?
5. How difficult and distractor are the MT items?

II. METHODOLOGY

This chapter is design to discuss methodology adopted in carrying out the study. The instrument used for this research is instrumentation. It is aimed at developing an instrument (MT) for evaluating SS1 students' Mastery of mathematics test. This research design has been used by other researchers who carried out similar studies including Opara and Magnus (2017) that conducted a research on development and validation of a test of geographic skills (TOGS) for senior secondary schools. Kpolovie (2010) defined instrumentation as the science of test development that is useful for test construction aimed at measurement and evaluation of psychological traits. Hence, instrumentation is a scientific research tool that is use to determine or measure human abilities and psychological evaluations.

A sample was drawn from the population of 3,326 SS1 mathematics students. This was done by dividing the population into three groupings or strata prior to sampling. Random samples of 823 students in five secondary schools were carried out. Specifically, table of instruments were constructed on:

i. Solving simple equations
ii. Bearing
iii. Statistics,
iv. Cartesian plan and co-ordinates
v. Set theory.

The instrument were use to collect data that was used for the determination of its characteristics and performance of students.

Table1. Test blueprint for a 54 items multiple-choice objective test in mathematics.

NO of Weeks	Contents Areas	Objects						
		Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation	Total
		15%	25%	20%	25%	10%	5%	100%

2	Solving simple equation	1	2	2	2	1	1	9
2	Bearing	1	2	2	2	1	1	9
3	Statistics	2	3	3	3	1	1	13.5
3	Cartesian co-ordinates	2	3	3	3	1	1	13.5
2	Set theory	1	2	2	2	1	1	9
12	Totals							54

Steps Involved are:

i. Determine the total number of weeks in teaching all the content areas.

i.e; 2+2+3+3+2=12 weeks

ii. Compute the number of items each content area would contribute to the 54 items needed. i.e; 2+2+3+3+2=12 weeks. Now in solving simple equations, 2 weeks was spent in convering the topic therefore,

$$\frac{2}{12}\text{X}\frac{54}{1} = 9$$

iii. Compute the number of items each cognitive level would contribute to each content area bearing in mind their respective percentages: That is Knowledge (15%), Comprehension (25%), Application (20%), Synthesis (10%), and Evaluation (5%).

For instance, knowledge level here is 15% for solving simple equations i.e;

$$\frac{15}{100}x\frac{9}{1} = 1.35 = 1$$

Results

The data collected at different stages of the research were analyzed and the result presented. These analysis were focused on the research question earlier stated.

What is the suitability of the MT items as related by the content experts?

Items were rated by five specialists. Thus, only the items with average rating 3.00 or equal to 3.00 are suitable while those average scores below 3.00 are unsuitable.

Table 2: The average rating score of MT items as determined from the rating of five content specialists are shown in the table below:

	RATE					
ITEMS	**1**	**2**	**3**	**4**	**5**	**Rater Average**
1	2	3	1	1	2	1.8
2	3	3	3	3	3	3
3	1	2	1	1	1	2
4	2	2	1	2	1	1.6
5	5	5	4	3	4	4.2
6	4	5	4	3	5	4.2
7	2	2	1	2	2	1.6
8	1	2	2	1	1	1.4
9	5	5	4	4	5	4.6
10	1	2	2	2	1	1.4
11	4	4	5	5	5	4.6
12	1	2	2	2	1	1.4
13	1	2	2	2	1	2.2
14	5	5	5	4	5	4.8
15	2	2	1	2	1	1.6
16	4	4	5	5	5	4.6
17	4	5	4	4	5	4.4

18	5	5	4	5	5	4.8
19	1	2	1	1	1	2
20	4	3	3	4	4	3.6
21	4	4	4	5	4	4.2
22	4	4	5	5	5	4.6
23	5	5	4	5	5	4.8
24	4	4	5	5	5	4.6
25	3	4	4	5	4	4
26	5	5	5	4	5	4.8
27	2	2	1	1	1	1.4
28	1	2	2	1	2	1.4
29	4	4	**5**	4	5	4.4
30	2	2	2	1	2	1.8
31	2	2	1	1	2	1.4
32	5	5	4	4	5	4.6
33	4	4	5	4	4	4.2
34	4	4	4	5	4	4.2
35	4	4	5	5	5	4.6
36	2	2	3	1	3	1
37	5	5	4	4	4	4.4
38	4	4	3	4	3	3.6
39	1	3	1	1	2	2.6
40	2	2	1	1	1	1.4
41	4	4	3	3	4	3.6
42	1	2	2	2	1	1.8
43	4	4	5	4	4	4.2
44	5	5	4	5	5	4.8
45	2	2	1	2	2	1.6
46	4	4	4	5	5	4.8
47	2	2	2	2	2	1.6
48	1	2	1	2	2	2.2
49	5	5	4	4	5	4.6
50	2	1	2	2	3	2
51	4	4	4	5	4	4.2
52	1	1	2	1	1	2.2
53	2	2	2	1	2	1.8
54	4	5	5	4	4	4.4
55	1	2	2	1	2	1.4
56	5	4	4	4	4	4.2
57	2	2	1	2	2	1.8
58	1	2	1	1	2	1.2
59	4	4	4	5	4	4.2
60	1	2	2	1	2	1.2
61	4	4	5	5	5	4.2
62	2	2	2	1	1	1.4
63	5	4	4	5	5	4.6
64	5	5	5	4	4	4.6
65	4	4	4	5	4	4.2
66	4	4	4	5	5	4.2
67	5	5	5	5	4	4.8
68	1	2	1	2	2	1.4

69	1	1	1	1	2	1
70	4	5	5	4	4	4.4
71	2	2	2	1	2	1.6
72	1	1	2	2	1	1.4
73	4	5	5	4	4	4.4
74	1	2	2	2	2	1.2
75	5	5	4	4	5	4.6
76	2	2	3	3	2	1.4
77	4	4	5	5	4	4.4
78	5	4	4	5	5	4.6
79	5	5	4	4	4	4.4
80	5	5	5	5	4	4.8
81	2	2	2	2	1	1.6
82	1	1	1	1	2	2.2
83	5	4	4	5	5	4.6
84	2	2	2	1	2	4.6
85	4	4	5	4	4	1.8
86	2	2	2	2	2	4.2
87	4	4	3	3	4	2.4
88	2	3	3	2	2	3.6
89	5	5	4	4	5	2.4
90	4	4	5	4	4	4.6
91	5	5	5	5	5	4.2
92	1	2	2	2	2	5
93	1	2	2	2	2	2.6
94	4	4	3	4	4	2.4
95	2	2	2	3	3	3.8
96	2	2	2	3	2	1.2
97	1	2	2	2	2	2.2
98	5	4	4	4	5	2.6
99	5	5	5	5	2	4.4
100	1	2	2	2	2	5
Total Item Average	315 **3.15**	318 **3.18**	311 **3.11**	310 **3.10**	319 **3.19**	315 **3.15**

A closer look at the table reveals that the average rating score of MT items range from 1.00 to 5.00.

How is the content validity of the MT established?
The validity index of the MT items were rated by two content experts as shown in table 3.

Table 3: Rating of two content experts.

	Specialist 1			
Specialist 2		**Item rated 1 and 2**	**Item rated 3, 4 or 5**	**Total**
	Items rated 1 and 2	**(a)** 9	**(b)** 5	A+b=14
	Items rated 3, 4 or 5	(c) 5	(d) 35	C+d=40
	Total	a+c=14	b+d=40	a+b+c+d=54

How is the cut off score of the MT determine?

The lower and upper group of the MT were established and a frequency distribution of the item score for the groups were plotted on the same graph as shown in figure 1 below. The test score where the two distributions intercept becomes the standard setting.

Table 4: Frequency distribution, for upper and lower group

Class Interval	Tally for U – Group	Frequency for U – Group	Tally for L – Group	Frequency for L – Group	Class Limit
0 – 5					0.5 – 5.5
6 – 10					5.5 – 10.5
11 – 15				5	10.5 – 15.5
16 – 20				10	15.5 – 20.5
21 – 25				25	20.5 – 25.5
26 – 30				32	25.5 – 30.5
31 – 35				45	30.5 – 35.5
36 – 40		10		25	35.5 – 40.5
41 – 45		15		15	40.5 – 45.5
46 – 50		20		5	45.5 – 50.5
51 – 55		23			50.5 – 55.5
56 – 60		30			55.5 – 60.5
61 – 65		25			60.5 – 65.5
66 – 70		22			65.5 – 70.5
71 – 75		12			70.5 – 75..5
76 – 80		5			75.5 – 80.5
81 – 85					80.5 – 85.5
		$\sum f = 162$		$\sum f = 162$	

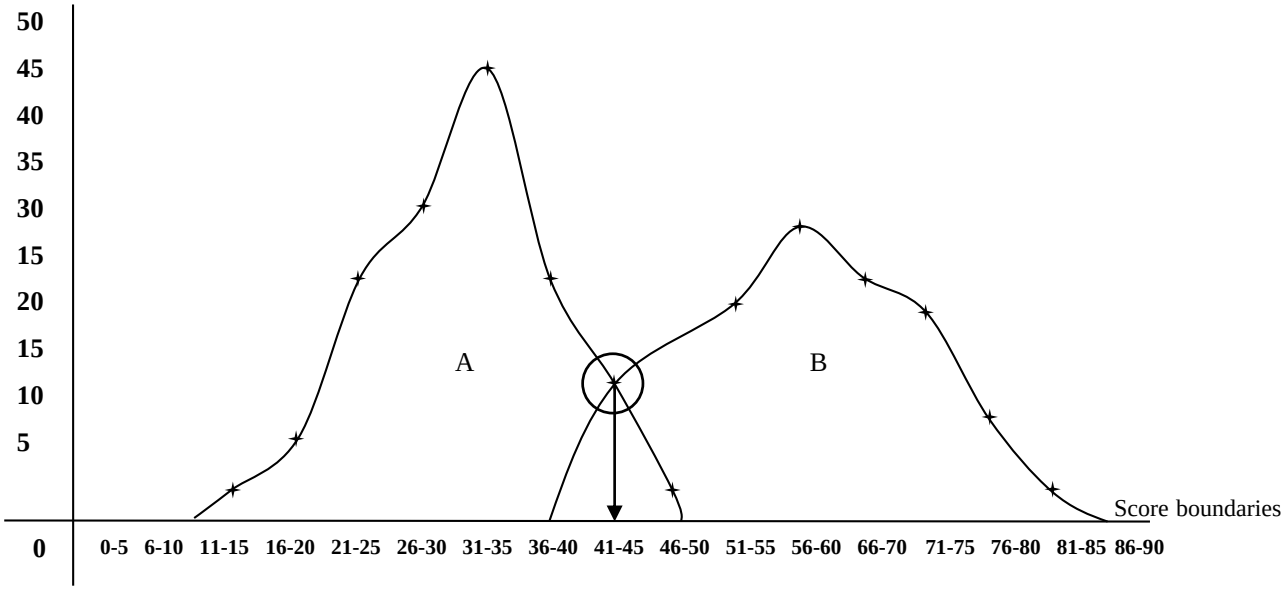

Fig. 1. Graphic presentation of data of lower (A) and upper (B) group of MT.

How is the reliability coefficient of the MT established?

The reliability coefficient of the MT were computed using test – retest technique. The reliability coefficient was base on three statistics , i.e. the Po, K, and Pc. Po measures the degree of agreement by the three administrations and, K measures the agreement uncontaminated by chance. Pc determine the agreement of the two administrations.

Table 5. Po, K and Pc reliability coefficient

	Test 1				**Po**	**K**	**Pc**
		Master	Non-mastery	Total			
Test 2	Mastery	A 108	B 106	214	0.55	0.12	0.49
	Non – Mastery	C 40	D 70	110			
	Total	148	176	324			

The analysis in the table above is shown as follows: Po, K and Pc

$$Po = \frac{a + d}{N} = \frac{100 + 70}{324} = \frac{170}{324} = 0.55$$

$$Pc = \frac{(a+b)(a+c) + (c+d)(b+d)}{N^2}$$

$$Pc = \frac{(108+106)(108+40) + (40+70)(106+70)}{324^2}$$

$$= \frac{(214)(148)+(110)(176)}{324^2} = \frac{31672+19360}{324^2} = \frac{51032}{104.976} = 0.486 \cong 0.49$$

$$K = \frac{Po-Pc}{1-Pc} = \frac{0.5-0.49}{1-0.49} = \frac{0.06}{0.51} = 0.117 \quad 0.12 =$$

K of 0.12 implies that 12% of the observed agreement in the decisions made by both test is uncontaminated by chance.

Difficult and distracter Index of the Mathematical Test (MT) items

Item difficulty means how difficulty an item is. Hence, this can be determined by the formula:

$$\text{Item difficult} = \frac{U+L}{N}$$

Where;
U = Number of candidate that passed the items in the master group
L = Number of candidate that passed the item in the non-mastery group
N = Number of candidate that sat for the test

III. DISCUSSION

In table 1 above, out of 100 items only 54 items was found suitable, this was achieved by deleting items that was rated below 3. Hence, only 54 mathematics test questions rated by the specialists as good items was administered to students. This indicates that the content specialists rated each item with 1, 2, 3, 4 or 5.

Thus, $\text{CVI} = \frac{d}{(a+b+c+d)} \quad \frac{40}{9+5+5+35} = \frac{40}{54} = 0.74$

To find the cut –off score of the mathematics test, 54 standardized items as administered to 324 students. These students was classified into lower and upper group i.e. 162 students in a group and their scores presented in a frequency distribution table as shown above (table 4). A graph shown frequency axis and score boundaries us presented. The two group (A & B) scores are plotted in the same graph. Their scores was established ranging from the least score to the highest score in the score boundaries axis. In the frequency axis, the number of time, these score occur is presented. In this case, in the frequency axis of the graph from the lower group, 5 students scored 11-15 marks, 10 students scored 16-20 marks, 25 students scored 21-23 marks, 32 students scored 26-30 marks, 4 students scored 31-35 marks, 25 students scored 36-40 marks, 15 students scored 41-45 marks, and 5 students scored 46-50 marks. In the upper group from the frequency table, 10 students scored 36-40 marks, 15 students scored 41-45 marks, 20 students scored 46-50 marks, 23 students scored 51-55 marks, 30 students scored 56-60 marks, and 25 students scored 61-65 marks.

22 students scored 66-70 marks, 12 students scored 71-75 marks and 5 students scored 76-80 marks. When these scores (A & B) was plotted in the same graph, two curve was obtained and the point where the two curve intercept is 45%. This become the accepted cut-off for the MT. Coefficient of agreement Po; above shows the measure of the proportion of students for whom the same decision, mastery or non-mastery is made on both administration, i.e. test 1 and test 2 by the raters. In the above table, 2 test was administered to a sample of 324 students and result classified for both test.

Computation of the Po by the formula: $\text{Po} = \frac{a + d}{N}$

Where;
Compute the Po by the formula:

$$\text{Po} = \frac{a+d}{N}$$

where Po = Agreement coefficient
a = Number classified as masters on both administration
d = Number classified as masters on both administration
N = Total number of subjects

$$\text{Po} = \frac{a+d}{324} = \frac{108 + 70}{324} = \frac{178}{324} = 0.55$$

Therefore, 0.55 implies 55% of the subjects were classified consistently in agreement, chance agreement K. This is called coefficient Kappa which takes chance agreement into consideration. It is referred to the proportion of consistent classifications observed beyond that expected by chance.

Expected chance agreement Pc: It measures the proportion of individual to have consistent classification even if there were no genuine relationship between the test. That is if the classification on the two administrations were completely independent. It is completed using the formula:

$$\text{Pc} = (a+b)\,(a+c) + (c+d) + (b+d)$$

N^2

Where Pc = proportion of agreement expected by chance, a, d and N retain their former interpretations
C = number classified as master and non master for test 1 and 2 respectively.
b = number classified as non-master and master for test 1 and 2 respectively.
Computing Pc using the above, $Pc = \frac{(214)(148) + (110) + (176)}{324^2}$

$$\frac{51,032}{104,976} = \frac{0.49}{}$$

How Difficult and distractor are the MT Items?

The difficulty indices and the distractor indices was computed using the respective formulae. 148 students were found to belong to mastery group since their scores range from 45% and above. While the remaining 176 students were found to belong to non-mastery group.

Formulae:

Item difficulty = $\frac{U + L}{N}$

Distracter (DI) = DI = $\frac{X_1 + V_L}{Nu + NL}$

Table 6.1. Lower and Upper group that answer item in the test

ITEM	U	L	N	D
1	216	102	324	0.981
2	132	24	324	0.41
3	210	96	324	0.944
4	96	30	324	0.38
5	144	42	324	0.574
6	186	42	324	0.703
7	210	60	324	0.833
8	198	54	324	0.777
9	198	84	324	0.877
10	120	18	324	0425
11	156	72	324	0.703
12	210	72	324	0.870
13	210	84	324	0.907
14	114	24	324	0.425
15	204	78	324	0.870
16	210	60	324	0.833
17	156	42	324	0.611
18	132	36	324	0.518
19	108	24	324	0.407
20	90	66	324	0.481
21	138	60	324	0.611
22	210	90	324	0.925
23	210	84	324	0.907
24	114	24	324	0.425
25	156	42	324	0.611
26	30	18	324	0.148
27	90	12	324	0.314
28	180	42	324	0.685
29	66	18	324	0.259
30	120	60	324	0.555
31	24	12	324	0.111

32	84	30	324	0.351
33	156	72	324	0.703
34	108	96	324	0.629
35	66	18	324	0.259
36	96	30	324	0.388
37	78	36	324	0.351
38	192	60	324	0.777
39	192	90	324	0.870
40	108	12	324	0.370
41	180	72	324	0.777
42	186	72	324	0.796
43	48	42	324	0.277
44	66	18	324	0.259
45	30	90	324	0.370
46	30	30	324	0.185
47	156	72	324	0.703
48	54	78	324	0.407
49	186	54	324	0.740
50	114	36	324	0.462
51	174	12	324	0.574
52	180	42	324	0.685
53	186	84	324	0.833
54	54	18	324	0.222

Table 6.1 shows a level of clarity for the interpretation of difficulty. It shows how the difficulty items were obtained from the table above, item 1 indicate that: 216 is the number of students in the upper (Mastery) class that got the option in item 1 correct and 102 students in the lower (non-mastery) got item 1 correct. 324 is the total number of candidates that set for the test. Now put the formula above into use:

$$D = \frac{U + L}{N} \text{ i.e. } \frac{216 + 102}{324} = 0.981$$

Therefore, 0.981 = difficult level of item.

Take item 6. 186 is the number of students in the upper (mastery) class that got the option in item 6 correct and 42 students in the lower (non-mastery) class got item 6 correct. 324 is the total number of candidates that sat for the test.
I.e. $\frac{186 + 42}{324} = 0.703$ Therefore, 0.703 = difficulty level of item 6, and in that order to 54 items administer in the mathematics test.

Distracter index: A distracter or distractor is an incorrect alternative option. It is a plausible wrong answer designed to be attractive to students who do not know the correct answer more than the students that know the correct answer.

Distracter index D can be computed by the formula: $d = \frac{L - U}{N}$

To get it right, a list of test items showing the number of students in mastery and non-mastery group and option tick in the test. Item 1 option is ''A'' i.e the key. The table 6.2 below, in the mastery class – 70 students tick ''A'' in item 1 while 45 students tick ''A'' in non-mastery class.

Option C item 1. 18 students tick C in the mastery class while 22 students tick C in the non-mastery class. Total candidate that attempt item 1 in the mastery class is 154 while total candidate in non-mastery class that attempt item 1 is 170. Option ''A'' is the correct answer (the key).

Table 6.2. Shows table of option tick by candidates in the mathematics test (distraction index)

ITEM		**A**	**B**	**C**	**D**	**E**	**Total**
1	**Mastery**	70*	33	18	20	13	154
	Non-Mastery	45	38	22	26	39	170
	M	*58	24	22	28	2	134
2	N	49	36	27	37	41	190

3	M	14	28	33	9	*60	144
	N	18	36	42	37	47	180
4	M	24	*59	26	18	28	155
	N	34	38	36	28	33	169
5	M	*80	22	19	17	18	156
	N	47	37	24	31	29	168
6	M	15	12	*77	10	15	129
	N	36	27	51	42	38	195
7	M	*68	23	15	23	22	151
	N	47	32	29	38	27	173
8	M	*78	10	8	12	12	120
	N	56	42	27	38	41	204
9	M	24	15	19	*68	21	147
	N	35	27	38	45	32	177
10	M	18	*70	15	16	18	137
	N	40	48	38	20	41	187
11	M	21	17	*69	14	19	140
	N	38	41	43	33	29	184
12	M	*71	19	18	19	17	144
	N	44	37	28	40	31	180
13	M	15	19	13	16	*81	144
	N	29	41	31	32	47	180
14	M	19	*66	18	20	20	143
	N	46	47	26	28	34	181
15	M	15	8	*73	19	18	133
	N	46	38	50	22	35	191
16	M	* 84	12	20	19	17	152
	N	48	21	40	36	27	172
17	M	*69	16	19	13	18	135
	N	48	21	40	36	27	172
18	M	58*	23	19	22	25	147
	N	46	37	21	33	40	177
19	M	20	21	*78	12	23	154
	N	27	40	45	20	38	170
20	M	*69	19	16	21	20	145
	N	47	25	31	40	36	179
21	M	22	14	*60	23	25	144
	N	38	27	41	38	36	180
22	M	21	23	21	16	*74	155
	N	27	35	30	28	49	169
23	M	18	*80	5	19	23	145
	N	29	48	27	35	40	179
24	M	25	*52	22	24	21	154
	N	41	46	31	28	24	170
25	M	6	12	*80	14	15	127
	N	36	31	59	27	44	197
26	M	19	18	12	20	*74	143
	N	39	38	27	36	41	181
27	M	19	*69	14	18	20	140
	N	38	48	21	41	36	184
28	M	*70	21	23	24	0	138
	N	49	33	37	42	25	186
29	M	22	17	*74	16	20	149
	N	39	24	50	21	41	175
30	M	15	21	4	*67	18	117
	N	48	37	28	53	41	207

	M	9	*79	20	27	25	160
31	N	15	48	23	34	44	164
	M	4	15	19	0	*88	126
32	N	35	44	39	27	53	198
	M	*66	18	23	22	14	143
33	N	49	27	42	38	25	181
	M	15	*68	16	6	21	126
34	N	36	55	38	29	40	198
	M	23	18	10	*80	19	150
35	N	43	28	26	47	30	174
	M	10	14	13	18	*75	130
36	N	23	39	33	41	58	194
	M	22	*69	13	25	26	155
37	N	28	46	26	31	38	169
	M	23	10	16	13	*80	142
38	N	52	28	44	34	36	194
	M	23	10	16	13	*80	142
39	N	40	37	25	22	58	182
	M	*88	13	21	24	5	151
40	N	53	29	30	38	23	173
	M	*72	24	26	20	12	154
41	N	48	35	27	41	19	170
	M	*76	12	15	18	15	136
42	N	49	25	38	42	34	188
	M	22	19	*67	24	20	152
43	N	40	28	45	32	27	172
	M	18	11	16	*77	18	140
44	N	42	26	31	50	35	184
	M	13	*83	16	1	18	133
45	N	31	60	37	22	41	191
	M	19	*68	20	2	22	131
46	N	32	58	41	26	36	193
	M	18	*77	6	19	23	143
47	N	29	50	26	33	43	181
	M	*68	10	18	22	18	158
48	N	48	29	36	45	40	198
	M	*86	14	18	22	18	158
49	N	45	21	34	40	26	166
	M	*72	16	20	22	17	147
50	N	45	22	36	41	33	177
	M	21	18	*88	13	4	144
51	N	42	24	55	28	31	180
	M	*77	11	20	18	22	148
52	N	48	26	39	23	40	176
	M	16	23	19	15	*69	141
53	N	33	41	25	38	46	183
	M	*81	16	8	18	15	138
54	N	47	36	29	40	34	186

Table 6.3. Shows difficulty and distracter index of MT items

ITEM	**D**	**D.I**				
1.	0.981	A *	B 0.02	C 0.01	D 0.02	E 0.08
2.	0.41	*	0.04	0.08	0.03	0.12
3.	0.944	0.01	0.02	0.03	0.09	*

4.	0.38	0.03	*	0.03	0.03	0.02
5.	0.574	*	0.05	0.02	0.04	0.03
6.	0.703	0.07	0.05	*	0.10	0.07
7.	0.833	*	0.03	0.04	0.05	0.02
8.	0.777	*	0.10	0.06	0.08	0.09
9.	0.877	0.03	0.04	0.06	*	0.03
10.	0.425	0.07	*	0.07	0.01	0.07
11.	0.703	0.05	0.07	*	0.06	0.03
12.	0870	*	0.06	0.03	0.06	0.04
13.	0907	0.04	0.07	0.06	0.05	*
14.	0.425	0.08	*	0.02	0.02	0.04
15.	0.870	0.10	0.09	*	0.01	0.03
16.	0.833	*	0.03	0.06	0.05	0.03
17.	0.611	*	0.06	0.08	0.05	0.05
18.	0.518	*	0.04	0.01	0.03	0.05
19.	0.407	0.02	0.06	*	0.02	0.05
20.	0.481	*	0.02	0.05	0.06	0.05
21.	0.611	0.05	0.04	*	0.05	0.03
22.	0.925	0.02	0.04	0.03	0.04	*
23.	0.907	0.03	*	0.07	0.05	0.05
24.	0.425	0.05	*	0.03	0.01	0.01
25.	0.611	0.09	0.06	*	0.04	0.09
26.	0.148	0.06	0.06	0.05	0.05	*
27.	0.314	0.06	*	0.02	0.07	0.05
28.	0.685	*	0.04	0.04	0.06	0.08
29.	0.259	0.05	0.02	*	0.02	0.07
30.	0.555	0.10	0.05	0.07	*	0.07
31.	0.111	0.02	*	0.01	0.02	0.06
32.	0.351	0.10	0.03	0.06	0.08	*
33.	0.703	*	0.08	0.06	0.05	0.03
34.	0.629	0.07	*	0.07	0.07	0.06
35.	0.259	0.06	0.09	0.05	*	0.03
36.	0.388	0.04	0.08	0.06	0.07	*
37.	0.351	0.02	0.05	0.04	0.02	0.04
38.	0.777	*	0.03	0.07	0.05	0.06
39.	0.870	0.05	0.04	0.03	0.03	*
40.	0.370	*	0.06	0.03	0.04	0.06
41.	0.777	*	0.07	0.01	0.07	0.02
42.	0.796	*	0.06	0.07	0.07	0.06
43.	0.277	0.06	0.03	*	0.03	0.02
44.	0.259	0.07	0.05	0.05	*	0.05
45.	0.370	0.06	*	0.07	0.07	0.057
46.	0.185	0.04	*	0.07	0.07	0.04
47.	0.703	0.03	*	0.06	0.04	0.06
48.	0.407	*	0.06	0.06	0.10	0.07
49.	0.740	*	0.02	0.06	0.06	0.03
50.	0.462	*	0.02	0.05	0.06	0.05
51.	0.574	0.07	0.02	*	0.05	0.08
52.	0.685	*	0.05	0.06	0.02	0.06
53.	0.833	0.05	0.06	0.02	0.07	*
54.	0.222	*	0.06	0.07	0.07	0.06

*= Key Option

http://dx.doi.org/10.29322/IJSRP.12.07.2022.p12741 www.ijsrp.org

In the table above, in item 1, distracter B is written 0.02, C is written 0.01, D is written 0.02 and E as 0.08. To obtain this;

Item 1 option B.

$\frac{38-33}{324} = 0.02$

i.e. 38 students of non-mastery tick option B and 33 students of mastery tick option B as well. Substituting these figures into the formula below, you obtained 0.02 and in that order.

Option C

$\frac{22-18}{324} = 0.01$

Option D

$\frac{26-20}{324} = 0.02$

Option E

$\frac{39-13}{324} = 0.08$

Formula is $d = \frac{L - U}{N}$

IV. CONCLUSION

Conclusion was drawn after data analyses as follows: The suitability of the MT items was determined after the rating score by the five specialists in term of appropriateness of each item. This finding is related to that of Rufina et al. (2015) who assessed teacher competence in test construction and content validity of teacher made examination questions in Borno state in Nigeria using questionnaires and found out that teachers in Borno state were not capable in test construction. The author also found that teachers questions has low content validity that emphasized lower levels of cognitive domain i.e. remembering, understanding and applying.

V. RECOMMENDATIONS

Based on the findings of the study, the following recommendations were made:

i. The MT should be used by mathematics teachers for both formative and summative evaluation of SS 1 students' mastery of the mathematics test. Students' results in MT should be used by teachers and guidance and counselors in advertising the students in their choice of mathematics award of scholarship.

ii. Again, MT should be use as a model test of mathematics test by mathematics teachers and future researchers. The mathematics especially for SS II and SS III students. Future researchers who will work on development and validation of tests of mathematics tests should also use MT as a reference point in constructing their own items. Those researching on the ''Level of SS II students mastery'' of the mathematics test in other states could administer MT to their subjects instead of taking pains in developing and validating their own test.

iii. Workshops, seminars and inter-services training programmes should be regularly organized for mathematics teachers which will acquaint them with the most effective instructional objectives for provoking students' acquisition skills.

AKNOWLEDGEMENT

The Author is particularly grateful to Mr. Authority Benson of the Niger Delta University who prodded me into producing the journal. I also wish to thank authorities in the discipline of measurement and evaluation whose contributions in form of proof-reading and advising led to the success of the journal.

REFERENCES

[1] [1]. Jayanthi, J. (2014). Development and validation of an achievement test in mathematics. International journal of mathematics and statistics invention (IJMSI), 2(4), 40-46.

[2] [2]. Opara, I.M. & Magnus-Arewa, E.A. (2017). Development and validation of mathematics achievement test for primary school pupils. British journal of education, 5(7), 47-57.

[3] [3]. Rafina, S.H. Abubakar, H.T. & Stephen, S.H. (2015). Assessing teaching competent in test construction and content validity of teacher made examination questions in commerce in Borno State, Nigeria. Education, 5(5), 123-128.

[4] [4]. Darling, L. H. (2002). Teacher quality and student achievement: a review of state policy evidence. Educational policy analysis archives, 8(1).

[5] [5]. Ololube, N.P. (2008). Evaluation competence of professional and non professional teacher in Nigeria. Studies in educational evaluation (SEE), 34(1), 44-51.

[6] [6]. UNESCO (1990). Guidelines on teachers' competence procedures. Code of Federal regulations. 29, (4), 1607.

[7] [7]. Akhter, N. & Bahoo, R. (2015). Development of a semi standardized test of education for intermediate level. Journal of education research, 18(2), 5-9.

[8] [8]. Ukeje, B.O. (2005). Production and retention of mathematical sciences teachers for Nigerian educational system. Reflective and intellective position papers on mathematics education. P. 80-102.

[9] [9]. Clement, D.H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the building blocks project. Journal for research in mathematics education, 38, 136-163.

[10] [10]. Unodiaku, S.S. (2009). Development and validation of mathematics readiness test for senior secondary school students. A Ph.D Thesis presented to the department of science education, University of Nigeria, Nsukka.

[11] [11]. Nasreen, A., Ahmad, A.U. & Sabiha, I. (2019). Development and validation of multiple-choice test geometry part of mathematics for secondary class. Global social sciences reviews, 4(2), 283-292.

[12] [12]. Odual, N.N. (2013). Relationship between mathematical ability and achievement in mathematics among female secondary school students in Bayelsa state. Procedia-social and behavioral sciences, 106, 2230-2240.

[13] [13]. Essein, I.T (2002). Development and validation of a test of geography skills (TOGS) for senior secondary schools. Unpublished Ph.D Education. Thesis, University of Port Harcourt.

AUTHORS

First Author – Otuoku Karl, HND (Enugu) , PG ED, M. ED (UPH). Depart of Psychology, Guidance and Counselling Isaac Jasper Boro College of Education, Sagbama, P.M.B 74 Yenagoa, Bayelsa State, Nigeria. Email: otuokukarl@gmail.com

Correspondence Author – Otuoku, Karl, Email: otuokukarl@gmail.com, Tel: 002348063740391