ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct1276

Application of the Split-Half Reliability (KR-20) Test in Validation of Multiple Choice-Type Testing Instrument to Assess Student Performance in Competence Based Education (CBE) System in Kenya

Dr. Demtila Nafula Wanjala[1]; Sibilike Khamala Makhanu[2]

[1]Lecturer in Education Management and Policy Studies, Department of Education Planning and Management/Education Foundations, Kibabii University, P. O. Box 1699-50200, Bungoma, Kenya.
[2]Professor of Civil Engineering, Department of Department of Civil and Structural Engineering, Masinde Muliro University of Science and Technology (MMUST), P. O. Box 190-50100, Kakamega, Kenya.

Publication Date: 2025/11/03

Abstract: Reliability is defined as the consistency of results from an administered test. Theoretically each statistical test contains some inherent error which is the portion of the score on the test that is not exactly relevant to the construct that is intended to be measured. Sit-in objective type questions have been embraced at basic and middle levels of education in Kenya for a long time. However universities have only recently embraced objective type questions by use of online platforms due to large number of students and the availability of reliable online platforms. Even when this is a positive development to address the immediate challenges facing universities the reliability of such instruments is paramount. The purpose of evaluation is to assess the essential life skills which include communication, problem solving and adaptability. In this paper the validation of objective type questions testing instrument at basic and tertiary level is validated by use of Split-Half Reliability Testing Instrument. The Split-half reliability determines how much error in a test score is due to poor test construction. The split-half reliability was administered to Grade 7 students at Ekato Junior Secondary School in Western Kenya in 9 subjects under the Competency Based Education (CBE). The computations were done by using Excel Spreadsheet to determine the reliability index through use the Kuder-Richardson formula 20 (KR-20).

Keywords: *Reliability; Split-Half Reliability; Kuder-Richardson Formula 20 (KR-20); Essential Skills; CBE.*

How to Cite: Dr. Demtila Nafula Wanjala; Sibilike Khamala Makhanu (2025). Application of the Split-Half Reliability (KR-20) Test in Validation of Multiple Choice-Type Testing Instrument to Assess Student Performance in Competence Based Education (CBE) System in Kenya. *International Journal of Innovative Science and Research Technology*, 10(10), 2174-2180. https://doi.org/10.38124/ijisrt/25oct1276

I. INTRODUCTION

The level to which knowledge has been passed to learners can be established through testing of the candidates. The purpose of a test is to assess the recipient's knowledge, skills and abilities regarding a specific discipline. Tests provide structured ways to determine how well the recipient has grasped the subject matter. Tests are important to measure progress of students and identify areas of weaknesses and strengths. Instructors can therefore use the test results as important feedback to enable in planning of the areas for further instruction. Tests are often used as important benchmark to make career decisions for learners such as course choices, career pathways, student placements and other educational issues. For policy makers, the performance in tests helps to evaluate the effectiveness of educational programs and in the identification of areas that require improvement. Tests help students to read and research more from the understanding that their performance will be used to determine their future academic progression.

The testing instruments have such profound significance. The instrument has to be reliable with minimal error. Reliability is the consistency of results from a test. It is the level to which the result of a measurement, calculation, or specification can be depended on to be accurate. Theoretically, each test contains some residual error, which is the portion of the score on the test construct to be measured. The residual error represents the difference between an observed value and a predicted value from the model. It

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct1276

signifies the unexplained variation in the data after model application. While often used interchangeably, errors refer to the difference between an observed value and the true, but often unknown, underlying value. Residuals on the other hand are differences between the observed value and the value predicted by the model or model answers. Error could be the result of initial poor test construction, distractions from the time the examinees participated in the measure, from the time the results from the administered assessment were evaluated.

Reliability is quantified through a reliability index which tries to determine the proportion of the test score that is due to error in the instrument. The reliability index provides a measure of how likely a given evaluation tool is to perform according to the intended purpose under the given constraints. The specific formula and interpretation of the reliability index is dependent on the field and specific system under analysis. Reliability indices are therefore valuable tools for assessing and comparing reliability of systems across various fields. Understanding the interpretation of the specific reliability index is pertinent for making informed decisions.

II. EVALUATION OF INSTRUMENT RELIABILITY

Evaluation of the reliability of an instrument involves assessing how consistently it produces the similar results under similar repetitive conditions in terms of past paper scenario. Reliability of a testing instrument can be measured through various methods which include split-half reliability, test-retest reliability, parallel forms reliability, inter-rater reliability and internal consistency reliability. Key Concepts in reliability evaluation are: Reliability: The extent to which an instrument yields consistent and dependable results. True Score: The underlying, accurate value of what is being measured. Measurement Error: The difference between the observed score and the true score. Consistency: Repeated measurements should yield similar results. There are five methods for evaluating reliability:

- *Test-Retest Reliability*

This involves administering the same instrument to the same group of individuals at two different time points and calculating the correlation between the two sets of scores to assess stability over time. For example in social survey one may give a questionnaire to a group of students, then repeating the same questionnaire to the same group a year later.

- *Parallel Forms Reliability*

In parallel forms reliability two different versions of the same instrument (parallel forms) that are designed to measure the same construct. This involves administering both forms to the same individuals and correlating the scores. For example two different but equivalent versions of an English test to the same group. The correlation of the test scores is then calculated.

- *Inter-Rater Reliability*

The inter-rater reliability involves evaluating the consistency of measurements made by different raters or observers. It is used when subjective judgments are involved. For example in assessing the agreement between two assessors scoring the same performance.

- *Split-Half Reliability*

This involves dividing the instrument into two halves and correlating the scores on each half. It determines how much error in a test score is due to poor test construction. The split-half reliability is computed by administering the test once and then calculating the reliability index by coefficient alpha, Kuder-Richardson formula 20 (KR-20) or Spearman-Brown formula.

- *Internal Consistency RR Cronbach's Alpha*

This is a statistical measure of internal consistency that assesses the average correlation between all possible item pairs within a scale. In popularity the split-half reliability and Cronbach's Alpha are mostly used.

III. CHOICE OF RELIABILITY TEST

In order to validate a measure focus may be in evaluating the split-half reliability of the measurement instrument. The split-half method will show how consistently the measuring instrument will assess the construct of focus. If one is measuring multiple constructs the split-half reliability will be considerably lower. In this case is recommended to separate the constructs that are to be measured into different parts and calculate the reliability separately for each construct. These are the so called too broad or subjective questions. Likewise if the calculated reliability is too low then the measure employed is probably measuring more constructs than it is designed to measure. It is recommended to revise the measure in order to focus on the construct of interest (Lenke et. Al., 1977). If the measure includes dichotomous items such as right-wrong answers or multiple type choices in examinations then the KR-20 formula is the best accepted statistic. If the Likert Scale or other types of items are used then the Spearman-Brown formula is best suited (Eun-Hyun Lee et. al., 2020).

- *Split-Half Reliability: KR-20*

An application of the KR-20 test was made on a class of students in Grade 7 of the Competency Based Education (CBE) education system in Kenya over a range of subjects. In each subject multiple type questions were used. The purpose was to evaluate the reliability of such an instrument across different subjects covering literacy, numeracy and innovations. Assessment of learners in Grade 4 by an NGO in 2021 showed that 60% of Grade 4 learners in Kenya could not read a Grade 3 appropriate test (Ayiera, 2023).

- *KR 20 Formula*

Kuder-Richardson formula 20 (KR-20) is only used if each item has a right answer, such as in multiple choice questions. It is not used in the case of Likert scale formulations.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct1276

- *KR-20 Formula is Given as:*

$$r_{KR20} = \left(\frac{k}{k-1}\right)\left(1 - \frac{\sum pq}{\sigma^2}\right) \quad \text{Equation – 1}$$

Where,

r_{KR20} = the Kuder-Richardson formula 20

k = the total number of test items

Σ = summation

p = the proportion of the test takers who pass an item

q = the proportion of the test takers who fail an item

σ^2 = the variation of the test

Grade 7 end term test results for students at Ekato Junior Secondary School were used in the analysis (Table 1).

Table 1 Raw Marks for Grade 7 Student Scores in End-Term Test at Ekato JSS

NO	NAME	ENG	MATH	KISW	SST	INTE SCI	CRE	PRETECH	AGR/HSCI	C/ARTS
1	ENGEPECE EGENYE	13	34	40	86	68	72	37	46	13
2	EUSCENE SHEGOGOCE	18	12	42	50	64	56	36	63	18
3	PREVEN CENYERENGE	24	18	30	52	49	38	12	18	24
4	PREGEC MEDEVE	4	12	30	28	29	28	22	34	4
5	CEPESCUS MEGUMPE	40	38	54	58	52	70	34	56	40
6	CEPVEN OCHEPO	6	8	8	4	17	20	2	43	6
8	FRENGPENO MEPEPE	54	20	42	60	87	58	36	12	54
9	GPECEOUS EDUVUGHE	56	22	44	56	24	80	34	65	56
10	GPOREE PUGEPEGE	24	18	14	6	25	2	0	43	24
11	ENNOCONC GEPGOMPOE	6	14	28	38	52	16	6	29	6
12	GONCREX SHEPESEE	6	4	24	28	35	6	4	8	6
13	GOPPEON MUGECE	8	12	24	20	46	12	18	14	8
14	PEVENO MEGHUNGU	20	28	38	54	52	50	46	38	20
15	MOPVEN MMPONO	24	10	42	72	49	72	46	40	24
16	MECCHOPPO GHEMOCO	32	32	48	60	72	50	64	59	32
17	NECHEN OPOCO	4	4	4	4	15	4	3	21	4
18	PEUPENO EYOCE	20	38	34	64	67	62	38	65	20
19	PROCEOUS MUCOPE	14	14	30	44	55	40	16	18	14
20	RODOMPCE GHECEPE	52	27	56	38	58	80	64	68	52
21	ROONOY SHESEGHE	6	6	6	10	14	2	0	6	6
22	SEMSON WONDO	12	26	16	72	62	42	36	57	12
23	SHENCOP OSCHOR	4	8	8	16	58	2	2	47	4
24	SHOEPE GENEEZE	56	20	38	60	48	62	44	38	56
25	CREZE MUCEMPE	32	28	46	52	58	48	32	45	32
26	VEONU GHUCEUPU	34	24	38	52	56	60	38	70	34
27	WEPSON MUHUCEU	12	22	42	60	42	52	46	58	12
	TOTAL	**589**	**503**	**832**	**1150**	**1254**	**1094**	3	**1061**	**589**
	AVERAGE	**21.81**	**18.63**	**30.81**	**42.59**	**48.23**	**40.52**	**26.59**	**40.81**	**21.81**

➢ *Coding of Raw Marks*

In order to avoid situations where the mark was equal to the set score, the mean scores were rounded to the nearest whole number as follows:

English – ENG (actual mean = 21.81%, rounded mean = 22%), Mathematics – MATH (actual mean = 18.63%, rounded mean =19%), Kiswahili – KISW (actual mean = 30.81%, rounded mean =31%), Social Studies – SST (actual mean = 42.59%, rounded mean =43%), Integrated Science – INTESCI (actual mean = 48.23%, rounded mean =48%), Christian Religious Education – CRE (actual mean = 40.52%, rounded mean =41%), Pre-technical – PRETECH (actual mean = 26.59%, rounded mean =27%), Agriculture/Home Science – AGR/HSC (actual mean = 40.81%, rounded mean =41%) and Creative Arts- C/ARTS (actual mean = 21.81%, rounded mean =22%).

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct1276

- *The Raw Marks Were Then Converted to Pass or Fail by Using the Criteria:*

If student mark was more than the rounded mean score it was awarded a PASS (1) and if the score was less than the rounded mean score then it is awarded FAIL (0).

- *This Criteria was Coded in Excel as Follows:*

IF (Score < RMean, "0", "1") Equation (2)

Where,

Score = Student raw mark in Table 1 and

RMean = Rounded-off mean score for the subject.

The conversion of Table 1 to Table 2 by use of coding in Equation (2) is as follows (Table 2).

Table 2 Coded Grade 7 Student Performance in End-Term Test at Ekato JSS

NO	NAME	ENG	MATH	KISW	SST	INTE SCI	CRE	PRETECH	AGR/HSCI	C/ARTS
1	ENGEPECE EGENYE	0	1	1	1	1	1	1	1	0
2	EUSCENE SHEGOGOCE	0	0	1	1	1	1	1	1	0
3	PREVEN CENYERENGE	1	0	0	1	1	0	0	0	1
4	PREGEC MEDEVE	0	0	0	0	0	0	0	0	0
5	CEPESCUS MEGUMPE	1	1	1	1	1	1	1	1	1
6	CEPVEN OCHEPO	0	0	0	0	0	0	0	1	0
7	DAMARES GHACEEVE	0	0	0	0	0	0	0	0	0
8	FRENGPENO MEPEPE	1	1	1	1	1	1	1	0	1
9	GPECEOUS EDUVUGHE	1	1	1	1	0	1	1	1	1
10	GPOREE PUGEPEGE	1	0	0	0	0	0	0	1	1
11	ENNOCONC GEPGOMPOE	0	0	0	0	1	0	0	0	0
12	GONCREX SHEPESEE	0	0	0	0	0	0	0	0	0
13	GOPPEON MUGECE	0	0	0	0	0	0	0	0	0
14	PEVENO MEGHUNGU	0	1	1	1	1	1	1	0	0
15	MOPVEN MMPONO	1	0	1	1	1	1	1	0	1
16	MECCHOPPO GHEMOCO	1	1	1	1	1	1	1	1	1
17	NECHEN OPOCO	0	0	0	0	0	0	0	0	0
18	PEUPENO EYOCE	0	1	1	1	1	1	1	1	0
19	PROCEOUS MUCOPE	0	0	0	1	1	1	0	0	0
20	RODOMPCE GHECEPE	1	1	1	0	1	1	1	1	1
21	ROONOY SHESEGHE	0	0	0	0	0	0	0	0	0
22	SEMSON WONDO	0	1	0	1	1	1	1	1	0
23	SHENCOP OSCHOR	0	0	0	0	1	0	0	1	0
24	SHOEPE GENEEZE	1	1	1	1	1	1	1	0	1
25	CREZE MUCEMPE	1	1	1	1	1	1	1	1	1
26	VEONU GHUCEUPU	1	1	1	1	1	1	1	1	1
27	WEPSON MUHUCEU	0	1	1	1	0	1	1	1	0

➢ *Calculation of P*

To calculate the proportion of students who passed the subject first the number of 1's for each subject were counted. This gave the total number of students who passed the subject by scoring above the rounded-off mean. The proportion p was obtained by dividing the total number of 1's with the total number of students which was 27 in this case. The calculated p is shown in Table 3.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct1276

Table 3 Proportion of Students who Passed, P

NO	NAME	ENG	MATH	KISW	SST	INTE SCI	CRE	PRETECH	AGR/HSCI	C/ARTS
1	ENGEPECE EGENYE	0	1	1	1	1	1	1	1	0
2	EUSCENE SHEGOGOCE	0	0	1	1	1	1	1	1	0
3	PREVEN CENYERENGE	1	0	0	1	1	0	0	0	1
4	PREGEC MEDEVE	0	0	0	0	0	0	0	0	0
5	CEPESCUS MEGUMPE	1	1	1	1	1	1	1	1	1
6	CEPVEN OCHEPO	0	0	0	0	0	0	0	1	0
7	DAMARES GHACEEVE	0	0	0	0	0	0	0	0	0
8	FRENGPENO MEPEPE	1	1	1	1	1	1	1	0	1
9	GPECEOUS EDUVUGHE	1	1	1	1	0	1	1	1	1
10	GPOREE PUGEPEGE	1	0	0	0	0	0	0	1	1
11	ENNOCONC GEPGOMPOE	0	0	0	0	1	0	0	0	0
12	GONCREX SHEPESEE	0	0	0	0	0	0	0	0	0
13	GOPPEON MUGECE	0	0	0	0	0	0	0	0	0
14	PEVENO MEGHUNGU	0	1	1	1	1	1	1	0	0
15	MOPVEN MMPONO	1	0	1	1	1	1	1	0	1
16	MECCHOPPO GHEMOCO	1	1	1	1	1	1	1	1	1
17	NECHEN OPOCO	0	0	0	0	0	0	0	0	0
18	PEUPENO EYOCE	0	1	1	1	1	1	1	1	0
19	PROCEOUS MUCOPE	0	0	0	1	1	1	0	0	0
20	RODOMPCE GHECEPE	1	1	1	0	1	1	1	1	1
21	ROONOY SHESEGHE	0	0	0	0	0	0	0	0	0
22	SEMSON WONDO	0	1	0	1	1	1	1	1	0
23	SHENCOP OSCHOR	0	0	0	0	1	0	0	1	0
24	SHOEPE GENEEZE	1	1	1	1	1	1	1	0	1
25	CREZE MUCEMPE	1	1	1	1	1	1	1	1	1
26	VEONU GHUCEUPU	1	1	1	1	1	1	1	1	1
27	WEPSON MUHUCEU	0	1	1	1	0	1	1	1	0
	Total of 1's	11	13	14	16	17	16	15	14	11
	Proportion who passed, p	0.407	0.484	0.519	0.593	0.63	0.593	0.556	0.519	0.407

➢ *Calculation of Q*

The value of q for each subject shows the proportion of students who failed the subject. Since the students either passed or failed the subject, the sum of p+ q = 1.0 as shown in Equation 3(a). Hence the value of q is obtained from Equation (3b), as q = 1 - p.

$$p + q = 1.0 \quad \text{Equation (3a)}$$

$$q = 1.0 - p \quad \text{Equation (3b)}$$

Table 4 Shows the Value of Q for Each Subject

NO	NAME	ENG	MATH	KISW	SST	INTE SCI	CRE	PRETECH	AGR/HSCI	C/ARTS
	Total of 1's	11	13	14	16	17	16	15	14	11
	Proportion who passed, p	0.407	0.484	0.519	0.593	0.63	0.593	0.556	0.519	0.407
	Proportion who failed, q	0.593	0.516	0.481	0.407	0.37	0.407	0.444	0.481	0.593

➢ *Summation of Product of P and Q*

From Table 4 the value of p is multiplied by the value of q. The values are then added to get the sum of product pq as shown in Table 5.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct1276

Table 5 Sum of Product of P and Q

NAME	ENG	MATH	KISW	SST	INTE SCI	CRE	PRETECH	AGR/HSCI	C/ARTS
Total of 1's	11	13	14	16	17	16	15	14	11
Proportion who passed, p	0.407	0.484	0.519	0.593	0.63	0.593	0.556	0.519	0.407
Proportion who failed, q	0.593	0.516	0.481	0.407	0.37	0.407	0.444	0.481	0.593
Product of p x q	0.241	0.250	0.250	0.241	0.233	0.241	0.247	0.250	0.241

Sum of the product of p and q, $\sum pq$, is computed as follows in Equation 4 using values in Table 5.

$$\sum pq = 0.241 + 0.250 + 0.250 + 0.241 + 0.233 + 0.241 + 0.247 + 0.250 + 0.241 = 2.194 \quad \text{Eqn (4)}$$

➢ *Calculation of Standard Deviation and Variance* □□
Table 6 shows how to calculate the variance, □□□.

- Step 1 – The total of1's for each student were calculated as shown in the TOTAL column. For example student No. 1 has a total of 7 and student No. 23 has a total of 2.
- Step 2 – The sum of all totals column is calculated to give 135.
- Step 3 – The mean is computed as 135/27 = 5.
- Step 4 – The DIFF column is then computed being the mean in step 3 subtract the total in Step 1 for each student. For example for student number 1, the DIFF = 5 – 7 = -2 and for student number 23 the DIFF = 5-2 =3.
- Step 5 – Square the values to obtain the last column of Table 5. For student number 1 the squared value is 4 while for student number 23 the squared value is 9.
- Step 6 – The sum of squares gives $\sum\left(\bar{X} - DIFF\right)^2 = 300$ Table 5. Thus, for ungrouped data the standard deviation, □□ is given by:

$$\sigma = \sqrt{\frac{\sum\left(\bar{X} - DIFF\right)^2}{n}} \quad \text{Equation 5(a)}$$

$$\sigma = \sqrt{\frac{300}{27}} \quad \text{Equation 5(b)}$$

$$\sigma = \sqrt{11.1111} \quad \text{Equation 5(c)}$$

$$\sigma = 3.3333 \quad \text{Equation 5(d)}$$

Table 5 Calculation of Variance□□

NO	NAME	ENG	MATH	KISW	SST	INTE SCI	CRE	PRETECH	AGR/HSCI	C/ARTS	TOTAL	DIFF	$\left(\bar{X} - DIFF\right)^2$
1	ENGEPECE EGENYE	0	1	1	1	1	1	1	1	0	7	-2	4
2	EUSCENE SHEGOGOCE	0	0	1	1	1	1	1	1	0	6	-1	1
3	PREVEN CENYERENGE	1	0	0	1	1	0	0	0	1	4	1	1
4	PREGEC MEDEVE	0	0	0	0	0	0	0	0	0	0	5	25
5	CEPESCUS MEGUMPE	1	1	1	1	1	1	1	1	1	9	-4	16
6	CEPVEN OCHEPO	0	0	0	0	0	0	0	1	0	1	4	16
7	DAMARES GHACEEVE	0	0	0	0	0	0	0	0	0	8	-3	9
8	FRENGPENO MEPEPE	1	1	1	1	1	1	1	0	1	8	-3	9
9	GPECEOUS EDUVUGHE	1	1	1	1	0	1	1	1	1	8	-3	9
10	GPOREE PUGEPEGE	1	0	0	0	0	0	0	1	1	3	2	4
11	ENNOCONC GEPGOMPOE	0	0	0	0	1	0	0	0	0	1	4	16
12	GONCREX SHEPESEE	0	0	0	0	0	0	0	0	0	0	5	25
13	GOPPEON MUGECE	0	0	0	0	0	0	0	0	0	0	5	25
14	PEVENO MEGHUNGU	0	1	1	1	1	1	1	0	0	6	-1	1
15	MOPVEN MMPONO	1	0	1	1	1	1	1	0	1	7	-2	4
16	MECCHOPPO GHEMOCO	1	1	1	1	1	1	1	1	1	9	-4	16
17	NECHEN OPOCO	0	0	0	0	0	0	0	0	0	0	5	25
18	PEUPENO EYOCE	0	1	1	1	1	1	1	1	0	7	-2	4
19	PROCEOUS MUCOPE	0	0	0	1	1	1	0	0	0	3	2	4
20	RODOMPCE GHECEPE	1	1	1	0	1	1	1	1	1	8	-3	9
21	ROONOY SHESEGHE	0	0	0	0	0	0	0	0	0	0	5	25
22	SEMSON WONDO	0	1	0	1	1	1	1	1	0	6	-1	1
23	SHENCOP OSCHOR	0	0	0	0	1	0	0	1	0	2	3	9
24	SHOEPE GENEEZE	1	1	1	1	1	1	1	0	1	8	-3	9
25	CREZE MUCEMPE	1	1	1	1	1	1	1	1	1	9	-4	16
26	VEONU GHUCEUPU	1	1	1	1	1	1	1	1	1	9	-4	16
27	WEPSON MUHUCEU	0	1	1	1	0	1	1	1	0	6	-1	1
	Total										135		300

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct1276

➢ *Calculation of KR-20 Reliability Index*

All the values in Equation 1 are now known.

$$r_{KR20} = \left(\frac{k}{k-1}\right).\left(1 - \frac{\sum pq}{\sigma^2}\right)$$

k = 9

$\sum pq$ = 2.194 from equation 4

□□□ = 3.3333 x 3.3333 = 11.111 from Step 6

By substituting the values in Equation 1we obtain the r_{KR20} as follows:

$$r_{KR20} = \left(\frac{9}{9-1}\right).\left(1 - \frac{2.194}{11.111}\right) \qquad \text{Equation 6(a, b, c, d)}$$

$$r_{KR20} = \left(\frac{9}{8}\right).(1 - 0.1975)$$

$$r_{KR20} = 1.125x0.8025$$

$$r_{KR20} = 0.9028$$

IV. DISCUSSION

Split-Half Reliability Test by calculating the Kuder-Richardson formula 20 (KR-20) was applied on Grade 7 end term test results for Ekato Junior Secondary School in Western Kenya in 9 subjects under the Competency Based Education (CBE) assessed by multiple-type questions. The purpose for the test was to evaluate the reliability of the testing instrument on the same group in 9 subjects. The result showed an index, rKR20 = 0.9028.

The KR-20 is a measure of internal consistency reliability for tests or assessments that have dichotomous items (items scored as either 0 or 1, (such as true/false, pass/fail, correct/incorrect). KR-20 reliability coefficients of less than 0.50 are considered low, between 0.50 and 0.80 moderate, and above 0.80 high (Tan, 2009).

Table 6 Interpretation of Results (Lenke et. al.,1977)

Reliability Range	Interpretation
0.90 and above	Excellent Reliability at the level of best standardized test.
0.80- 0.89	Very good for a classroom test
0.70 – 0.79	Good for classroom test, in the range of most. There are probably a few items which could be improved.
0.60 – 0.69	Somewhat low. The test needs to be supplemented by other measures (such as more tests) in order to determine grades.There are probably some items which could be improved.
0.50 – 0.59	Suggests need for revision of the test, unless it is quite short (ten or fewer items). The test definitely needs to be supplemented by other measures such as more tests for grading.
0.49 and below	Questionable reliability. This test should not contribute heavily to the course grade and it needs revision.

V. CONCLUSIONS

The results of KR-20 test on the multiple-type questions used for the Grade 7 end-term test at Ekato JSS, indicate a high reliability estimate of 0.9028 which is just above the thresh hold of 0.90 (Table 6). A generally accepted threshold for high reliability is 0.8 or above (Tan, 2009). The value of 0.9028 indicates that the multiple-type questions as a test is consistent in measuring the same construct, which is a desirable property for a reliable assessment.

RECOMMENDATIONS

It is recommended that the KR20 test of reliability should be applied to subjects in categories of literacy, numeracy and creativity to assess the appropriateness of the adopted test instruments in the determination of competencies in the different career pathways under the CBE system of education in Kenya.

REFERENCES

[1]. Ayiera Arlene (2023), Kenya needs to invest in advancing lieracy and numeracy in early grades,
[2]. Teaching at the Right Level Africa (TaRL Africa), https://teachingattherightlevel.org
[3]. Eun-Hyun Lee, Eun Hee Kang, Hyun-Jung Kang (2020), Evaluation of Studies on the
[4]. Measurement Properties of Self-Reported Instruments, Asian Nursing Research, Volume 14, Issue 5, December 2020, Pages 267-276
[5]. Lenke, Joanne M. et.al. (1977), Differences Between Kuder-Richardson Formula 20 and Formula 21 Reliability Coefficients for Short Tests, SOB DAM Apr-77, IOTI 10p.j Paper presented at the Annual fleeting of –the American Educational Research Association (61st, BexJork, Be* Tort* April 4-8, 1977)
[6]. Tan, H.H (2009), Redescription of Betta anabatoides Bleeker; and a new species of Betta from West Kalimantan, Borneo (Teleostei: Osphronemidae). Zootaxa 2165:59-68.