Original Research Article

Determining the Psychometric Properties of Non-Verbal Intelligence Test (NVIT) Among Pupils Using Classical Test Theory in Rivers State

Ogidi Reuben C[*]

Department of Educational Psychology, Guidance and Counselling Faculty of Education, Ignatius Ajuru University of Education Rumuolumeni, Port Harcourt, Nigeria

Abstract: In test development, there is the need for accurate estimation of item difficulty, discrimination and distracter indices of the test especially in the classical test theory. The study adopted instrumentation research design. The sample of the study was 1000 drawn from students in public junior secondary school one in Rivers West Senatorial District, Rivers State. The sample was drawn through multi stage sampling procedure. The instrument for data collection was a 50 item non-verbal intelligence test developed and standardized by the researcher in 2007. The instrument yielded a reliability coefficient of .92 and was determined through the use of Kuder-Richardson estimate$_{21}$. Results of the study showed that most of the items of the non-verbal intelligence test possessed moderate difficulty index and discrimination index and distracter index. Based on the results, the researcher recommended that there should be more awareness on the development of non-verbal intelligence test and the regular training of test developers on intelligence testing generally and non-verbal intelligence test in particular.

Keywords: Non-verbal intelligence test (NVIT), intelligence test, difficulty index, discrimination index, distracter index.

INTRODUCTION

A test is essentially an instrument that indicates the presence or absence of a trait, behaviour, quality, attribute, characteristics etc possessed by an individual or group of individuals that are of interest to the tester. It is a series of task presented to an individual(s) in order to determine the behaviour sought for by the examiner. The world book encyclopedia in Ukuije and Opara (2012) assert that test is an attempt to measure a person's knowledge, intelligence and other characteristics in a systematic way. Shertzer and Stone in Inko-Tariah and Ogidi (2017) posit that test is an objective sample of some aspect of behaviour. In this regard, the objectivity refers to the requirements that are involved in its administration; scoring and interpretation are independent of the individual examiners' subjective judgement. Also, Ipaye in Inko-Tariah and Ogidi (2017) view test as an act of exposing an individual or group of individuals to series of questions, exercises, tasks etc to measure their skills, knowledge, intelligence, general ability and aptitude in an attempt to obtain a score. It is a measurement and evaluation instrument or device for measuring skills, knowledge, intelligence, achievement, ability, aptitude, attitude, interest, motivation, competence etc.

Anikweze (2018) posit those good tests are useful in discovering both the unseen and the un-seeable potentials in individuals as well as their deficiencies. He further explained that the potentials may be attitudes, developmental levels, social values, intelligence or certain capabilities possessed by the individuals. Test are relatively structured or controlled. This explains why testing conditions rules and regulations are often stated including timing of the test, specified sitting arrangement, materials that are disallowed and other rubrics of the test.

One of the classifications of test is the intelligence test. Ushie in Chikwe (2017) explained that intelligence test are tests used in determining students' overall brain power or intellectual or general mental ability to learn. Such test measure learners' general mental ability, abstract reasoning or ability to understand the complex relationship between objects, views, situations etc. Ogidi (2007) posit that intelligence test also called mental ability test are psychological test designed to measure an individual's aptitude for scholastic work or other kinds of occupations requiring reasoning, verbal ability or other kinds of occupations requiring reasoning, verbal ability or the manipulation of objects. The following are the

uses of intelligence test: (i) The items on intelligence tests represent attempts to assess individual differences in the effects of experiences common to nearly everyone in the culture. (ii) It enable people to resolve difficulties and make decisions about their lives. (iii) The information gathered from intelligence test could play a useful role in diagnosing strengths and weaknesses, thus suggesting ways for the individual to improve his or her performance. (iv) Intelligence test is used in selecting people for a particular station in life. This may involve an academic decision such as identifying students who have the aptitude for a particular course of study. (v) It can also be used for placement in schools. (vi) Scores from intelligence tests are also used for research purposes.

One of the categorization of intelligence test is the non-verbal intelligence test. Inko-Tariah and Ogidi (2017) assert that non-verbal intelligence test refer to test that require the ability of an individual to process information, reason abstractly and solve problems without word expression. Walsh (2014) explains that non-verbal intelligence test are test that do not require language. Onunkwo (2002) informed that in such test, figures of various forms are involved. Non-verbal test do not require verbal response (Aminu, 2015). They are made up of figures and shapes of different forms. Ogidi (2007) assert that non-verbal test is developed without verbal symbolism. However, the figural content of non-verbal intelligence test are concrete material such as perceived through the senses. It does not represent anything except itself. Non-verbal intelligence test have a lot of benefits which are (i) it eliminates cultural bias (ii) it is a more valid long term measure of school potential (iii) it is also a legitimate aid in determining the range and strength of an individual's cognitive abilities (iv) it taps a set of thinking skills basic to intellectual functioning and measures general intelligence (v) it is handy while testing children (vi) it is useful in studying children having language limitations.

Amadi *et al*. (2019) assert that testing procedures and tools should be reliable within the framework of a test theory or model of application. Testing procedures determines the credibility of the scores provided by items and the validity aids in estimating the abilities of the testees.

Test developers make use of two models in analyzing the characteristics of the responses of testees to the test items. These are the classical test theory and the item response theory. However, there are differences in determining the quality of items using classical test theory and item response theory – two approaches commonly used to determine the quality of test items. Mittee (2019) assert that classical test theory also called (true score model) is a test theory which postulates that the observed or obtained score of an examinee on a test is the sum of two unobserved scores (true score which is error free and an error score). On the other hand, the item response theory is a test theory that is interested in the relationship between the ability of an examinee (latent or hidden variable and the probability that the testee will answer an item correctly. IRT links the latent trait in the testees to some observable characteristics giving each testee a numerical value or score on an ability scale. This study considered item analysis of the non-verbal test using classical test theory.

It is important that test possess good psychometric properties including the non-verbal intelligence test. The psychometric properties of a test include validity, reliability and usability. Shimberry (2014) assert that validity is the level of confidence with which an examinee's test score could be used to infer the ability under measurement possessed by the examinee. A valid test must satisfy both the eudiometric and the psychometric function of evaluation. Anikweze (2018) posit that reliability indicates the degree of accuracy with which a test measures what it is designed to measure. Thus, it implies consistency of test results over time and item. Usability refers to the fitness of a test to be applied on a particular set of individuals. Asuru (2015) suggest that usability refers to the extrinsic or practical or non-technical factors that affect the quality of a test. Related to validity, reliability and usability of a test is the item analysis of test.

Item analysis is the process of determining the psychometric qualities or features of each item that make up the entire test. In the classical test theory, the item analysis involves determining the following coefficients: (i) item difficulty which is also called item facility (IF). It refers to the percentage or proportion of persons answering each item correctly. It gives an estimate of the proportion of students likely to answer an item correctly in subsequent testing periods, assuming that the original group of students from which the index was computed is identical to the present group taking the test (Onunkwo, 2002). The higher the difficulty index, the easier the item and vice versa.

Item discrimination index is the difference between the proportion of students who passed on item in the upper ability groups and the proportion of those who passed the item in the lower ability group? It is the ability of a test to discriminate between high ability students and low ability students. Micheal and Karnes in Onunkwo (2002) explain that this power of a test is to be used for ranking students on the basis of achievement. Kpolovie (2010) posit that discrimination index is the extent to which an item of a test accurately differentiates between learners in the upper ability and low ability group. A negative discrimination index indicates items that tend to penalize students in the upper ability group. Amadi *et al*. (2019) explain that discrimination index is the best measure to determine the extent to which an item differentiate between the

performance of students in the upper ability group and those from the low ability group. Distracter index refer to the extent to which the options of a test are plausible or appeal to the dull students rather than the brilliant students. This indicates that dull students tend to the options or distracters while the brilliant students tend towards the key.

Anyanwale (2018) suggested that under CTT items with difficulty (D) of less than 0.3 and greater than 0.8 and discrimination indices (R) of less than 0.2 should be deleted. This is because items with difficulty index of 0.8 are easy while items with index of 0.2 are too difficult. However, item difficulty index (D) of 0.50 is considered moderate. Also Onunkwo (2002) discrimination index (R) of 0.3 is acceptable. Opera (2016) posit that a distracter index with positive value is acceptable.

Most of the studies that determined the item analysis of test focused on achievement test. Amadi *et al*. (2019) estimated the psychometric properties of 50 multiple choice Mathematics item for senior secondary II students in Ikwerre Local Government Areas of Rivers State using classical test theory. Also, Mittee (2019) carried out a comparative study of classical test theory and item response theory using item analysis results of quantitative chemistry achievement test. However, this present study determined the psychometric properties of non-verbal intelligence test (NVIT) among students in junior secondary schools one (Basic 7) using classical test theory in Rivers State.

OBJECTIVES OF THE STUDY
The major objective of this study is to determine the psychometric properties of non-verbal intelligence among pupils in Rivers State. Specifically, the study sought the following:
1. Determine the difficulty index of the non-verbal intelligence test.
2. Determine the discrimination index of the non-verbal intelligence test.
3. Determine the distracter index of the non-verbal intelligence test.

METHODS
The study adopted instrumentation research design. This design involves the development, estimation of difficulty, discrimination and distracter indexes of research instrument in non-verbal intelligence test. Multi stage sampling procedure was adopted in the selection of the sample for the study. Simple random sampling method was used in selecting five local government areas out of the eight local government area that make up the Rivers West senatorial districts. Forty (40) JSS1 students were selected from five public senior secondary schools in each local government in the district, totaling 1000 students for the study. The instrument for data collection was developed and standardized by the researcher in 2007. The instrument contained 50 items, scored over 100 percent (each item scored right attracted 2 marks). The instrument was pilot tested on forty students who did not take part in the study. The internal consistency of the instrument was determined using Kuder-Richardson21. The instrument yielded a reliability coefficient of .92, indicating that the instrument was quite reliable for use in the study. The Administration of the test lasted for 30 minutes. The performance (scores) of the students were arranged 51-100 for the upper ability group while 1-50 for the lower ability group. 27% of the scripts from the upper and lower ability group was randomly selected and consisted of 42 scripts. For the difficulty index, items with difficulty index of .3 and .70 were deemed to be good items. Also, for the discrimination index, items with discrimination index ranging from .3 and .49 was deemed to discriminate effectively while for the distracter index, items with positive value was deemed to have distract effectively. Manual method was used in determining the difficulty, discrimination distracter index of the instrument.

RESULTS
Research Question One
Which items estimate effectively the difficulty index on students' performance in the non-verbal intelligence test?

Table-1: Difficulty Index of the Students Performance on the Non-Verbal Intelligence Test

S/N	Upper	Lower	Difficulty Index (D)	Total
1	26	10	0.86	42
2	21	8	0.69	42
3	22	12	0.80	42
4	8	10	0.42	42
5	16	9	0.60	42
6	18	7	0.60	42
7	20	9	0.69	42
8	24	12	0.86	42
9	10	4	0.33	42
10	19	13	0.76	42
11	16	12	0.67	42
12	25	11	0.87	42
13	22	10	0.76	42
14	8	6	0.33	42
15	5	4	0.21	42
16	22	9	0.74	42
17	15	10	0.60	42
18	18	8	0.62	42
19	22	11	0.79	42
20	24	12	0.86	42
21	8	5	0.31	42
22	21	8	0.69	42
23	20	9	0.69	42
24	18	10	0.67	42
25	22	8	0.71	42
26	19	12	0.74	42
27	18	11	0.69	42
28	21	10	0.74	42
29	23	12	0.83	42
30	19	9	0.67	42
31	21	8	0.69	42
32	6	3	0.21	42
33	7	5	0.29	42
34	18	10	0.67	42
35	20	8	0.67	42
36	22	9	0.74	42
37	21	10	0.74	42
38	8	2	0.23	42
39	9	4	0.31	42
40	26	12	0.90	42
41	18	9	0.64	42
42	17	11	0.67	42
43	16	10	0.62	42
44	5	3	0.19	42
45	22	13	0.83	42
46	21	8	0.69	42
47	18	12	0.71	42
48	17	13	0.71	42
49	16	15	0.74	42
50	18	9	0.64	42

Table 1 shows that item 2, 4, 5, 6, 7, 9, 10, 11, 13, 14, 16, 17, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 34, 35, 36, 37, 39, 41, 42, 43, 46, 47, 48, 49, 50 were retained. On the other hand, items were 1, 3, 8, 20, 29, 40 and 45 were rejected because, it was too easy. In addition, items 15, 32, 33, 38, 44 were rejected because the difficulty level of these items was too high. Thus while 38 items were retained, 12 items were rejected because it was either too difficult or too easy.

Research Question Two

Which items discriminate on students' performance in non-verbal intelligence test?

Table-2a: Items Discrimination Index of the Students Performance on the Non-Verbal Intelligence Test

S/N	Upper	Lower	Discrimination Index (R)	Total
1	26	10	0.76	42
2	21	8	0.62	42
3	22	12	0.48	42
4	8	10	-0.10	42
5	16	9	0.33	42
6	18	7	0.52	42
7	20	9	0.52	42
8	24	12	0.57	42
9	10	4	0.29	42
10	19	13	0.29	42
11	16	12	0.19	42
12	25	11	0.67	42
13	22	10	0.57	42
14	8	6	0.10	42
15	5	4	0.05	42
16	22	9	0.62	42
17	15	10	0.24	42
18	18	8	0.48	42
19	22	11	0.52	42
20	24	12	0.57	42
21	8	5	0.14	42
22	21	8	0.62	42
23	20	9	0.52	42
24	18	10	0.38	42
25	22	8	0.67	42
26	19	12	0.33	42
27	18	11	0.33	42
28	21	10	0.52	42
29	23	12	0.52	42
30	19	9	0.48	42
31	21	8	0.62	42
32	6	3	0.14	42
33	7	5	0.10	42
34	18	10	0.38	42
35	20	8	0.57	42
36	22	9	0.62	42
37	21	10	0.52	42
38	8	2	0.29	42
39	9	4	0.24	42
40	26	12	0.67	42
41	18	9	0.43	42
42	17	11	0.29	42
43	16	10	0.29	42
44	5	3	0.10	42
45	22	13	0.43	42
46	21	8	0.62	42
47	18	12	0.29	42
48	17	13	0.19	42
49	16	15	0.05	42
50	18	9	0.43	42

Table-2b: Discrimination Index (R) of the Items on Students Performance in Non-Verbal Intelligence Test

Items on very high positive discrimination 0.6 and above	items on high positive discrimination 0.50 -0.59	Items on moderate positive discrimination 0.30-0.49	Items on borderline positive discrimination 0.20-0.29	Items on lo positive discrimination 0.00-0.19	Items on negative dsicriminaiton- 0.1 and below
10	11	10	8	10	1

Tables 2a and 2b reveals that 10 items are on very high positive discrimination, 11 items are on high positive discrimination, 10 items are on moderate positive discrimination, 8 items are on borderline positive discrimination, 10 items are on low positive discrimination while 1 item was on negative discrimination. Thus, 31 items had acceptable discrimination power, 8 items are on borderline positive discrimination power while 11 items are rejected based on its discrimination power.

Research Question Three

Which items had good distracter index on the students' performance on the non-verbal intelligence test?

Table-3: Distracter Index of the Students' Performance on the Non-Verbal Intelligence Test

S/N	A			B			C			D			E			N
	U	L	DI	U	L	DI	U	L	DI	U	L	DI	U	L	DI	
1			*	0	2	0.04	0	0	0.00	0	1	0.02	0	0	0.00	42
2	2	3	0.02	0	0	0.00	1	1	0.00	0	0	0.00			*	42
3	0	0	0.00	1	2	0.02	0	0	0.00			*	0	1	0.02	42
4	0	1	0.02	0	3	0.06	0	3	0.06			*	0	5	0.10	42
5	1	3	0.04	1	2	0.02	0	0	0.00	0	0	0.00			*	42
6	0	0	0.00	0	2	0.04	0	1	0.02			*	0	1	0.02	42
7	1	2	0.02	0	1	0.02			*	0	0	0.00	2	2	0.00	42
8	0	1	0.02	0	1	0.02			*	1	2	0.02	1	2	0.02	42
9	0	6	0.12	0	9	0.18			*	0	1	0.02	0	3	0.06	42
10	1	2	0.02	0	0	0.00	0	1	0.02	0	0	0.00			*	42
11	0	0	0.00			*	4	12	0.16	1	8	0.14	0	4	0.08	42
12	4	8	0.08			*	3	6	0.06	0	0	0.00	1	1	0.00	42
13	1	2	0.02	2	4	0.04	2	1	-0.02			*	0	0	0.00	42
14	0	4	0.08	1	8	0.14	0	5	0.10			*	0	6	0.12	42
15	0	3	0.06	2	10	-.16			*	0	7	0.14	1	6	0.10	42
16	0	12	0.24	0	2	0.04			*	0	0	0.00	0	8	0.16	42
17	1	4	0.06	0	4	0.08	0	3	0.06	2	14	0.24			*	42
18	6	7	0.02	0	2	0.04	8	5	-0.06			*	0	0	0.00	42
19	1	1	0.00	3	1	-0.04	6	4	-0.04	2	1	-0.02			*	42
20	1	1	0.00	1	5	0.08	1	5	0.08			*	2	10	0.16	42
21	4	5	0.02	3	4	0.02	3	5	0.04			*	2	1	0.02	42
22			*	3	12	0.18	0	3	0.06	0	2	0.04	0	3	0.06	42
23	0	5	0.10			*	1	7	0.12	0	4	0.08	0	5	0.10	42
24	3	9	0.12	0	1	0.02			*	1	6	0.10	1	4	0.06	42
25	0	7	0.14	0	3	0.06	0	1	0.02	0	0.10	0.20			*	42
26	0	2	0.04			*	2	11	0.18	1	6	0.10	0	7	0.14	42
27	7	3	-0.08	9	5	-0.08			*	1	1	0.00	1	1	0.00	42
28	0	1	0.02	0	2	0.01	1	12	0.22			*	0	5	0.10	42
29	0	0	0.00	4	7	0.06	5	9	0.08			*	5	11	0.12	42
30	0	2	0.04	0	4	0.08	0	12	0.24			*	0	3	0.06	42
31	1	7	0.12	0	2	0.04	2	11	0.18	0	3	0.06			*	42
32	0	6	0.12	0	1	0.02	1	8	0.14			*	0	10	0.20	42
33	5	5	0.00			*	0	0	0.00	0	0	0.00	0	0	0.00	42
34	0	2	0.04	0	9	0.18	0	5	0.10			*	0	3	0.06	42
35			*	4	5	0.02	6	3	-0.06	4	5	0.02	5	4	-0.06	42
36	0	6	0.12	1	9	0.16			*	0	0	0.00	0	3	0.06	42
37	0	5	0.10	0	6	0.12			*	0	7	0.14	0	5	0.10	42
38	0	4	0.08			*	0	2	0.04	0	2	0.04	0	8	0.16	42
39	0	2	0.04			*	2	9	0.14	1	8	0.14	1	4	0.06	42
40	2	7	0.10	0	3	0.06	0	2	0.04	5	15	0.20			*	42
41	1	6	0.10			*	0	3	0.06	1	7	0.12	0	3	0.06	42
42	0	6	0.12			*	1	8	0.14	0	2	0.04	0	1	0.02	42
43	0	8	0.16			*	0	4	0.08	4	15	0.22	0	1	0.02	42
44	0	5	0.10			*	0	6	0.12	0	2	0.04	1	8	0.14	42
45	4	16	0.24			*	0	2	0.04	3	7	0.08	0	2	0.04	42
46	1	8	0.14	0	2	0.04			*	0	1	0.02	0	8	0.16	42
47	1	12	0.22	0	4	0.08			*	0	6	0.12	0	3	0.06	42
48			*	0	5	0.10	0	6	0.12	0	7	0.14	0	3	0.06	42
49			*	2	3	0.02	0	0	0.00	0	0	0.00	10	4	-0.12	42
50	3	6	0.06	1	4	0.06	1	5	0.08	0		*	0	5	0.10	42

The asterisk indicates the key of the items.

Positive distracter	Negative distracter	Zero Distracter
Item	Item	Item
1b, 1d, 2a, 3b, 3E, 4A, 4B, 5A, 5B, 6B, 6C, 6E, 7A, 7B, 8A, 8B, 8D, 8E, 9A, 9B, 9C, 9D ETC	13C, 15B, 18C, 19C, 27C, 35C, 49E	1C, 1E, 2b, 2c, 2d, 3c, 4c, 4E, 5C, 5D, 6A, 7D, 7E ETC

DISCUSSION

The result of the study showed that for the 50 item non-verbal intelligence test, the difficulty index ranges from 0.19 of item 44 to 0.90 of item 40. Thus, 38 items were retained while 12 items were rejected due to either too high difficulty level or been too easy. Such items should therefore be discarded or modified. This result is in agreement with Ayanwale (2018) that items below or above 0.2 – 0.8 should be deleted.

The result of the study also showed that 31 items discriminated positively while 8 items were on the borderline. However, while 10 had low discrimination, 1 item showed negative discrimination. Thus, the items that showed low discrimination and negative discrimination should be deleted or modified. This is in agreement with Onunkwo (2002) that discrimination index of 0.30 and above is acceptable.

The result of the study revealed that the distracter index of the non-verbal intelligence test ranges from 0.00 to 0.24. In addition, while some of the distracters have positive values, some had zero while others showed negative values. The distracters with negative value are not effective. However, the distracters with negative values showed that it attracted students from the upper ability group rather than students from the lower ability group. Thus, the distracters with zero and negative values ought to be deleted or substituted. This is in agreement with Opara (2016) that distracters with zero or negative values should be deleted.

CONCLUSION

The major purpose of item analysis is to determine the characteristics of the test items. In the classical test theory (CTT), the item analysis involves determining the difficulty index, discrimination index and distracter index. This is because it is important that students who answered a test item rightly or not should be based on their mastery of such test item and nothing else. The information from the item analysis showed that some of the items of the non-verbal intelligence test are good while other items need to be deleted or modified.

RECOMMENDATIONS

Based on the results of the study, the researcher recommended as follows:

1. That test developers should verify the psychometric properties of test items in order to determine its characteristics and ascertain that it is of high quality.
2. Test developers should be trained and retrained in the procedures necessary in determining the difficulty, discrimination and distracter index of test items.
3. Test developers should be encouraged to develop indigenous standardized test. This will enhance and enriched the pool of information about testees when such tests are administered.

REFERENCES

* Amadi, B.D, Ogunka, R.I., & Long John, I.T. (2019). Estimating the psychometric properties of 50 multiple choice Mathematics item for senior secondary school II students in Ikwerre Local Government Area of Rivers State using classical test theory *AJB-SDR*, *1*(2), 24-33.
* Anikweze, C.M. (2018). *Measurement and evaluation for teacher education*. Shiloh Areas Associates.
* Asuru, V.A. (2015). *Measurement and evaluation in education and psychology* (2nd edition). Pearl publishers.
* Ayanwale, M.A. (2018). *Efficacy of item response theory in the validation and score ranking of Dichotomous and polychromous response Mathematics Achievement Test in Osun State*, unpublished Ph>.D thesis: University of Ibadan.
* Chikwe, C.K. (2017). *Fundamentals of test, measurement and evaluation in education*. Emmanest ventures.
* Inko-Tariah, D.C., & Ogidi, R.C. (2017). *Fundamentals of psychological testing for psychologists, counsellors and educationists*. Rodi printing and publishing company.
* Kpolovie, P.J. (2010). *Advanced research method*. Springfield publishers.
* Mittee, T.L. (2019). Comparative study of classical test theory and item response theory using item analysis results of quantitative chemistry achievement test. *AJB-SDR*, 1(1), 26-36.
* Ogidi, R.C. (2007). *Development and standardization of non-verbal intelligence test (NVIT)*. A thesis presented to the Department of Educational Psychology, Guidance and Counselling, Faculty of Education, University of Port Harcourt, Port Harcourt, Nigeria.

- Onunkwo, G.I.N. (2002). *Fundamentals of educational measurement and evaluation*. cape publishers.
- Opera, I.M. (2016). *Test construction and measurement: Concepts and applications.* Career publishers.
- Shimberg, G.R. (2014). *Flux (computer programme)*. McGraw-Hill.
- Ukwuije, R.P.I., & Opara, I.M. (2012). *Test and measurement for teachers.* 3rd edition, Chadik printing press.

Cite This Article: Ogidi Reuben C (2021). Determining the Psychometric Properties of Non-Verbal Intelligence Test (NVIT) Among Pupils Using Classical Test Theory in Rivers State. *EAS J PsycholBehavSci, 3*(4), 50-57.