International Journal of Education, Learning and Development
Vol. 9, No.6, pp.57-77, 2021
Print ISSN: 2054-6297(Print)
Online ISSN: 2054-6300 (Online)

DETERMINATION OF THE PSYCHOMETRIC PROPERTIES AND STANDARDIZATION OF REYNOLDS ADAPTABLE INTELLIGENCE TESTS-NONVERBAL (RAIT-NV) USING ITEM RESPONSE THEORY IN NIGERIA

Magnus-Arewa Eseose Anthonia

Department of Educational Psychology, Guidance & Counselling, Faculty of Education, University of Port Harcourt, PMB 5323, Choba, Rivers State, Nigeria.

Dr. G. W. Orluwene

Department of Educational Psychology, Guidance & Counselling, Faculty of Education, University of Port Harcourt, PMB 5323, Choba, Rivers State, Nigeria

ABSTRACT: *The study focused on the determination of the psychometric properties and Standardization of Reynolds Adaptable Intelligence Test (RAIT-NV) using Item Response Theories in Nigeria. The study used the triangulation research design. Five research questions guided the study. A sample of 2120 students was randomly drawn using multistage sampling techniques from a population of 14,107,456 of all the undergraduate, secondary and upper primary students in Nigeria. The instrument for the data collection is Reynolds Adaptable Intelligence Test (RAIT-NV). RAIT-NV reliability coefficient was 0.92 using Cronbach Alpha and the construct validity was 0.88. To answer the five research questions, data generated were analyzed using The X-Calibre 4.2, EIRT, SPSS, Microsoft excel, data was also analyzed using the IRT logistic model (2PLM), TIF, Factor analyses, Correlations, Analysis of variance, (ANOVA) z-score, T-score, normalized standard score and percentile ranks. Result showed that, under the IRT framework that RAIT-NV item difficulty graduates from very easy to very difficult, RAIT-NV items were able to discriminate between examinee high and those low on the trait been measured. RAIT-NV had satisfactory factor structure and had local independent, TIF was satisfactory. Percentile ranks, z scores, T scores, and Normalized standard score of RAIT-NV were established. The finding of the study showed that using IRT in revalidation and standardization of instrument provide reliable and valid instrument for measuring intelligence. Based on the findings, it was recommended among others that Teachers, test developers, psychologist, researchers and relevant educational agencies, should ensure that they establish the Z scores, T scores, Percentile rank and Normalized standard score of any instrument they are revalidating and standardizing for measuring intelligence. This is recommended to even out all the difference from age, gender, ethnicity and educational level associated with raw scores.*

KEYWORDS; intelligence test, item response theory, Reynolds adaptable intelligence test – nonverbal

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

INTRODUCTION

The relevance of establishing the extent of availability, magnitude, presence or absence of physical, social and psychological attributes has always remained a reoccurring decimal in our collective evolution. Mostly, this has resulted in varied speculations, theories and hypotheses on the appropriate terms, conceptualizations and operationalization of these attributes. While the physical domain of knowledge has achieved a fairly stable and relatively agreed upon standard of measurement, the study of psychological constructs, including intelligence and other non-cognitive abilities, do not follow similar rules of measurement. This has further resulted in the proliferation of theories aimed at capturing not only the essence, but also the applicability, of psychological constructs in everyday practice.

Despite, the debates and disagreements regarding the true nature of psychological constructs, there is a consensus that the measurement of psychological constructs must be reduced to specific indicators that represents the observation and documentations of such constructs. This feeble agreement, has informed, and continues to reforms, the field of psychology referred to as psychometrics. Within the field of psychology, and across related disciplines of education, economics, sociology, political science and management sciences, the need for the measurement and assessment of abstract constructs, has not only become timely, but pressing, especially with the advanced pace of globalization and involving trend of digital, economic and social evolution. Addressing this needs, requires a new set of tools, thinking, paradigms and frameworks, which might not only challenge existing assumptions, but also create new domains of knowledge for the ever dynamic world of the 21st century. It has therefore become pertinent that individuals, societies, and nations develop a robust system of education that not only equips them for future opportunities, but also identify their areas of strength and weakness.

Developing and reforming the educational system of a society, and helping students benefit optimally from the instruction, depends to a large extent on what the students know, as well as what they can gain proficiency on. It is therefore on this premise that globally, the measurement and assessment of intellectual skills has been a regular feature of education. Various skills and variables have been assessed within the purview of providing education for students including family background, school variables, teachers characteristics, motivation, self-efficacy, learned helplessness etc. While most of the variables have addressed the importance and perspective of ensuring optimal academic achievement among students and improving the general standard of education, investigations into the aspects that is most important, and how well to improve on such aspects have generated more questions than answers.

Within general research paradigm, it is clearly required that presentation of research output integrate a section for method of data collection, which often involves the use of valid and reliable instruments. Corroborating this position, Emekene (2017) stated that it is a well-established truism that no effective research can succeed without accurate, valid and reliable instruments. The implication of this position is that effective educational policy can take shape when the policy is

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

based on empirically verifiable and scientifically valid and reliable instruments, including that of intelligence.

With the ever increasing plethora of empirical and informal inquiry into the factors influencing educational development of individuals and societies, one common denominator that has received sustained interest is the concept of intelligence, ancient philosophers, such as Aristotle posits that all individuals possess general reasoning and judgmental intelligence, albeit to different degrees. Furthermore, Aristotle is of the opinion that the larger the brain size, the greater the intelligence of an individual. According To Kpolovie (2017), long before psychology became a distinct scientific field of study, intelligence was conceived by Galton as a product of hereditary and based on genetic factors. Many centuries later, Alfred Binet, the man credited with developing the first modern conception of intelligence testing, argued that intelligence is synonymous with the commonsense by stating that intelligence refers to judgment, good sense, and the faculty of adapting to one's self to circumstances. From his conceptualization of intelligence, Binet developed a test to measure the intellectual skills of French school children in 1904, a practice that has become the standard of intellectual assessment till this day (Orluwene, 2012).

Following the tradition of Binet and other scholars of the empirical study of intelligence, there is a surplus of various instruments designed to assess the intellectual skills of individuals generally and students specifically. While this might be considered superfluous, the process of scientific development, including that of education, demands that no practice is accepted as sacrosanct. Rather there should be a need for the periodic review of practices and assumptions guiding the policy and application of education. It is for this reason that the development and standardization of valid and reliable instrument for the assessment of intelligence has become a *sine qua non* for the advancement of education (Kpolovie & Emekene, 2016). However, following the standard method, it is important that the definition of intelligence is operationally situated for any meaningful instrument to be drawn from it.

The original RAIT was designed as an intelligence test that is feasible for administration for both individuals and groups. It has been standardized for administration with examinees ages 10 to 75 years. It is composed of seven subtest assessing crystalized, fluid, and quantitative intelligences. The full battery requires a total testing time of 50 minutes, with each subset having a maximum time limit. However, RAIT was prone to some other challenges that characterized -other tests including the challenge of administering on individuals with speech, visual and learning disability. In addition, the many subsets of original RAIT was hypothesized as having a confounding effect on the generally performance of students.

The developmental process involved in the standardization of the RAIT-NV was extensively robust which involve only statistical approaches, but also, expert judgment to identify items that were ambiguous, offensive, or bigotry. Statistical approaches adopted both classical test theory and item response theory for the identification and elimination of gender and ethnic item bias. The items in the instrument had ample comprehension flexibility which enables examinees to read

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

them or to be read to by others. The items also allow use of hand gestures for those who have hearing impairment (Reynolds, 2016).

The nature of the RAIT-NV was designed to provide an assessment known as Nonverbal Intelligence Index (NVII). The nonverbal intelligence provides a summary index of fluid intelligence (i.e., problem solving in the absence of requisite factual knowledge) assessed through nonverbal reasoning tasks that tend to invoke deductive rather than inductive reasoning. The NVII is calculated as the scaled sum of the T scores for the two nonverbal subsets which are Nonverbal Analogies (NVA) and Sequence (SEQ). The NVA subset requires examinees to complete a pictorial display of the relationship between two pictures when one picture is missing. Examinees are required to deduce the principle of the relationship and choose the picture that best completes the pictorial analogy this subset contains 52 items and expected to last for seven minutes only. Similarly, the SEQ subset seeks examinees understanding of verbal reasoning and sequencing along with deductive skills. The items in these subsets, totally 43, require examines to choose from a set of pictures, the picture that best completes a series or progression of change. The test has a time limit of 10 minutes.

While, IRT seems to be trend for the psychometric analysis of various items, it has been recommended by Petrillo, Cano, McLeod, and Coon (2005) that both frameworks (IRT and CTT) leads to better improved psychological assessments. According to them, IRT has the advantage of providing more detailed diagnostic information on how scales can be improved, CTT has the advantage of identifying problematic test items which threatens the validity of the overall scale scores, provides data on sets of redundant items and elimination of skewed response categories. Considering the benefits of both frameworks, the current study adopts the eclectic approach by integrating elements of Classical Test Theory and Item Response Theory in the Revalidation and Standardization of the Reynolds Adaptable Intelligence Test in Nigeria.

It has also been observed that many test users or researchers pay less attention to the importance of revalidation of an instrument, especially those developed in other culture. This unfortunately has yielded results which are celebrated erroneously instead of being questioned. It is also observed that majority of local researchers that lack the knowledge, pace and capacity to revalidate the psychometric properties of such instrument for use locally, adopts an existing instrument and use them, these abnormalities consciously or unconsciously exhibited by local researchers forms the gap through which the present study is based. Therefore considering the inherent challenges of intelligence testing in Nigeria and the serious implications involved in making decision based on intelligence testing in multiple domains, as well as expert recommendations on cross-cultural testing, this present study seeks to empirically determine the psychometric properties and as well the standardization of the Reynolds Adaptable Intelligence Test – Nonverbal for possible suitability, adaptability and utility within Nigeria using IRT.

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

Research Questions

The following research questions were answered to further guide the conduct of this study:

1. To what extent is local independence of the Reynolds Adaptable Intelligence Tests – Nonverbal (RAIT-NV) maintained in Nigeria using factor analysis.
2. What is the item difficulty and item discrimination index of the Reynolds Adaptable Intelligence Tests – Nonverbal (RAIT-NV) in Nigeria using the Item Response Theory?
3. What is the internal consistence of the Reynolds Adaptable Intelligence Tests – Nonverbal (RAIT-NV) using Cronbach alpha reliability?
4. What is the test information of the Reynolds Adaptable Intelligence Tests – Nonverbal (RAIT-NV) using Item Response Theory?
5. What is the z scores, T scores, Percentile rank and Normalized standard Nonverbal intelligence index of RAIT-NV subset in Nigeria?

METHODOLOGY

Triangulation research design was used for the study. It allows for a multi-method approach to studying related or intertwined phenomena. Reynolds Adaptable Intelligence Tests – Nonverbal were revalidated and using a multi-method approach using the multiple triangulations research design (RAIT-NV). The study was carried out in Nigeria across four geo-political zones. The study's population included all university undergraduates (1,794,989) in Nigeria's 92 public universities, as well as all students (4,758,739) in Nigeria's (upper primary) and secondary schools, for a total population of 7,553,728, bringing the total population to 14,107,456. (FRN: National Population Commission, 2019; Federal Ministry of Education, 2019).

The sample size of 2200 undergraduates, upper primary and secondary school students, males and females, whose ages ranged from ten to forty years old were spread across four main cultural groups (Hausa, Igbo, Yoruba, and Minorities), were used for the study. The current study's instrument is based on the Reynolds Adaptable Intelligence Examination (RAIT), which is a quick, accurate, and reliable test of nonverbal intelligence. It was developed utilizing the RAIT's two nonverbal subtests to provide a reliable assessment of fluid intelligence. Despite the fact that the RAIT-NV has a time limit, it is still a power test, not a timed test. Individuals with hearing problems, minimal language skills or no reading skills, motor coordination, or no visual-motor abilities can use the RAIT-NV, which reduces the confounds that might occur when manipulated items are used to test nonverbal intelligence. The RAIT-NV can be given to a person or a group. It can be employed in human resource and associated industrial settings, as well as in schools, juvenile and adult justice systems, and clinical settings.

The exam is intended to demonstrate consistency across a broad age range. Gender and ethnic bias were carefully tested, minimizing gender and ethnicity as confounds, which is especially essential for use with English as a second language (ESL) students and adults. Fluid intelligence is assessed

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

using two subtests. The Nonverbal Analogies and Sequences subtests feature vivid, enticing images to keep examinees interested. The RAIT Nonverbal Intelligence Index (NVII), which is scaled to the common IQ metric, is created by adding the two subtests together. The RAIT-NV is divided into two subtests, each of which is timed separately. The two subtests take a total of 17 minutes. Except when employing specified alternate administration instructions, the proctor must use a timer to keep track of the time restrictions for each subtest.

The validity of the above-mentioned instrument is well-known and widely acknowledged. The RAIT-NV was standardized using a population-proportionate, stratified random sampling plan based on 2010 U.S. Census population statistics on a sample of 2,124 people from 39 states. The test has a construct validity of 0.75 to 0.95 based on correlation with other tests (RAIT), (WISC-IV), (WAIS-IV), (RIAS), Wonderlic, (Beta III), (WRAT), and others (TIWRE).

According to Reynold (2016), test retest reliability ranges from 0.74 to 0.99 from ages 10 to 75, Cronbach alpha reliability ranges from 0.87 to 0.94 from ages 10 to 75, and alternate form reliability ranges from 0.85 to 0.94. Fifty pupils were used in a pilot test to ensure the instrument's dependability. The test retest reliability approach was used to determine the instrument's reliability. The test retest coefficients for RAIT-NV were r =.872, which was significant at the 0.05 level. This demonstrates that the RAIT-NV has a consistent test score. The 50 respondents' RAIT-NV scores were also exposed to Cronbach's Alpha, yielding a reliability coefficient of.794, indicating that the RAIT-NV has a high coefficient of stability and internal consistency.

Mean, standard deviation, IRT, CTT, 2PLM, Factor analysis, correlations (Cronbach alpha), normalized standard score, z- scores, k- scores, percentile rank, and other statistical procedures were used to analyze the data and answer research questions using statistical software packages such as X-Calibre, Statistical Package for Social Science (SPSS), and Microsoft EXCEL.

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

RESULTS

Research Question 1: To what extent is local independence of the Reynolds Adaptable Intelligence Tests – Nonverbal (RAIT-NV) maintained using factor analysis in Nigeria?

Table 1.1: Item local independence using Eigen values of factor analysis.

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	11.881	12.507	12.507	11.881	12.507	12.507
2	8.361	8.801	21.308	8.361	8.801	21.308
3	4.842	5.097	26.405			
4	4.372	4.602	31.007			
5	3.717	3.913	34.920			
6	3.084	3.246	38.166			
+		+	+			
+	+	+	+			
+	+	+	+			
+	+	+	+			
+	+	+	+			
+	+	+	+			
90	.199	.209	99.084			
91	.195	.205	99.289			
92	.187	.197	99.485			
93	.178	.187	99.673			
94	.159	.167	99.840			
95	.152	.160	100.000			

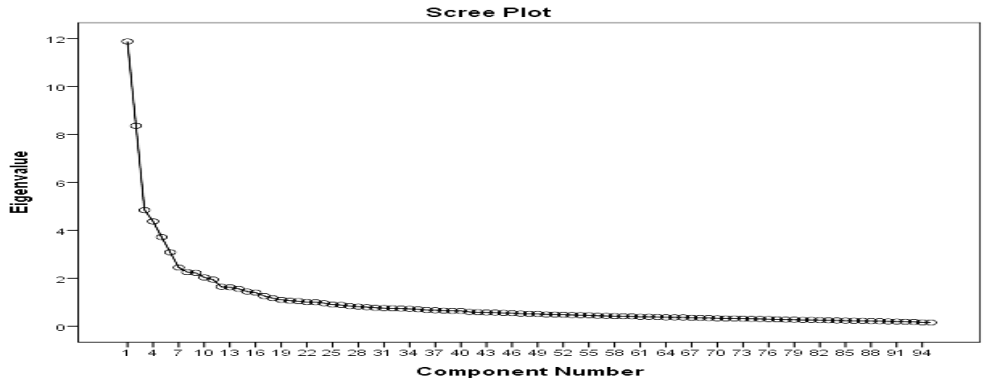

Figure 1.1: Scree plot of local independence of RAIT-NV

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

From Table 1.1 indicated that the highest eigenvalue is 11.88 and is for component one. This shows that the largest component explains 12.50% of the variance; it explains that the items contained in the test hang together on one distinct factor which is Intelligence.. This result revealed that the RAIT-NV fulfilled the assumption of unidimensionality and by extension local independence as the factor analysis results were in line with the set condition for assessing unidimensionality and local independence of items in a test. Also, dichotomous test items are unidimensional when the first factor loading for all items is significantly greater than 1 and when the first eigenvalue is substantially greater than the next, the result reveals the distance between the first eigen value of 11.88 and the next 8.36 eigen value as substantially greater than 1. This value therefore suggest that the assumption of unidimensionality is met by this model, which means that RAIT-NV has local independence.

In addition, a close scrutiny of the scree plot shown below shows that there is only one construct before the breaking point or elbow joint. This therefore succinctly shows the RAIT-NV is measuring intelligence. Therefore the underlining construct is effectively examined by the scale and it ensures its unidimensionality. Since the assumption of unidimensionality is met by this model, it invariably means that local independence holds.

Research Question 2: What is the item difficulty index and item discrimination indices of the Reynolds Adaptable Intelligence Tests – Nonverbal (RAIT-NV) using the Item Response Theory in Nigeria?

Table 1.2 Item difficulty and the item discrimination indexes of RAIT-NV within the IRT framework.

NVA Item	A	B	SEQ Items	A	B
1	0.138	-2.276	1	0.144	-2.120
2	0.212	-0.647	2	1.377	-0.197
3	0.185	-1.004	3	1.238	-0.419
4	0.243	-1.154	4	1.166	-0.441
5	0.826	-0.825	5	1.338	-0.179
6	0.279	-0.393	6	1.048	-0.229
7	0.746	-1.465	7	1.237	0.145
8	0.164	-1.619	8	0.848	-0.013
9	0.781	-1.159	9	1.128	-0.435
10	0.905	-0.975	10	0.854	-0.304
11	0.689	-0.828	11	0.930	-0.299
12	1.116	-1.053	12	0.504	0.276
13	1.091	-0.949	13	0.523	0.271
14	0.929	-1.004	14	0.744	0.637

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

15	1.168	-0.777	15	0.403	1.487
16	0.571	-0.516	16	0.523	0.703
17	0.740	-0.990	17	0.604	0.562
18	0.738	-1.011	18	1.119	0.223
19	1.093	-0.456	19	0.815	-0.154
20	0.347	-0.532	20	0.472	0.660
21	0.496	0.024	21	0.847	0.117
22	0.275	2.248	22	0.946	-0.016
23	0.486	-0.402	23	0.455	0.786
24	0.465	-0.646	24	0.560	0.648
25	0.165	3.000	25	0.448	1.706
26	0.439	0.333	26	0.276	2.246
27	1.128	-0.524	27	0.468	1.488
28	1.263	-0.621	28	0.329	1.589
29	0.948	-0.816	29	0.357	1.561
30	0.240	1.630	30	0.566	0.855
31	0.288	0.131	31	0.533	1.725
32	0.393	0.704	32	0.401	0.490
33	0.664	0.154	33	0.285	1.588
34	0.405	0.359	34	0.221	3.043
35	0.509	0.599	35	0.234	2.366
36	0.323	1.955	36	0.463	1.238
37	0.571	-0.057	37	0.355	2.231
38	0.489	0.732	38	0.199	3.000
39	0.783	0.015	39	0.186	1.906
40	0.479	0.156	40	0.655	0.668
41	0.548	0.229	41	0.238	2.845
42	0.364	2.583	42	0.470	1.396
43	0.487	1.332	43	0.201	3.000
44	0.153	3.000			
45	0.203	3.000			
46	0.339	1.777			
47	0.220	3.646			
48	0.298	1.303			
49	0.267	1.332			

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

50	0.291	1.366			
51	0.465	1.276			
52	0.144	-2.120			

The IRT parameters table presents the IRT item parameters. The "a" parameter index shows the discrimination of the items, as larger values for "a" will result in a greater slope of the IRF and indicate the item differentiates examinees well. A careful examination of the *a* parameter column both for the NVA and SEQ subset reveals that most of the test items discriminated well. The a values ranged from 0.1 to 1,4 for NVA and 0,1 to .1.9 for SEQ.

The "b" parameter is the item difficulty parameter and equals the location on the theta continuum where the probability of a correct response equals .50. It follows that multiple choice items with more positive "b" parameters are more difficult for examinees, as a higher trait level is required to endorse the keyed response 50% of the time. Higher *b* parameters (> 1.0) indicate that the item is more difficult; a value below -1.0 indicates that the item is very easy. According to X-Calibre manual, the difficulty index "ranges in theory from negative to positive infinity, but in practice from -3.0 (very easy) to +3.0 (very difficult)."

Just like in the CTT analysis, the table reveals for both the NVA and the SEQ subset of the RAIT-NV item difficulty that graduates from very easy to very difficult which reflects the design of the test by its constructor. A closes scrutiny of the *b* parameter column shows that the values of *b* for the NVA ranged from -2.12 to 3.00 and same for SEQ. The b parameter kept graduating in difficulty for the whole test.

Research Question 3; What is the internal consistence of the Reynolds Adaptable Intelligence Tests – Nonverbal (RAIT-NV) using Cronbach alpha reliability in Nigeria?

The table 1.3: The internal consistence of RAIT-NV using Cronbach alpha reliability

Scale	Cronbach's Alpha	N of Items
RAIT-NV	.920	95
RAIT-NVA	.873	.52
RAIT–SEQ.	.896	43

The table 1.3 above shows a Cronbach's Alpha reliability coefficient of .896 for the RAIT SEQ subset which is a very high coefficient, a reliability coefficient of .873 was obtained for the NVA subset which was also very high while for the whole test RAIT-NV, a reliability coefficient of .920. Another very high correlation coefficient. This reveals the internal consistency of the items and by implication the reliability of the test. In order words since a high reliability coefficient was obtained, it implies the items are internally consistent with themselves therefore reliability and by

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

extension has construct validity.

Research Question 4 What is the test information of the Reynolds Adaptable Intelligence Tests – Nonverbal (RAIT-NV) using Item Response Theory in Nigeria?

.

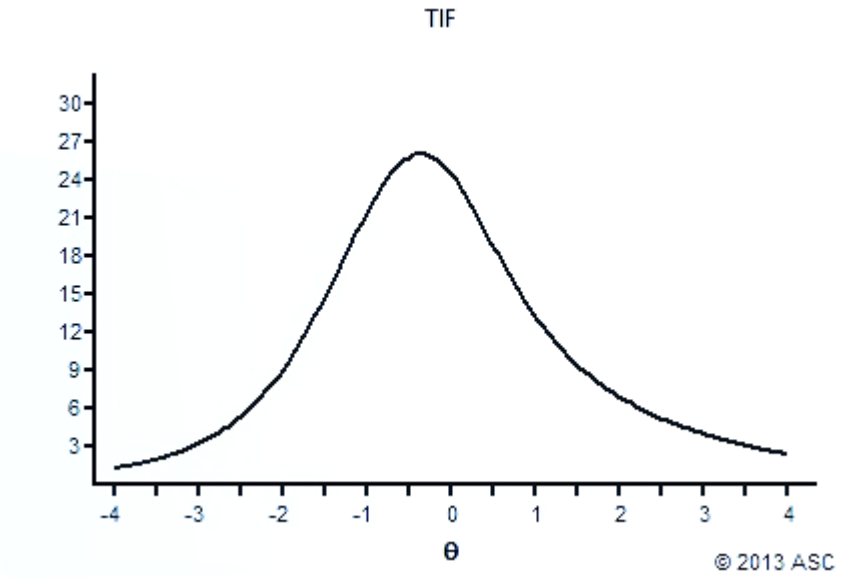

Figure 1.2: Test Information Function for all calibrated RAIT-NV items.

The figure above displays a graph of the Test Information Function for all calibrated items. The test information function sums up the item information functions to summarize where the test is providing information. The TIF is a graphical representation of how much information the test is providing at each level of theta. Maximum information was 26.097 at theta = -0.350. At the cut point of theta = 0.200 (EPC = 0.500) the TIF equaled 22.582. The test information function shows that the maximum amount of information provided by the RAIT-NV was 22.582 at a theta that is an ability level of -0.350 (i.e. the point at which the curve peaks). Which means the maximum information was 22.582 at a theta i.e an ability level of -0.350. At the cut point of theta = 0.200 (EPC = 0.500) the TIF equaled 22.582. In this case, thus, the TIF provides satisfactory information over the ability trait range since it takes the shape of a normal distribution curve.

Figure 1.3 displays a graph of the Conditional Standard Error of Measurement (CSEM) Function. The CSEM is an inverted function of the TIF, and estimates the amount of error in theta estimation for each level of theta. The minimum CSEM was 0.196 at theta = -0.350. At the cutpoint of theta

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

= 0.200 (EPC = 0.500) the CSEM equaled 0.210.

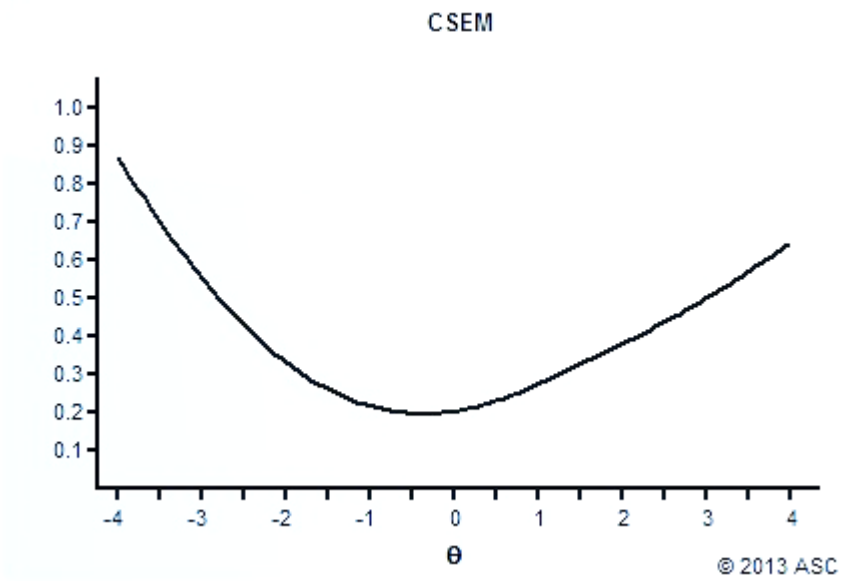

Figure 1.3: The Conditional Standard Error of Measurement (CSEM)

As earlier mentioned, the test information function sums up the item information functions to summarize where the test is providing information. The item information function shows or gives the contribution of each items to a test. An item provides more information about examinees where it provides more slope.

Research Question 5; What is the Z scores, T scores, Percentile rank and Normalized standard Nonverbal Analogies intelligence index of RAIT-NVA subtest in Nigeria

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

Table 1.5: The Raw Score, Z scores, T scores, Percentile rank and Normalized score of the RAIT-NVA Subtest in Nigeria.

RawScore	Freq	Per	Cum Per	%ile Rk	Zscores	Tscores	Stand NVI
5	10	0.5	0.5	0	-2.56	24.38	61.56
6	3	0.1	0.6	1	-2.45	25.48	63.61
7	9	0.4	1	1	-2.34	26.58	64.86
8	23	1.1	2.1	1	-2.23	27.68	66.52
9	31	1.5	3.6	2	-2.12	28.78	68.1
10	34	1.6	5.2	4	-2.01	29.88	69.8
11	13	0.6	5.8	5	-1.9	30.98	71.4
12	47	2.2	8	6	-1.79	32.08	73.1
13	25	1.2	9.2	8	-1.68	33.19	74.7
14	37	1.7	10.9	9	-1.57	34.29	76.4
15	19	0.9	11.8	11	-1.46	35.39	78
16	21	1	12.8	12	-1.35	36.49	79.7
17	14	0.7	13.5	13	-1.24	37.59	81.3
18	53	2.5	16	14	-1.13	38.69	83
19	51	2.4	18.4	16	-1.02	39.79	84.6
20	34	1.6	20	18	-0.91	40.89	86.3
21	44	2.1	22.1	20	-0.8	41.99	87.9
22	60	2.8	24.9	22	-0.69	43.1	89.6
23	41	1.9	26.8	25	-0.58	44.2	91.2

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

24	72	3.4	30.2	27	-0.47	45.3	92.9
25	59	2.8	33	30	-0.36	46.4	94.5
26	47	2.2	35.2	33	-0.25	47.5	96.2
27	58	2.7	38	35	-0.14	48.6	97.7
28	102	4.8	42.8	38	-0.03	49.7	99.5
29	139	6.5	49.3	43	0.08	50.8	101.2
30	117	5.5	54.9	49	0.19	51.9	102.8
31	96	4.5	59.4	55	0.3	53.01	104.5
32	118	5.5	65	59	0.41	54.11	106.1
33	107	5	70	65	0.52	55.21	107.8
34	112	5.2	75.3	70	0.63	56.31	109.4
35	98	4.6	79.9	75	0.74	57.41	111.1
36	75	3.5	83.4	80	0.85	58.51	112.7
37	58	2.7	86.2	83	0.96	59.61	114.4
38	54	2.5	88.7	86	1.07	60.71	116
39	60	2.8	91.6	89	1.18	61.81	117.7
40	48	2.2	93.8	92	1.29	62.92	119.3
41	25	1.2	95	94	1.4	64.02	121
42	37	1.7	96.7	95	1.51	65.12	122.6
43	15	0.7	97.5	97	1.62	66.22	124.3
44	10	0.5	97.9	97	1.73	67.32	125.9

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

45	7	0.3	98.3	98	1.84	68.42	127.6
46	17	0.8	99.1	98	1.95	69.52	129.2
47	8	0.4	99.4	99	2.06	70.62	130.9
48	4	0.2	99.6	99	2.17	71.72	132.5
49	2	0.1	99.7	100	2.28	72.83	134.2
51	6	0.3	100	100	2.5	75.03	137.5

The table shows the score, z-score, t score, percentile rank and the normalized standard index of the NVA subset of RAIT-NV. The transformation of the RAIT-NV items in Nigeria into these scores helps in giving a detailed description of examines performance on the RAIT-NV subsets as this process evens out the differences associated with raw scores that may have hindered proper description of scores as earlier mentioned. For instance a raw score of 45 has a standard score Nonverbal intelligence index of 127, a z score of 1.84, a t score of 68.42 and a percentile rank of 98 in the subtest. The table 4.7 therefore compares the reasoning of a person with that of all other person in Nigeria, irrespective of age, gender, ethnicity or educational qualification to know his/her relative standing.

Table 1.5.1: The Raw Score, Z scores, T scores, Percentile rank and Normalized standard score of the RAIT-SEQ subtest in Nigeria

RawScore	Freq	Per	Cum Per	%ile Rk	Zscores	Tscores	Stand SEQ
0	41	1.9	1.9	0	-3.11	18.87	53.3
1	6	0.3	2.2	1	-3	19.97	54.9
2	6	0.3	2.5	2	-2.89	21.07	56.6
3	21	1	3.5	3	-2.78	22.18	58.2
4	31	1.5	5	4	-2.67	23.28	59.9
5	22	1	6	5	-2.56	24.38	61.5
6	24	1.1	7.1	6	-2.45	25.48	63.2
7	26	1.2	8.3	7	-2.34	26.58	64.8
8	33	1.6	9.9	8	-2.23	27.68	66.5
9	61	2.9	12.8	10	-2.12	28.78	68.1
10	26	1.2	14	13	-2.01	29.88	69.8
11	51	2.4	16.4	14	-1.9	30.98	71.4

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

12	52	2.5	18.9	16	-1.79	32.08	73.1
13	48	2.3	21.1	19	-1.68	33.19	74.7
14	62	2.9	24.1	21	-1.57	34.29	76.4
15	53	2.5	26.6	24	-1.46	35.39	78
16	56	2.6	29.2	27	-1.35	36.49	79.7
17	81	3.8	33	29	-1.24	37.59	81.3
18	62	2.9	35.9	33	-1.13	38.69	83
19	105	5	40.9	36	-1.02	39.79	84.6
20	75	3.5	44.4	41	-0.91	40.89	86.3
21	91	4.3	48.7	44	-0.8	41.99	87.9
22	102	4.8	53.5	49	-0.69	43.1	89.6
23	109	5.1	58.7	54	-0.58	44.2	91.2
24	97	4.6	63.3	59	-0.47	45.3	92.4
25	125	5.9	69.2	63	-0.36	46.4	94.5
26	64	3	72.2	69	-0.25	47.5	96.2
27	70	3.3	75.5	72	-0.14	48.6	97.9
28	58	2.7	78.2	75	-0.03	49.7	99.5
29	135	6.4	84.6	78	0.08	50.8	101.2
30	71	3.3	87.9	85	0.19	51.9	102.8
31	46	2.2	90.1	88	0.3	53.01	104.5
32	31	1.5	91.6	90	0.41	54.11	106.1
33	38	1.8	93.3	92	0.52	55.21	107.8
34	33	1.6	94.9	93	0.63	56.31	109.4
35	36	1.7	96.6	95	0.74	57.41	111.1
36	18	0.8	97.5	97	0.85	58.51	112.7
37	18	0.8	98.3	97	0.96	59.61	114.4
38	8	0.4	98.7	98	1.07	60.71	116
39	10	0.5	99.2	99	1.18	61.81	117.7
40	17	0.8	100	99	1.29	62.92	119.3
41	1	0	100	100	1.4	64.02	121

1.51 shows the score, z-score, t score, percentile rank and the normalized standard index of the SEQ subtest of RAIT-NV. The transformation of the RAIT-NV items in Nigeria into these scores helps in giving a detailed description of examines performance on the RAIT-NV subtest as this process evens out the differences associated with raw scores that may have hindered proper description of scores as earlier mentioned. For instance a raw score of 40 has a standard score Nonverbal intelligence index of 119, a z score of 1.29, a t score of 62 and a percentile rank of 100

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

in the SEQ subtest. The table 4.7.1 therefore compares the reasoning of a person with that of all other person in Nigeria, irrespective of age, gender, ethnicity or educational qualification to know his/her relative standing.

DISCUSSION OF FINDINGS

Local independence of Reynolds Adaptable Intelligence Test- Nonverbal (RAIT-NV) in Nigeria.

The findings from research question one, indicated that the local independence of the Reynolds Adaptable Intelligence Test- Nonverbal (RAIT-NV) in Nigeria using factor analysis ,showed that component one had the highest eigenvalue which is 11.88. This shows that the largest component explains 12.50% of the variance; it explains that the items contained in the test hang together on one distinct factor which is Intelligence. This result revealed that the RAIT-NV fulfilled the assumption of unidimensionality which is local independence as the factor analysis results were in line with the set condition for assessing local independence of items in a test by Thissen, D & Orlando (2001). According to them dichotomous test items are unidimensional when the first factor loading for all items is significantly greater than 1. Since the assumption of unidimensionality is met by this model, it invariably means that local independence holds .and by extension Local Independence (Ubi 2006). This finding is in line with that of Orluwene and Asiegbu 2016 who were able to check the assumptions of unidimensionality and local independence using factor analysis. That is the result of study on the results revealed that the test items met the assumptions of local independence when they investigated bias in attest using IRT model. This is also in tandem with the findings of Kpolovie and Emekene 2016, Emekene 2016,

Item Difficulty and Item discrimination index under the Item Response Theory (IRT) framework of Reynolds Adaptable Intelligence Test – Nonverbal (RAIT-NV) in NA careful examination of the *a* parameter column both for the NVA and SEQ subtest reveals that most of the test items discriminated well. The a values ranged from 0.1 to 1,4 for NVA and 0,1 to .1.9 for SEQ. Higher *b* parameters (> 1.0) indicate that the item is more difficult; a value below -1.0 indicates that the item is very easy. According to X-Calibre manual, the difficulty index "ranges in theory from negative to positive infinity, but in practice from -3.0 (very easy) to +3.0 (very difficult)." Just like in the CTT analysis, the table reveals for both the NVA and the SEQ subtest of the RAIT-NV item difficulty that graduates from very easy to very difficult which reflects the design of the test by its constructor. A closes scrutiny of the *b* parameter column shows that the values of *b* for the NVA ranged from -2. 12 to 3.00 and same for SEQ. The b parameter kept graduating in difficulty for the whole test.. From the tables, it was observed that the IRT framework items gave the estimates of parameters for all 95 items in the RAIT-NV scale subjected to its 2PLM calibration process using the Xcalibre. This study has been able to sufficiently establish the qualities of RAIT-NV Test in Nigeria which is a first. This is in line with the result of Ojerinde

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

2013 in the study,Petrillo, J., Cano, S. J., McLeod, L. D., & Coon, C. D. (2005)and Emekene (2017)

Establishment of Reliability of Reynolds Adaptable Intelligence Test-Nonverbal(RAIT-NV) in Nigeria.

Findings from this study showed that Cronbach's alpha reliability coefficient of 0.920 for the whole 95 RAIT-NV items. , a reliability coefficient of 0.873 was obtained for the NVA subset and 896 for the a reliability coefficient SEQ subset which was also very high.. This reveals the internal consistency of the items and by implication the reliability of the test. In order words since a high reliability coefficient was obtained, it implies the items are internally consistent with themselves therefore reliability and by extension has construct validity. It was also further able to establish the internally consistent of the test using the subtest total correlation of the RAIT-NV. This shows that the RAIT-NV has items that are internally consistent and as such valid and reliable. The reliability of the RAIT-NV scale as indicated by the alpha value is very much in tandem and consistent with the reliability published by RAIT-NV Developer and publishers. This finding is also in line with the study of Emekene (2017) where he was able to establish the construct validity and reliability using diverse methods like the cronbach alpha and others). Internal consistency is usually measured with Cronbach's alpha, a statistic calculated from the pairwise correlations between items. Internal consistency ranges between negative infinity and one. Coefficient alpha will be negative whenever there is greater within-subject variability than between-subject variability. Internal consistency assesses the consistency of results across items within a test. According to Patrick (2011) who used the cross sectional design to validate The Schutte Self-report Emotions Intelligence Test (SSEIT) using Nigerian secondary school adolescents. Observed internal consistency of SSIET showed a Cronbach's alpha coefficient of .90,measures reveal that SSIET had significant reliability. Abdalgadr (2009) also worked on "Standardization of Raven's standard progressive matrices test for a Libyan sample"**.** Quantitative research designs (descriptive and comparative survey, correlational and cross-sectional) were used. The aim of this study was to standardize the SPM test to a Libyan setting to develop norms for the classic form of the SPM test to identify the distribution of IQ scores within Libyan students. . The result of the findings indicated that SPM reliability was 0.94, validity and item analysis indicated that the SPM test may be considered as an appropriate measure of mental ability for Libyan students. Kpolovie and Emekene 2018 had Cronbach's alpha that was moderately high at 0.78." This was similar to the alpha reported in the first part of their work *(a* = .81). The authors and promoters of the Raven's APM test-the Raven and Raven's company, have reported consistency coefficients between of r=0.83 and r=0.87 These results obtain from various study is in line with the alpha reliability coefficient of RAIT-NV in Nigeria.

Establishment of the Test Information Function of Reynolds Adaptable Intelligence Test-Nonverbal (RAIT-NV) in Nigeria

This study was able to go further and produce the test information function of RAIT-NV. It displayed a graph of the Test Information Function for all calibrated items. The test information

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

function sums up the item information functions to summarize where the test is providing information. The TIF is a graphical representation of how much information the test is providing at each level of theta. Maximum information was 26.097 at theta = -0.350. The test information function shows that the maximum amount of information provided by the RAIT-NV was 22.582 at a theta i.e an ability level of -0.350 (i.e. the point at which the curve peaks). I.e maximum information was 22.582 at a theta i.e an ability level of -0.350. At the cut point of theta = 0.200 (EPC = 0.500) the TIF equaled 22.582. In this case, thus, the TIF provides satisfactory information over the ability trait range since it takes the shape of a normal distribution curve. Also displayed was a graph of the Conditional Standard Error of Measurement (CSEM) Function. The CSEM is an inverted function of the TIF, and estimates the amount of error in theta estimation for each level of theta. The minimum CSEM was 0.196 at theta = -0.350. So the maximum amount of information provided by the RAIT-NV was 22.582 at a theta i.e an ability level of -0.350 (i.e. the point at which the curve peaks). The maximum information was 22.582 at a theta i.e an ability level of -0.350. In this case, thus, the TIF provides satisfactory information over the ability trait range since it takes the shape of a normal distribution curve. Study by Oku and Iweka 2018 revealed same. In the same vein Kpolovie and Emekene 2018 showed same when they carried out their study.

The Z scores, T scores, Percentile rank and Normalized standard score of RAIT-NV in Nigeria

The score, z-score, t score, percentile rank and the normalized standard index of the RAIT-NV NVA & SEQ showed the transformation of the RAIT-NV items in Nigeria into these scores helps in giving a detailed description of examines performance on the RAIT-NV subtest as this process evens out the differences associated with raw scores that may have hindered proper description of scores as earlier mentioned. For instance a raw score of 40 has a standard score Nonverbal intelligence index of 119, a z score of 1.29, a t score of 62 and a percentile rank of 100 in the SEQ subtest. The tables 4.7 & 4.7.1 compares the reasoning of a person with that of all other person in Nigeria, irrespective of age, gender, ethnicity or educational qualification to know his/her relative standing. Unfortunately, there are no norms reported on any study or (studies) of RAIT-NV from or about Nigeria. The RAIT-NV has now been normed and standardized in Nigeria. Therefore the outcome of any future test on the RAIT-NV by any University undergraduate secondary school and upper primary student in Nigeria can easily be transformed from raw score to the normalized standard score IQ. The percentile rank table generated in this research work is suitable for University undergraduate, secondary school and upper primary students in Nigeria. In a related study carried out by Kpolovie and Emekene 2018 on the application of psychometric analyses on the Advanced Progressive Matrices (APM) scale using a validation and standardization sample randomly drawn in Nigeria was transformed into normalized standard score IQ for use in Nigeria, which is in line with normalized standard score IQ of RAIT-NV for use in Nigeria.

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

CONCLUSION

The following conclusions were drawn on the basis of the problem investigated and the results of the data analysis conducted in the study.

1. It was concluded that RAIT-NV fulfills the assumption of local independence as items holds. As result showed that the items contained in the test hang together on one distinct factor which is Intelligence.
2. Under the IRT framework It is concluded that RAIT-NV had item difficulty ranging from very easy to very difficult while RAIT-NV was able discriminate between those that have high and low intelligence ability.
3. In conclusion the RAIT-NV items had a high construct validity coefficient of 0.887, while the construct validity of RAIT-NVA was 0.885 and RAIT-NV SEQ was 0.887. The validity coefficient was significant at 0.000,($p<0.005$).
4. In conclusion the RAIT-NV items had a high Cronbach alpha reliability coefficient of 0.920, RAIT-NV subtest where significantly high RAIT-NVA had a reliability coefficient 0.896 while 0.873 was obtained for the RAIT –NV SEQ subtest.
5. It was concluded that the test information function (TIF) provides satisfactory information over the ability trait range since it takes the shape of a normal distribution curve.
6. The Z scores, T scores, Percentile rank and Normalized standard score of RAIT-NV in Nigeria was established, to even out the differences associated with raw scores.

Implications of the study

The following are implication of the study

1. Teachers, test developers, psychologist, researchers and relevant educational agencies, should ensure that they establish the Z scores, T scores, Percentile rank and Normalized standard score of any instrument they are revalidating and standardizing for measuring intelligence. This is recommended to even out all the difference from age, gender, ethnicity and educational level associated with raw scores.
2. Another implication is for test developers, psychologist, researchers and relevant testing bodies, should ensure adopt IRT approach in the psychometric analysis of their test as this give detailed parameter estimate of test items that are sample independent

Suggestions for Further Research

(i) Application of Generalizability Theory in the revalidation and Standardization of Reynolds Adaptable Intelligence Test-Nonverbal (RAIT-NV) in Nigeria

(ii) Validation and Standardization of Reynolds Adaptable Intelligence Test-Nonverbal (RAIT-NV) in identify academically talented students in Nigeria.

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

References

Emekene, C.O. (2017). Psychometric analysis of the advanced progressive matrices scale for use in Nigeria. *Unpublished Ph.D Thesis*, University of Port Harcourt. 2017.

Kpolovie, P. J., & & Emekene, C. O. (2016). Psychometric advent of advanced progressive matrices- smart version (APM-SV) for use in Nigeria. *European Journal of Statistics and Probability*. *4*(3), 20-60. https://www.eajournals.org/wp-content/uploads/Psychometric-Advent-of-Advanced-Progressive-Matrices-%e2%80%93-Smart-Version-Apm-Sv-For-Use-in-Nigeria.pdf

Kpolovie, P. J. (2017). Intelligence and academic achievement: A longitudinal survey. *International Journal of Recent Scientific Research, 7*(5), 11423-11439.

Miron, M. (2014). A validation study of a transferred group intelligence test. *International Journal of Psychology*. *12*(3), 193-205. https://doi.org/10.1080/00207597708247389

Orluwene G.W & Asiegbu C. (2017). Detecting Item Bias in Rivers State JSSCE Business Studies Using item Response Theory (IRT) Approach. *European journal of statistics and probability*, 5(2), 17-27

Oku, K. & Iweka F. (2018). Development, Standardization and Application of Chemistry Achievement Test Using the One-Parameter Logistic Model (1-Plm) of Item Response Theory (Irt). *American Journal of Educational Research*, 6(3), 238-257

Ojerinde, D. (2013). Classical test theory (CTT) vs item response theory (IRT): an evaluation of the comparability of item analysis results. Retrieved from https://ui.edu.ng/sites/.../PROF%20OJERINDE'S%20LECTURE%20(Autosaved).pdf'

Orluwene, G. W. (2012). *Introduction to test theory and development process*. Harey Publications Coy.

Petrillo, J., Cano, S. J., McLeod, L. D., & Coon, C. D. (2005). Using classical test theory, item response theory, and Rasch measurement theory to evaluate patient-reported outcome measures: A comparison of worked examples. *Patient-Reported Outcomes 18*(1), 25-34. https://doi.org/10.1016/j.jval.2014.10.005.