International Research Journal of
SCIENCE, TECHNOLOGY, EDUCATION, AND MANAGEMENT
P-ISSN: 2799-063X | E-ISSN: 2799-0648
ISO 9001:2015 Certified
Volume 3, No. 4 | December 2023

Revalidating the test of general reasoning ability in the African context: Evidence from Rivers State Nigeria

Stella Eteng-Uket
Department of Educational Psychology, G/C, Faculty of Education, University of Port Harcourt, Rivers State, Nigeria
stella.eteng-uket@uniport.edu.ng

ABSTRACT

As the world becomes increasingly complex, the demand for individuals with strong general reasoning abilities has never been more important. To meet this demand, educators and employers need a reliable way to measure the general reasoning ability of their students and employees. The Test of General Reasoning Ability (TOGRA) has emerged as a promising tool for measuring general reasoning ability. However, it has not been used in developing nations like Africa, including Nigeria. One possible reason for this is the lack of revalidation of the TOGRA in Nigeria. Therefore, to gain recognition in Nigeria's research community and also be used with confidence, it is necessary to revalidate the test in the local context. This prompted the study. Eight research questions were drawn, a triangulation research design was adopted, and a sample of 400 persons was drawn using a multi-stage sampling approach from the Port Harcourt metropolis in Rivers State, Nigeria. Two instruments were used for data collection: the Test of General Reasoning Ability (TOGRA) and the Reynold Adaptable Intelligence Test Nonverbal (RAIT-NV). The data were analysed using various statistical tools and software. Results revealed that TOGRA possesses adequate difficulty and discrimination indices, high-reliability indices, and validity. The conclusion was that TOGRA is a reliable and valid measure of reasoning ability in Nigeria. It was recommended that it be used for research and practical purposes in educational and occupational settings.

ARTICLE INFO

Received : Apr. 22, 2023
Revised : Dec. 5, 2023
Accepted : Dec. 30, 2023

KEYWORDS

Item response theory, Reliability, Test of General Reasoning Ability (TOGRA), Validation, Validity

Suggested Citation (APA Style 7th Edition):

Eteng-Uket, S. (2023). Revalidating the test of general reasoning ability in the African context: Evidence from Rivers State Nigeria. *International Research Journal of Science, Technology, Education, and Management*, 3(4), 136-152. https://doi.org/10.5281/zenodo.10516381

INTRODUCTION

As the world becomes increasingly complex, the need for individuals with strong general reasoning abilities has never been more important. In today's fast-paced and ever-changing world, individuals are required to possess a range of skills and competencies that enable them to solve complex problems, adapt to new situations, and think critically. General reasoning ability is one of such skill that has been identified as a key predictor of success across a range of domains, including academics, employment, and everyday life.

General reasoning ability is a fundamental cognitive skill that allows individuals to solve problems, make decisions, and draw logical conclusions based on available information. General reasoning ability encompasses a range of cognitive processes, including abstract reasoning, problem-solving, and critical thinking (Carroll, 1993), these skills enable individuals to analyse and evaluate information, recognize patterns, and make sound judgments based on evidence.

Studies has revealed that persons with high degree of general reasoning ability tend to perform better in academic, professional, and social contexts, they are also better able to adapt to changing circumstances and navigate complex situations (Ackerman et al., 2002, 2005; Ackerman 2022; Gómez-Veiga et al.,2018; Engelhardt et al., 2016; Klauer & Phye 2008; Rani 2017; Ree & Earles 1992; Schmidt & Hunter 1998; Stanovich & West 1998; 2000)

In Africa and Nigeria in particular, the demand for individuals with strong general reasoning abilities has risen in recent years, as the country seeks to build a workforce that can compete in the global market. This need has been further fuelled by the increasing reliance on technology and automation, which require individuals to be able to think creatively and adapt to novel circumstances. In addition, the Nigerian government has made significant investments in education, with a focus on improving the calibre of instruction and learning in the country. As a result, educators and employers are now seeking reliable ways to measure the general reasoning ability of their students and employees. To meet this demand, educators and employers need a reliable way to measure general reasoning ability in their students and employees. The Test of General Reasoning Ability (TOGRA) has emerged as a promising tool for measuring overall capacity for reasoning.

TOGRA was developed by Reynolds (2014). Verbal, nonverbal, quantitative, and problem-solving thinking abilities are evaluated by the TOGRA (Reynold 2014). In a review of the test by Ramos et al., (2017), they stated that the items as multiple-choice tests address auxiliary cognitive functions such attention, spatial abilities, numeric knowledge, auditory and visual perception, language processing, deductive and inductive reasoning, and visual imagery. Additionally, they proposed that the TOGRA include items in item formats that use both inductive and deductive processes to evaluate numeric problem solving, crystallized and fluid abilities, and verbal and nonverbal thinking. When an expedited assessment of reasoning skills and under-pressure problem solving is deemed beneficial in a selection process, TOGRA is a valuable tool for examiners in a variety of contexts. (Reynolds 2014). This measuring scale is well suitable and is a workable choice for application in educational institutions, juvenile and adult justice systems, clinical settings, human resources departments, and associated industrial contexts. (Reynold 2014).

The TOGRA instrument was developed for use in the US and has proven to be a valuable tool. However, it has not been widely used in developing nations like Africa in general and Nigeria in particular. There is no evidence of samples being drawn from any part of Africa, and there have been no studies from or about Nigeria using the instrument. One possible reason for this is the lack of revalidation of the TOGRA in Nigeria. Thus, given the low level of attention the TOGRA scale has received in Africa and Nigeria, it is essential to explore its potential by investigating its psychometric properties thoroughly. It is crucial to raise awareness of the scale and incorporate it as a reliable tool for measuring general reasoning ability. Therefore, to gain recognition in Nigeria's research community and also be used with confidence, it is necessary to revalidate the test in the local context. By providing a revalidated tool for measuring general reasoning ability, this study’s result can help to identify individuals who possess the skills and competencies required for success in today's world. This can inform educational policy and practice, as well as employment policies, and can ultimately contribute to the development of a skilled and competitive workforce in Nigeria.

This research aims to bridge the identified gap by investigating whether the scale maintains the same psychometric properties when administered to respondents in Nigeria so it can easily be used. In specific terms, it hopes to provide response to the following research questions below.

1. What is the item difficulty index using the ‘b-parameter’ and the item discrimination index using the ‘a-parameter’ of TOGRA within the IRT framework?
2. What is the test bias of TOGRA using the Mantel-Haenszel Method?
3. What is the stability and equivalence coefficient of TOGRA using the alternate form reliability?
4. What is the internal consistency coefficient of TOGRA using the Cronbach alpha method?
5. What is the internal consistency coefficient of TOGRA using the Split Half method?
6. What is the construct validity of TOGRA using the convergent correlation of TOGRA with RAIT-NV?
7. What is the construct validity (Unidimensionality) of TOGRA using factor analysis?
8. What is the classification and Description of TOGRA General Reasoning Index (GRI) in Rivers state, Nigeria?

METHODS

Research Design

The triangulation research design was used. This design involves multiple research methodologies, measurement instruments, and statistical tools to comprehensively investigate a particular phenomenon Kpolovie (2010). The aforementioned design permits a variety of methods to be used to studying an instrument's psychometric qualities. Various methods such as the two-parameter model, Differential Item Functioning (DIF), factor analysis, correlation, split-half estimates, test-retest estimates, equivalent form estimate, Cronbach alpha estimate, and correlational method analysis were applied.

Population, Sampling and sampling Technique

The study population comprised 57,800 students from universities and secondary schools in Port Harcourt metropolis, River’s state, Nigeria. A sample size of 400 students was calculated from the population using Slovin's formula for minimum sample size as a guide. The work adopted a multi stage sampling procedure. In the first stage of the sampling, a simple random sampling procedure by balloting was used to randomly select one federal university from the three government owned universities in rivers state. To randomly select secondary schools from the 86 secondary schools in Port Harcourt metropolis, simple random sampling (throw of coin) was used to draw five (5) schools from the entire eighty-six (86) schools in Port Harcourt metropolis. In the second stage of the sampling, a stratified random sampling technique based on age (Age 10- 20yrs, Age 21- 30yrs, Age 31 above) was used to draw a total of 400 (Age 10- 20yrs = 138, Age 21- 30yrs = 169, Age 31 above = 73) students from the university and secondary schools in Port Harcourt metropolis.

Instrument for Data Collection

Two instruments were used for data collection, Test of General Reasoning Ability (TOGRA) and Reynold Adaptable Intelligence Test Non-Verbal (RAIT-NV). The Test of general reasoning ability (TOGRA) was developed by Reynold (2014), it is a standardized, timed assessment of reasoning and problem-solving abilities designed for use with test takers between the ages of 10 and 75. Its items evaluate reasoning abilities—verbal, nonverbal, and quantitative—as well as problem-solving techniques through both inductive and deductive problems. It takes 16 minutes to administer and 2-3 minutes to score each of the 60 items. Examinees can choose from options by supplying the answers that best finish an unfinished progressive series that is offered. For convenience, additional derived scores in the form of z scores, stanines, percentiles, and, for lower ages, age equivalents, are provided. It is given, either manually with paper and pencil or electronically, to individuals or groups of any size. There are two equivalent other forms of it: the green form and the blue form. The TOGRA items were developed from a pool of original items that assessed a wide range of ability constructs centered around the abilities of reasoning and problem solving. Using population-proportionate, stratified random sampling based on 2010 U.S. Census population figures, the TOGRA was standardized on a sample of 3,013 people selected from 39 states. The second instrument was the Reynold Adaptable Intelligence Test NonVerbal (RAIT-NV) by Reynold (2016). This is a rapid test of nonverbal intelligence. It was created using the two nonverbal subtests from the Reynold Adaptable Intelligence Test (RAIT) to give an accurate assessment of fluid intelligence. The RAIT-NV is not a speeded test; rather, it is a power test with a maximum duration constraint. Both individual and group administration are possible. It can be applied in education, juvenile and adult justice systems, healthcare settings, human resources and related industrial settings, and schools.

Reliability and validity of the Instrument

According to Reynold's (2014) report, the reliability of TOGRA varies between 0.74 and 0.99 for test-retest reliability, 0.87 and 0.94 for Cronbach alpha reliability, and 0.85 and 0.94 for alternate form reliability for individuals aged 10 to 75. According to correlations with other tests (RAIT, WISC-IV, WAIS-IV, RIAS, Wonderlic, (Beta III), (WRAT), and (TIWRE), the test's construct validity ranges from 0.75 to 0.95. The Reynold Adaptable Intelligence Test NonVerbal (RAIT-NV), the second instrument, has construct validity ranging from 0.75 to 0.95 based on correlation with other tests such as RAIT, WISC-IV, WAIS-IV, RIAS, Wonderlic, (Beta III), WRAT, and TIWRE. According to Reynold's (2016) report, the reliability varies between 0.74 and 0.99 for test-retest reliability, 0.87 and 0.94 for Cronbach alpha reliability, and 0.85 and 0.94 for alternate form reliability for individuals aged 10 to 75.

Method of Date Analysis

The data was analysed and research questions were answered using mean, standard deviation, IRT 2PLM, DIF, Factor analysis, correlations (split half, test retest and correlation with other measure) and parallel form utilising statistical software programs like the Item Response Theory application for Excel add-in, the Statistical Package for Social Science (SPSS) and Microsoft EXCEL.

RESULTS AND DISCUSSION

Results

The Research Question 1: What are the item difficulty index using the b-parameter and the item discrimination index using the a-parameter of TOGRA within the CTT and IRT frameworks?

Table 1a. Item difficulty index using the p-value using and the item discrimination index using the item total correlation of TOGRA within the CTT

Item	(difficulty)	Correlation (discrimination)
ITEM 1	0.793	0.203
ITEM2	0.778	0.370
ITEM3	0.703	0.297
ITEM4	0.620	0.292

ITEM5	0.410	0.360
ITEM6	0.593	0.475
ITEM7	0.690	0.419
ITEM8	0.725	0.241
ITEM9	0.760	0.194
ITEM10	0.655	0.345
ITEM11	0.503	0.342
ITEM12	0.649	0.371
ITEM13	0.608	0.335
ITEM14	0.530	0.398
ITEM15	0.360	0.334
ITEM16	0.520	0.419
ITEM17	0.538	0.337
ITEM18	0.550	0.412
ITEM19	0.518	0.475
ITEM20	0.505	0.418
ITEM21	0.535	0.416
ITEM22	0.400	0.499
ITEM23	0.258	0.314
ITEM24	0.330	0.407
ITEM25	0.475	0.493
ITEM26	0.343	0.520
ITEM27	0.330	0.466
ITEM28	0.293	0.337
ITEM29	0.420	0.465
ITEM30	0.278	0.487
ITEM31	0.300	0.499
ITEM32	0.210	0.413
ITEM33	0.305	0.561
ITEM34	0.255	0.493
ITEM35	0.230	0.446
ITEM36	0.225	0.528
ITEM37	0.155	0.431
ITEM38	0.233	0.511
ITEM39	0.215	0.518
ITEM40	0.170	0.468
ITEM41	0.135	0.566
ITEM42	0.135	0.524
ITEM43	0.120	0.511
ITEM44	0.100	0.521
ITEM45	0.080	0.431
ITEM46	0.085	0.461
ITEM47	0.113	0.540

ITEM48	0.093	0.582
ITEM49	0.083	0.374
ITEM50	0.065	0.367
ITEM51	0.115	0.550
ITEM52	0.088	0.510
ITEM53	0.083	0.392
ITEM54	0.090	0.526
ITEM55	0.065	0.478
ITEM56	0.060	0.422
ITEM57	0.063	0.479
ITEM58	0.050	0.184
ITEM59	0.065	0.331
ITEM60	0.068	0.283

Table 1a shows the p-value which indicates the difficulty indexes for the 60 TOGRA items within the CCT framework. Higher p values (> 1.0) shows that the task is extremely simple, and a value less than -1.0 shows that the task is very simple. The creator of the TOGRA created it with the intention of having the difficulty index progress from very easy to very challenging items. This design is seen in the progression of the item difficulty as the first 30 items were seen to have items that were easy (higher p-values) which then started graduating to items that are difficult (lower p-value) for the next ten items and then for the last twenty items. Precisely, the first 10 items had p-values close to 1.0 and while the last 10 items had values less than 0.01

The table shows again the discrimination index indicated by item total correlation within the CTT framework. These reveal that the 60 TOGRA items were able to discriminate between examinee high on the trait being measured and those low on the trait. The table shows that just 2 items had discrimination value less than .2; no item had a negative value or a value of 0.00 which is an indication of an item which is not able to discriminate. This reveals that the TOGRA has good discrimination index within the CTT framework

Table 1b item difficulty index using the b-parameter using and the item discrimination index using the a parameter of TOGRA within the IRT framework

Item	discrimination(*a*) Slope	difficulty (*b*)Threshold
ITEM 1	0.333	-3.000
ITEM2	0.447	-3.000
ITEM3	0.282	-2.800
ITEM4	0.134	-2.060
ITEM5	0.145	1.255
ITEM6	0.393	-1.208
ITEM7	0.361	-2.329
ITEM8	0.263	-3.000
ITEM9	0.282	-3.000
ITEM10	0.223	-2.397
ITEM11	0.062	-0.087
ITEM12	0.223	-2.292
ITEM13	0.162	-1.849
ITEM14	0.074	-0.486

ITEM15	0.209	1.802
ITEM16	0.106	-0.431
ITEM17	0.063	-0.527
ITEM18	0.109	-0.907
ITEM19	0.416	-0.455
ITEM20	0.107	-0.199
ITEM21	0.171	-0.724
ITEM22	0.402	0.707
ITEM23	0.321	2.639
ITEM24	0.316	1.736
ITEM25	0.519	-0.101
ITEM26	0.541	0.998
ITEM27	0.443	1.321
ITEM28	0.333	2.127
ITEM29	0.537	0.342
ITEM30	0.526	1.631
ITEM31	0.619	1.234
ITEM32	0.500	2.426
ITEM33	0.793	0.966
ITEM34	0.598	1.682
ITEM35	0.559	2.027
ITEM36	0.726	1.716
ITEM37	0.591	2.780
ITEM38	0.711	1.674
ITEM39	0.738	1.790
ITEM40	0.677	2.370
ITEM41	0.954	2.281
ITEM42	0.788	2.540
ITEM43	0.809	2.687
ITEM44	0.878	2.830
ITEM45	0.751	3.000
ITEM46	0.819	3.000
ITEM47	0.882	2.656
ITEM48	0.990	2.746
ITEM49	0.678	3.000
ITEM50	0.705	3.000
ITEM51	0.941	2.537
ITEM52	0.853	2.995
ITEM53	0.712	3.000
ITEM54	0.887	2.960
ITEM55	0.853	2.899
ITEM56	0.773	3.000
ITEM57	0.863	2.900

ITEM58	0.607	3.000
ITEM59	0.660	3.000
ITEM60	0.621	3.000

The item difficulty of the 60 TOGRA items within the IRT framework, or the b-parameter, is displayed in Table 1b. The difficulty index, according to the e-IRT result, "ranges in practice from -3.0 (very easy) to +3.0 (very difficult), but in theory from negative to positive infinity." larger negative b parameters suggest that the item is extremely easy, while larger positive b parameters imply that the item is more challenging. As previously stated, the creator of the TOGRA test devised and designed the test so that the difficulty index progresses from easy to difficult items. Also, precisely just like in the CTT analysis, this design is seen in the progression of the item difficulty as the first 30 items were seen to have items that were easy (high negative b parameter) which then started graduating to items that are difficult (high positive b parameter) for the next ten items and then for the last twenty items. Also, precisely, the first 10 items had b parameter estimate close to -3.00 and 3.000 (indicating easy items) and while the last 10 items had b parameter estimate close to +3.00 and 3.000 (indicating difficult items)

Table 1b displays the discrimination index within the IRT framework, as measured by the slope (a parameter). The a-parameter indicates that the 60 TOGRA items were able to differentiate between those high and low on the trait being measured. The table indicates that only 7 items had a parameter less than 0.2, suggesting that all items had a good fit. Overall, using the 2 logistic model, the estimation was performed for all 60 items, resulting in a good fit (χ^2=4691.948, df=1710, p<0.005, log 6589.656.). This demonstrates that the TOGRA has a good discrimination and difficulty indices fit within the IRT framework.

Research Question 2: What is the DIF (test bias) of TOGRA using the Mantel-Haenszel Method?

Table 2. DIF (item bias) using the mantel Haenszel method

ITEM	GENDR	WRONG	RIGHT	Mantel-Haenszel χ^2	SIG
1	MALE	21	117	3.416	0.065
	FEMALE	62	200		
2	MALE	35	103	1.102	0.294
	FEMALE	53	209		
3	MALE	31	107	4.821	0.028*
	FEMALE	88	174		
4	MALE	44	94	2.953	0.086
	FEMALE	108	154		
5	MALE	82	56	0.000	0.986
	FEMALE	154	109		
6	MALE	55	83	0.025	0.875
	FEMALE	108	154		
7	MALE	41	97	0.89	0.771
	FEMALE	83	178		
8	MALE	39	99	0.017	0.897
	FEMALE	71	191		

9	MALE	34	104	0.009	0.926
	FEMALE	62	200		
10	MALE	329	334	4.714	0.030*
	FEMALE	422	338		
11	MALE	88	575	7.252	0.007*
	FEMALE	142	618		
12	MALE	394	269	6.381	0.012*
	FEMALE	502	258		
13	MALE	55	83	1.005	0.943
	FEMALE	102	160		
14	MALE	69	69	0.587	0.444
	FEMALE	119	443		
15	MALE	94	44	1.285	0.257
	FEMALE	162	100		
16	MALE	67	71	12.467	0.000*
	FEMALE	125	137		
17	MALE	70	68	1.430	0.232
	FEMALE	115	147		
18	MALE	75	202(30.5)	4.007	0.933
	FEMALE	117	145		
19	MALE	62	76	0.738	0.390
	FEMALE	131	131		
20	MALE	74	64	1.353	0.245
	FEMALE	123	139		
21	MALE	73	65	0.341	0.068
	FEMALE	112	150		
22	MALE	89	49	1.494	0.222
	FEMALE	151	111		
23	MALE	101	37	0.054	0.817
	FEMALE	196	66		
24	MALE	86	52	0.773	0.183
	FEMALE	182	80		
25	MALE	68	70)	0.690	0.406

	FEMALE	142	120		
26	MALE	85	53	0.1.343	0.247
	FEMALE	178	84		
27	MALE	86	52	1.772	0.183
	FEMALE	180	80		
28	MALE	89	49	3.220	0.073
	FEMALE	193	69		
29	MALE	71	67	3.304	0.069
	FEMALE	161	101		
30	MALE	99	39	0.002	0.962
	FEMALE	190	72		
31	MALE	87	51	4.352	0.037*
	FEMALE	193	69		
32	MALE	111	27	0.146	0.703
	FEMALE	205	57		
33	MALE	94	44	0.104	0.748
	FEMALE	184	79		
34	MALE	97	41	1.638	0.201
	FEMALE	201	61		
35	MALE	107	31	1.004	0.953
	FEMALE	201	61		
36	MALE	105	33	1.133	0.715
	FEMALE	205	57		
37	MALE	113	25	1.636	0.425
	FEMALE	224	38		
38	MALE	108	30	1.155	0.693
	FEMALE	199	63		
39	MALE	110	28	1.090	0.765
	FEMALE	204	58		
40	MALE	112	26	1.325	0.568
	FEMALE	220	40		
41	MALE	113	25	3.226	0.071
	FEMALE	233	29		

42	MALE	113	25	3.226	0.071
	FEMALE	233	29		
43	MALE	119	19	1.393	0.531
	FEMALE	233	29		
44	MALE	126	12	0.207	0.649
	FEMALE	234	28		
45	MALE	132	6	3.090	0.079
	FEMALE	236	26		
46	MALE	125	13	0.084	0.772
	FEMALE	241	21		
47	MALE	118	20	1.746	0.186
	FEMALE	137	25		
48	MALE	123	15	0.396	0.529
	FEMALE	240	22		
49	MALE	125	13	0.181	0.670
	FEMALE	242	20		
50	MALE	128	10	0.051	0.821
	FEMALE	246	16		
51	MALE	118	20	1.429	0.232
	FEMALE	236	26		
52	MALE	120	18	3.477	0.062
	FEMALE	244	18		
53	MALE	121	17	3.814	0.051
	FEMALE	246	16		
54	MALE	126	12	1.001	0.977
	FEMALE	238	24		
55	MALE	125	13	2.263	0.133
	FEMALE	249	13		
56	MALE	126	10	2.029	0.154
	FEMALE	250	12		
57	MALE	128	10	0.144	0.704
	FEMALE	247	15		
58	MALE	130	8	0.084	0.772
	FEMALE	250	12		

59	MALE	126	12	1.162	0.281
	FEMALE	248	14		
60	MALE	128	10	0.006	0.938
	FEMALE	245	17		

The table 2, shows the Mantel-Haenszel analysis of the TOGRA 60 items which assess for item bias of the items based on gender. Mantel-Haenszel (M-H) analysis was conducted using a 2 X 2 factorial analysis to identify items with bias with the indication of the presence of DIF. The factors were gender (Males Versus Females) and response (Wrong Versus Right). When the associated Chi-Square $_{\chi}2$ level is significant, there is DIF, when it is not significant there is no DIF. Table 2 shows the M-H DIF statistics. This statistic tested the hypothesis that examinees performance on item i (i = 1, 2, 3, 4 ……60) is independent of gender (at $p < 0.05$ for two tailed). A level of significance value for an item that is less than 0.05 indicates that the performance of the examinees on the item is dependent on their gender and that bias towards gender may exist and such item displays DIF. On the other hand, a level of significance greater than 0.05 indicates that examinees performance on the item is independent of their gender. From the above, items 3, 10, 11, 12, 16 and 31 exhibits DIF and by implication may have gender bias.

Research Question 3: What is the test score stability coefficient of TOGRA using the test re-test reliability method?

Table 3. Stability coefficients of TOGRA using the test retest

Test-Retest	N	$\bar{x}$	SD	R	Sig	P
1^{st} Testing	46	16.78	10.1	.649	.000	0.05
2^{nd} Testing	46	20.00	8.64			

The table 3 shows that the mean and standard deviation for first testing was 16.78 and 10.1 respectively, while second testing had a mean of 20.00 and standard deviation of 8.64. On correlation, a correlation coefficient of .649 was obtained which was significant at 0.05 levels. This shows that the TOGRA has test score stability and thus reliable.

Research Question 4: What is the stability and equivalence coefficient of TOGRA using the alternate form reliability?

Table 4 Stability and equivalence coefficients of TOGRA using the parallel form reliability

P. Form	N	$\bar{x}$	SD	R	Sig	P
Blue form	43	19.83	10.9	.617	.000	0.05
Green Form	43	16.72	8.09			

The table 4 shows that the mean and standard deviation for the blue form was 19.83 and 10.9 respectively, while the green form had a mean of 16.72 and standard deviation of 8.09. On correlation, a correlation coefficient of .617 was obtained which was significant at 0.05 level. This shows that the two forms Blue and Green of TOGRA are equivalent and stable and therefore reliable.

Research Question 5: What is the internal consistency coefficient of TOGRA using the Cronbach alpha method?

Table 5 internal consistency coefficient of TOGRA using the Cronbach's Alpha method

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	No Items	No of Participant
.904	.929	60	400

Table 5 shows a Cronbach's Alpha coefficient of .904 which is a very high coefficient. This shows that the items are internally consistent and the TOGRA very reliable.

Research Question 6: What is the internal consistency coefficient of TOGRA using the Split Half Method?

Table 6. Internal consistency coefficient of TOGRA using the Split Half method

Spearman-Brown Coefficient	.677
Number of Items	60

The Split-half reliability analysis for TOGRA shows a reliability estimate of Spearman-Brown coefficient of .677. Therefore, a split half coefficient of .677 was obtained for TOGRA

Research Question 6: What is the construct validity of TOGRA using the convergent correlation of TOGRA with RAIT-NV

Table 6. Construct validity of TOGRA using the convergent correlation of TOGRA with RAIT-NV

Test	N	$\bar{x}$	SD	R	Sig	P
TOGRA	50	19.12	8.75	.600	.000	0.05
RAIT-NV	50	28.10	11.6			

The TOGRA scale shows a strong positive relationship with RAIT-NV with a correlation value of 0.60. The RAIT-NV scale in itself is a measure of intelligence and reasoning and its positive correlation shows that TOGRA has construct validity for use in Rivers state Nigeria.

Research Question 7: What is the construct validity (Unidimensionality) of TOGRA using factor analysis?

Table 7. Construct validity of TOGRA using screen plot and Eigen values of factor analysis

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	12.675	21.126	21.126	12.675	21.126	21.126
2	4.295	7.158	28.284			
3	2.149	3.581	31.865			
4	1.661	2.769	34.634			
5	1.590	2.650	37.284			
+	+	+	+			
+	+	+	+			
+	+	+	+			
55	.273	.456	98.208			
56	.268	.447	98.655			
57	.245	.408	99.063			
58	.212	.353	99.416			
59	.193	.322	99.737			
60	.158	.263	100.000			

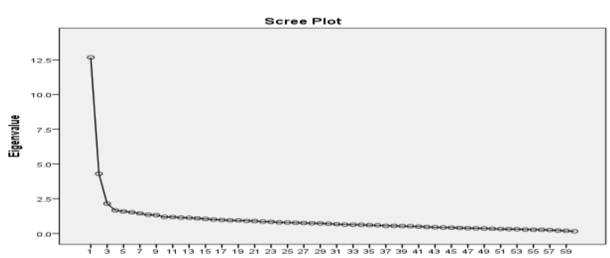

Figure 1. Scree plot

From Table 7 the highest eigenvalue is 12.67 and is for component one. This shows that the largest component explains 21.12% of the variance. It explains that the items contained in the test hang together on one distinct factor. This fact is corroborated by the result of other construct analysis carried out and internal consistency reliability analysis of the 60 TOGRA items which gives a very high value of 0.92. This shows that the test has high internal consistency and it can be inferred that it measured only one trait, reasoning ability. The factor analysis results were consistent with the predetermined criteria for evaluating unidimensionality, indicating that the TOGRA scale fulfilled is unidimensional. Also, when the first eigenvalue is significantly greater than 1 and the first factor loading for every item in the dichotomous test set is greater than 1, the test items are considered unidimensional. The difference between the first component, 12.6, and the second component, 4.2, can be observed to be quite substantial. Given that there was a discernible and substantial difference between the two components, more than 1. The TOGRA items appear to be unidimensional based on this value.

Figure 1's scree plot, when closely examined, reveals that there is just one construct before the elbow joint or breaking point. Thus, this clearly illustrates that the TOGRA is testing only the fundamental concept of reasoning.

Research Question 8; What is the classification and Description of TOGRA GRI in Rivers state, Nigeria

Table 8. Classification and Description of TOGRA GRI in Rivers state Nigeria.

GRI	Percentage of population	Brief description
130 and above	9.5	Significantly above average
120 -129	6	Moderately above average
110 -119	9.25	Above Average
90-109	41	Average
80-89	20.75	Below Average
70-79	11.25	Moderately below average
69 and below	2.25	Significantly below average

As shown in Table 8, examinees with significant above average reasoning form approximately 9.5% of Rivers state examinees on their TOGRA reasoning ability. About 6% fall into the group identified with moderately above average reasoning, 9,5% are above average, 41% which is the majority occupy the average group, 20.75% are within the moderately below average, while 2.25% significantly below average.

Discussion

The item discrimination and item difficulty of the TOGRA exam employing the CTT and IRT frameworks, respectively, are reported in Tables 1a and 1b. It was noticed from the tables that the CTT provided estimates of all the item parameters of the entire TOGRA scale, in the same way that the IRT framework items provided estimates of parameters for all 60 TOGRA scale items after undergoing its 2PLM calibration procedure. Values of b for the TOGRA in this study ranged from -3.00 (extremely easy) to +3.00 (very difficult), and for the CTT framework, from 0.77 (very simple) to 0.050 (very tough), according to a detailed analysis of the b parameter column of the 2-PLM utilized to examine the TOGRA. As a result, the difficulty scale (b parameter) increased from extremely simple to difficult. This finding obtained in River state Nigeria is consistent with the result obtained by TOGRA test developers which produced items that graduated in difficulty. From the table also, it was observed that the CTT gave the estimates of all the item parameters of the whole RAIT-NV just like the IRT framework items gave the estimates of parameters for all 60 items in the TOGRA scale subjected to its 2PLM calibration process. The estimation of item parameters under both IRT and CTT framework is to have comparable data since CTT is simple and flexible, while under IRT though it been difficult and tasking because of the heavy mathematical operation and fulfilment of stringent assumptions required by its models is more powerful, has greater computing power and analytical potency and possess a more robust mathematical precision.

The tables 1a and 1b show again the discrimination index the slope (a parameter) which is the discrimination index within the IRT framework and the item total correlation within the CTT framework. From the tables within the IRT and CTT framework, it shows that the 60 TOGRA items were able to discriminate between examinee high on the trait being measured and those low on the trait. This reveals that the TOGRA has to some extent good discrimination index within the IRT framework, just like it did in the CTT framework. Table 2 supplied information about item bias of the TOGRA items via the Mantel-Haenszel statistics. Six (6) items under the Mantel-Haenszel statistics had P values that was less than 0.05 ($P < 0.05$) of all the 60 TOGRA items when bias was examined based on gender representing 10% of the items that composed the full TOGRA scale.

Tables 3, 4, 5 and 6 reveal the reliability coefficients of TOGRA. The reliability of the TOGRA scale as indicated by the Cronbach Alpha value by this researcher is 0.92 which is consistent with the reliability coefficient of 0.90 obtained by the test developer as recorded in the TOGRA manual, for test retest a reliability coefficient of .64 as against the coefficient ranges of.74 and above obtained by the test developer, The Split half reliability estimate of the first half of the TOGRA test yielded a value of 0.82 and that of the second part of the test was 0.92. When corrected to the full test, using the Guttman Split-half coefficient, the TOGRA scale had a split-half reliability (r) of 0.66 and using the Spearman Brown was 0 .67 in Nigeria. The reliability was also computed using the Parallel form and it was found to be 0.61. This is in tandem and consistent with the reliability obtained from the test developer.

The construct validity of the TOGRA scale was also established through correlation evidence via a correlation with the RAIT-NV which gave an r of 0.60 as seen by table 7. Also construct validity was established through factorial evidence of factor analysis. Table 8 shows that reasoning, the single construct evaluated by the TOGRA scale in this study, accounted for 21.12 percent of the overall variation. This constitutes a substantial portion of the model. This confirms its unidimensionality, demonstrating that the instrument reliably assesses test-takers' reasoning skills. This affirmed the construct validity of TOGRA in Rivers state Nigeria. Table 8.1 shows reports of the classification and qualitative description of TOGRA in Port Harcourt metropolis, River's state, Nigeria.

CONCLUSION AND SUGGESTIONS

The validation of the TOGRA for use in Nigeria demands establishment of the suitability of the items by conducting item by item analysis (establishing item difficulty, discrimination and item bias), it's reliability and validity. Which this study has done in establishing the above-mentioned statistical operations and procedures using a multiple triangulation approach. This has established that in Port Harcourt metropolis in River's state Nigeria, TOGRA can be applied in a variety of contexts where an accelerated assessment of reasoning and problem solving under pressure is required.

Grounded on the results obtained, it can be decided that the TOGRA is a reliable and valid measure of reasoning ability in the population of Rivers state, Nigeria. The item difficulty and discrimination parameters were estimated using both the CTT and IRT frameworks, which showed that the test items are well-graded in terms of difficulty and can differentiate between individuals with different levels of reasoning ability. The high reliability coefficients obtained from a variety of internal consistency metrics, including Cronbach's alpha, split-half, and test-retest reliability estimates indicate that the TOGRA test has high internal consistency and stability over time. The reliability coefficients suggest that the TOGRA is a consistent and dependable tool of reasoning capacity. The construct validity of the TOGRA test was determined through correlation and factor analysis evidence. The correlation between TOGRA and RAIT-NV indicates that the TOGRA is positively related to other measures of reasoning ability, providing evidence of its concurrent validity. The factor analysis results indicate that the test measures a single construct, reasoning ability, and explain a substantial portion of the overall variation, supporting the unidimensionality of the test. However, it should be noted that the item bias analysis revealed some items with bias based on gender, which is an important issue to be addressed in future revisions of the test.

In conclusion, the TOGRA test is a dependable and accurate measurement of reasoning capacity in River's state, Nigeria, and can be used for research and practical purposes such as educational and occupational counselling. The findings of this research contribute to the body of knowledge on psychometric evaluation of reasoning and cognitive ability tests and highlight the significance of considering cultural factors in the development and validation of such tests.

Based on the results, the following suggestions were made:

The TOGRA test can be considered a reliable and valid instrument for measuring reasoning ability, particularly in Nigeria. The high reliability coefficients and evidence of construct validity suggest that the test can provide accurate and consistent results.

The use of both CTT and IRT frameworks can provide complementary information about the performance and characteristics of test items. Also, The Mantel-Haenszel statistics revealed item bias based on gender in 10% of the items. To ensure fairness and equity, these items should be reviewed and modified if necessary to eliminate any form of bias. Further investigation could be done to identify the sources of bias and address them appropriately.

The TOGRA test could be used in educational and occupational settings to assess reasoning ability and inform decision-making processes such as admissions, hiring, and promotions. Furthermore, it is recommended that the TOGRA test be used in conjunction with other measures to obtain a comprehensive assessment of an individual's reasoning ability.

Further research could be conducted to explore the generalizability of the TOGRA test in other contexts and populations, as well as its predictive validity for academic and job performance.

REFERENCES

Ackerman, P.L. (2022). Intelligence … moving beyond the lowest common denominator. *American Psychologist*. Advance online publication. https://doi.org/10.1037/amp0001057

Ackerman, P.L., Beier, M.E., & Boyle, M.O. (2005). Working memory and intelligence: the same or different constructs?. *Psychological bulletin*, *131*(1), 30–60. https://doi.org/10.1037/0033-2909.131.1.30

Ackerman, P.L., Beier, M. E., & Boyle M.O. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General*, *131*(4), 567-589. https://doi.org/10.1037/0096-3445.131.4.567.

Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.

Engelhardt, L. E., Mann, F. D., Briley, D. A., Church, J.A., Harden, K. P., & Tucker-Drob, E. M. (2016). Strong genetic overlap between executive functions and intelligence. *Journal of experimental psychology. General*, *145*(9), 1141–1159. https://doi.org/10.1037/xge0000195

Gómez-Veiga, I., Vila Chaves, J.O., Duque, G. & García Madruga, J.A. (2018) A New Look to a Classic Issue: Reasoning and Academic Achievement at Secondary School. *Front. Psychol*. 9,400. https://doi.org/10.3389/fpsyg.2018.00400

Klauer, K.J. & Phye G.D. (2008). Inductive reasoning; a training approach. *Rev Educ. Res*.78, 85-123. https://dpi.org/10.3102/0034654307313402

Kpolovie, P. J. (2010). *Advanced research methods*, New Owerri: Springfield Publishers Ltd.

Ramos, H., Ramos, S.J., & Juntune, J.E. (2018). Test Review: Test of General Reasoning Ability (TOGRA), by C. R. Reynolds. *Journal of Psychoeducational Assessment,* 36(8), 869–872. https://doi.org/10.1177/0734282917709741

Rani K.V. (2017) Reasoning Ability and Academic Achievement Among Secondary School Students In Trivandrum. *Journal on School Educational Technology*. 13(2), 20-30. https://doi.org/10.26634/JSCH.13.2.13828

Ree, M. J. & Earles, J.A. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science,1,* 86- 89. https://doi.org/10.1111/1467-8721.EP10768746

Reynold, C.R. (2014) *Test of General Reasoning Ability*. Lutz, FL: PAR.

Reynold, C.R. (2016) *Reynold Adaptable Intelligence Test Non Verbal*. Lutz, FL: PAR

Stanovich, K.E. & West, R.F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*(5), 645- 665. https://doi.org/10.1017/S0140525X00003435

Stanovich, K.E. & West, R.F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General, 127*(2), 161–188. https://doi.org/10.1037/0096-3445.127.2.161

Schmidt, F.L. & Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262–274. https://doi.org/10.1037/0033-2909.124.2.262