Application Of Two- Parameter Latent Trait Model in The Construction and Validation of Agricultural Science Achievement Test in Rivers State

OKORONKWO, MICHAEL KANU[1], DR. AMADIOHA A[2]
[1, 2]*Department of Educational Psychology Guidance and Counselling University of Port Harcourt, Port Harcourt, Nigeria.*

Abstract- The central objective of this study was the application of two-parameter latent trait model(2-PLM) in the construction and validation of Agricultural science achievement test. The study was theoretically anchored item response theory while adopting the instrumentation research design. Two objectives and two research questions guided the study and a validation sample of 1065 students out of a total population of 61045 senior secondary two students in Rivers state was drawn through a multistage sampling procedure ranging from cluster to balloting to proportionate stratified. The data analyses conducted showed the ASAT items had suitable difficulty index ranged from -3.733 to +2.112 and discrimination indices values ranging from 0.120 to +5.676.. Several recommendations were made, amongst them was that the 2-PLM be frequently employed to determine the credibility of Agricultural Science achievement tests taken by secondary school students.

I. INTRODUCTION

Agriculture has been defined as the art and science or management that deals with the cultivation of crops and the rearing of animals for man's use (Iwena,2018).It plays a crucial role in the economy of most nations of the world. Only recently, has Agriculture scaled up its food production capacity in response to the global food crisis which resulted in higher food prices and growing hunger, triggered by the multiple shocks of the Covid-19 pandemic and the on –going war in Ukraine.Agricultureis the mainstay of the Nigerianeconomy and the main source of livelihood for most Nigerian citizens.It provides food for the teeming Nigerian population;provides materials for sheltersuch as windows,doors,chairs,tables etc; provides materials for making clothes and shoes such as fibre and cotton,hides and skin,wool, silk etc; provides raw materials for agro-allied industries; provides income for farmers ; provides foreign exchange earnings- putting aside the existence of Crude oil, Nigeria rely on the exported agricultural products to generate and create most of its national revenue and finally Agriculture remains the biggest employer of labour in Nigeria of all sectors reported at 36.66% of the total employment (world bank, 2023)

The Nigerian Educational Research and Development Council (NERDC) in its recent curriculum review emphasizes vocational education in Nigerian secondary schools, with the aim of reducing the high rate of youth unemployment in the nation (NERDC, 2011). Undoubtedly, Agricultural science is a vocational subject in secondary school which provides students with skills, knowledge and attitudes necessary for future employment in Agricultural occupations. The main objectives of teaching Agricultural science in Nigerian secondary schools are to stimulate and sustain students' interest in Agriculture; impart functional knowledge and practical skills in Agriculture to students and prepare students for further studies and occupation in Agriculture.

Despite the enormous importance and lofty objectives behind the study of Agricultural Science, it has been observed that the rate at which students still fail the subject in internal and external examinations is record-high.Experience has shown that students may perform poorly in tests not just because of their low ability but also due to faults inherent in the tests as most tests administered in Nigeria are being poorly constructed and therefore lack validity and reliability (Onunkwo cited in Orluwene, 2012)

Psychological tests have been defined as any psychometrically derived measurement instrument that assesses the psychological construct in which a structured sample of an examinee's behaviour in a specified domain is obtained and subsequently quantified , scored, interpreted and synthesized using a standardized process for the purpose of evaluative conclusion or recommendation(American Educational Research Association ,American psychological association &National council on Measurement in education,2004)

Tests can therefore be regarded as instruments used to determine the relative presence or absence of the trait measured for.It could also be a measurement instrument or device administered to someone to determine the relative value of the traits or skills to which the test relates(Orluwene,2012a;2012b).

However, the crucial nature of tests in education cannot be over emphasized. Tests form the integral part of teaching and learning and serves a variety of functions which are classified broadly into four: instructional functions, administrative functions, guidance functions and research functions(Opara,2016)

There are various types of tests.Based on traits, tests are classified as intelligence tests, specific aptitude test and achievement test.

Achievement Tests are tests designed to measure the degree of attainment of educational objectives in a content, subject or series of subjects It refers to tests designed to measure the knowledge, skills, and abilities attained by a test taker in a field, in a subject area, or in a content domain in which the test taker has received training or instruction. (orluwene,2012)

The purpose of achievement testing is to measure some aspect of the intellectual competence of human beings: what a person has learned to know or to do. Teachers use achievement tests to measure the attainments of their students. Employers use achievement tests to measure the competence of prospective employees. Professional associations use achievement tests to exclude unqualified applicants from the practice of the profession. In any circumstances where it is necessary or useful to distinguish persons of higher from those of lower competence or attainments, achievement testing is likely to occur.

No matter the nature or purpose of a particular achievement test, it is pertinent that it should be assessed forthe presence of error,reliability and validity.Douglas,Khavari & Ferber in Orluwene(2012) stated that regardless of the purpose of psychometric tests,it is commonly believed that psychometric instruments should be examined for internal psychometric properties. More so, that it is deemed necessary to determine the primary underlying dimension being measured by the items and to delete those items, which fail to measure this dimension adequately. To achieve this , the development of all tests anchors on some core test theories such as the classical test theory and latent trait theory. However, this study, Application of two-parameter latent trait model in the construction and validation of Agricultural science achievement test is anchored on the Latent trait test theory.

II. STATEMENT OF THE PROBLEM.

Agricultural science is a vocational subject in secondary school which provides students with skills, knowledge and attitudes necessary for future employment in Agricultural occupations. The main objectives of teaching Agricultural science in Nigerian secondary schools are to stimulate and sustain students' interest in Agriculture; impart functional knowledge and practical skills in Agriculture to students and prepare students for further studies and occupation in Agriculture. In spite of these lofty objectives, the records of the West African Examination Council (WAEC) and National Examination Council(NECO) reveal that Agricultural Science examination results are generally poor in Nigeria. This foretells doom for the future of agriculture in the nation's economy. It is suspected that the poor performance in Agricultural science examinations or tests is partly linked with lack of interest in the subject. In Nigeria, the society places emphasis on certain professions as more prestigious and this tend to influence the decisions of students about their career choices. Agriculture is perceived as a low-status occupation, and this perception has dissuaded a lot of students from investing their interest and energy in the study agricultural science or taking up a career or future studies in agriculture. Some students believe that studying agriculture limits their career options. They think there is lack of diverse and lucrative job opportunities in the field of agriculture compared to other fields, such as medicine, law, engineering or finance. These misconceptions about the field can be attributed to lack of awareness about the diverse and dynamic nature of modern agriculture such as agribusiness, precision farming, and agricultural technology.

Experience has also shown that students may also fail to perform or do well in tests or examinations sometimes due to the nature of the questions that make up the test. That is, some of the items may fall outside the scope of the scheme of work, syllabus or curriculum or may be too difficult depending on the average ability level of the students in the class. This implies that students may fail due to their low or poor ability or due to faulty instruments.

At the moment there is little or no validated instrument through which achievement in agricultural science among senior secondary two students in Rivers state can be assessed. This study was therefore set out to solve this problem by applying the two- parameter latent trait model (2-PLM) to construct and validate appropriate instrument(ASAT) through which that can be effectively achieved.

III. AIM AND OBJECTIVES OF THE STUDY.

The main aim of the study was to construct and validate the Agricultural Science achievement test(ASAT) based on two-parameter latent trait model. Specifically the objectives of the study were to ;

1. Examine the extent the items of ASAT complied with the IRT assumptions of unidimensionality and local independence.

2. Estimates the item difficulty of the ASATconstituting items based on the two-parameter latent trait model.

Research Questions

The following research questions guided the study:

1: What are the estimates of item difficulty of the ASAT?

2: What are the estimates of item discrimination of the ASAT?

IV. LITERATURE REVIEW.

Psychological tests have been defined as any psychometrically derived measurement instrument that assesses the psychological construct in which a structured sample of an examinee's behaviour in a specified domain is obtained and subsequently quantified , scored, interpreted and synthesized using a standardized process for the purpose of evaluative conclusion or recommendation(American Educational Research Association ,American psychological association &National council on Measurement in education,2004)

Tests can therefore be regarded as instruments used to determine the relative presence or absence of the trait measured for. It could also be a measurement instrument or device administered to someone to determine the relative value of the traits or skills to which the test relates (Orluwene,2012a;2012b).

Every test development exercise is anchored on two test theories; the classical test theory and the Item response theory or the Latent trait theory. The present work is anchored on Latent trait theory.According to Kpolovie (2014) Latent trait theory(LTT) is a modern theory of item development that emanated from extensive item analysis on the basis of computer-adaptive tests or tailored tests which hypothesizes that a test taker's score is rather dependent on the level of difficulty of specified range of items that he can correctly answer from the population of omnibus or infinite items that measure the latent trait or ability. It is anchored on the relationship between the individual examinee's latent or underlying psychological trait and his response to an item on a test which measures that specific attribute.

This is partly due to the principle of LTT that each individual possess a certain unknown or hypothetical level of any psychological attribute which is reflected to an extent in the person's response to any of the items on a test theory that trait.

DeMars(2010) asserted that Latent trait theory is a general framework for specifying mathematical functions that characterize the relationship between a person's ability or trait as measured by an instrument and the person's responses to the separate items in the instrument. It is a set of latent variable techniques especially designed to model the interaction between a subject's "ability" and the item level stimuli(Chalmers,2012).The latent trait framework emphasizes how responses can be thought of in probabilistic terms In LTT the item responses are considered the outcome(dependent) variables and the examinee's ability and the items' characteristics are the latent predictor(independent) variables.

Latent Trait Models (LTM)

Latent trait theory includes of a set of models that describe the interactions between a person and the test items.(Reckarse,2009).

There are two main variants of latent trait models. One is Gaussian or "normal Ogive" Model (O-give refers to the characteristics "S" shape of an item response function) derive from the assumption of normally-distributed measurement error. The other main variation consists of logistic-Ogive and Rasch models, which take advantage of some mathematical convenient properties of logarithmic relations (Anastasi &Urbina cited in Orluwene, 2012).

In general, all the models assume that the principle of local independence applies and that the items in the test being fitted by a model measure a common ability. Latent trait models are usually differentiated by the number of parameters estimated for the items and the nature of the item characteristics curve (ICC). In other words, Lazarsfeld and Henry cited in Orluwene (2012) posited that a significant distinction among the models is in the mathematical form taken by the ICC. However, the IRT model in focus here is the 2-PLM.

Two-parameter Logistic Model: This model proposed by Birnabaum (1968) contains two item parameter (difficult and discrimination indices) and latent ability omitting the guessing the parameter. Since latent trait model is a special kind of factor analysis mode, the discrimination parameter maybe interpreted as an adjusted factor loading indicating the relationship between the item and latent ability. The discrimination parameter is adjusted by dividing the factor loading by a measurement error parameter while the definition of difficulty indices in the one-parameter model holds. The two parameter logistic model contains values for each item difficulty b and the discrimination index ai but assumes the vulnerability of guessing c is zero. Here , b is still location on ability .

In two parameter logistic model (2PLM) the probability of correct response is mathematically given as follows:

$$P_i(\theta) = P_r\left(X_i = \frac{1}{\theta}, ai, bi\right) = \frac{exp[ai\ (\theta - bi)]}{1 + exp\ [ai(\theta - bi)/]}$$

Where $\times_i$ is the response to a single item i;

θ is the ability level of the test taker

ai is the discrimination level

bi is the difficultly level.

V. METHODS

This study adopted the instrumentation research design. The population for this study is 61049 senior secondary school two(SSS2)students in all the 286public senior secondary schools in Rivers State.Out of the 61049 population, 2. The sample for this study consisted of 1065 students who offer Agricultural science in public Senior Secondary School II in Rivers state. The multistage sampling technique, consisting of three stages was used to draw the sample for the study. At the first stage ,the study employed the cluster method to split Rivers state into three senatorial districts known as clusters. These were the Rivers East, Rivers South/East and Rivers West. After that,, the simple random method by balloting was applied to draw two(2) local government areas from each of the clusters in stage one. Rivers East(Emohua and Ikwerre);Rivers West (Ahoada-East and Onelga) and then Rivers South East (Oyigbo and Eleme). Finally,Proportionate stratified random method was applied at the third stage, to draw the study sample of 1065 students as shown in the table below.

Table 1: Proportional Stratified Random Sampling Table

S/N	Sample LGA	Population of Students	Proportion	Number Sampled
1	Emohua	3,571	0.17	179
2	Port Harcourt City	6,052	0.28	303
3	Eleme	1,756	0.08	88
4	Oyigbo	1,884	0.09	95
5	Ahoada East	2,762	0.13	139
6	Onelga	5,125	0.25	261
	Total			1065

This sample of 1065 is adequate for an item response theory anchored study such as this. Item response theory requires that sample size to be large (around 1,000) in order to obtain accurate item-parameter estimates that results in accurate estimates of ability, upon which some high-stakes decisions are made (Hambleton cited in Sahin& Anil 2017)

The instrument that will be used for collection of data in this study is titled "Agricultural achievement test".It is an120-item researcher-made instrument whose items were developed from the approvedSenior Secondary School two Agricultural science textbooks, based on the National Educational research and development Council(NERDC)curriculum. The instrument consists of two sections- sections A and B. Section A elicitspersonal information from the respondents such as class of study,age,gender, school type etc. while section B is made up of 120 multiple choice Objective Test items with 4 options A-D.

The planning stage of the ASAT started with a review of the Agricultural science textbooks such as Iwena(2018) and Erebor(2015). Subsequently, the Agricultural science teachers at the senior secondary school level were engaged in an interview which facilitated the formulation of the ASAT outlines and objectives. It equally facilitated the assignment of the ASAT items into the six different levels of the cognitive domain such as Knowledge, Comprehension, Application, Analysis, Synthesis and Evaluation following the Bloom's taxonomy of educational objectives for cognitive domain as put forward by Dr.Benjamin Bloom in 1956 (kpolovie,2014)

The second stage of the ASAT items development had to do with specifying the content areas, preparing the test specification, writing the items and ascertaining the right options(key) of the ASAT items.

The content areas of the senior secondary schooltwo Agricultural Science curriculumand their corresponding number of weeks covered in this test construction exerciseinclude:

Agricultural Ecology (6)weeks;.Agricultural Engineering(2)weeks.Forestry
(2) weeks ;Ornamental plants(2) weeks; Crop protection(4) weeks; Animal science(13) weeks; Agricultural economics and extension(5) weeks.

A test specification was prepared based on the content outlines, number of weeks, test length and percentage ratings of the instructional objectives for cognitive domain.

Table.2: Table of specification for the 120 items on Agricultural Science Achievement Test.(ASAT)

		Objectives						
Number of weeks	Content Areas	K 25%	C 35%	Ap 25%	An 5%	S 5%	E 5%	Total 100%
6	1	5	8	5	1	1	1	21
2	2	2	3	2	0	0	0	7
2	3	2	3	2	0	0	0	7
2	4	2	3	2	0	0	0	7
4	5	3	5	3	1	1	1	14

13	6	12	16	12	2	2	2	46
5	7	4	7	4	1	1	1	18
34	Total	30	45	30	5	5	5	120

Note: K= Knowledge, C=Comprehension, Ap=Application,An=Analysis,S=Synthesis and E=Evaluation

Copies of the instrument ,the scheme of work and the test specification were given to three(3) experts, one(1) in educational measurement and evaluation who was the researcher's supervisor and other two(2) subject specialists who were practicing Agricultural science teachers, for meticulous vetting, editing and critical review of the wordings of the items. The reasons for the exercise were to ensure non-inclusion of irrelevant or defective items in the test and to establish the face and the content validities of the instrument. The measurement expert matched the ASAT items to the appropriate instructional objective in the cognitive domain and content area in order to determine the content validity and equally commented on the appropriateness and arrangement of the stems and alternatives respectively. The Agricultural science teachers served to ascertain the correct options (keys), ensure clarity of words and suitability of the items to the comprehension level of the students. Eventually, the test items were revised and modified by the researcher based on the shortcomings identified by the expert and subject specialists in order to ensure a high degree objectivity. Consequently 12 items were deleted having 108 items left. These items constituted the ASAT for Trial testing.

Also, copies of the 108 itemed ASAT were administered to 100 Senior Secondary School two(SSS2)students who offer Agricultural Science. These students attended some other schools outside the ones which formed the original sample for the main study. The researcher administered, supervised and retrieved copies of the written test on the spot with the assistance of the Agricultural science class teachers who taught these students in their respective schools. The trial testing served to determine the adequacy of the duration for the test; to identify items whose language seemed ambiguous for the comprehension of the students and to evaluate the students' performance.

The copies of the written test were marked and the scores used to determine the performance of the students on the specific items that made up the ASAT. And based on the 2-PLM benchmarks, eight (8) items found wanting were also deleted. Hence, the ASAT for the main study was composed of 100 items. The analysis for the research questions was done using Ex-Calibre 4:2:2, a software for Item Response Theory developed by Assessment System Corporation.

VI. RESULTS

After the data analysis the result obtained for the research question 1 and 2 are presented in table 3 below.

Table 3: Estimates of Item Difficulty and Discrimination parameters of the ASAT.

Seq.	Item ID	P	R	A	B	Flag(s)
1	1	0.283	0.096	0.423	1.211	K
2	2	0.141	0.571	3.464	0.454	
3	3	0.523	0.367	1.466	-0.142	
4	4	0.591	-0.070	0.468	-0.636	K
5	5	0.241	0.191	0.552	1.164	K
6	6	0.566	-0.013	0.153	-1.140	K, La
7	7	0.285	0.602	3.665	0.190	
8	8	0.623	-0.012	0.166	-1.893	K, La
9	9	0.542	0.378	1.799	-0.158	
10	10	0.390	0.438	1.474	0.138	
11	11	0.499	0.458	1.813	-0.078	

12	12	0.484	0.508	1.839	-0.051	
13	13	0.160	0.062	0.491	1.936	K
14	14	0.561	-0.097	0.136	-1.150	K, La
15	15	0.590	-0.286	0.120	-1.805	K, La
16	16	0.662	0.129	0.230	-1.867	K, La
17	17	0.469	0.067	0.172	0.252	K, La
18	18	0.759	0.418	1.771	-0.675	
19	19	0.318	-0.104	0.379	1.069	K
20	20	0.412	-0.109	0.141	1.254	K, La
21	21	0.688	0.517	3.425	-0.297	
22	22	0.466	0.017	0.459	0.040	
23	23	0.292	0.071	0.378	1.272	K
24	24	0.647	0.353	1.129	-0.512	
25	25	0.226	0.465	1.475	0.532	
26	26	0.583	0.155	0.196	-1.142	K, La
27	27	0.475	0.581	4.064	-0.019	Ha
28	28	0.605	0.151	0.199	-1.389	K, La
29	29	0.838	0.021	0.288	-3.505	K, La, Lb
30	30	0.414	0.405	1.454	0.087	
31	31	0.309	0.502	2.001	0.248	
32	32	0.418	0.353	1.226	0.094	
33	33	0.418	0.373	1.280	0.090	
34	34	0.449	0.263	0.841	0.041	
35	35	0.700	0.497	2.506	-0.393	
36	36	0.320	0.561	4.972	0.125	Ha
37	37	0.363	0.537	2.829	0.117	
38	38	0.541	0.028	0.179	-0.680	K, La
39	39	0.718	-0.424	0.145	-3.733	K, La, Lb
40	40	0.276	0.574	3.389	0.212	
41	41	0.545	0.411	1.140	-0.224	
42	42	0.126	0.118	0.531	2.112	K
43	43	0.532	0.062	0.178	-0.571	K, La
44	44	0.449	0.325	1.314	0.017	
45	45	0.367	0.350	0.914	0.289	
46	46	0.279	-0.086	0.345	1.503	K
47	47	0.221	0.575	2.803	0.336	
48	48	0.317	0.471	1.692	0.266	
49	49	0.471	-0.062	0.153	0.270	K, La
50	50	0.356	0.304	0.969	0.308	
51	51	0.367	0.102	0.458	0.591	K
52	52	0.445	0.347	1.381	0.023	
53	53	0.612	-0.504	0.114	-2.316	K, La
54	54	0.540	0.434	1.219	-0.203	
55	55	0.489	-0.056	0.152	0.008	K, La
56	56	0.632	0.523	2.823	-0.242	
57	57	0.100	0.208	0.650	2.001	K
58	58	0.338	0.581	4.523	0.113	Ha

59	59	0.239	0.517	1.915	0.402	
60	60	0.480	0.164	0.177	0.108	K, La
61	61	0.230	0.536	1.788	0.448	
62	62	0.324	-0.057	0.325	1.204	K
63	63	0.268	0.281	0.882	0.650	
64	64	0.347	-0.075	0.382	0.854	
65	65	0.473	-0.159	0.146	0.259	K, La
66	66	0.611	-0.063	0.174	-1.643	K, La
67	67	0.276	0.531	1.914	0.323	
68	68	0.612	-0.276	0.135	-2.024	K, La
69	69	0.391	-0.001	0.152	1.490	K, La
70	70	0.295	0.537	2.143	0.261	
71	71	0.249	0.110	0.522	1.179	K
72	72	0.247	0.597	3.622	0.243	
73	73	0.534	-0.274	0.124	-0.762	K, La
74	74	0.341	0.371	1.151	0.303	
75	75	0.374	0.030	0.456	0.555	K
76	76	0.665	0.010	0.194	-2.181	K, La
77	77	0.427	0.245	0.687	0.146	
78	78	0.270	0.030	0.485	1.125	K
79	79	0.491	0.664	5.676	-0.026	Ha
80	80	0.309	0.509	2.373	0.218	
81	81	0.561	0.096	0.185	-0.911	K, La
82	82	0.437	0.446	1.107	0.055	
83	83	0.485	0.547	4.129	-0.028	Ha
84	84	0.205	0.594	4.588	0.266	Ha
85	85	0.462	0.365	1.186	-0.012	
86	86	0.421	0.360	1.069	0.101	
87	87	0.285	-0.014	0.430	1.166	K
88	88	0.268	0.477	1.795	0.358	
89	89	0.498	-0.083	0.145	-0.117	K, La
90	90	0.258	0.039	0.429	1.358	K
91	91	0.195	0.620	3.859	0.311	
92	92	0.333	0.422	1.033	0.356	
93	93	0.215	0.484	2.186	0.417	
94	94	0.231	0.467	1.994	0.409	
95	95	0.371	0.497	2.272	0.125	
96	96	0.511	0.373	1.133	-0.137	
97	97	0.448	0.333	1.318	0.019	
98	98	0.561	0.389	1.733	-0.198	
99	99	0.646	0.616	5.676	-0.176	Ha
100	100	0.480	0.478	2.478	-0.036	

Table 4.1 presents the classical statistics, the item parameters and any flag for each calibrated item. The K flag indicates that the keyed alternative did not have the highest correlation with total score. The F flag indicates that the item fit statistics (Z-residual) was significant and the item did not fit the model. The La, Lbflap indicate that the a/b parameters were lower than the minimum acceptable value. The Ha, Hb flags

indicates that the parameters were higher than the maximum acceptable value. From the table above, P is the probability of correct response while R is the item statistics for dichotomously scored item and which also indicates the data fit, 'a' is the discrimination index while 'b' is the difficulty index. Item 15 has the lowest 'a' value of 0.120, while items 79 and 99 have the highest values of 5.676 each. For the 'b' parameter, item 39 has the lowest 'b' value of -3.733 while item 42 has the highest 'b' value of 2.112.

Item 39 was flagged with K,La and Lb flag. This indicates that the item fit statistics was significant, thus the items did not fit the 2 parameter IRT model.

VII. DISCUSSION OF FINDINGS.

The b-parameter is the item difficulty parameter and indicates the location on the theta (θ) continuum where the probability of a correct response equals c/2 + .50. 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). The b parameter shows the examinee's θ value for which the item is appropriate. Higher difficulty indices or b-parameters greater than one(> 1.0) indicate that the item is more difficult; a value below minus one(< -1.0) indicates that the item is very easy. The b parameter shows the examinee's θ value for which the item is appropriate. From the results, it could be observed that 11% ASAT items comprising 13,20,23,42,46,57,62,69,71,87,90 were difficult; 23% ASAT items comprising items 6,8,14,15,16,17,26,28,29,38,39,43,49,53,55,60,65,66,68,73,76,81,89 were easy while 66% comprising 1,2,3,4,5,7,9,10,11,12,18,19,21,22,24,25,27,30,31,32,33,34,35,36,37,40,41,44,45,47,48,50,51,52,54,56,58,59,61,63,64,67,70,72,74,75,77,78,79,80,82,83,84,85,86,88,91,92,93,94,95,96,97,98,99,100 were slightly and moderately difficult. These slightly or moderately difficult items and were classified as good items in accordance with Opara (2016) who asserted that a good item should not be very difficult or very easy but moderate. In addition, the difficulty index or 'b' parameters of these items ranged from -0.012 to +1.40. This result is similar and falls in range with the findings of Nworgu and Agah(2012) who found difficult index range of -0.4 to 1.79 for Mathematics achievement test developed with the 3-PLM.

The ASAT discrimination index 'a' values ranged from 0.120 of item 15, which was the lowest, to +5.676 of items 79 and 99, which was the highest.

On the overall, 30% ASAT items comprising items 1,6,8,13,14,20,26,28,29,38,39,42,43,46,49,51,53,55,57,60,62,64,65,66,68,69,71,73,76,78,81,87,89,90 had low discrimination indices while 18 ASAT items comprising items 4,16,17,19,22,23,30,31,32,33,34,37,45,50,61,63,74,77 had their discrimination indices fall within +0.39 to +0.40 moderate discrimination indices indicative of good items in line with Orluwene (2012) who stated that discrimination indices or 'a' parameters within+0.39 to +0.40 are appropriate. Furthermore, 52 ASAT items comprising items 2,3,4,5,7,9,10,11,12,15,18,21,24,25,27,35,36,40,44,47,48,52,54,56,58,67,70,72,75,79,82,83,84,85,86,88,91,92,94,96,97,98,99,100 had high discrimination indices with the 'a' parameter ranging from +1.00 to +5.67 and were classified as good items. This result is consistent with Kpolovie and Emekene(2016) who embarked on item response theory validation of advanced progressive matrices for use in Nigeria and also found the discrimination indices or 'a' parameters of good items of the APM to be equal to or greater than one.

Item Selection

The final version of the ASAT was made up of 50 items which were selected from the items which had moderate difficulty and those which had high discrimination indices, falling within the range -0.500 to +0.532 & -1.000 to +1.00 respectively. These include:items 2,3,,5,7,9,10,11,12,18,21,24,25,27,30,31,32,33,35,36,37,40,41,44,47,48,52,54,56,58,59,61,67,70,72,79,80,82,83,84,85,88,91,92,93,94,95,96,97,98,99 &100

CONCLUSION

The ASAT had suitable difficulty index ranged from -3.733 to +2.112 and discrimination indices values ranging from 0.120 to +5.676 and therefore should employed to determine the performance of senior secondary two students in Agricultural Science. It has also been confirmed that two-parameter latent trait model is a viable model for the construction and validation of achievement tests.

REFERENCES

[1] American Educational Research Association, American psychological association & National council on Measurement in education (2004)

[2] Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. Journal of statistics software,48(6),1-29. Retrieved from http://www.jststsoft.org/v48/i06

[3] DeMars,c.(2010). Item response theory. Understanding statistics measurement.Oxford University press.

[4] Erebor,O.,(2013) Comprehensive Agricultural Science for senior secondary schools.A.Johnson publishers Ltd .Lagos Nigeria.

[5] Eze,P.O,Agbodion,V.V,Ossai,O.V,Ugwuezuonu ,A.U,Egenti,N,T&Eseadi,C(2017)Procedures for the construction and Development of Psychological Test.European Journal of scientificResearch.ISSN 1450-216x//1450-202,145(2)146-151. Retrieved from http://www.europeanjournalofscienific research.com

[6] Iweka,F (2015).Basic principles of educational measurement and evaluation .Omoku;Chifas Nigeria.

[7] Iweka.F.(2014). Comprehensive guide to test construction and administration. Omoku;Chifas Nigeria.

[8] Iweka.F.(2020).Measurement Theory and Advanced Multivariate Analyses.

[9] Jayanthni, J.(2014). Devlopment and Valdation of an achievement test in mathematics .International Journal of Mathematics and Statistics Invention (IJMSI), 2(4) 40-46.

[10] Kpolovie, P. J and Emekene C. O. (2016)Item Response Theory Validation Of Advanced Progressive Matrices In Nigeria. British Journal of psychology research.C Vol.4, No.1, pp.1-32, May 2016 Published by European Centre for Research Training and Development UK (www.eajournals.org)ISSN 2055-0863(Print), ISSN 2055-0871(Online)T

[11] Kpolovie,J.P.(2010) *Advanced Research Methods, New Owerri: Springfield*

[12] Kpolovie,J.P.(2014) *Tests, measurement and Evaluation in Education. Second*

[13] *National Examination Council(2014).Regulations and syllabuses for senior school certificate Examination(Revised Edition).pp52-80,Enugu Otuson Nigeria Limited.*

[14] *Nigerian Education Research and DevelopmentCouncil(2011).Senior Secondary Education Curriculum in Agricultural science for SSS1-3(Revised Edition) Lagos,NERDC press.*

[15] Nworgu B.G. and Agah J.J. (2012). Application of three parameter logistic model in the Calibration of a mathematics Achievement Test. Journal of Educational Assessment in Africa 29 (7) 162 – 172.

[16] Obilor,E.I.& Akpan, U. I. (2020) *Construction and Validation of Basic Science Achievement Test for Junior Secondary Three (JS3) Students in Public Secondary Schools in Akwa Ibom State.International Journal of Innovative Education Research 8(4):124-135, Oct.-Dec., 2020 © SEAHI PUBLICATIONS, 2020 www.seahipaj.org ISSN: 2354-2942*.

[17] *Ogbonna,J.U.(2018).Application of three-parameter latent trait model in the development and validation of mathematics achievement test. Unpublished ph.D thesis of university of Port Harcourt,Rivers State.*

[18] Opara,I.M.(2016).*Test construction and measurement concepts and applications.Owerri career publishers.*

[19] Orangi,A.M &DoramiK.(2010). Developing a social study achievement test for high school students based on Item response theory. Journal of psychological models and methods fall 2010.1(1).

[20] Orluwene, G.W.(2012).*Fundamentals of Testing and Non-Testing tools in*

[21] Orluwene,G.W.(2012).*Introduction to Test Theory and Development Process. Port Harcourt:Lucky Press Ltd.*

[22] Port Harcourt;Iyke communication. *Publishers Ltd.*

[23] Sahin,A.& Anil, D (2017).The effects of Test length and sample size on item parameters in item response theory.Educational Sciences:Theory and practice,17,321-335.http//dx.doi.org/10.12738/estp.1.0270.

[24] Samejima, Fumiko. Handbook of Modern Item Response Theory, ch. 6 (pp.85-100). Retrieved from https://www.researchgate.net/publication/279369018_Graded_Response_Model on April 21, 2018.