ASSESSMENT OF STANDARD ERROR ESTIMATES OF ITEM RESPONSE THEORY PARAMETERSOF BASIC SCIENCE IN BASIC EDUCATION CERTIFICATE EXAMINATION IN TARABA STATE, NIGERIA

BAKARI Iliyasu[1]
mayowando@mail.com

AMUCHE Christian Igomu[2]
ciamuche@gmail.com
Department of Educational Foundations,
Taraba State University, Jalingo

ATUMBA Bako Joseph[3]
Atumba.josephbako@tsu.edu.ng
Department of Guidance and Counselling,
Taraba State University, Jalingo

MOHAMMED Muktar[4]
muktarjal@gmail.com
Registry Department, Federal Polytechnic Bali,
Taraba State

ALI Hassan[5]
aleeeeuhassan@gmail.com
Department of Science Education[1&5]
Federal University Lokoja, Kogi State

Abstract

This study examined the Standard Error Estimates of Item Response Theory parameters of Basic Science in Basic Education Certificate examination in Taraba State, Nigeria. The study adopted a descriptive research design of survey type. The population of the study was 19,455 Upper Basic III Students, out of which a sample of 622 students was selected using proportional random sampling, from a total of 304 public Upper Basic Schools in Taraba State. The study adopted three sampling techniques at different stages in selecting the schools and students. Data were obtained using 2021 and 2022 Basic Education Certificate Examination Basic Science multiple choice test items developed by the Taraba State Education Resource Centre. The instrument was validated by three experts from Faculty of Education Taraba State University, Jalingo. The reliability of the examination was determined using Split-half formula which yield coefficients of 0.758 and 0.884 respectively. Data collected were analyzed using IRTPRO 4.2 to determine the Standard Error of Estimates parameter indices using the Three Parameter Logistic Model. All hypotheses were tested at 0.05 significance level using an independent sample t-test. Findings showed that not less than 70% of the items were considered

good as having their parameters SEEs which lie within the acceptable limit. It was concluded that Basic Science BECE questions for the years 2021 and 2022 were good based on the standard error of estimates. It was recommended among others that items that failed to meet the minimum threshold value of SEE should be eliminated or revised by the examination body.

Keywords: Standard Error Estimate, BECE, Difficulty, Discrimination, Guessing

Introduction
Science is any system of knowledge concerned with the physical world and its phenomena and entails unbiased observations and systematic experimentation. Science is a special type of knowledge, having a very strict methodology. There is no general scientific method as is commonly believed by many members of the general public (Enemarie, Ogbeba & Ajayi, 2019). Basic science formally known as Integrated Science is the first form of science a child comes across at the Upper Basic Education level. Hence, Basic Science prepares students at the Junior Secondary School (JSS) level for core science subjects (Physics, Chemistry, Biology and Mathematics) at the senior secondary school level (Ogunjobi, 2016). This highlighted the fact that for students to be able to study single science subjects at the senior secondary school level successfully such students had to be well-grounded in Basic Science at the Upper Basic Education level. The subject has been made compulsory for all students because of its importance to everyday life.
Despite the importance of the subject to students and national development, Oguguo and Lotobi (2019) lamented that there has been a downward trend in students' performance. This could be attributed to many factors including assessment of students' scores, measurement errors and inadequate exposure of students to activities, inadequate preparation, inability to comprehend questions, lack of effective methods of teaching science subjects, gender insensitivity and lack of qualified science teachers (Balarabe, 2016).

Educational assessment is central to the practice of education as it determines whether or not the goals of education are being met. The feedback from the student's educational assessment, not only provides information about the educational progress of learners in a systematic way but is also indicative of the quality of teaching and learning that take place in the classroom (Obioma, Junaidu & Ajagun, 2013). Therefore, despite the stressful nature of assessment, there is the need for Test developers of Basic Science Basic Education Certificate Examination in Taraba State to assess learners professionally to improve instruction.

It is observed that the Taraba State Education Resource Centre 2021 and 2022 BECE Basic Science multiple -choice test items developed by the Taraba State Education Resource Centre, has been determining test items before now using Classical Test Theory. Unfortunately, it is fraught with shortcomings. Among these are what Fan (1998) summarized as circular dependency in terms of the observed scores is sample dependent and the item statistics (item difficulty and item discrimination) are sample dependent. Moreover, the difficulty and discrimination indices are not taken into consideration in the estimation of observed scores. These shortcomings of Classical Test Theory informed the need to shift ground to the use of Item Response Theory (IRT).

Classical Test Theory (CTT) aims to measure a person's underlying ability or trait through a test score.

The basic intents of CTT are:

Reliability: Consistency of test scores across different administrations, forms, or raters.

Validity: Accuracy in measuring the intended construct or trait.

Objectivity: Minimizing bias and ensuring scores are free from personal opinions or influences.

Generalizability: Extending test results to broader populations or contexts.

Item Response Theory (IRT), also known as the latent response theory refers to a family of mathematical models that attempt to explain the relationship between latent traits (unobservable characteristics or attributes) and their manifestations (i.e., observed outcomes, responses or performance). They establish a link between the properties of items on an instrument, individuals responding to these items and the underlying trait being measured. IRT assumes that the latent construct (e.g., stress, knowledge, attitudes) and items of a measure are organized in an unobservable continuum. Therefore, its main purpose focuses on establishing the individual's position on that continuum.

Under IRT, the primary interest is in whether an examinee gets an item correctly or not, rather than in the raw test scores (Ayanwale, 2017). More importantly, to effectively estimate the ability of the examinee from his/her response to a particular test item, the item parameters of the test should be taken into consideration. The values of the item parameters and ability parameters depend on the type of parameter model used. In IRT, for test items that are dichotomously scored, there are three parameter models. These are one, two and three parameter logistic models. These models provide mathematical equations for the relation of the probability of correct response to ability (Baker, 2001). The standard error gives researchers an indication of how accurately their sample data represents their intended population (Agresti & Finlay, cited in Ogbonna & Opara, 2018). In other word, the standard error considers the variability of all means from samples of the same population and provides a way to measure the average distance between a sample mean and a population mean. In IRT, concept of standard error of estimates is adopted similar to the reliability concept in CTT. Standard Errors of Estimates (SEE) emphasize how confident we can be at each ability level (McAlphine 2013). The concept (SEEs) is derived from the item information. It emphasizes how much we can learn about the latent trait from an item and each item has an item information curve. In a three-parameter latent trait (3PLM) each item is described by three parameters (difficulty, discrimination and guessing) and each item parameter has its standard error of estimate.

Ogbonna and Opara (2018) studied standard error of estimation using 3PLM of item response theory of test items of achievement test in junior secondary school mathematics in Rivers state of Nigeria. Finding showed that the standard errors of the discrimination parameters (SEEa) ranges from 0.08 of item 37 to 0.115 of item 1, the difficulty index (SEEb) ranges from 0.030 of items 37 and 39 to 0.103 of item and the guessing parameter (SEEc) ranges from 0.039 of item 25 and 32 to 0.076 of item 1. Similarly, Matilda & Amen (2018) investigated flawed items of the West African Senior School Certificate Examination (WASSCE) Agricultural Science multiple choice items across 2013 to 2014 to determine the level of flaw in the item parameters across the stated examination years using Item Response Theory (IRT). The study established that there was no significant difference between standard errors of estimation (SEEs) of

parameters of objective test items across 2012 and 2014. There was no significant difference between the flawed items and the standard items with regards to the item difficulty, item discrimination and guessing parameters across the years.

The researchers observed that Basic Science of BECE questions in recent years are often fraught with reliability-related challenges as the standard process for constructing instruments seems not followed and the mode of administration is not properly monitored as examination materials are kept in custody of school's days before commencement of the examination. All these lead to significant measurement errors in the measurement process. When a measurement is poor, then there will be an inaccurate data-base inference, which in turn leads to wrong decision-making. The use of a poorly designed instrument is a major problem as it affects students' interests, career choices and achievement. When an instrument like BECE does not possess the necessary characteristics, it ought to, this means that the examination may not be reliable, and the effort to achieve the educational objectives will be in futility.

One criteria for rejecting an item is when an item's difficulty SEE is equal to or greater than a predetermined value (Toland, 2008). The purpose of the study was to examine standard error of estimates of IRT parameters of Basic Science in Basic Education Certificate conducted examination and specifically, to determined standard error estimates of difficulty indices, discrimination indices, and guessing chances of objective test items for the year 2021 and 2022. To further the understanding of item response theory parameters of basic science in BECE this study raised and answered the following questions:

i. What are the standard error of estimates (SEEs) of difficulty indices of Multiple-choice objective test items of BECE Basic Science of the years 2021 2022 in Taraba State?

ii. What are the standard error of estimates (SEEs) of discrimination indices of Multiple-choice objective test items of BECE Basic Science of the years 2021 and 2022 in Taraba State?

iii. What are the standard error of estimates (SEEs) of guessing chances of Multiple-choice objective test items of BECE Basic Science of the years 2021 and 2022 in Taraba State?

Research Hypotheses

The following null hypotheses were tested at 0.05 significance level.

Ho[1]: There is no significant difference between standard errors of estimations (SEEs) of difficulty indices of Multiple-choice objective test items of BECE Basic Science of the years 2021 and 2022 in Taraba State.

Ho[2]: There is no significant difference between standard errors of estimation (SEEs) of discrimination indices of Multiple-choice objective test items of BECE Basic Science of the year 2021 and 2022 in Taraba State.

Ho[3]: There is no significant difference between standard errors of estimation (SEEs) of guessing chances indices of Multiple-choice objective test items of BECE Basic Science of the year 2021 and 2022 in Taraba State.

Methodology

This study employed a descriptive research design of survey type. The population consisted of 19,455 Upper Basic III students of public schools and a total number of 304 public Upper Basic Schools in Taraba State, Nigeria. (Taraba State Education Resource Centre, Jalingo, 2023). A sample of 622 students were randomly selected using multi-stages sampling procedure. The sample size was obtained using Krejcie and Morgan 1970 probability sampling table at a confidence level of 99 per cent. The instruments used for data collection was adapted achievement test titled Basic Science Achievement Test (BSAT) "A" and "B". The instrument contains 2021 and 2022 Multiple-choice objective test items of BECE Basic Science, each contained 60 multiple-choice test items with options A-D.

The 2021 and 2022 were tagged "Type 'A' and 'B' One mark was given for each correct response and zero for incorrect response and omission. Each item in the instruments was expected to be answered within one minute whereas; the entire time allotted for the entire section is 60 minutes. The scoring guide which contained all keys to the sixty (60) multiple-choice items for each year was obtained from Education Resource Centre, Jalingo.

The instrument for data collection was validated by three experts from Faculty of Education, Taraba State University, Jalingo. The experts checked the relatedness of the items to the subject matter, appropriateness and content relevance to the issue under investigation. A pilot test was carried out to determine the reliability of BSAT types A and B. Data obtained from the pilot test were scored and analyzed using Split-half formula to determine the internal consistency of the instrument. The results of the reliability analyses revealed reliability coefficients of 0.785 and 0.884 respectively. Data collected was analyzed using Item Response Theory Modeling for Patient-Reported Outcomes (IRTPRO 4.2) to determine the parameter indices of test items difficulty, discrimination and probability of guessing values and standard errors of estimate of the item parameters.

The basis for interpretation of all the parameters was when an item's threshold value is less than 0.05 the item is regarded as a good item, while an item threshold value is equal or greater than 0.05 such item is considered as a poor item. All hypotheses were tested at 0.05 significance level using independent samples t-test.

Results

Research Question One: There is no significant difference between standard errors of estimations (SEEs) of difficulty indices of Multiple-choice objective test items of BECE Basic Science of the years 2021 and 2022 in Taraba State?

To answer this research question, the threshold values Standard Errors of Estimate of Difficulty Index of test items in Basic Science BECE for Year 2021 and 2022 based on Three Parameter Logistic Model (3PLM) are presented in table.

Item	Year 2021 SEE b	Item	Year 2021 SEE b	Item	Year 2022 SEE b	Item	Year 2022 SEE b
iv.	.04	31.	.03	i.	.01	31.	.03
v.	.03	32.	.03	ii.	.07*	32.	.03
vi.	.05*	33.	.04	iii.	.37*	33.	.04
vii.	.04	34.	.01	iv.	.04	34.	.04
viii.	.04	35.	.02	v.	.31*	35.	.04
ix.	.02	36.	.08*	vi.	.04	36.	.02
x.	.02	37.	.04	vii.	.03	37.	.05*
xi.	.04	38.	.05*	viii.	.01	38.	.04
xii.	.07*	39.	.03	ix.	.04	39.	.06*
xiii.	.06*	40.	.02	x.	.02	40.	.01
xiv.	.08*	41.	.04	xi.	.04	41.	.04
xv.	.03	42.	.06*	xii.	.03	42.	.04
xvi.	.04	43.	.04	xiii.	.04	43.	.03
xvii.	.04	44.	.01	xiv.	.04	44.	.04
xviii.	.35*	45.	.03	xv.	.31*	45.	.03
xix.	.03	46.	.03	xvi.	.02	46.	.02
xx.	.01	47.	.02	xvii.	.02	47.	.04
xxi.	.04	48.	.04	xviii.	.03	48.	.04
xxii.	.01	49.	.04	xix.	.04	49.	.05*
xxiii.	.02	50.	.03	xx.	.04	50.	.03
xxiv.	.03	51.	.06*	xxi.	.02	51.	.02
xxv.	.05*	52.	.04	xxii.	.04	52.	.01
xxvi.	.04	53.	.01	xxiii.	.47*	53.	.04
xxvii.	.03	54.	.03	xxiv.	.02	54.	.04
xxviii.	.01	55.	.04	xxv.	.04	55.	.05*
xxix.	.45*	56.	.03	xxvi.	.01	56.	.03
xxx.	.06*	57.	.02	xxvii.	.03	57.	.01
xxxi.	.04	58.	.04	xxviii.	.01	58.	.04
xxxii.	.03	59.	.02	xxix.	.04	59.	.03
xxxiii.	.04	60	.04	xxx.	.01	60	.02

Table 1, showed the standard error of estimates of difficulty index for Basic Science BECE examination for the years 2021 and 2022. The result showed that 48 items representing 80% of the items in year 2021 are good with threshold values of less than .05. While 12 items representing 20% of the items in the year 2021 are considered poor items with threshold values of ≥ 0.05, the items flagged as poor are items 3, 9, 10, 11, 15, 20, 22, 26, 27, 37, 38, 42 and 51. The result in the table also shows that 51 items representing 85% of the items in year 2022 are good items with threshold values of less than .05. While 9 items representing 15% of the items in the year 2022 are considered poor items with

threshold values of ≥ 0.05, the items flagged as poor are items 2, 3, 5, 15, 23, 37, 39, 49 and 55. The result shows that the items are good.

Research Question Two: There is no significant difference between standard errors of estimation (SEEs) of discrimination indices of Multiple-choice objective test items of BECE Basic Science of the year 2021 and 2022 in Taraba State?

To answer this research question, the threshold values Standard Errors of Estimate of Discrimination Index of Test Items in Basic Science BECE for years 2021 and 2022 based on Three Parameter Logistic Model are presented in the table.

Item	Year 2021 SEE a	Item	Year 2021 SEE a	Item	Year 2022 SEE a	Item	Year 2022 SEE a
i.	.03	31.	.02	i.	.03	31.	.01
ii.	.02	32.	.06*	ii.	.03	32.	.02
iii.	.05*	33.	.03	iii	.01	33.	.03
iv.	.03	34.	.01	iv	.04	34.	.05*
v.	.04	35.	.04	v.	.03	35.	.03
vi.	.01	36.	.03	vi	.10*	36.	.02
vii.	.04	37.	.01	vi	.02	37.	.04
viii.	.04	38.	.03	vi	.04	38.	.03
ix.	.04	39.	.01	ix	.03	39.	.05*
x.	.01	40.	.04	x.	.01	40.	.03
xi.	.07*	41.	.15*	xi	.03	41.	.09*
xii.	.52*	42.	.03	xi	.04	42.	.03
xiii.	.01	43.	.02	xi	.04	43.	.04
xiv.	.03	44.	.04	xi	.03	44.	.02
xv.	.04	45.	.02	xv	.01	45.	.03
xvi.	.04	46.	.06*	xv	.92*	46.	.04
xvii.	.03	47.	.07*	xv	.05*	47.	.05*
xviii.	.04	48.	.02	xv	.03	48.	.02
xix.	.03	49.	.04	xi	.02	49.	.04
xx.	.01	50.	.03	xx	.05*	50.	.02
xxi.	.04	51.	.01	xx	.03	51.	.03
xxii.	.05*	52.	.04	xx	.03	52.	.03
xxiii.	.04	53.	.02	xx	.02	53.	.03
xxiv.	.85*	54.	.03	xx	.06*	54.	.01
xxv.	.03	55.	.05*	xx	.01	55.	.04
xxvi.	.01	56.	.02	xx	.03	56.	.03
xxvii.	.04	57.	.75*	xx	.03	57.	.06*
xxviii.	.09*	58.	.03	xx	.04	58.	.04
xxix.	.04	59.	.07*	xx	.04	59.	.01
xxx.	.01	60	.01	xx	.08*	60	.03

Table 2, showed the standard error of estimates of discrimination index for Basic Science BEC examination for the years 2021 and 2022. The result showed that 47 items representing 78% of the items in year 2021 are good with threshold values of less than .05. While 13 items representing 22% of the items in the year 2021 are considered poor items with threshold values of ≥ 0.05, the items flagged as poor are items 3, 11, 12, 22, 24, 28, 32, 41, 46, 47, 55, 57 and 59. The result in the table also shows that 49 items representing 82% of the items in year 2022 are good items with threshold values of less than 0.05. While 11 items representing 18% of the items in the year 2022 are considered poor items with threshold values of ≥ 0.05, the items flagged as poor are items 6, 16, 17, 20, 24, 30, 34, 39, 41, 47 and 57. The result shows that the items are good.

Research Question Three: There is no significant difference between standard errors of estimation (SEEs) of guessing chances indices of Multiple-choice objective test items of BECE Basic Science of the year 2021 and 2022 in Taraba State?

To answer this research question, The Threshold values Standard Error of Estimate of Guessing Chance of test items in Basic Science BECE for Year 2021 and 2022 based on Three Parameter Logistic Model are presented in the table.

Item	Year 2021 SEE c	Item	Year 2021 SEE c	Item	Year 2022 SEE c	Item	Year 2022 SEE c
i.	.04	31.	.04	1.	.03	31.	.03
ii.	.03	32.	.01	2.	.02	32.	.02
iii.	.01	33.	.03	3.	.04	33.	.03
iv.	.56*	34.	.03	4.	.03	34.	.07*
v.	.03	35.	.04	5.	.03	35.	.01
vi.	.02	36.	.03	6.	.01	36.	.03
vii.	.04	37.	.01	7.	.02	37.	.04
viii.	.05*	38.	.02	8.	.02	38.	.04
ix.	.04	39.	.04	9.	.04	39.	.02
x.	.02	40.	.03	10.	.05*	40.	.05*
xi.	.05*	41.	.04	11.	.09*	41.	.04
xii.	.04	42.	.07*	12.	.01	42.	.03
xiii.	.03	43.	.03	13.	.03	43.	.02
xiv.	.01	44.	.04	14.	.04	44.	.04
xv.	.04	45.	.06*	15.	.01	45.	.40*
xvi.	.03	46.	.03	16.	.04	46.	.08*
xvii.	.05*	47.	.04	17.	.03	47.	.03
xviii.	.02	48.	.01	18.	.04	48.	.05*
xix.	.03	49.	.04	19.	.04	49.	.01
xx.	.01	50.	.05*	20.	.03	50.	.04
xxi.	.03	51.	.04	21.	.01	51.	.03
xxii.	.04	52.	.03	22.	.74*	52.	.03
xxiii.	.04	53.	.02	23.	.02	53.	.04
xxiv.	.09*	54.	.01	24.	.03	54.	.01
xxv.	.04	55.	.04	25.	.01	55.	.03
xxvi.	.01	56.	.03	26.	1.60*	56.	.03
xxvii.	.31*	57.	.06*	27.	.01	57.	.33*

xxviii.	.03	58.	.01	28.	.04	58.	.04
xxix.	.05*	59.	.01	29.	.04	59.	.06*
xxx.	.04	60	.07*	30.	.03	60	.02

Table 3, showed the standard error of estimates of guessing chances for Basic Science BEC examination for the year 2021 and 2022. The result showed that 48 items representing 80% of the items in year 2021 are good with threshold values of less than 0.05. While 12 items representing 20% of the items in the year 2021 are considered poor items with threshold values of ≥0.05, the items flagged as poor are items 4, 8, 11, 17, 24, 31, 39, 42, 45, 50, 57 and 60. The result in the table also shows that 49 items representing 82% of the items in year 2022 are good items with threshold values of less than .05. While 11 items representing 18% of the items in the year 2022 are considered poor items with threshold values of ≥0.05, the items flagged as poor are items 10, 11, 22, 26, 34, 40, 45, 46, 48, 57 and 59. The result shows that the items are good.

Hypothesis One: There is no significant difference between standard errors of estimations (SEEs) of difficulty indices of Multiple-choice objective test items of BECE Basic Science of the years 2021 and 2022 in Taraba State.

Table 4: Independent Samples t-test of Significance of Difference between Standard Errors of Estimate (SEEs) of Difficulty Indices of Basic Science BECE Questions for Year 2021 and 2022

	Levene's Test for Equality of Variances		t-test for Equality of Means			
	F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference
Equal variances assumed	1.215	.273	-.456	118	.649	-.00650
Equal variances not assumed			-.456	112.119	.649	-.00650

Result of independent-samples t-test in Table 4 shows that there is no significant difference between standard error of estimation (SEEs) of difficulty indices of objective test items of Basic Science BECE questions for year 2021 and for year 2022 in Taraba State(t(118) = -.456, p = 0.649). Thus, the null hypothesis which states that there is no significant difference between standard errors of estimation (SEEs) of difficulty indices of objective test items of Basic Science BECE questions for year 2021 and for year 2022 in Taraba State is retained. This implies that the examinations did not differ significantly in their standard error estimation of difficulty indices of objective test items of Basic Science BECE questions.

Hypothesis Two: There is no significant difference between standard errors of estimation (SEEs) of discrimination indices of Multiple-choice objective test items of BECE Basic Science of the year 2021 and 2022 in Taraba State.

Table 5: Independent Samples t-test of Significance of Difference between Standard Errors of Estimation (SEEs) of Discrimination Indices of Basic Science BECE Questions for Year 2021 and 2022

	Levene's Test for Equality of Variances		t-test for Equality of Means			
	F	**Sig.**	**T**	**Df**	**Sig. (2-tailed)**	**Mean Difference**
Equal variances assumed	2.149	.145	.802	118	.424	.01983
Equal variances not assumed			.802	110.009	.424	.01983

Result of independent-samples t-test in Table 5 showed that there is no significant difference between standard errors of estimation (SEEs) of discrimination indices of objective test items of Basic Science BECE questions for year 2021 and for year 2022 in Taraba State(t(118) = -.802, p = 0.424).Thus, the null hypothesis which states that there is no significant difference between standard errors of estimation (SEEs) of discrimination indices of objective test items of Basic Science BECE questions for year 2021 and for year 2022 in Taraba State is retained. This implies that the examinations did not differ significantly in their standard error estimation of discrimination indices of objective test items of Basic Science BECE questions.

Hypothesis Three: There is no significant difference between standard errors of estimation (SEEs) of guessing chances indices of Multiple-choice objective test items of BECE Basic Science of the year 2021 and 2022 in Taraba State.

Table 6: Independent Samples t-test of Significance of Difference between Standard Errors of Estimation (SEEs) of guessing chances of Basic Science BECE Questions for Year 2021 and 2022

	Levene's Test for Equality of Variances		t-test for Equality of Means			
	F	**Sig.**	**T**	**Df**	**Sig. (2-tailed)**	**Mean Difference**
Equal variances assumed	4.727	.032	-1.095	118	.276	-.03400
Equal variances not assumed			-1.095	72.653	.277	-.03400

Result of independent-samples t-test in Table 6 showed that there is no significant difference between standard errors of estimation (SEEs) of guessing chances of objective test items of Basic Science BECE questions for year 2021 and for year 2022 in Taraba State(t(118) = -1.095,

p = 0.276).Thus, the null hypothesis which states that there is no significant difference between standard errors of estimation (SEEs) of guessing chances of objective test items of Basic Science BECE questions for year 2021 and for year 2022 in Taraba State is retained. This implies that the examinations did not differ significantly in their standard error estimation of guessing chances of objective test items of Basic Science BECE questions.

Discussion of the findings
Finding of this study revealed, that 48 items representing 80% of the items in year 2021 were good with difficulty threshold values of less than .05. While 12 items representing 20% of the items in the year 2021 are considered poor items with threshold values of ≥ 0.05. In the year 2022, 51 items representing 85% of the items were good items with threshold values of less than .05. The finding is similar to Ogbonna and Opara (2018) whose study shows that the standard errors of the difficulty index (SEEb) ranges from 0.030 of items 37 and 39 to 0.103 of item. Also, the finding is in agreement with Oguguo and Lotobi (2019) whose finding of their study revealed that 40 and 45 items satisfied the IRT difficulty, discrimination and guessing parameter respectively. While 38 items satisfied the combined three IRT parameter estimates in BECE Basic Science.

Finding of this study also revealed that the research question two with 47 items representing 78% of the items in year 2021 were good with discrimination threshold values of less than .05. While 13 items representing 22% of the items in the year 2021 were considered poor items with threshold values of ≥ 0.05. The result also showed that 49 items representing 82% of the items in year 2022 were good items with threshold values of less than .05. While 11 items representing 18% of the items in the year 2022 are considered poor items with threshold values of ≥ 0.05. The finding is similar to Ibiene, Njigwum and Ibeh (2021) who revealed that 55 items were considered good while 43 items were marked for elimination. The finding is contrary to Ogbonna and Opara (2018) whose study showed that the standard errors of the discrimination parameters (SEEa) ranges from 0.08 of item 37 to 0.115 of item 1.

Finding revealed that the standard error of estimates of guessing chances for Basic Science BECE examination for the year 2021 and 2022. The result showed that 48 items representing 80% of the items in year 2021 were good with threshold values of less than .05. While 12 items representing 20% of the items in the year 2021 were considered poor items with threshold values of ≥ 0.05. The result also showed that 49 items representing 82% of the items in year 2022 were good items with threshold values of less than .05. While 11 items representing 18% of the items in the year 2022 were considered poor items with threshold values of ≥ 0.05. The result showed that the items were good. The finding correlates with Ani (2014) whose study revealed that 49 items of the multiple-choice question in Economics were reliable. This finding is also similar to Ogbonna and Opara (2018) whose study showed that the standard errors of the guessing parameter (SEEc) ranges from 0.039 of item 25 and 32 to 0.076 of item 1.

Conclusion
In line with the findings of this study, the following conclusion were made:
    a. Basic Science BECE for years 2021 and 2022 were considered and their parameters of SEEs lie within the acceptable limit.

b. It was also observed that the derivation of the standard errors of estimates depend on the item parameters.

c. There is a strong relationship between the item defaulting and its standard error of estimate. The lower the standard error of an item parameter, the more information the items provide.

d. Basic Science BECE Multiple-choice objectives for the years 2021 and 2022 were good based on the standard error of estimates.

Recommendations

Based on the findings of the study, the following recommendations were made:

• Items that failed to meet the minimum threshold value of SEE should be eliminated or revised by the examination body.

• The Taraba State Education Resource Centre (ERC) saddled with the responsibility of constructing Basic Science BECE test items should pilot test generated items before their administration to students for the purpose of examinations.

• Standard error of estimates of item response theory parameters be used to identify or differentiate between good and bad items in any test development process.

References

Ani, E. N. (2014). Application of Item Response Theory in the development and validation of multipile-choice test in Economics. Unpublished M.Ed Thesis, Department of Science Education, University of Nigeria, Nsukka.

Ayanwale, M.A. (2017). Efficacy of Item Response Theory in score ranking and concurrent validity of dichotomous and polytomous response mathematics achievement test in Osun State, Nigeria. Unpublished Ph.D thesis. Institute of Education. University of Ibadan.

Baker, F.B. (2001). *The basics of item response theory.* ERIC Clearinghouse on Assessment and Evaluation, College Park, MD.

Balaraba, M. (2016). Effects of psycho-social factors on the students' academic achievement in Basic Science. *Journal of Research in Science Teaching,5(1),* 234-240.

Enemarie, V., Ogbeba, J. & Ajayi, V.O. (2019). Students' Achievement in Basic Science in Basic Education Certificate Examination as a Predictor of their Performance in Biology in Senior Secondary Certificate Examination. Department of Curriculum and Teaching, Benue State University, Makurdi, Nigeria. Research Gate. A Discourse on Educational Issues, 281 – 295. 281 – 295.

Fan, X. (1998). Item Response Theory and Classical Test Theory: An Empirical Comparison of Their Item/Person Statistics. *Educational and Psychological Measurement, 58,* (32), 357-38.

Ibiene, T. L., Njigwum, A. S. & Ibeh, C. P. (2021). validation of Basic Science Achievement test using innovative techniques (R software). *International Journal of Innovative Research and Development. 10*(9), 77-87.

Matilda U. O. & Amen, V. U. (2018). Investigating flaw items of WASSCE Agricultural Science multiple choice items in Edo South Senatorial District, Nigeria. *African Journal of Interdisciplinary Studies. 11*(1), 1-7.

NECO.(2018). Annual Report 2018. Federal Ministry of Education https://education.gov.ng

Ogbonna, G. K. & Opara, N. R. (2018). Comparison of classical test theory and item response theory in terms of item parameters. *European Journal of Research on Education, 2*(1), 1-6.

Oguguo, B. C. E., & Lotobi, R. A. (2019). Parameters of basic science test items of 2011 Basic Education certificate examination using item response theory (IRT) approach in Delta State, Nigeria. *European Scientific Institute, 1*(6), 22 – 36.

Osterlind, S. J & Everson H. T (2019). Differential Item Functioning. 2455 Teller Road, Thousand Oaks California 91320 United States of America: SAGE Publications, Inc.; 2019

Toland, M.D. (2008). Determining the accuracy of item transactions,19:3,1028-1029. Retrieved from http://www.rasch.org/rmt/rmt193c. htm *Trials Based on Item Response Theory*, Controlled Clinical Trials, 24,390–410.Trust Publishers, pp. 66-83.

Thissen, D. & Wainer, H. (2018). Some standard errors in item response theory. *Psychometrika* ,47, 397- 412.