NATIONAL EXAMINATIONS COUNCIL SENIOR SCHOOL CERTIFICATE PHYSICS PRACTICAL EXAMINATION MEASURE OF PRACTICAL SKILLS PROFICIENCY

BY

Alaba Adeyemi Adediwura: Department of Educational Foundations, Faculty of Education, Obafemi Awolowo University, Ile-Ife; E-mail: yemtoy20002000@gmail.com or alawura@oauife.edu.ng

Yusuf Olayinka Shogbesan: Department of Arts and Social Sciences Education, Faculty of Education, Al-Hikmah University, Ilorin; E-mail: yusufolayinka813@yahoo.com
&
Oluseyi Peter Adeoye: Department of Educational Foundations and Management, Osun State College of Education, Ila-Orangun; E-mail: seyipeter680@gmail.com

Abstract
The study determined the precision with which Senior School Certificate (SSC) Physics practical examination items measure proficiency in Physics practical and estimated the proportion examinees' ability precisely measured. It also determined the extent to which the examination measure the same theoretical construct among students as well as its reliability The study adopted the descriptive survey research design. The study's population comprised all Senior Secondary Three (SS III) students offering Physics in Osun State during the second term of 2017/2018 academic session. The study sample consist of 631 students chosen using a multistage sampling technique. The NECO 2018 SSC physics practical examination items were used to collect relevant data for the study. The Standard Error of Estimate (SE), limited-information fit measure, and empirical reliability were used to analyze the collected data. The results showed that NECO physic practical examination measure examinees proficiency in Physics practical skills with low precision. However, the ability scores of the examinees were estimated high precision. The results also showed that 88.6% of the examinees fitted the model used in the calibration of the data. Further results showed that the construct validity of the NECO test was low but with a high reliability. The study concluded that the practical examination measures examinees' practical skills proficiency with low precision and low construct validity. across sub-groups. It is therefore, recommended that examination bodies ensure construct validity and measurement precision during item construction.
Keywords: Precision, Proficiency, measure, Reliability, Theoretical construct and Validity

Introduction

Physics is known to many as a difficult subject as it involves understanding of abstract concepts and complex formulae. Practical work plays an important role in the learning of Physics. Practical work in the laboratory enables phenomena to be reproduced and thus enhance students' understanding of abstract concepts (Deacon & Hajek, 2011). Practical work also enables students to check on theories and principles and help to correct misconceptions and confusion between what is learnt in schools and what is experienced in daily life, and thus create links between theory and practice (Prades & Espinar, 2010). The nucleus of all technology is commonly known as physics. Simply put, physics is in control of all aspects of technology. It also demonstrates that physics is at the heart of all technological advancements. At the senior secondary school level, physics is classified as a field of science that deals with matter, energy, their interrelation, and their measurements. Practical approach is a teaching method that teachers can use to effectively teach physics in senior secondary schools. Practical work can be planned experiment, observation, specimen collection and even field work carried out during or after a physics lesson.

According to Opong (1981) on way of differentiating a scientist from a non-scientist is from experimentation. Hence, Abdullahi (1982) defines experiment as operations or procedure used for the purpose of testing a supposition, confirming the known and discovering the unknown. In discovering the unknown, new theories are involved hence experimentation and theory are inter-dependent; each nourishing the other. According to Solomon (1980) "Science belongs in the laboratory in the same way as cooking belongs in the kitchen and gardening belongs in the garden." these portray the value of hands-on learning in school laboratories. Experiment is outstanding feature of practical work; it helps to prepare the mind to think creatively, rationally and objectively without bias because science by its nature deals with truth. Any statement made has to be tested to confirm its validity and reliability. Physics teachers should try to inculcate this idea of objectivity into their students. If students are allowed to collect specimen by

ISSN 2756-4444
E-ISSN 2756-4452

themselves, it can bring about increased motivation and meaningful learning, student tends to understand better during practical class that they did themselves. In essence, it aids comprehension.

According to Azar and Şengülec (2011), practical activities in Physics, improve students' learning, help develop positive attitude towards Physics and more importantly engender permanence of knowledge. Tamir (1977) listed the following goals for using laboratories in science education: to assist students in (a) comprehending abstract and complex scientific concepts through the use of concrete materials, (b) developing problem-solving and analytical skills, (c) developing practical abilities, and (d) developing good attitudes toward science. The Federal Government of Nigeria's National Policy on Education (FGN, 2013 advocates inquiry-based learning for all science instruction and classroom activities. Science experiments that promote the learning of scientific concepts and help secondary school students acquire process and problem-solving skills are recommended as one of the most effective ways to promote the learning of scientific concepts and help secondary school students acquire process and problem-solving skills. The role of functional Physics education in technological advancements of nations cannot be over emphasized. Highly developed Physics education, no doubt, is needed to keep on producing future engineers, technologists and scientists for driving the economic engines of the nation. Any country who decides to pay lip service to the development of her Physics education will surely lag behind others in the comity of nations. It is against this background that the major stakeholders (e.g., Physics teachers, Parents, public examining bodies and government officials) in Physics education in Nigeria continue to express worries about the declining level of enrolment and performance of Nigerian students in Physics, especially at secondary school level.

The statistics of enrolment of students for senior secondary school certificate examination (SSSCE) being conducted by National Examination Council (NECO) shows that on the average, in the last three years, less than 40% of the total candidates who registered for SSSCE in Nigeria actually registered for Physics. On achievement, the level of achievement of candidates between 2016 and 2018 was below 50%. In fact, the percentage of candidates, who passed at the minimum credit level (C6), on the average, was below 45%. A pertinent question that arises is what went wrong? From the NECO Chief Examiners‟ Report for 2016, 2017, and 2018 attention is being drawn to the fact that students are not doing well in Physics practical. Some of the aspects of Physics practical where students are exhibiting problems include: optics (use of protractor, parallax error arising from sighting of images, reading of values and measurement of angles from laboratory instruments) and Simple Harmonic Motion (Measurement of time, taking of period of oscillation, extension of spiral spring and determination of acceleration due to gravity). Other related problems identified included: wrong responses to questions bordering on the theory of the experiment, inability to plot graphs involving small values and make deductions from the graphs (NECO Chief Examiners' Report). These findings as enumerated in the NECO Chief Examiners‟ Report indicate that students do not have requisite skills in Physics practical. Thus, if this is the case, it could be one of the reasons why students continue to perform poorly in Physics. Therefore, there is the need to establish the extent to which Senior School Certificate (SSC) Physics practical examination measure practical skills proficiency in the students.

Purpose of the Study

The broad objective of this study is to establish SSC Physics practical examination measure of practical skills proficiency in the students. Specifically, the objectives of the study are to:

1. determine the precision with which SSCE Physics practical test items measure the examinees proficiency in Physics practical;
2. estimate the proportion of the examinees' ability precisely measured by SSCE Physics practical test;
3. determine the extent to which SSC Physics practical examination measure the same theoretical construct among: (a) male and female students, (b) Students from schools located in rural and urban areas and (c) Students from private and public schools; and
4. estimate the reliability of SSC Physics practical examination conducted by NECO.

Research Questions

1. What is the precision with which SSCE Physics practical test items measure the examinees proficiency in Physics practical?
2. What proportion of the examinees' ability were precisely measured by SSCE Physics practical test?

ISSN 2756-4444
E-ISSN 2756-4452

3. To what extent does the SSCE Physics practical test measure the same theoretical construct among: (a) male and female students, (b) Students from schools located in rural and urban areas and (c) Students from private and public schools?
4. How reliable is the SSCE Physics practical test conducted by NECO?

Methodology

The study adopted the descriptive survey research design. The study's population comprised all Senior Secondary Three (SS III) students offering Physics in Osun State during the second term of 2017/2018 academic session. The study sample consist of 631 students chosen using a multistage sampling technique. From each of the three senatorial district that made up the State, five Local Government Areas (LGAs) were chosen at random and from each of the LGAs, stratified random sampling was used to choose four schools, with school ownership as the basis for stratification. The study made use of intact class of physics students in each of the selected schools that therefore made up the sample size of 661 students. The NECO 2018 SSC physics practical examination items were used to collect relevant data for the study. The Standard Error of Estimate (SE), limited-information fit measure, and empirical reliability were used to analyze the collected data.

Results

Research Question 1: What is the precision with which NECO Physics practical test items measure the examinees proficiency in Physics practical?

According to DeMars (2010), Standard Error of estimate (SE) is obtained as the reciprocal of the square root of the amount of item or test information. Mathematically,

$$SE = \frac{1}{\sqrt{I(\theta)}}$$

Where $I(\theta)$ = item or test information at a given ability level. In this study the ability level of each of the examinees that took the test were used. According to rule of thumb, Standard error of Estimate < 0.5 for an item or a test to be considered informative and precise for the measurement of the parameter estimate. An informative item or test is one whose precision of measuring the underlying estimate is high. The estimate of the SE of NECO Physics practical test items is presented Table 1.

Table 1: Standard Error of Measurement of NECO SSSCE Physics Practical Test Items

S/N	Item	a Est	Remark	b1 Est	b2 Est	b3 Est	b4 Est	b5 Est	b6 Est	Remark
1	Q1_ai	0.23	Precise	0.33						Precise
2	Q1_aii	0.08	Precise	0.35	0.67					Not Precise
3	Q1_aiii	0.13	Precise	0.4	1.6					Not Precise
4	Q1_aiv	0.06	Precise	0.48	0.99					Not Precise
5	Q1_av	0.11	Precise	0.43	2.75					Not Precise
6	Q1_avi	0.06	Precise	1	1.68					Not Precise
7	Q1_avii	0.12	Precise	0.55						Not Precise
8	Q1_Gra	0.17	Precise	0.11	0.11	0.21	0.21	0.19	0.13	Precise
9	Q1_Slo	0.42	Precise	0.07	0.07					Precise
10	Q1_Eva	0.27	Precise	0.07	0.37					Precise
11	Q1_Prec	0.12	Precise	0.13	0.12					Precise
12	Q1_bi	0.06	Precise	1.5	0.9					Not Precise
13	Q1_bii	0.08	Precise	1.56	0.59					Not Precise
14	Q2_ai	0.06	Precise	0.5	0.52	0.38				Not Precise
15	Q2_aii	0.08	Precise	0.26	0.41	1.36				Not Precise
16	Q2_aiii	0.07	Precise	0.22	0.25					Precise
17	Q2_aiv	0.07	Precise	0.27	0.37					Precise

ISSN 2756-4444
E-ISSN 2756-4452

18	Q2_av	0.11	Precise	0.21	1.31					Not Precise
19	Q2_avi	0.09	Precise	0.35	1.08					Not Precise
20	Q2_Gra	0.12	Precise	0.13	0.13	0.38	0.38	0.29	0.14	Precise
21	Q2_Slo	0.14	Precise	0.15	0.14					Precise
22	Q2_Prec	0.14	Precise	0.12	0.12					Precise
23	Q2_bi	0.1	Precise	0.3	0.28					Precise
24	Q2_bii	0.08	Precise	0.46	0.4	0.62				Not Precise
25	Q3_ai	0.1	Precise	6.62						Not Precise
26	Q3_aii	0.1	Precise	50644.56	149407.5					Not Precise
27	Q3_aiii	0.09	Precise	12.53	21.89					Not Precise
28	Q3_aiv	0.05	Precise	3.02	25.87					Not Precise
29	Q3_av	0.09	Precise	22.88	84.29					Not Precise
30	Q3_avi	0.09	Precise	54.09	251.16					Not Precise
31	Q3_Gra	0.08	Precise	0.16	0.15	0.2	0.27	0.28	0.19	Precise
32	Q3_Slo	0.18	Precise	0.13	0.12	0.15				Precise
33	Q3_Inter	0.26	Precise	0.07						Precise
34	Q3_Eva	0.46	Precise	0.06	0.12					Precise
35	Q3_Pre	0.07	Precise	0.3	0.27					Precise
36	Q3_bi	0.1	Precise	0.21	0.19					Precise
37	Q3_bii	0.06	Precise	0.47	0.37					Precise

Table 1 showed that the estimated SE for the discrimination of all the task in the NECO test was less than 0.5. Furthermore, the table showed that for: item 1, the only step difficulty returned SE less than 0.5, implying that the precision with which the difficulty of the item was estimated was high; item 2, while step 1 was less than 0.5, step difficulty 2 was greater than 0.5, implying that the precision with which the difficulty of the item was estimated was low. The precision with which the remaining items were estimated and assessed in similar manner. The assessment showed that the precision with which difficulty parameters of 19 (2, 3, 4, 5, 6, 7, 12, 13, 14, 15, 18, 19, 24, 25, 26, 27, 28, 29 and 30) items were estimated was low and the precision with which the difficulty parameters of 18 (1, 8, 9, 10, 11, 16, 17, 20, 21, 22, 23, 31, 32, 33, 34, 35, 36 and 37) items of the task in the practical test was high. The implication of the result is that the precision with which NECO Physics practical test items measure the examinees proficiency in Physics practical was just fair.

Research Question 2: What proportion of the examinees' ability was precisely measured by NECO SSSCE Physics practical test?

To answer this research question, the standard error with which the ability score of the examinees in NECO Physics practical test were estimated was obtained and the distribution of the SE is presented in Table 2

Table 2: Distribution of Examinees' Ability Standard Error of Measurement

SE	Frequency	Percent
0.2	343	54.4
0.3	236	37.4
0.4	52	8.2
Total	631	100

Table 2 showed the distribution of the SE with which the ability of the examinees who took the NECO test. The table showed that the standard error with which the ability score of all the examinees in NECO Physics practical test were estimated were less than the 0.5 benchmark. The result showed that the precision with which the ability scores of the examinees were estimated was high.

ISSN 2756-4444
E-ISSN 2756-4452

Research Question 3: What proportion of the examinees fit the model adopted in the calibration of the NECO SSSCE Physics practical test?

To answer this research question, the person-fit of each of the examinees that took the test was extracted from the calibrated NECO SSSCE Physics practical test. In the context of IRT, person fit refers to the alignment between an examinee's response pattern and the IRT model selected for modelling the response data.

Table 3: Distribution of Examinees' fit Estimate to GPCM

Zh	Frequency	Percent	Cumulative Percent	zh	Frequency	Percent	Cumulative Percent
-5	1	0.2	0.2	-0.5	13	2.1	31.2
-4.9	1	0.2	0.3	-0.4	15	2.4	33.6
-4.2	1	0.2	0.5	-0.3	14	2.2	35.8
-3.9	2	0.3	0.8	-0.2	21	3.3	39.1
-3.8	1	0.2	1	-0.1	20	3.2	42.3
-3.7	1	0.2	1.1	0	12	1.9	44.2
-3.5	2	0.3	1.4	0.1	20	3.2	47.4
-3.4	2	0.3	1.7	0.2	10	1.6	49
-3.3	1	0.2	1.9	0.3	17	2.7	51.7
-3.1	1	0.2	2.1	0.4	15	2.4	54
-3	2	0.3	2.4	0.5	19	3	57.1
-2.8	1	0.2	2.5	0.6	29	4.6	61.6
-2.7	5	0.8	3.3	0.7	20	3.2	64.8
-2.6	2	0.3	3.6	0.8	13	2.1	66.9
-2.5	4	0.6	4.3	0.9	22	3.5	70.4
-2.4	6	1	5.2	1	19	3	73.4
-2.3	1	0.2	5.4	1.1	20	3.2	76.5
-2.2	3	0.5	5.9	1.2	10	1.6	78.1
-2.1	7	1.1	7	1.3	12	1.9	80
-2	6	1	7.9	1.4	23	3.6	83.7
-1.9	4	0.6	8.6	1.5	18	2.9	86.5
-1.8	10	1.6	10.1	1.6	13	2.1	88.6
-1.7	8	1.3	11.4	1.7	11	1.7	90.3
-1.6	9	1.4	12.8	1.8	10	1.6	91.9
-1.5	12	1.9	14.7	1.9	16	2.5	94.5
-1.4	7	1.1	15.8	2	7	1.1	95.6
-1.3	9	1.4	17.3	2.1	5	0.8	96.4
-1.2	9	1.4	18.7	2.2	3	0.5	96.8
-1.1	5	0.8	19.5	2.3	6	1	97.8
-1	12	1.9	21.4	2.4	11	1.7	99.5
-0.9	15	2.4	23.8	2.5	1	0.2	99.7
-0.8	8	1.3	25	2.6	1	0.2	99.8
-0.7	10	1.6	26.6	2.8	1	0.2	100
-0.6	16	2.5	29.2	Total	631	100	

Table 3 showed the distribution of the fitness of the person to the Generalized Partial Credit Model (GPCM) IRT model used in the calibration of the NECO Physics practical test. The table showed that

ISSN 2756-4444
E-ISSN 2756-4452

7.0% of the examinees returned zh values that were less than the -2 large negative value benchmark and 11.4% of the examinees returned zh values that were greater than the 2 large positive value benchmarks. Furthermore, the table shows that 88.6% of the students returned zh values that was within the the -2 and 2 benchmark. The result showed that most of the examinees fitted the IRT model, GPCM used in the calibration of the data. The implication of the result is that the validity of the selected model at the examinee level and the meaningfulness of the estimated proficiency in Physics practical were high.

Research Question 4: To what extent does the NECO SSSCE Physics practical test measure the same theoretical construct between**:** (a) male and female students, (b) Students from schools located in rural and urban areas and (c) Students from private and public schools?

To answer this research question, the response of the students to the NECO test was subjected to test calibration along male and female student's samples using multiple-group IRT modelling. To judge the functionality of the test among identifiable subgroup samples, the reduced M2, limited-information fit measure (Cai & Hansen, 2013) was used. The reduced M2 returned a significant value when an instrument is invariant with respect to identifiable subgroup. However, the result provided M2 significant value is not always dependable. This is because it has been established that it usually affected by sample size. To ensure accuracy of the measure, the ratio of degree of freedom to M2 value and other measures such as: root mean square error of approximation (RMSEA), root mean square error of approximation (RMSEA), Tucker-Lewis Fit Index (TLI) and comparative fit index (CFI), standardized root mean square residual (SRMR) have been suggested to overcome the weakness of the M2 significance value (Kline, 2016). As a general rule, the ratio of degree of freedom to M2 value less than 3:1 (Kline, 2016), values of TLI and CFI over 0.90, and values of RMSEA lower than 0.08 (McDonald and Ho, 2002) and values of SRMR lower than 0.08 (Hu and Bentler, 1999) indicate an acceptable model-data fit. The result is presented as follows:

Table 4a: Consistency of the Functionality of NECO SSSCE Physics Practical Test between Male and Female Students

	M2	df	P	RMSEA	SRMSR.FEMALE	SRMSR.MALE	TLI	CFI
Stats	6087.334	1172	0.000	0.08	0.140	0.129	0.84	0.85

Table 4 presented the assessment of the consistency of the functionality of the NECO Physics practical among male and female students. The table showed that the reduced M2 was significant (M2 (1172) = 6087.334, $p < 0.05$). Also, the ratio of the degree of freedom (df) was greater than 3:1 (5.2:1). Furthermore, TLI, CFI and the SRMSR except RMSEA for the two sub-groups for the model were below the acceptable standard (TLI = 0.84, CFI = 0.85, SRMSR for female = 0.140 and SRMSR for male = 0.129, RMSEA = 0.08). The result showed the consistency of the functionality of the NECO test between male and female was low construct validity.

Table 4b: Consistency of the Functionality of NECO SSSCE Physics Practical Test between Students of Schools Located in Rural and Urban Areas

	M2	df	P	RMSEA	SRMSR.RURAL	SRMSR.URBAN	TLI	CFI
Stats	6267.967	1172	0.000	0.08	0.142	0.131	0.83	0.84

Table 4b presented the assessment of the consistency of the functionality of the NECO Physics practical among students of schools located in rural and urban setting. The table showed that the reduced M2 was significant (M2 (1172) = 6267.967, $p < 0.05$). Also, the ratio of the degree of freedom (df) was greater than 3:1 (5.4:1). Furthermore, TLI, CFI, SRMSR except RMSEA for the two sub-groups for the model did not meet the acceptable standard (TLI = 0.83, CFI = 0.84, SRMSR for rural = 0.142 and SRMSR for urban = 0.131, RMSEA = 0.08). The result showed the consistency of the functionality of the NECO test among students of schools located in rural and urban settings was quite low.

Table 4c: Consistency of the Functionality of NECO Physics Practical Test between Students from Private and Public Schools

	M2	df	p	RMSEA	SRMSR.PRIVATE	SRMSR.PUBLIC	TLI	CFI
Stats	6067.174	1172	0.000	0.08	0.130	0.136	0.84	0.85

Table 4c presented the assessment of the consistency of the functionality of the NECO Physics practical among students of private and public schools. The table showed that the reduced M2 was significant (M2 (1172) = 6067.174, $p < 0.05$). Also, the ratio of the degree of freedom (df) was greater than 3:1 (5.1:1).

ISSN 2756-4444
E-ISSN 2756-4452

Furthermore, TLI, CFI, SRMSR except RMSEA for the two sub-groups for the model did not meet the acceptable standard (TLI = 0.84, CFI = 0.85, SRMSR for rural = 0.130 and SRMSR for urban = 0.136, RMSEA = 0.08). The result showed the consistency of the functionality of the NECO SSSCE test among students of public and private schools was quite low. The results showed that the consistency of the functionality of the test among male and female students, private and public students and among students from rural and urban settings respectively was consistently low. The implication of the result is the extent to which the NECO Physics practical test measures the same theoretical construct among male and female, rural and urban schools' students, and students of private and public schools was consistently low. Therefore, the construct validity of the NECO test was low

Research Question 5 How reliable is the NECO SSSCE Physics practical test?

To answer this research question, the internal consistency of the ability scores obtained on the NECO Physics practical test was obtained. To achieve this under IRT framework, the responses of the examinees to the tasks in the test was subjected to empirical reliability. The analysis was conducted with mirt package. The result is presented in table 5.

Table 5: Empirical Reliability Estimate of NECO Physics Practical Test

Empirical Reliability
0.93

Table 5 showed the internal consistency of the ability scores obtained from the NECO SSSCE Physics practical test. The table showed that the reliability estimate of the SSSCE test was 0.93. The result showed that the reliability of the SSSCE test was high. The implication of the result is that the NECO SSSCE Physics practical test was reliable.

Discussion

The assessment of the precision with which NECO Physic practical measure the examinees' practical proficiency showed that the precision with which difficulty parameters of 19 items were estimated was low while it was high for 18 items. An indication that the precision with which NECO Physics practical test items measure the examinees proficiency in Physics practical was just fair. However, the standard error of the ability score of all the examinees in NECO SSSCE Physics practical test was estimated to be less than the benchmark. This showed that the precision with which the ability scores of the examinees were estimated was high.

The person-fit of each of the examinees that took the test was extracted from the calibrated NECO Physics practical test. In the context of IRT, person fit refers to the alignment between an examinee's response pattern and the IRT model selected for modelling the response data. Person-fit indices are used to assess the validity of the selected model at the examinee level and the meaningfulness of the estimated latent trait levels (Embretson & Reise, 2000). In the literature, there are various person-fit indices proposed for the IRT models. the Most prominent of the indices is the standardized fit index called Zh, originally proposed by Drasgow, Levine, and Williams (1985). The indices, Zh is a standardized statistic that follows a standard normal distribution, thus, the expected value for Zh is zero when the examinee's response pattern is aligned with the selected IRT model. Large negative Zh values (e.g., $Zh < -2$) indicate person misfit. Large positive Zh values indicate that the likelihood for the examinee's response pattern is higher than the predicted likelihood based on the selected IRT model (Embretson & Reise, 2000), which might also be problematic for latent trait estimation.

The response of the students to the NECO SSSCE Physics test items was subjected to test calibration along male and female student's samples using multiple-group IRT modelling. To judge the functionality of the test among identifiable subgroup samples, the reduced M2, limited-information fit measure (Cai & Hansen, 2013) was used. The reduced M2 returned a significant value when an instrument is invariant with respect to identifiable subgroup. However, the result provided M2 significant value are not always dependable. This is because it has been established that it usually affected by sample size. To ensure accuracy of the measure, the ratio of degree of freedom to M2 value and other measures such as: root mean square error of approximation (RMSEA), root mean square error of approximation (RMSEA), Tucker-Lewis Fit Index (TLI) and comparative fit index (CFI), standardized root mean square residual (SRMR) have been suggested to overcome the weakness of the M2 significance value (Kline, 2016). As a general rule, the ratio of degree of freedom to M2 value less than 3:1 (Kline, 2016), values of TLI and CFI over

0.90, and values of RMSEA lower than 0.08 (McDonald and Ho, 2002) and values of SRMR lower than 0.08 (Hu and Bentler, 1999) indicate an acceptable model-data fit.

The internal consistency of the ability scores obtained on the NECO Physics practical test was obtained. To achieve this under IRT framework, the responses of the examinees to the tasks in the test was subjected to empirical reliability. The analysis was conducted with mirt package. The result shows the internal consistency of the ability scores obtained from the NECO SSSCE Physics practical test. The result showed that the reliability of the SSSCE test items was high at 0.93. The implication of the result is that the NECO Physics practical test items was reliable.

Conclusion

The study concluded that NECO physics practical examination measures examinees' practical skills proficiency with low precision and low construct validity across sub-groups.

Recommendations

It is therefore, recommended that

1. NECO and other examination bodies that conduct such practical test should during the construction of examination items be conscious of the construct validity and precision with which the items will measure the target trait(s) in the examinees no matter the group they belong to.
2. Items functionality between groups should be ascertained by examination bodies to ensure equity and test fairness to all examines.
3. Examination bodies at all times should ensure the validity of the selected model atexaminee level as well as the meaningfulness of the estimated proficiency.

References

Abdullahi, A. (1982). *Science Teaching in Nigeria*. Ilorin: Atoto Press.

Azar, A., and Şengülec, Ö. A. (2011). Computer-Assisted and Laboratory Assisted teaching methods in physics teaching: The effect on student achievement and attitude towards physics. *Eurasian Journal of Physics and Chemistry Education* (Special Issue) 43 –5

Cai, L., & Hansen, M. (2013). Limited-information goodness-of-fit testing of hierarchical item factor models. *British Journal of Mathematical and Statistical Psychology*, 66, 245-276.

Deacon, C., and Hajek, A. (2011). Student perceptions of the value of physics laboratories, *International Journal of Science Education*, 33 (7), 943–977.

DeMars, C. (2010) *Item Response Theory: Understanding Statistics Measurement*. Oxford University Press, Oxford. https://doi.org/10.1093/acprof:oso/9780195377033.001.0001

Drasgow, F., Levine, M. V., & Williams, E. A. (1985). Appropriateness measurement with polychotomous item response models and standardized indices. *British Journal of Mathematical and Statistical Psychology*, 38, 67- 86.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates.

Federal Republic of Nigeria (2013). *National Policy on Education*. 4th – edition. Lagos: NERDC Press.

Hu, L.T. and Bentler, P.M. (1999), "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives," *Structural Equation Modeling*, 6 (1), 1-55.

Kline R. B. (2016*). Principles and Practice of Structural Equation Modeling*, 4th Edn. New York, NY: The Guilford Press.

McDonald, R.P. and Ho, M.-H.R. (2002), "Principles and Practice in Reporting Statistical Equation Analyses," *Psychological Methods*, 7 (1), 64-82

NECO. (2015). *Chief examiner repor*t. Minna: NECO Press.

NECO. (2016). *Chief examiner report.* Minna: NECO Press.

NECO. (2017). *Chief examiner report*. Minna: NECO Press.

Opong, I. K. (1981). Science education in Nigeria, the product of science. *Science Teachers Association of Nigeria*, 19(2).

Prades, A., & Espinar, S. R. (2010). Laboratory assessment in chemistry: an analysis of the adequacy of the assessment process. *Assessment & Evaluation in Higher Education*, 35(4), 449-461.

Solomon, J. (1980). *Teaching children in the laboratory*, unpublished Dissertation, University of Port Harcourt Nigeria.

Tamir, P. (1977). How Are the Laboratories Used? *Journal of Research in Science Teaching*, *14*(4), 311-6.