



Estimating Standard Error of Measurement in BECE 2022 Basic Science Multiple Choice Test Items “in Etsako West Local Government Area of Edo State” Using Bootstrap Methods

PRAISE KEHINDE ADEOSUN
University of Benin, Benin City, Nigeria.

IYABO AYODEJI UKANAH
Adekunle Ajasin University, Akungba-Akoko, Ondo State, Nigeria

Abstract. Evaluation is an essential concept in educational setting which is usually carried out through measurement process to ascertain each testakers abilities in a population during testing and this can be done by determining the standard error of measurement. This study used Bootstrap methodologies to assess the standard error of measurement for the 2021 Basic Education Certificate Examination. This study was conducted using a descriptive research approach. The population consisted of 3,504 Upper Basic Secondary School three (JSS3) pupils from 14 public and 20 private institutions. A multistage selection approach was used to pick 350 students from five (5) and ten (10) public and private Upper Basic Secondary schools using stratified, simple random, and proportional sample sizes. The data collecting instrument consisted of sixty (60) multiple choice items from the 2022 BECE, Basic Science, and Technology multiple choice test. The instrument's dependability was tested using the Kuder-Richardson 20 (KR 20) to determine internal consistency. It returned a result of 0.91. The acquired data was examined using bootstrap statistics and the bootstrap t-test for independent samples. The results of this research indicated that there was moderate diversity in the test scores of pupils in public as well as private educational institutions for the 2022 BECE of Basic Science and Technology multiple choice exam questions. It also revealed that the scores of private Junior Secondary School three (3) students had more variability than those from the public Junior Secondary School. It was recommended that test should be administered fairly and consistently in both private and public schools which will help to create equitable testing environment for all students,

regardless of whether they attend private or public schools.

1. Introduction

The development of a nation is the function of how science and technology is handled because science and technology has become the live wire fostering development in the world including Nigeria. Technological and scientific advances have a significant part in the formation and growth of a nation. Nations that are deemed developed and civilized have reached this position via intentional scientific teaching of their inhabitants (Ajayi, 2017). Science and technology is an indispensable tool that any nation that wishes to progress in the socio-economic sphere cannot afford not to include the learning of science in the school curriculum. As a result, in order for Nigeria to attain scientific and technological growth, it is critical that the country begin preparing for a solid fundamental scientific education foundation for its citizens as early as possible, as children's exploration commences at a young age. According to Enemarie (2016), Nigeria is poised to become one of the world's most scientific and technologically sophisticated nations as a consequence of significant scientific and technical contributions to human growth. As a result, fundamental concepts in technology and science are taught in primary schools to capture students' attention early on. As additional instruction, basic science is also taught in upper basic secondary schools in Nigeria to assist students in developing a tangible as well as solid foundation for scientific and technological advancement.

To determine the level of achievement and extent to which basic skills have been acquired by the pupils, test must be conducted to provide useful information about their performance and this can be accomplished through educational testing. One of such tests conducted to know the level of achievement of the students “is the Basic Education Certificate Examination.” The Edo State Ministry of Education conducts the Basic Education Certificate Examination on an annual basis for students who would be transitioning from the junior secondary school category to the senior secondary school category after three years. So, anytime a test, such as the BECE, is conducted, the test user wants to be certain that the results of the students can be repeated if the same group of students have been evaluated again under comparable conditions. The expected constancy of test scores is referred to as reliability.

To some degree, all psychological measures are erroneous. Each tester and developer is concerned about the level of unreliability in a collection of observations. So, anytime a student replies to a set of test items, his or her score reflects a small sample of behaviour. As a result, scores recorded under the above circumstances are prone to measurement error. Miller and Liun, as reported in Esomunu and Okeaba, 2021, defined error of measurement as a condition in which a testee's result might be understated or overstated. Errors are significant sources of construct-irrelevant variation in assessment findings (Adeosun & Ekwere, 2022). There are two types of measurement errors: systematic and random. Systematic shortcomings are those that constantly affect a person's performance as a result of certain particular features about that person or when the test is unconnected to the construct being measured, whereas random errors affect an individual's score as a result of purely unforeseen circumstances such as guessing, become distracted during test oversight, etc. Systematic errors produce erroneous test results, limiting their practical use (Crocker & Algina 2008), whilst random errors diminish test score consistency and usefulness.

Due to the inconsistencies in measurement, there is need to for estimating errors. An error according to the CTT model is the discrepancy or difference between a test takers observed scores and his /her true score. According to Ado, referenced in Imasuen and Ogbomwan (2014), CTT is a basic model that illustrates how measurement errors affect observed scores. Meanwhile, it is impossible to calculate the precise amount of inaccuracy in a specific score, but the CTT technique may be used to characterise how individual test takers' seen results differ from their real

scores. The actual score is the average score of a test taker's observed score based on an increased number of repeated tests. These repeated tests provide a distribution of several test takers' likely observed scores around their real scores within a broader range of scores for the entire group on a single exam. As a result, there is a standard deviation for the entire group. In theory, each test taker's own distribution of potential observed scores around their real score has a standard deviation. Consequently, when individual error standard deviations are averaged for the group, the result is known as standard error of measurement (SE). As a result, standard error of measurement may be defined as the standard deviation of mistakes in a measurement related with the test scores of a specific group of test takers. (AERA, APA, & NCME, 1985) cited in (Adewuni & Busari, 2021).

The standard error of measurement is the most straightforward and suitable measure of variability. It evaluates how repeated measures of a person on the same instrument likely to spread about his or her real score. The SEM is based on the inaccuracy in predicting a person's observed score given their real score on a randomly selected parallel test (Price, 2017). Furthermore, SEM is directly connected to test reliability, which implies that the greater the standard error, “the lower the test's reliability, and the smaller the standard error of measurement, the higher the test's reliability.” SEM is used to calculate a confidence limit and interval for observed scores. Several non-parametric approaches have been developed to estimate the extent of the standard error of measurement. Examples include the Feldt technique, Bionomial, and Bootstrap approaches.

The Bootstrap technique is a non-parametric resampling approach that does not rely on the normalcy assumption. According to Efron, as described in Imasuen and Ogbomwan (2014), bootstrapping is a computationally costly approach for estimating standard errors of a wide range of statistics. It allows for estimation of standard errors utilising sample size and simulation. The bootstrap method is a statistical approach for estimating the sampling distribution of a statistic. The goal of employing bootstrap methods is to do many forms of statistical inference using contemporary computing capacity. Bootstrapping is a resampling process that leverages data from a single sample to produce a sampling distribution by picking random samples from the known sample along with replacing them. In the bootstrap technique, replacement and repetition are the most appropriate ways to take a sample from the population. The bootstrap technique uses repeated sampling with replacement from a data set to estimate

the standard error and other features of a sampling distribution.

private Upper Basic Secondary School in Etsako Local Government Area of Edo State?”

1.1 Statement of the Problem

In conducting an evaluation, measurement is an important component that must be taken. The success of learning can be seen from the size of the measurement of the standard error that occurs in testing process. The standard error of measurement is useful for estimating how widely the real score may differ from the observed score of an average test taker in the community. There have been several methods created for calculating standard error of measurement, which are divided into three categories. We now have those that use classical test theory, item response theory, and those that make no assumptions. Those method based on assumptions usually provide a point estimate which may not reflect the actual value of the test taker. This now pave way for other methods to be considered, which one of them is Bootstrap methods. It is also observed that the use of bootstrap methods in estimating standard error of measurement has not given adequate attention in Nigeria and Edo State in particular. Therefore, this study estimated standard error of measurement of the 2022 Basic Education Certificate using Bootstrap methods in Etsako West Local Government Area of Edo State.

1.2 Research Questions

- What is the standard error of measurement of the 2022 BECE in public Upper Basic Secondary School in Etsako “Local Government Area of Edo State?”
- What is the standard error of measurement of the 2022 BECE in private Upper Basic Secondary School in Etsako Local Government Area of Edo State?”
- What is the difference in the standard error of measurement of the 2022 BECE in public and

1.3 Hypothesis

Research question three was hypothesized “and tested at 0.05 level of significance:

- There is no significance difference in the standard error of measurement of the 2022 BECE between public and private Upper Basic Secondary Schools in Etsako Local Government Area of Edo State.”

1.4 Research Methodology

This study used a descriptive research methodology and included 3,504 pupils from 14 public and 20 private junior secondary schools in Etsako West, Edo State. The study employed a multistage sampling approach that included stratified, simple random, and proportional procedures. The pupils were classified depending on school ownership, namely public and private upper basic secondary schools. Three students used simple random selection to choose five public schools from 14 and ten private schools from 20 Upper Basic Secondary schools. Also, proportionate sampling techniques was used to select 250 students from 5 public schools and 100 students from the 10 private schools. The research instrument used was 2022 Basic science and Technology multiple test items of Edo State Basic Certificate Examination conducted by Edo State Ministry of Education. The Kuder Richardson formular 20 was utilized to establish the reliability coefficient and it gave an index of 0.82. To address the study questions, the data was analysed using Bootstrap statistics, as well as the hypothesis was tested using a t-test with independent statistics.

2. Results

Research Question 1: What is the standard error of measurement of the 2022 BECE in public Upper Basic Secondary School in Etsako West Local Government Area, Edo State?”

Table 1: Bootstrap Standard Error of Measurement of the 2022 BECE in Public Upper “Basic Secondary School in Etsako West Local Government Area, Edo State”

Statistic	Bootstrap				
	Bias	Std. Error	BCa 95% Confidence Interval		
			Lower	Upper	
Mean	41.94	-.01	.80	40.49	43.60
N	250	0	0	.	.
Std. Deviation	12.371	-.007	.452	11.507	13.275

Table 1 shows mean value of 41.94, standard deviation of 12.37, and standard error of measurement as 0.45, with a bootstrap accentuated 95% confidence interval BC_a (11.51 – 13.28). This is an indication that the variability of the scores of the 2022 BECE in public Upper Basic Secondary School in Etsako West “Local Government Area, Edo State was moderate.

Research Question 2: What is the standard error of measurement of the 2022 BECE in private Upper Basic Secondary School (3) in Etsako West Local Government Area, Edo State?”

Table 2: Bootstrap Standard Error of Measurement of the 2022 BECE in Private Upper Basic “Secondary School in Etsako West Local Government Area, Edo State”

	Statistic	Bootstrap			
		Bias	Std. Error	BCa 95% Confidence Interval	
				Lower	Upper
Mean	37.69	-.08	.94	36.03	39.23
N	100	0	0	.	.
Std. Deviation	9.553	-.066	.406	8.864	10.125

Table 2 shows mean value of 37.69, standard deviation of 9.55, and standard error of measurement as 0.41, with a bootstrap accentuated 95% confidence interval BC_a (8.86 – 10.13). This is an indication that the variability of the scores of the 2022 BECE in private Upper Basic “Secondary School in Etsako West Local Government Area, Edo State was moderate.

Hypothesis One: There is no significant difference in the standard error of measurement of the 2022 BECE in public and private Upper Basic Secondary School (3) students in Etsako West Local Government Area, Edo State.”

Table 3: Bootstrap for Independent Samples Test of the “Difference in the Standard Error of Measurement of the 2022 BECE in Public and Private Upper Basic Secondary School in Etsako West Local Government Area, Edo State”

Statistic	Schools	N	Mean	Standard deviation	t-value	Mean Difference	Bootstrap				
							Bias	Std. Error	Sig. (2-tailed)	BCa 95% Confidence Interval	
										Lower	Upper
	Public	250	41.94	12.37	-3.083	-4.246	.053	1.268	.002	-6.810	-1.650
	Private	100	37.69	9.55							

Table 3 indicates a computed value of -3.083 with a p-value of 0.002. At the 0.05 level of significance, the p-value is smaller than the alpha level. Therefore, the null hypothesis is that “there is no significant difference in the standard error of measurement of the 2022 BECE in public and private Upper Basic Secondary Schools in Etsako West Local Government Area, Edo State” is rejected. The standard error of measurement for the 2022 BECE differs significantly across both government and private Upper Basic Secondary Schools in Etsako West Local Government Area, Edo State. This is an indication that the scores of public Upper Basic Secondary schools’ three (3) students had more variability than those from the private schools.

3. Discussion of Finding

The results obtained from the analysis of research question 1 show that there is a high level of uniformity between the noticed and actual scores, implying that the variability of 2022 BECE scores in public Upper

Basic Secondary Schools in Etsako West Local Government Area of Edo State is moderate. When scores have a moderate level of variability, it provides a more accurate picture of the underlying distribution of scores which means that the individual scores are spread out but not too much (not wildly different from each other) within the population. It allows meaningful comparisons between individuals and groups. Also the moderate level of variability means that the test is not too easy or too difficult and able to differentiate between students various ability level. As a result, the exam is dependable and gives valuable information about the students' performance. This is consistent with the findings of Adewuni and Busari (2021), who discovered that students' performance in WASSCE May/June economics multiple choice objective tests 2019 had a SEM of 4 (+ or -4), indicating that the test was more accurate, reliable, as well as correct.

The findings of research question 2 means that the variability of the scores of the 2022 BECE in private

Upper Basic Secondary Schools in Etsako West Local Government Area, Edo State was also moderate, it implies that they provide a reliable and accurate measure of student performance without being too easy or too difficult. Additionally, it indicates that the test is fair and that all students have a chance to demonstrate their abilities. This is consistent with the research findings of Adewuni and Busari (2021), who discovered that students' performance on the WASSCE May/June economics multiple choice objective exams 2019 had a SEM of 4 (+ or -4), indicating that the test is more exact, dependable, and right.

The findings of hypothesis 1 showed that the private Upper Basic Secondary school score had more variability than public school scores. The larger variability in private schools may be due to the fact that there is more diversity in the types of schools that fall under the private category. For instance, some private schools may be highly selective and have a more homogenous student body, while others may be less selective and have a more diverse student body. This diversity in the types of private schools may contribute to the observed difference in variability. Also, private schools have more flexibility in their curricular, teaching methods, and policies, this greater flexibility may lead to more variation in how the schools operate, which in turn could lead to more variability in student outcomes while public schools tend to have a more standardized curriculum, so there may be less variability in the content covered and in the way that it is taught.

In addition, public schools often have more standardized testing procedures which may lead to less variability in the scores. Finally, private schools may have more resources, such as smaller class sizes which could contribute to the observed variability in student outcomes while public schools often have a larger number of students taking the same test, which could also lead to less variability in the scores. This difference is statistically significant, this implies that public Upper Basic Secondary schools score had less variability than private Upper Basic Secondary schools scores. The implication of this finding is that the SEM for BECE is not equal for public and private school students and this should be taken into consideration when interpreting the scores. This is in line with the study of Satya (2021) The results of the research demonstrated that private schools provide higher-quality education than public schools. It also agreed with Olanigan et al. (2022), who discovered that applicants students at private educational institutions surpassed those in public educational institutions in regards to academic success. This

nevertheless differed with Agah (2013), who reported in the research he conducted that the mean ability assessments of students in both public and private educational institutions did not change substantially, nor did the mean ability estimates alongside additional moments (standard deviation and variance).

4. Conclusion

According to the findings of the present research, the examination demonstrated moderate variability in BECE scores in both private and public institutions. This indicates that the exam given to students in both private and public schools is dependable since it provides a pretty accurate picture of the students' performance with a decent degree of variation between individual scores. It also suggests that the test is neither too simple nor too challenging. The data also revealed a statistically significant difference between kids' scores in private and public schools, although a modest one. However, the greater variability in private schools indicates that there is more variation in the students' scores which may be due to a number of factors, such as the quality of instruction, the resources available to students, or the characteristics of the individual students themselves. Therefore, it is important to consider this variability when interpreting the results of the BECE.

5. Recommendations

Based on the findings of this study:

- It was recommended that other measures of knowledge or skills such as teacher observation or classroom performance should be considered in addition to the test scores.
- It was recommended that there should be provision for additional support for private schools such as professional developments for teachers, increased access to resources and improved consistency in test administration.
- Further research should be carried out into specific causes of the variability in scores between private and public schools to help to identify and address any underlying issues which will provide more accurate and comprehensive assessment of student's performance as well allow for more effective interventions and improved educational outcomes for all students.

References

- Adewuni, D. A. & Busari, Y. T. (2021). Analysis of Standard Error of Measurement (SEM) of 2019 West African Senior School Certificate Examination Multiple-Choice Objective Tests in Economics. *Kashere Journal of Education*, 2(2): 25-33.
- Adeosun, P.K. & Ekwere N.S. (2022). Application of Many- Fact Rasch Measurement in complex Problem Solving Skills in 2019 NECO Physics Essay Items. *Asian Journal of Management Sciences & Education (AJMSE)*, 11(1), 13-21.
- Agah, J.J, (2013). Relative Efficiency of Test Scores Equating Methods in Comparison of Students Continuous Assessment Measures. Ph.D Thesis submitted to the Department of Science Education, University of Nigeria, Nsukka. <https://www.unn.edu.ng/publication/s/files/AGAH%20JOHN%20JOSEPH.pdf>
- Ajayi, V. O. (2017): Effect of Hands-on- Activity-based method on interest of Senior Secondary students in Organic Chemistry. *Scholarly Journal of Education*, 6(1), 1-5.
- Enemarie Veronica (2016) Achievement in basic science as a predictor of performance in senior secondary sciences in Benue State Education Zone B. Unpublished M.Ed. Dissertation, Benue State University, Markudi.
- Esomonu, N. P., & Okeaba, J. U. (2021) Estimating Measurement Error and Score Dependability of the Inventory for Students' Intergration into University Academic Culture (ISIUAC) Using Generalizability Theory. *Rivers State University Journal of Education (RSUJOE)*, 24(1):35-46.
- Imasuen, K., & Ogbomwan, S. M. (2014). Bootstrap estimates for non- linear regression analysis. *Nigerian Annuals Journal of Natural Sciences*, 5(1):101-118.
- Olanigan, N.S.,Alaba, A., & Olawale A.(2022)Linear and Separate Calibration Methods of Equating Continuous Assessment Scores of Public and Private Elementary Schools. *Journal of Integrated Elementary Education*,2(2),117129Website:<https://journal.walisongo.ac.id/index.php/jieed>.
- Satya, K. & Nikita, G. (2021). A Comparative Study in Public and Private Sector. *International Journal of Innovative Science and Research Technology*. 6(7). ISSN No:2456-2165.