Kashere Journal of Education 2021, 2(2): 25-33. ISSN: 2756-6021 (print) 2756-6013 (online)
DOI: https://dx.doi.org/10.4314/kje.v2i2.4 Creative Commons Attribution License (CC BY 4.0)
Adewuni, D. Adebayo and Busari Y. Taiwo ©2021 Federal University of Kashere

Analysis of Standard Error of Measurement (SEM) of 2019 West African Senior School Certificate Examination Multiple-Choice Objective Tests in Economics

[1]Adewuni, D. Adebayo and [2]Busari, Y. Taiwo
[1]Department of Educational Foundations, Faculty of Education, Federal University of Kashere, Gombe State ekerinonibiyo@gmail.com +2347064641192
[2]Department of Business Education, School of Vocational and Technical Education, Crestfield College of Education, Erin-Osun, Osun State, Nigeria taiwobusariyusuf@yahoo.com

Abstract
This study analyzed the Standard Error of Measurement (SEM) of 2019 May/June West African Senior School Certificate Examination Multiple-choice Objective Tests in Economics at three different confidence Interval (CI). A quantitative research design of the descriptive type was adopted for the study. The sample of the study was Three hundred and Two (302) Senior Secondary School Three (SSS.3) students that offered Economics selected from twelve (12) schools in the three senatorial districts in Osun State, Nigeria and Multi-stage sampling technique was adopted in this study. 2019 May/June and Nov/Dec (GCE/WASSCE) Multiple-Choice Objective Tests in Economics were adopted as instruments for the study. Data collected were analyzed using Descriptive statistics. The findings of this study revealed that the performance of students in WASSCE May/June Economics Multiple choice objective test of 2019 flagged the SEM of 4 (+ or - 4). Also, the performance of students in 2019 GCE WASSCE Economics Multiple-choice objective test flagged the SEM of 12 (+ or - 12) both at 68% confidence interval. It was concluded that 2019 May/June WASSCE Economics Multiple-choice objective test is more precise, reliable and correct than 2019 GCE WASSCE Economics Multiple-choice objective test at all the confidence intervals. The study recommended that educators should consider the magnitude of SEMs for students across the achievement distribution. It was also recommended that test practitioners should adopted classical Test Theory (CTT) in test scoring and test precision.
Keywords: Standard Error of Measurement (SEM), Confidence Interval, Economics

Citation: Adewuni, D. A. and 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.
Submitted: 17/6/2021 **Accepted**: 20/9/2021 **Published**: 1/12/2021

Introduction

The reliability of an examination provides useful information about its performance (Tighe, McManus, Dewhurst, Chis, & Mucklow, 2010). The mere fact that an examination has a high reliability does not ensure that it is necessarily functioning effectively, because the reliability is heavily dependent upon the ability, range of the candidates who are taking it (Tighe et al. 2010). When examinations have very small numbers of candidates, there is a greater risk that the reliability will be distorted by an unusually high or low spread of candidate abilities (Tighe et al. 2010)

To measure the improvement of testees, it is vital that the assessment used is designed bearing this in mind. And to do this, such assessment must be with precision to get a clearer view of these testees whether they are on, above or below certain level. To track testees' progress over time, it's critical to use an assessment that provides one with accurate estimates of their achievement-assessments with a high level of precision. When measures of precision are referred to, Standard Error of Measurement (SEM) is one of the concepts referenced to. According to Jensen (2015) the range in which an examinee true score likely falls can be estimated; in general, the smaller the range, the greater the precision of the assessment. SEM, put in simple terms, is a measure of the precision of the assessment-the smaller the SEM, the more precise the measurement capacity of the instrument. Consequently, smaller standard errors translate to more sensitive measurements of students' progress (Jensen 2015). So, to this point it is learned that smaller SEMs are related to greater

Kashere Journal of Education 2021, 2(2): **25-33. ISSN: 2756-6021 (print) 2756-6013 (online)**
DOI: https://dx.doi.org/10.4314/kje.v2i2.4 **Creative Commons Attribution License (CC BY 4.0)**

precision in the estimation of student achievement, and, conversely, that the larger the SEM, the less sensitive is their ability to detect changes in student achievement. Thus, the SEM is more formally defined as: "the standard deviation of errors of measurement that is associated with the test scores for a specified group of test-takers" (AERA, APA, & NCME, 1985). The simplest and often most appropriate measure of variability is the standard error of measurement (SEM). It is the standard deviation of several measurements made on the same person, indeed Bland and Altman (1996) prefer the term within-subject standard deviation. Most literature on assessment suggest it is calculated as a derivative of the intra-class correlation coefficient (ICC) and as a consequence many researchers do not appreciate just how simple a measure the SEM is and Bland and Altman (1996) equally opined that the simplest and often most appropriate measure of variability is the Standard Error of Measurement (SEM).

The observed score and its associated SEM can be used to construct a "confidence interval" to any desired degree of certainty. For example, a range of ± 1 SEM around the observed score is the range within which there is a 68% chance that a student's true score lies, with the observed representing the most likely estimate of this testee's score. Intuitively, if a larger range is specified around the observed score—for example, ± 2 SEM, examiners or researchers would be much more confident that the range encompassed the examinee's true score, as this range corresponds to a 95% confidence interval. So, to this point it could be learned that smaller SEMs are related to greater precision in the estimation of student achievement, and, conversely, that the larger the SEM, the less sensitive is our ability to detect changes in student achievement (Bland & Altman,1996).

According to AERA, APA, and NCME (1985) The SEM is usually accompanied by confidence interval, or a range around the estimated "true" score. The measurement unit is the same as the original test scores. For example, if measuring is done in points, the SEM will be in points and if measuring is done in percentages, the SEM will be in percentages. Common SEM confidence intervals and their formulas according to AERA, APA, and NCME (1985) are 68%CI = Score ±SEM, 95%CI = Score ±(1.96*SEM) and 99%CI = Score ±(2.58*SEM).

Educators should consider the magnitude of SEMs for students across the achievement distribution they are using to make educational decisions is highly accurate for all students, regardless of testees achievement. According to Jenson, (2015(. SEM isn't the only the factor that impacts the accuracy of a test. Accuracy is also impacted by the quality of testing conditions and the energy and motivation the testees put to the test. In fact, an unexpected low test score is more likely to be caused by poor conditions or low motivation than a problem of testing instrument. To ensure an accurate estimate of student achievement, it is important to use a sound assessment, administer assessments under conditions conducisive to high test performance, and have students ready and motivated to perform.

A Rasch Model is appropriate when a researcher wishes to use the total score on a test or questionnaire to summarize each person's response. Responses are added across items to give each person a total score. This total score summarizes the responses to all the items. Summing the scores of the items to give a single score, for a testee implies that the items are intended to measure single variable, often referred to as a unidimensional variable. Rasch Model is the only Item Response Theory (IRT) Model in which the total score across items characterizes an examinee. It is also the simplest of all models having the minimum of parameters for the person (just one) and just one parameter corresponding to each category of an item (Adedoyin, Nenty & Chilisa,2008; Hambleton, Swaminathan & Roger, 1991).

Classical Test Theory (CTT) and Item Response Theory (IRT) are widely perceived as representing two very different measurement frameworks. Fan (1998) however, stated that few studies have empirically examined the similarities and differences in the parameters estimated using the two frameworks. CTT which is the main focus of this study focuses on test-level information, item statistics (item difficulty and item discrimination) are also an important part of the CTT model. At the item

Kashere Journal of Education 2021, 2(2): 25-33. ISSN: 2756-6021 (print) 2756-6013 (online)
DOI: https://dx.doi.org/10.4314/kje.v2i2.4 Creative Commons Attribution License (CC BY 4.0)

level, the CTT model is relatively simple; CTT does not involve a complex theoretical model to relate an examinee's ability to succeed on a particular item. Instead, CTT collectively considers a pool of examinees and empirically examines their success rate on an item, (assuming it is dichotomously scored). This success rate of a particular pool of examinees on an item, well known as the P-value of the item, is used as the index for the item difficulty (actually, it is an inverse indicator of item difficulty, with higher a value indicating an easier item). The ability of an item to discriminate between a higher ability examinee and a lower ability examinee is known as item discrimination.

Any high-stakes examination is expected be as accurate, and hence as repeatable as possible. One of the usual measure of reliability in an assessment is Cronbach's coefficient alpha (Cronbach,1951), with alpha values ranging from 0 to 1, where 1 indicates perfect reliability and 0 indicates a test that is no better than marks awarded at random. There are many other statistical parameters, Standard Error of Measurement (SEM) is mainly seen as useful only in determining the accuracy of a pass mark or the observed score of the testee(s). However, the alpha coefficient depends both on SEM and on the ability range (Standard Deviation, SD) of candidates taking an examination.

The role of high stake examination cannot be overemphasized as it is instrumental in helping the country in capacity building by preparing examinees for advanced education to equip them with right type of knowledge needed in man power development. There are various examination bodies in Nigeria that conduct different types of examinations. These are the Joint Admissions and Matriculation Board (JAMB) which conducts the Unified Tertiary Matriculation Examination (UTME) for secondary school leavers seeking admission to any institutions of higher learning in Nigeria. The West African Examinations Council (WAEC) conducts the West African Senior School Certificate Examination (WASSCE), The National Examinations Council (NECO) is saddled with the responsibilities of conducting the Senior Secondary School Certificate (SSCE), Business Studies and Common Entrance Examinations, and other examinations such as Basic Education Certificate Examination (BECE), the Junior Secondary Certificate Examination (JSCE), the National Common Entrance Examination (NCEE), the Gifted Examination into Suleja Academy and National Business and Technical Examination Board (NABTEB) handed both ordinary and advance certificates examinations in Business and Technical Subjects. Others are the examinations conducted in collaboration with or on behalf of other examining bodies, such as City and Guilds of London, the Royal Society of Arts, University of London GCE examination for non-west Africans, Scholastic Aptitude Test and Graduate Record Examinations for Educational Testing Services, Princeton, United States of America (USA)

The place of Economics as a subject in the WASSCE in Nigeria is very essential as it is one of the subjects that must be taken by all commercial students who want to obtain admission for higher education to study accounting and most social sciences courses. The quality of WASSCE must be assessed and maintained. It is against this backdrop this study investigated the Standard Error of Measurement (SEM) of 2019 May/June WASSCE Economics Multiple choice objective test and 2019 GCE WASSCE Economics Multiple-choice objective test.

Purpose of the Study

The main purpose of this study was to examine the SEM of 2019 WASSCE Economics Multiple-choice objective tests. Specifically, the study intended

a. to compare the academic performance (True Score) of students in the 2019 May/June WASSCE in Economics and the 2019 Nov/December WASSCE in Economics with respect to Confidence intervals.
b. to determine the precision in 2019 May/June WASSCE Economics Multiple-choice objective test and the 2019 Nov/December WASSCE Economics Multiple-choice objective test.

Research Questions

The following questions were generated to guide the study:

a. What is the academic performance (True Score) of students in the 2019 May/June WASSCE Economics Multiple choice

Kashere Journal of Education 2021, 2(2): **25-33. ISSN: 2756-6021 (print) 2756-6013 (online)**
DOI: https://dx.doi.org/10.4314/kje.v2i2.4 Creative Commons Attribution License (CC BY 4.0)
Adewuni, D. Adebayo and Busari Y. Taiwo

objective test and 2019 Nov/Dec WASSCE Economics Multiple-choice objective concerning confidence intervals?

b. Is there any precision in the 2019 May/June WASSCE Economics Multiple-choice objective test and the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test?

Methodology

This study adopted a quantitative research design of the descriptive type because it allowed the researcher to answer questions about the correlation between measured variables, to explains, predicts, compare and control certain phenomena (Mutodi & Ngirande, 2014)

The population of this study consisted of all senior secondary school students (SSS) in Two Hundred and Forty (240) public senior secondary schools in Osun State (Ministry of Education, State of Osun, 2014) The sample of the study was Three hundred and Two (302) Senior Secondary School Three (SSS.3) students that offered Economics selected from twelve (12) schools in the three senatorial districts in Osun State, Nigeria and Multi-stage sampling technique was adopted in this study. 12 schools out of 240 public secondary schools were chosen based on convenience in transportation, finance and administration of instruments. Senior secondary schools in the 3 senatorial Districts in state (Osun West, Osun Central and Osun East). 4 public secondary schools were randomly selected from each senatorial district.

The 2019 June/July of the Senior School Certificate Examination (SSCE) and 2019 Nov/Devc Senior School Certificate Examination (SSCE) series was adopted and used as instruments in this study. The instrument consisted of the 2019 past questions of Senior School Certificate Examination (SSCE) conducted by the West African Examination Council and consisted of Fifty (50) items each. The researchers believed that both the validity and reliability of these tests have been determined by the said examination body before administration, hence the issue of validity and reliability estimation of these tests/ test items were not addressed.

The researchers visited all the selected schools and administered the instruments to the selected students on two occasions each with 40 minutes' duration for each administration. This was made possible through the help of research assistants (Subject teachers). The two tests were administered on different days. Data collected from the administration of the instruments were subjected to analysis with Descriptive Statistics with the use of SPSS software.

Results

Research Question 1: What is the academic performance (True Score) of students in the 2019 May/June WASSCE Economics Multiple-choice objective test and the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test concerning Confidence intervals?

Table 1 shows the academic Performance in the 2019 WASSCE May/June Economics Multiple-choice objective test and the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test concerning Confidence Intervals (CI).

Table 1: Academic performance of examinees concerning Confidence Intervals

	2019 May/June WASSCE				2019 Nov/Dec. WASSCE			
	68%	95%	99%	Total	68%	95%	99%	Total
High	188(62.3%)	113(37.4%)	92(30.4%)	393(43.4%)	106(35.1%)	51(16.9%)	21(7%)	178(19.6%)
Average	26(8.6%)	71(23.5%)	19(6.3%)	116(12.8%)	76(25.2%)	43(14.2%)	26(8.6%)	145(16%)
Low	88(29.1%)	188(39.1%)	191(63.2%)	397(43.8%)	120(39.7%)	208(68.9%)	255(84.4%)	583(64.3%)
TOTAL	302	302	302	906	302	302	302	906

The table1 shows the comparison between the two examinations (The WASSCE May/June and November) under the three most common Confidence Intervals (CI) with a pass mark of 50%. For 2019 May/June WASSCE, it shows that under 68% confidence interval 188 students constituting 62.3% of the sampled group performed well, 26 (8.6%) performed averagely and 88 (29.1%) performed poor in the 2019

Kashere Journal of Education 2021, 2(2): 25-33. ISSN: 2756-6021 (print) 2756-6013 (online)
DOI: https://dx.doi.org/10.4314/kje.v2i2.4 Creative Commons Attribution License (CC BY 4.0)

May/June WASSCE Economics Multiple-choice objective test. Under 95% confidence Interval it shows that 113 students constituting 37.4% of the sampled group performed well, 71 (23.5%) performed averagely and 188 (39.1%) performed poorly. In 99% confidence interval 92 students constituting (30.4%) of the respondents performed well. 19 (6.3%) performed averagely well while 191 (63.2%) performed below average.

For the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test with a pass mark of 50% as well, it shows that under 68% confidence interval 106 students constituting 35.1% of the sampled group performed well, 76 (25.2%) performed averagely and 120 (39.7%) performed poor in the 2019 November/December WASSCE Economics Multiple-choice objective test. Under 95% confidence Interval it shows that 51 students constituting 16.9% of the sampled group performed well, 43 (14.2%) performed averagely and 208 (68.9%) performed poorly. In 99% confidence interval 21 students constituting 7% of the respondents performed well. 26 (8.6%) performed averagely well while 255 (84.4%) performed below average.

Under 68% confidence interval (Normal confidence or chance) of SEMs of ±12 and ±4 for the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test and May/June WASSCE Economics Multiple choice objective test respectively, the 2019 May/June WASSCE Economics Multiple-choice objective test has a more high-performance rate of 188 (62.3%) than that of 2019 Nov/Dec WASSCE Economics Multiple-choice objective test with 106 (35.1%). The 2019 May/June WASSCE Economics Multiple-choice objective test has an average performance rate of 26(8.6%) less than that of the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test with 76(25.2%). Meanwhile a percentage rate of failure of the 2019 May/June WASSCE Economics Multiple-choice objective test is low with 88(29.1%) compared to that of the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test of 120(39.7%).

Under 95% confidence interval (Moderate confidence or chance) of SEMs of ±23.52 and ±7.84 for the 2019 Nov/Dec WASSCE Economics Multiple choice objective test and the May/June WASSCE Economics Multiple-choice objective test respectively, the 2019 May/June WASSCE Economics Multiple-choice objective test has a more high-performance rate of 113 (37.4%) than that of the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test with 51 (16.9%). The 2019 May/June WASSCE Economics Multiple-choice objective test has an average performance rate of 71 (23.5%) higher than that of the 2019 GCE WASSCE Economics Multiple-choice objective test with 43(14.2%). Meanwhile percentage rate of failure of the 2019 May/June WASSCE Economics Multiple-choice objective test is low with 118(39.1%) compared to that of the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test of 208(68.9%).

Under 99% confidence interval (High confidence or chance) of SEMs of ±12 and ±4 for the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test and May/June WASSCE Economics Multiple-choice objective test respectively, the 2019 May/June WASSCE Economics Multiple-choice objective test has more high performance rate of 92 (30.4%) than that of the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test with 21 (7%). The 2019 May/June WASSCE Economics Multiple-choice objective test has an average-performance rate of 19(6.3%) less than that of the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test with 26(8.6%). Meanwhile a percentage rate of failure of the 2019 May/June WASSCE Economics Multiple-choice objective test is low with 191(63.2%) compared to that of the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test of 255(84.4%).

Furthermore, it could be seen that under 68% confidence interval of ±4 for the 2019 May/June WASSCE Economics Multiple-choice objective test, respondents have 188 (62.3%) high-performance rate than that of 95% confidence interval with 113(37.4%) of SEM±7.84 of the same examination and more than 99% confidence interval of 92 (30.4%) with SEM ±10.32 of the same examination. More so, it could be seen that under 68% confidence interval of ±4 for the 2019 May/June WASSCE Economics

Multiple-choice objective test, respondents have 26 (8.6%) average performance rate less than that of 95% confidence interval with 71(23.5%) of SEM±7.84 of the same examination and more than 99% confidence interval of 19 (6.3%) with SEM ±10.32 of the same examination.

And finally for the 2019 May/June WASSCE Economics Multiple-choice objective test, it could be seen that under 68% confidence interval of ±4 for the 2019 May/June WASSCE Economics Multiple-choice objective test, respondents have 88 (29.1%) low-performance rate less than that of 95% confidence interval with 118(39.1%) of SEM±7.84 of the same examination and also less than 99% confidence interval of 191 (63.2%) with SEM ±10.32 of the same examination.

Meanwhile, it could be seen that under 68% confidence interval of ±12 for the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test, respondents have 106 (35.1%) high-performance rate than that of 95% confidence interval with 51(16.9%) of SEM±23.52 of the same examination and also more than 99% confidence interval of 21 (7%) with SEM ±30.96 of the same examination. Furthermore, it could be seen that under 68% confidence interval of ±12 for the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test, respondents have 76 (25.2%) average performance rate than that of 95% confidence interval with 43(14.2%) of SEM±23.52 of the same examination and also more than 99% confidence interval of 26 (8.7%) with SEM ±30.96 of the same examination.

Meanwhile, it could be seen that under 68% confidence interval of ±12 for the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test, respondents have 120 (39.7%) low-performance rate less than that of 95% confidence interval with 208(68.9%) of SEM±23.52 of the same examination and also less than 99% confidence interval of 255 (84.4%) with SEM ±30.96 of the same examination.

Aggregately, the 2019 May/June WASSCE Economics Multiple-choice objective test records 393 (43.4%) total high-performance rate under 68%, 95% and 99% confidence intervals than that of the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test with total high-performance rate of 178 (19.6%) under 68%, 95% and 99% confidence intervals. Therefore, respondents passed the 2019 May/June WASSCE Economics Multiple-choice objective test than the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test.

Aggregately, the 2019 May/June WASSCE Economics Multiple-choice objective test records 116 (12.8%) total average performance rate under 68%, 95% and 99% confidence intervals slightly less than that of the 2019 Nov/Dec WASSCE economics Multiple-choice objective test with total average performance rate of 145 (16%) under 68%, 95% and 99% confidence intervals. And finally, the 2019 May/June WASSCE Economics Multiple-choice objective test records 397 (43.8%) total poor performance rate under 68%, 95% and 99% confidence intervals greatly less than that of the 2019 Nov/Dec WASSCE economics Multiple-choice objective test with a total poor performance rate of 583 (64.3%) under 68%, 95% and 99% confidence intervals. Therefore, respondents failed less in the 2019 May/June WASSCE economics Multiple-choice objective test than the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test.

Research Question 2: Is there any precision in 2019 May/June WASSCE Economics Multiple-choice objective test and the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test?

Table 2 show the SEM Summary for 2019 May/June WASSCE Economics Multiple-choice objective test and the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test concerning Confidence Intervals (CI).

Table 2: SEM of examinees concerning Confidence Intervals

Examination	68%	95%	99%
May/June Series	±4 (Precision)	±7.84(Moderate precision)	±10.32(Partial Precision)
Nov/Dec. Series	±12 (No Precision)	±23.52(No Precision)	±30.96(No Precision)

From the summary table above it is shown that the performance of students in the WASSCE May/June Economics Multiple-choice objective test of 2019 in Economics is more reliable and precise when compared with the 2019 WASSCE Nov/Dec Economics examination in terms of their test scores across three administrations with a small SEM of 4 (+ or - 4). Also, the performance of students in the2019 Nov/Dec WASSCE Economics Multiple-choice objective test has shown variation (No Precision) in terms of their test scores across three administrations with a larger SEM of 12 (+ or - 12).

± 12 SEM around the observed scores is the range within which there is a 68% chance that a student's true score lies. This is a greater SEM compared to that of the 2019 May/June WASSCE Economics Multiple-choice objective test which shows how the students' overall true scores spread out or away from the observed scores (with a greater deviation from the mean score either to the right or to the left). This will take a toll on the precision and the accuracy of these students' scores.

± 4 SEM around the observed scores is the range within which there is a 68% chance that a student's true score lies. This is a smaller SEM compared to that of the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test which shows how the students' overall true scores spread about or around (to the right or left) with a low deviation from the mean score. This shows relatively more precision and the accuracy of these students' scores.

To be 95% confident of the student's true score, for the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test, the SEM is ± 23.52. SEM of 23.52 around the observed scores is the range within which there is a 95% chance that a student's true score lies. This is an even greater SEM compared to that of 68% confidence interval of the same 2019 GCE WASSCE Economics Multiple-choice objective test which shows how the students' overall true scores spread out or away from the observed scores (with greater deviation from the mean score either to the right or to the left).

Also for the 2019 May/June WASSCE Economics Multiple-choice objective test. To be 95% confident of the student's true score, the SEM is ± 7.84. SEM of 7.84 around the observed scores is the range within which there is a 95% chance that a student's true score lies. This is a relatively small but higher SEM compared to that of 68% confidence interval of the same the 2019 May/June WASSCE Economics Multiple-choice objective test which shows how the student's overall true scores relatively spread out or away from the observed scores (with a relatively great deviation from the mean score either to the right or to the left).

To be 99% confident of the student's true score, For the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test, the SEM is ± 30.96. SEM of 30.96 around the observed scores is the range within which there is a 95% chance that a student's true score lies. This is a huger SEM when compared to that of 68% and 95% confidence intervals of the same the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test which shows how the students' overall true scores spread hugely out or away from the observed scores (with greatly huge deviation from the mean score either to the right or to the left).

To be 99% confident of the students' true score, For 2019 May/June WASSCE Economics Multiple-choice objective test, the SEM is ±10.32. SEM of 10.23 around the observed scores is the range within which there is a 99% chance that a student's true score lies. This is an even greater SEM compared to that of 68% and 95% confidence interval of the same 2019 May/June WASSCE Economics Multiple-choice objective test which shows how the students' overall true scores spread out or away from the observed scores (with relatively high deviation from the mean score either to the right or to the left).

From the analysis above, the 2019 May/June WASSCE Economics Multiple-choice objective test even flagged a smaller SEM of ±10.32 at the maximum confidence interval of 99% compared to that of the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test with SEM of ±12 at a minimum confidence interval of 68%. This shows that the 2019 May/June WASSCE Economics Multiple-choice objective test is more of a precise, correct and reliable examination with SEMs ±4, ±7.84 and ±10.32 at lowest, moderate and highest confidence intervals of 68%, 95% and 99%

respectively (assurance of correctness) than 2019 GCE WASSCE Economics Multiple-choice objective test with SEMs ±12, ±23.52 and ±30.96 at lowest, moderate and highest confidence intervals of 68%, 95% and 99% respectively (assurance of correctness).

Discussion of the findings

Finding in this study revealed that the 2019 May/June WASSCE Economics Multiple-choice Objective Test flagged a smaller SEM of + or - 4 and as stated by AERA, APA, & NCME (1985) the smaller the SEM, the more precise the measurement capacity of the instrument. Consequently, smaller standard errors translate to more sensitive measurements of student progress. AERA, APA, & NCME (1985) On MAP assessments, student RIT scores are always reported with an associated SEM, with the SEM often presented as a range of scores around a student's observed score. The result of the 2019 May/June WASSCE Multiple-Choice objective test in Economics is in agreement with the findings of PMETB (2009,2008,2007,2004) in their study showing the SEM has not changed at all, despite being used on a much-restricted sample that is of much greater average ability than the total sample

The findings of this study were inconsistent regarding the precision of both tests. It was revealed that both tests flagged certain SEMs which has shown that there is variability in the two tests. It was revealed that the 2019 May/June WASSCE Economics Multiple choice objective test is more precise, reliable and correct than the 2019 Nov/Dec WASSCE Economics Multiple choice objective test at all the confidence intervals. However, according to Baker (2016) it was stated that if series of measurements of the same clinical variable is made, we expect those measurements to show some variability. they further claimed that this may be because the measurement process is imperfect, or depend on exactly how the test is performed or because the underlying property being measured may vary from test to test. Therefore, the study agreed with the statement of Baker (2016). The findings of this study are also evident in the statement of PMETB (2009,2008,2007,2004) where they stated that SEM is a better measure of the quality of an assessment and is recommended for routine use.

Conclusions

Based on the findings of this study it was concluded that testees passed the 2019 May/June WASSCE Economics Multiple-choice objective test more than they passed the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test at all the confidence intervals. It was also concluded that the 2019 May/June WASSCE Economics Multiple-choice objective test is more precise, reliable and correct than the 2019 Nov/Dec WASSCE Economics Multiple-choice objective test at all the confidence intervals.

Recommendations

In view of the findings and conclusion of this study, it was recommended that educators should consider the magnitude of SEMs for students across the achievement distribution to ensure that the information they are using to make educational decisions is highly accurate for all students, regardless of their achievement level. It was also recommended that classical Test Theory (CTT) should be adopted in test scoring and test precision because the responses of the students will be treated using their composite scores., thus test practitioners should endeavour to perform precision test assessment analysis (Standard Error of Measurement) of their tests during scoring to help them to gain insight in to students' true scores and correct interpretation of the score.

References

Adedoyin, O. O., Nenty, H. J. & Chilisa, B. (2008). Investigating the variance of Item difficulty parameter estimates based in CTT and IRT. *Educational Research and Review*. 3(2), 83-93.

AERA, APA, & NCME (1985). Standards for educational and psychological testing. Washington, D. C.: *American Psychological Association. p. 94.*

Bland, J. M. and Altman, D. G. (1996). Measurement Error. *British Medical Journal*, 313: 744 - 753.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika.* 16: 297-334.

Fan, X. (1998). Item Response Theory and Classical Test Theory: an empirical comparison of their item/person statistics. *Educational and Psychological Measurement*. Sage Publication, Inc. V58 N3 P357 (25).

Hambleton, R. K.., Swaminathan, H. & Roger, H. J. (1991). Fundamentals of Item Response Theory, Newbury Par, CA. Sage

Jensen, N. (2015). *Making Sense of Standard Error of Measurement*, Research and Thought Leadership, MAP Growth.

Mutodi, P. & Ngirande, H. (2014). Influence of students' perceptions on mathematics performance. A case study of a selected high school in South Africa. *Mediterranean Journal of Social Sciences*, 5(3):431-445.

Postgraduate Medical Education and Training Board (2004). Principles for an assessment system for postgraduate training: A working paper from the Postgraduate Medical Education Training Board, London: PMETB Google Scholar.

Postgraduate Medical Education and Training Board (2007). Developing and maintaining an assessment system - a PMETB guide to good practice, London: PMETB Google Scholar.

Postgraduate Medical Education and Training Board (2008). Standards for curricula and assessment systems., London: PMETB Google Scholar.

Postgraduate Medical Education and Training Board (2009). Reliability issues in the assessment of small cohorts (Guidance 09/1)., London: PMETB, [http://www.pmetb.org.uk] Google Scholar.

Richard, B. (2016). Calculating the Standard Error of measurement. Retrieved from www.worldpress.com.

Tighe, J., McManus, I. C., Dewhurst, N. G., Chis, L. & Mucklow, J. (2010). The Standard Error of Measurement is a More Appropriate Measure of Quality for Postgraduate Medical Assessments than is reliability: An Analysis of MRCP (UK) Examinations; *BMC Medical Education 10 P.40.*