

GENDER DIFFERENCE ON IRT ANALYSIS OF PSYCHOMETRIC PROPERTIES OF THE 2015 KANO STATE MATHEMATICS SENIOR SECONDARY CERTIFICATE QUALIFYING EXAMINATION IN DALA EDUCATION ZONE, KANO STATE.

¹Nafisa Badamasi Tahir, ²Dr Mukhtar Muhammad
 Department of Educational Psychology and Counselling,
 Federal College of Education (Technical) Bichi
 Email: nafisabmt@gmail.com
 Phone Number:08036079644
 &
³Sadiya Halima Mustapha
 Department of Education
 Federal University, Dutse

Abstract

"Gender Difference on IRT Analysis of psychometric properties of the 2015 Kano State Mathematics Senior Secondary Qualifying Examination in Dala Education Zone " is a study that sought to investigate the difficulty, discrimination and guessing-indices of the multiple-choice items of the examination. Three (3) hypotheses stating no significant differences in the mean difficulty, discrimination and guessing indices of items in relation to gender were tested at 0.05 level of significance. The study employed the Item Response Theory (IRT) and Item Analysis as theoretical and conceptual framework respectively. Fourteen Thousand, Five Hundred and Twelve (14,512) (student's scripts were the total population of the study. The sample size used was 370 using the Research Advisors (2006) table and multi stage cluster sampling as the sampling technique. The instrument used for the study was the Scripts of students in Mathematics multiple choice examinations set and administered by KERD in 2015. The study revealed that there was no significant difference in the mean difficulty and discrimination indices of the 2015 Mathematics SSCQE items for male and female students, while there is a significant difference on the guessing indices for male and female students. The paper finally recommended that Further studies on item analysis of SSCQE on mathematics and other subjects should be carried out in other zones of the state. In addition, Further studies on item analysis should be investigated on boarding school students and on mixed schools students to determine the effect of gender on the variables.

Keywords: Gender, Item Response Theory (IRT), Psychometric, Mathematics, Qualifying.

Introduction

It is a common practice for schools to report the ability in terms of student's raw score, which is operationalized as the number of correct response for objective item test. High percentage of correct response is associated with upper ability students whereas the lower ability students are to obtain the least number of correct responses. In interpreting raw

scores, it is assumed that a particular score represents the amount of ability. For example, if a student obtains 90% score, he or she is assumed to acquire the same amount of ability. However, as rightly observed by Wright and Masters (1982), the assumption is rather skewed. Rather than having 90% ability, a student who scored 90% simply means he or she has higher ability than those who scored below 90%. Similarly, when a student fails to score any marks, it doesn't mean he or she has zero ability. It is more accurate to say that all items were more difficult for him or her. In education, theories that address the above mentioned shortcoming of measurement are called test theories. There are three (3) types of test theories in education measurement; these are classical test theory, item response theory and generalizability theory.

For multiple choice items to be an effective evaluation tool, it has to possess some good qualities which include reliability, validity, objectivity and usability. Good quality items are determined by analysing the items. Items can be analysed both qualitatively and quantitatively. Qualitative analysis is done in terms of their content and form, as well as the evaluation of item/ in terms of effective item-writing procedures. Quantitative analysis on the other hand is done in terms of their Psychometric properties. It includes the analysis for item difficulty and item discrimination. Difficulty index indicates how difficult an item is, while discrimination power indicates how an item differentiates between upper ability and lower ability students.

In Kano state, like in some other states in Nigeria, a Qualifying Examination is usually conducted at senior secondary two (SS2) mostly to prepare the candidates for the coming Senior Secondary Certificate Examination. This examination is termed as Senior Secondary Certificate Qualifying Examination (SSCQE) and is conducted by Kano Educational Resource Department (KERD). It was introduced in 1994 for the following reasons:

- i. To serve as mock examination
- ii. To include parents involvement in the education of their children
- iii. To reduce unwanted government expenditure to external examination bodies on unserious students
- iv. To prepare schools for the task of re-addressing identified weaknesses in terms of syllabus coverage and training
- v. To serve as yardstick for government sponsorship.

Even though the Kano State government has made a commendable effort of preparing its students for national examinations like Senior Secondary Certificate Examination (SSCE) through conducting the Senior Secondary Certificate Qualifying Examination (SSCQE), the results of all these examinations are still nothing to write home about. This is especially true with Mathematics results. Despite the importance place on Mathematics, researchers such as Obioma (2005) Okereke, (2006); Okigbo and Osuafor (2008) had observed that a greater majority of the students fail the subject each year. As a result, stakeholders have continued to trade blames on the causes of this mass failure. A great deal of researches has been done and the results shifted the blame on parents, some on society, and others on students themselves, still with others blaming the teachers who have the lion share of the blame. As accusations and counter-accusations on who to blame on the mass failure of students-persist, it is a clear indication that there are still problems yet unsolved. It is against this background that this study investigated the problem of mass failure in

Mathematics, from a psychometric perspective. Accordingly, it sought to examine the behaviour of the items on difficulty, discrimination and guessing parameters, as well as the behaviour of these parameters in relation to gender. Bearing in mind, the knowledge of the findings of the analysis will reveal a clearer picture of the quality of test items taking by students. And shall answer such questions as to whether the test items are appropriate for the student's level and are really suitable for distinguishing between the upper ability students and the lower ability students.

Statement of the problem

Data from West African examination council (WAEC) and Kano Educational Resource Department (KEDR) have shown a consistent poor performance in Mathematics in the Senior School Certificate Examination (SSCE) and Senior School Certificate Qualifying Examination (SSCQE) respectively. For example, in a study by Zalman and Wonu (2017), it was revealed that as far back as twenty-six (26) years ago (1991 - 2016), only 27.31% of students in Nigeria obtained credit and above (Ai-C6) while 72.6% had pass and below (D7-F9) in the MAY/JUNE 2016 Secondary Certificate Examination (SSCE) in general Mathematics. In the same year, studies by Salman (2012) and Adebule (2009), revealed the poor performance. If this is allowed to continue, the fear is that the country is at risk of so many problems. Some of which include; the country may not achieve its vision of 20:20, which is basically anchored on education. The vision promises a large, strong, diversified, sustainable and competitive economy that effectively harnesses the talent and energies of its people and responsibly exploits, its natural endowments to guarantee a high Standard of living and quality of life of its citizens.

Secondly, the importance of Mathematics as prerequisite for entry into any higher institution and the rate at which students failed the subject had made it difficult for majority of students to gain admission into the higher institution of learning in recent times. For instance, Obemeata as cited by Adepoju (2002) found that about 93% of secondary school leavers in any given year fail to qualify for university education and so end up forfeiting the pursuit of many careers that should have benefitted them and the country better and this is very threatening to national development.

Thirdly, the year under investigation, 2015, is a year that Kano state witnessed an exceptional mass failure in Mathematics SSCQE. Records have shown that over 80,000 senior secondary school (SS2) students sat for the examination but only 18,000 passed at credit level it while the rest failed. This translated to only 12 per cent of the total students who got the required 5 credits including mathematics and English (KEDR).

The importance of Mathematics and its trend of poor performance and mass failure is what necessitated this research. Many studies were conducted about the problem, from different perspectives. This researcher investigated the problem from the perspective of analysis of the psychometric properties of the items on 2015 Mathematics SSCQE. Therefore, this research work focussed on difficulty, discrimination, and guessing parameters of each item. The purpose was to find out whether or not the behaviour of the items contributed to the problem.

Bichi (2015) analyses the difficulty and discrimination indices of Chemistry test items used in Kano State qualifying examination in July, 2014. His work titled "Item Analysis using a Derived Science Achievement Test Data" demonstrated the use of classical item analysis to

evaluate quality of the items. Although, the study failed to provide the total population used, and how the sample was selected, it stated the use of 530 student's scripts as sample. Survey design was employed to collect the relevant data. SPSS version 20 was used for the analysis to determine the difficulty and discrimination indices of the items and the classification of the items according to their item characteristics. Findings revealed that out of 40 items, 12 (30%) needs to be revised or improve and 28 (70%) items were good items. Similarly, findings indicated a significant positive *correlation between item difficulty and item discrimination indices. It was recommended. that items off Chemistry tests should be made to pass through all the processes of standardization and validation by conducting psychometric analyses to improve their quality.

Sabri (2013), also analyses difficulty and discrimination indices, as well as the effects of distractors in the Beginner String Ensemble Test among Music Students in Public Universities, India. Results revealed good, distribution of difficulty throughout the test with 78% of the items in the moderate level of difficulty. PPMC was used to calculate the discriminating power of each item, 59% of which were acceptable and 41 % were classified as poor discriminating items. A high percentage of 70% from the total distractors were regarded as implausible. The researcher further established the quality of the test as a whole by using the Kuder-Richardson Formula (KR20 and KR21). The computed KR20 of 0.717 and KR21 of 0.703 indicated that the test is a reasonably reliable instrument in producing consistent scores.

Olutola (2015) tried to obtain the empirical data on the item difficulty and discrimination indices of senior secondary certificate examination (SSCE) multiple choice biology tests used by the West African Examinations Council (WAEC) and National Examinations Council (NECO) in Nigeria. Instruments used were the 2008 NECO and WAEC multiple choice biology test papers. The population of the study was all senior secondary school students in Ekiti state. Descriptive survey research design was used. Results indicated that with a mean difficulty of 0.42 for WAEC and 0.40 for NECO, and mean discrimination of 0.43 for WAEC and 0.39 for NECO, the 2008 WAEC SSCE multiple choice biology test have more discriminating items and more moderate difficult items than the 2008 NECO SSCE biology test.

Mansour S, (2012), carried out analysis on the quality of 2008-2009 multiple choice items for final year students in nursing and midwifery faculty in the Kashan University of Kashan, IR Iran. It was a cross-sectional study where the results were compared to the results of other faculties in the same university, and the same year. The analysis was done from the classical test theory to determine the item difficulty, item discrimination, internal consistency and the distractors role using Laboratory of Education Research Test Analysis Package (LERTAP) version 5.0d. Results revealed that the average of difficulty index was 0.5, discrimination index was 0.36 and 19.3% of all the distractors were sufficiently attractive to be selected. This indicated that the items were of considerable quality. And, their measured indexes were better compared to the rest of the university.

Objectives of the Study

This research work had the following as its objectives, to;

- i. Find out whether or not there is difference in the mean difficulty parameter of the items for students on the basis of gender.

- ii. Find out whether or not there is difference in the mean discrimination parameter of the items for students on the basis of gender.
- iii. Find out whether or not there is a difference in the mean guessing parameter of 2015 SSCQE Mathematics items on the basis of gender.

Research Hypotheses

The following hypotheses were formulated and tested at 0.05 level of significance:

- i. There is no significant difference in the mean difficulty parameter of 2015 SSCQE Mathematics items on the basis of gender.
- ii. There is no significant difference in the mean discrimination parameter of 2015 SSCQE Mathematics items on the basis of gender.
- iii. There is no significant difference in the mean guessing parameter of 2015 SSCQE Mathematics items on the basis of gender.

Design

Bearing in mind that the focus of this research is to analyse the difficulty and discrimination indices of Mathematics Qualifying Examinations as well as guessing indices, the study employed the ex-post facto design. The design is defined as a systematic empirical enquiry in which the researcher does not have direct control of the independent variable(s) because their manifestations have already occurred or because they are inherently not manipulable (Ary, 1979).

Population of the Study

The population of this study comprised all the 14512 Senior Secondary Two (SS2) students who sat for the 2015 Mathematics Qualifying Examination in Dala Education Zone of Kano State, Nigeria. The zone comprised fifty eight (58) senior secondary public schools spread in two (2) Local Government Areas, namely, Gwale and Dala. Thirty - four (34) of the schools were in Gwale while the remaining 24 were in Dala. Fifty five (55) of the schools were Day Schools with only 3 Girls Boarding Schools, 19 were Mixed Schools, 16 are Boys schools and 23 are Girls schools (KERD, 2016).

Sample Size

The population under study is characteristically large and is spread over a large geographical area. Considering these facts, the researcher made reference to Research Advisors (2006) revised version table for determining sample size of a known population. Therefore, 370 SS2 students were drawn from 15 schools in Dala Local Government with the exception of one (1) Boarding School and eight (8) Mixed Schools.

S/N	School Name	School Type	Population	Proportionate Sample
1	A	Girl	494	44
2	B	Boys	995	88
3	C	Boys	537	48
4	D	Boys	141	12
5	E	Boys	93	8
6	F	Girl	107	10
7	G	Boys	277	24
8	H	Girl	244	22
9	I	Boys	85	8
10	J	Girls	333	30
11	K	Boys	220	20
12	L	Girls	232	20
13	M	Girls	190	17
14	N	Boys	152	14
15	O	Girls	56	5
Total			4156	370

Table 3. 1: Summary of the population by gender and the proportional sample size.

Sampling Technique

The researcher selected the sample using the multi-stage cluster sampling technique. This method according to Asika (2008) is used when the researcher recognizes that some populations are distributed in clusters or pockets of settlement and he wants to use these clusters as a basis for selection.

Data Analysis Procedure

The hypotheses generated were tested through the use of inferential statistics, specifically dependent sample t —test at alpha level of 0.05. A t-test is a type of statistical test that is used to compare the means of two groups (kim, 2015).

Results

Test of Hypothesis One (1)

There is no significant difference in the mean difficulty level of 2015 SSCQE items between male and female students.

Table 4.8 Difficulty indices across gender.

Gender	N	Mean	SD	df	P
Male difficulty indices	40	-0.04450	6.8476	39	-1.325
Female difficulty indices	40	1.40375	1.94578		0.193

From the above table, the average mean for the male students is -0.04450 and 1.40375 for the female students. With our p value as 0.193 being > 0.05 , the hypothesis is

accepted, thus, there is no significant difference in the mean difficulty level of 2015 SSCQE mathematics test items on the basis of gender.

Test Hypothesis Two (2)

There is no significant difference in the mean discrimination power of 2015 SSCQE items between male and female students.

Table 4.9 Discrimination indices across gender.

Gender	N	Mean	SD	df		
Male discrimination indices	40	0.0543	1.23935	39	-0.0271	0.788
Female discrimination indices	40	0.1285	1.87536			

From the above table, the mean was 0.0543 for male students and 0.1285 for female students. 0.788 was the p value which is 0.05. Therefore, the hypothesis is accepted. Thus, there is no significant difference in the mean discrimination power of 2015 SSCQE items on the basis of gender.

Test of Hypothesis Three (3).

There is no significant difference in the mean guessing index of 2015 SSCQE items between male and female students.

Table 4.10 Guessing indices across gender.

Gender	N	Mean	SD	Df		
Male guessing indices	40	0.2458	1.76592	39	2.883	0.006
Female guessing indices	40	-0.5157	1.09001			

From the above table, the mean index for male students is 0.2458 and -0.5157 for female students. The p value is 0.006 < 0.05 which means the hypothesis is rejected. Thus, there is a significant difference in the mean guessing index of 2015 SSCQE items on the basis of gender.

Findings

- i. There is no statistically significant difference in the mean difficulty level of 2015 SSCQE Mathematics test items on the basis of gender.
- ii. There is no statistically significant difference in the mean discrimination power of 2015 SSCQE Mathematics items on the basis of gender.
- iii. There is statistically significant difference in the mean guessing index of 2015 SSCQE Mathematics items on the basis of gender.

Discussions

With regards to the hypotheses tested, the first hypothesis which stated that there is no significant difference in the mean difficulty level of 2015 mathematics SSCQE

items on the basis of gender, was tested and revealed p value as 0.193 which is > 0.05 . This implies that the hypothesis was upheld, thus, there is no significant difference in the mean difficulty level of 2015 Mathematics SSCQE items for male and female students. This study is in line with that of Bandele and Adewale (2013) that examined the item difficulty levels of WAEC, NECO and NABTEB Mathematics examination. They also found out that although there was differences in the mean difficulty levels of the tests, with NECO producing the most difficult test, followed by NABTEB and then WAEC, these differences were not significant, thus, the study equated their difficulty. This finding is also in agreement with the findings of Anagbogu, Ihejiamaiwu and Uba (2014) who also revealed that there was no significant difference in student's performance on the basis of sex in both WAEC and NECO Mathematics test items. The second hypothesis which states that there is no significant difference in the mean discrimination power of 2015 Mathematics SSCQE items on the basis of gender was tested, and revealed a p. value of $0.788 > 0.05$, thus the hypothesis was upheld. This by implication means there was no difference in the manner the items discriminated between the male and the female students in. the 2015 Mathematics SSCQE items. These findings contradict the findings of Olatunji and Owolabi (2009) in which the study revealed that number of options significantly affect the discrimination of NECO multiple choice items in Economics.

The third hypothesis was also tested at 0.05 level of significance and results revealed p value as 0.006 this led to the rejection of the hypothesis. Thus, there is a significant difference in the mean guessing indices of 2015 Mathematics SSCQE items on the basis of gender. With the mean guessing indices as 0.24 and -0.51 for male and female students respectively, this by implication means that the number of items answered correctly by mere guessing were more with male students than with female students.

Conclusion

It was also concluded that there is no significant difference in the mean difficulty and discrimination indices of the items for male and female students, while there is a difference on the guessing indices of the items for male and female students.

Recommendations

- i. Studies should be carried out on the essay part of 2015 mathematics SSCQE to generate more empirical data and have a clearer picture of the overall test items.
- ii. Further studies on item analysis of SSCQE on mathematics and other subjects should be carried out in other zones of the state, also to have clearer picture of the test.
- iii. Further studies on item analysis should be investigated on boarding schools students and on mixed schools students to determine the effect of the variables.
- iv. Other forms of analysis such as qualitative analysis and differential item functioning (DIF) should be carried out on SSCQE on different subjects, locations and time frame so as to pin point lapses and improve the overall quality of the test.

References

Abiodun, A.I &Modupe, O.F. (2013). Mass failure in West African Senior School Certificate Examinations in Nigeria: The Teachers ' perspective. Case Studies Journal, (2), 2305509X.

Adebule, S.O (2009). R?tiability and levels of difficulty of objective test items in a mathematics test:' A study often senior secondary schools in five local government areas ofAkure, endo state. Educational Research and Review, 4(11) 585-587.

Adegoke, B.A. (2014). The Role of Item Analysis in Detecting and Improving Faulty Physics Objective Test Items. Journal of Education and Practice, 5(21), 2222-1735.

Ahmad Z.K &Nordin A. (2015). Modelling a multiple choice mathematics test with the Rasch model. 8(12).

Akpan, A.A. (1987). Correlation of mathematical problem-solving ability among secondary schools students in the Cross River State ofNigeria. Unpublished Ph.D. Thesis, University of Ibadan.

Ale, S.O &Lawal, A.O. (2010). The national mathematical center and the mathematics improvement project in nation building. Journal of Mathematical Sciences Education,

Alkharusi, H. (2012). Generalizability Theory: i4n analysis of variance approach to measurement problems in educational assessment. Journal of studies in education 2(1), 2162-6952.

An, X. & Yung, Y. (2014). Item Response Theory: What it is and Flow you can use the IRT Procedure to Apply it, SAS364. SAS Institute Inc.

Anaduaka, U.S & Okafor, C.F. (2013). Poor Performance of Nigerian Students in Mathematics in Senior Secondary Certificate Examination (SSCE): What is not working? JORIND, 11(2), 1596-8303.

Anagbogu, G. E. ,Ihejamaizu, C. C &Uba, E.J. (2014). Analysis of mean score values Qf examinees in WAEC and NECO mathematics objective test items and students ability parameters in Cross Rivers Nigeria. British Journal of Humanities and Social Science, 12 (1), p37;42.

Anasasi, A & Urbina, S. (2010). Psychological Test)ng.7th Edition. New Delhi: PHI Learning Private Ltd.

Ary, D, Lucy, C.J, Ashar, R. (1979). Introduction to research in educa(ion, New York; Holt Rinehaft and Winstom.

Asika, N. (2008). Research Methodology in the Behavioural Science. Lagos: Longman.

Baker,F.B.(2001). The Basics of Item Response Theory. Retrieved from <http://www.ericae.net/irt/>

Bichi, A.A (2015). Item Analysis using a Derived Science Achievement Test Data. International Journal of Science and Research (IJSR) ISSN (Online): 4(5), 2319-7064. Retrieved from www.ijrs.net on 19/4/2017

Bich, M. Y. (2004). Introduction to Research Methods and Statistics. Kano: DEBIS-CO.

Birnbaum, A. (1968). Some latent trait models and -their use in inferring an examinees ability. In Baker, F.B.(2001). The Basics. ofItem Response Theory.

Boopathiraj, C. &Chellamani, K. (2013).Analysis of test items on difficulty level and discrimination index in the test ofResearch in Education. International Journal of Social Science and Interdisciplinary Research, 2(2), 189-193.

Brennan, R.L (2001). Generalizability Theory. New York: Springer-verlag.

Cohen, S. (2009). Psychological Testing and Assessment: An Introduction to Test and Measurement. (7thed.). McGraw-Hill Primis.

Doerr, A. (2011). Gender issues in Mathematics Education. Retrieved from <https://www.unr.edu> on 18/09/2018.

Fajemidagba, M.O. (1986). Mathematical word problem solving: An analysis of errors committed by students. The Nigerian journal of Guidance and Counselling,2(i) 23-30.

Haladyna, T.M: (1999). Deyeloping and Validating Multiple-choice Exam Items. (2nded.). Mahwah, NJ: Erlbaum Association.

Hannah, C. (2014, August 11). Borgen magazine, global education. Retrieved from <http://www.borgenmagazine.com/>

Harris, D. (1989). Comparison of 1, 2 and 3 parameter IRT Models. Educational Measurement Issues and Practice, 8, 35-41.

Koul, L. (2009). Methodology ofEducational Research. New Delhi: Vikas.

Krejcie, R.V. & Morgan, D. (1970).Educational and Psychological Measurement, 30, 607-610.

McCowan, R &McCowan, S. (1999). Item Analysis for Criterion Referenced Test. Center for Development of Human Services. (CDHS)

Morizot, J.; Ainsworth, A. T; &Reise, S.P (2007). Towards modern psychometrics: Application of item response theory models in personality research.

Musa, M &Dauda, E.S. (2014). Trends Analysis of Students' Mathematics Performance in West African Senior Secondary Certificate Examination from 2004 to 2013: Implication for Nigeria 's vision 20:2020. British Journal of Education, 2(7), 50-64.

Nosa, N.E &Ohenhen, N.O. (1998). Science, technology and mathematics: The tools for technological development in Nigeria. Journal of Education, 3(1) 350-356.

Nuhfer, F.B. (1996). The place of Formative Evaluation in Assessment and ways to Reaps their Benefits. Denver: University of Colorado.

Obemeater, J.O. (1995). Education: An unprofitable industry in Nigeria. A postgraduate school interdisciplinary Research Discourse. University of Ibadan.

Oguntuase, D.M, Awe, 0.0, &Ajayi, I.A, (2013).Empirical Nexus between Teaching/Learning Resources and Academic Performance in Mathematics among Pre-University Students in Ile-Ife, South-West Nigeria. International Journal of Scientific and Researcjh Publications, 3(3), 2250-3153.

Ojerinde, D. Pepoola, K. Qjö, F &Onyeneho, P. (2012). Introduction to Item Response Theory: Parameter Models, Estimation and Application. Abuja Marvelous Mike.

Okurut, O. In Sam, W.B, Joshua M. T, & Alice E.A (2005) Gender differences and Mathematics Achievementofrural senior secondary students in cross river state, Nigeria.

Olatunji, D. & Owolobi, H.O. (2009). Difficulty and discrimination of Economics test items with various option formats among secondary schools in Ilorin, Nigeria. *Ilorin Journal of Education*, 28, 49-63.

Olutola, A. T (2015) Empirical Analysis of Item Difficulty and Discrimination Indices of Senior Secondary Certificate Multiple Choice Biology Tests in Nigeria.

Osunde, A. (2009). Essay and Multiple Choice Tes : Bridging the Gap. Workshop paper on multiple choice test item writing procedures for Academic Staff, University of Ilorin. pp. 14-24.1n Thomas, O.A. (2015). fmpirical Analysis of Item Difficulty and Discrimination Indices Q/Senior School Certificate Multiple Choice Biology Test in Nigeria.

Oyejide., A.P. (1991). Effects of Confidence Scoring ProcedureS on the Psychometric Properties of Three Multiple Choice Test Formats,

Sabri, S. (2013). Item Analysis of student comprehensive Test for Research in teaching Beginner String Ensemble using Model Based Teaching Among Students in public Universities. *International Journal of Education and Research*, 1 (12), 2201-6333.

SalmannM.F, Mohammed, A.S, Ogunlade, A.A & Ayinla, J.O (2012). Causes of mass failure in Senior School Certificate Mathematics Examination as viewed by Secondary School Teachers and Students in Ondo, Nigeria. *Journal of Education and Practice*, 3(8), 22221735.

Sam, W.B, Joshua M. T, & Alice E.A (2005) Gender differences and Mathematics Achievement of rural senior secondary students in cross river state, Nigeria.

Sidhu, K.S. (2012). IYévv Approaches To Measurement and Evaluation. New Delhi: Sterling.

Stata Item Response Theory Reference Manual. (2015). Release 14. Texas: STATA Press Publishers.

Susan, A.N. (2003). Formative Evaluation Tools for Campus Conflict Resolutions and MediationProgrammes: Overviews. Denver University of Colorado.

Swerdlik, C. (2009). Psychological Testing and Assessment: An Introduction to Test and Measurement. (7t ed.). McGraw-hillPrimis.

Tachie,S.A&Chireshe R. (2013). High Failure Rate in Mathematics Examinations in Rural Senior Secondary Schools in Mthatha District, Eastern Cape: Learners' Attributions. *Stud Tribes Tribals*, 1 1(1) 67-73.

Wright, B.D. (1982) Masters GN Rating Scale Analysis. Embrestson, SE: MESA press. www.kerdonline.net www.washington.edu www.ums.edu <https://explorable.com>