

**LEVINE LINEAR AND CHAINED EQUIPERCENTILE
EQUATING OF SENIOR SCHOOL CERTIFICATE
EXAMINATION CHEMISTRY MULTIPLE-CHOICE
PAPERS IN SOUTH-WEST, NIGERIA**

By

AROKOYO, Oluwatosin Adebowale

MATRIC NO: 02/25OC188

**A Ph.D. RESEARCH REPORT PRESENTED TO THE DEPARTMENT OF
SOCIAL SCIENCES EDUCATION, FACULTY OF EDUCATION, UNIVERSITY
OF ILORIN, ILORIN, NIGERIA IN PARTIAL FULFILMENT OF THE
REQUIREMENTS FOR THE AWARD OF DOCTOR OF PHILOSOPHY
DEGREE (Ph.D.) IN EDUCATIONAL RESEARCH, MEASUREMENT AND
EVALUATION**

**MAY, 2021
DECLARATION**

I hereby declare that this Thesis entitled “Levine linear and Chained Equipercntile Equating of Senior School Certificate Examination Chemistry Multiple-choice Papers in South-west, Nigeria” was written by me and it has been the record of my own work. I also proclaim that neither the whole work nor any part of it has been, is being or is to be submitted for another degree at this or any other University or examining body.

Signature:.....

AROKOYO, Oluwatosin Adeboale

May, 2021

CERTIFICATION

This is to certify that this study titled “Levine linear and Chained Equipercentile Equating of Senior School Certificate Examination Chemistry Multiple-choice Papers in South-west, Nigeria” was carried out by AROKOYO, Oluwatosin Adebawale (02/25OC188), and has been read and approved as meeting part of the requirements of the Department of Social Sciences Education, Faculty of Education, University of Ilorin, Ilorin, Nigeria for the award of Doctor of Philosophy (Ph.D.) Degree in Educational Research, Measurement and Evaluation.

Prof. O. E. Abdullahi
Supervisor

Date

Prof. R. W. Okunloye
Head of Department

Date

Prof. Felicia A. O. Olasehinde-Williams
Dean Faculty of Education

Date

External Examiner

Date

DEDICATION

This work is dedicated to my God, my Saviour, the Giver of life, with whom there is no shadow of turning;

And to my husband and child.

ACKNOWLEDGEMENTS

My most profound gratitude goes to Almighty God, my father, my Saviour, the only wise God who has kept me this far, and enormously made provision for this programme, all praise to Him.

My sincere appreciation goes to my supervisor, Prof. O. E. Abdullahi, for his relentless effort in ensuring that this work finally comes to its completion. I really appreciate him and wish him and his family God's blessings.

I also appreciate the Head of Department, Prof. R.W Okunloye for his support and encouragement throughout my programme.

I want to sincerely appreciate Prof. H. O. Owolabi who has been one of the key players in the completion of this thesis and has imparted great wisdom in me.

I am very grateful to these lecturers – Prof. O. B. Olawuyi, Dr. H. T. Yusuf , Dr. M. I. Jimoh, Dr. Balogun, Dr. (Mrs) M. B. Bello and Mrs Balogun, I appreciate their encouragement, guidance and co-operation. I also thank every other academic and non-academic staff of the Department of Social Sciences education. Faculty of Education, University of Ilorin, Ilorin.

I will also want to specially appreciate the efforts of Dr. S.D Daramola (a big sister) and Dr. Jumoke Oladele (a dear friend). Their effort from the beginning to the end of this work is deeply appreciated.

I want to deeply and sincerely appreciate my parents, Prince. J. A. Arokoyo of blessed memory who made sound and quality education to be my right, and the best mom, Mrs. Eunice Olusegun Arokoyo, a woman of inestimable value. Words cannot quantify her enormous support, prayers and constant help, she gives large doses of each. I pray she will live long to eat the fruits of her labour in Jesus name. My adorable siblings whose love, encouragement

and supports kept me going, Sister Temitope and My brother Oluwaremilekun, I cannot thank them enough.

I also appreciate my mother in-law, mummy Adeosun. I thank her for her prayers and great support all the way.

This study would not have been possible without daily loads of love and support I receive from my most treasured beings, my family. My husband, Dr. Abimbola Adeosun deserves an extraordinary recognition for his unwavering love, support, encouragement, patience and his amiable sense of understanding. He was always supporting me to read and read. These have given me courage to keep going on and my adorable daughter, Tijesunimi Adeosun who came in the course of this study, thanks for being so patient, studying was not difficult with her. I appreciate the duo for their immeasurable sacrifices which has brought me this far. I cannot love you less.

To all my colleagues, Tolulope Sowunmi, Dr. Banjo Lawal, Dr. Osho whom I have learnt so much from, evaluators' members, family friends, the Filanis, Igbayiloyes, Adeotis, Akejus, Sowunmis, Adegobolas and Full Gospel Business Men's fellowship (University Road Chapter, Ado-Ekiti), I appreciate you all for your supports one way or the other. Thank you all and God bless you.

All Glory, praise and adoration to my God, the King of Kings.

TABLE OF CONTENTS

Content	Page
Title Page	i
Declaration	ii

Certification	iii
Dedication	iv
Acknowledgement	v
Table of Contents	vii
List of Tables	viii
List of Figures	ix
Abstract	x

CHAPTER ONE: INTRODUCTION

Background to the Study.....	1
Statement of the Problem.....	7
Purpose of the Study	10
Research Questions	11
Scope of the Study	11
Operational Definition of Terms.....	12
Significance of the Study	23

CHAPTER TWO: REVIEW OF RELATED LITERATURE

Standardization of Tests	15
Test Score Equating and Methods of Equating	19
Levine Linear Equating	49
Chained Equipercentile Equating	52
Equating Designs	54
Test Equating in Public Examination	64
Equating of Senior Secondary School Chemistry Scores	74
Theoretical Framework.....	82
Appraisal of the Reviewed Literature	86

CHAPTER THREE: RESEARCH METHODOLOGY

Research Design.....	89
Population, Sample and Sampling Techniques.....	91
Instrumentation	93
Procedure for Data Collection	94
Ethical Considerations	95
Data Analysis Techniques.....	96

CHAPTER FOUR: DATA ANALYSIS AND RESULTS

Descriptive Statistics	97
Answers to Research Questions	98
Summary of Findings	106

CHAPTER FIVE: DISCUSSION, CONCLUSION AND RECOMMENDATIONS

Discussion of Findings	107
Conclusion	111
Implications of the findings of the Study.....	111
Recommendations	111
Suggestion for Further Studies	111
References	105
APPENDICES.....	126
Appendix A.....	126
Appendix B.....	130
Appendix C.....	136

LIST OF TABLES

Table 1: Number and Percentages of Students who Obtained Grades 1- 6 in West African Senior School Examination May/ June in Chemistry from 2007 – 2019	3
Table 2: Single Group (SG) Design	56
Table 3: Single Group Counterbalanced (CB) Design	56
Table 4: Random Group Design	57
Table 5: Nonequivalent Groups Anchor Test (NEAT) Design	58
Table 6: Nonequivalent Groups with Anchor Test (NEAT) Design	91
Table 7: Sample Size Table	93
Table 8: Schedule for Data Collection	95
Table 9: Sample Size of Respondents According to Gender and Test Forms Answered	97
Table 10: Profile of Respondents' Performance in Common Items	98
Table 11: Profile of Respondents' Performance in Unique Items	99
Table 12: Levine Linear Equating of WAEC, NECO and NABTEB SSCE Chemistry Multiple-choice Papers.....	101
Table 13: Summary Showing Chained Equipercetile Equating of WAEC, NECO and NABTEB SSCE Chemistry Multiple-choice Papers.....	103
Table 14: Summary Showing the Invariance in the Score Obtained on the Forms from the Two Equating Methods.....	105

LIST OF FIGURES

Figure 1: Diagrammatic Representation of Test Score Equating Methods and Designs	63
--	----

Figure 2: Levine linear equating of scores from test forms A, B and C (WAEC, NECO and NABTEB) SSCE Chemistry multiple choice papers 102

Figure 3: Percentile rank of WAEC, NECO and NABTEB SSCE Chemistry multiple-choice papers 104

ABSTRACT

Senior School Certificate Examination (SSCE) is a standardised test conducted by three examination bodies in Nigeria: West African Examinations Council (WAEC), National Examinations Council (NECO) and National Business and Technical Examination Board (NABTEB). Test experts are interested in the equivalence of grades obtained in subjects taken by candidates across the different subjects. There is however a dearth of literature comparing results of different equating methods for examination bodies in Nigeria especially in SSCE Chemistry. This study was designed to investigate Levine linear and Chained equipercentile equating of Senior School Certificate Chemistry multiple-choice papers in South-west Nigeria. The objectives of the study were to: (i) examine students' performance on the common items of SSCE Chemistry Multiple-Choice papers; (ii) determine students' performance in the unique items; (iii) determine the results of Levine linear method of equating tests with the use of standard score deviates; (iv) determine the results of chained equipercentile method of equating tests with the use of percentile rank; and (v) examine the invariance of SSCE equated scores obtained by Levine linear and Chained equipercentile equating methods.

The research design adopted for the study was Non Equivalent Groups Anchor Test design. The study was carried out among 322,484 Senior Secondary III students in the South-west geopolitical zone of Nigeria. A sample of 1,461 students was selected using multi-stage sampling procedure. The 2017 WAEC, NECO and NABTEB SSCE multiple choice Chemistry papers were used as instruments for data collection. The WAEC, NECO and NABTEB papers yielded coefficients of content validity of 0.70, 0.67 and 0.69 and reliability of 0.81, 0.79 and 0.78 respectively. Data obtained were analysed using mean, standard deviation, standard score deviates of T-score, percentile rank and coefficient of variation.

The findings of this study were that:

- i. mean performance of students on the common items of WAEC, NECO and NABTEB were 5.46, 6.04 and 7.36 respectively;
- ii. mean performance of students in unique items across the test forms (WAEC, NECO and NABTEB) were 17.32, 18.03 and 20.48 respectively;
- iii. the Levine linear method indicated that a score of 20 was found equivalent to scores of 20.17, 20.47 and 22.97 in WAEC, NECO and NABTEB respectively with a standard score deviate of 21;
- iv. chained equipercentile score of 20 in WAEC was equivalent to scores of 20 and 23 in NECO and NABTEB respectively which corresponded to the percentile rank of 75; and
- v. levine linear equating yielded a coefficient of variation of 35.1% while chained equipercentile equating yielded 45.5%.

The study concluded that Levine linear equating was more efficient for equating of Chemistry SSCE scores. This implies that methods to be used for equating test scores should first be determined by test experts before equating. The study recommended that test developers should adopt the Levine linear equating method that yielded lower coefficient of variation in equating test scores in Chemistry multiple choice items.

Word Count: 484

REFERENCES

- Ababio, A.Y. (2000). *New school Chemistry for Senior Secondary School Series*. Onitsha African-FEP Publishers Ltd. 3-5
- Abiri, J. O. O. (2007). *Elements of measurement, evaluation and statistical techniques in education*. Ilorin: Library and Publication Committee, University of Ilorin, Nigeria.
- Achor, E.E., Agogo, P. O. & Orokpo, C. A. (2001). *Some Nigeria Students' Performance in Practical and Theoretical Chemistry Tests as Predictors of their Performance in MOCK-SSCE Chemistry Examinations*. Retrieved March 27th, 2017 from <http://www.sciencepub.net>
- Adewale, J. G. (2015). *Equating two Year BECE Results in Basic Science and Technology in Oyo State, Nigeria*. International centre for educational evaluation (ICEE) Institute of education, university of Ibadan - Nigeria.
- Adokoniyi, (2014). *Equating of Kwara State Joint Senior Secondary School mock multiple-choice Economics papers*. An unpublished M.Ed thesis, University of Ilorin.
- Adonu, I. I. (2014). *Psychometric analysis of WAEC and NECO practical physics tests using partial credit model*. An unpublished PhD thesis, University of Nigeria, Nsukka.
- Agah, J. J. (2013). *Relative efficiency of test scores equating methods in the comparison of students' continuous assessment measures*. An unpublished PhD thesis, University of Nigeria, Nsukka.
- Aguinis, H, Henle, CA, Ostroff, C. (2001). *Measurement in work and organizational psychology* in N. Anderson, D. S. Ones, Sinangil HK, Viswesvaran C. (Eds.), *Handbook of industrial, work, & organizational psychology*. 2001;27-50. London/Amsterdam: Sage.
- Agusiobo, B. C. (2003). Early childhood care curriculum development. *Journal of Curriculum Studies*, 10, (2), 287 – 292
- Akinlele, P.B. (2004). *The Development of Item Bank for Selection Test into Nigeria Universities*. Unpublished Ph.D Thesis Institute of Education. University of Ibadan.
- Alabekee, E.C. (2013). *Estimating Measurement Error and Score Dependability in Examinations Using Generalizability Theory*. University of Nigeria, Nsukka, Nsukka: An unpublished PhD Thesis.
- Albano, A. (2016). *Equate: Observed- Score Linking and Equating*. R package version 2.0-5, URL. Retrieved May 17th, 2017 from <https://CRAN.R-project.org/package=equate>.

- Albano, A. D. (2011). *Equate: Statistical methods for test equating* [Computer software manual]. Available from <http://CRAN.R-project.org/package=equate> (R package)
- Aleluya, D. (2012). *Interpreting test scores: criterion and norm-referenced framework*. Retrieved February 8th, 2018 from <https://www.slideshare.net>
- Alfred, O. (2011). *Assessment of the Equivalence of 2009 Senior School Certificate Multiple Choice Economics Test Items in Ilorin, Nigeria*. M.Ed thesis, University of Ilorin, Ilorin
- Aşiret, S., & Sünbül, S. Ö. (2016). Investigating test equating methods in small samples through various factors. *Educational Sciences: Theory & Practice, 16*, 647-668.
- Baghaei, M. (2010). Test Score Equating and Fairness in Language Assessment. *JELS, 1*, No. 3, 113-128
- Bandele, S.O. & Adewale, A. E. (2013). Comparative analysis of the reliability and validity coefficients of WAEC, NECO and NABTEB construct Mathematics examinations. *Journal of education and Social Research, 397-402*.
- Bartlett, J. E., Kotrlik, J. W. & Higgins, C. C. (2001). Organizational research: determining appropriate sample size in survey research. *Learning and Performance Journal, 19*, 43-50.
- Brennan, R.L. (1990). *Congeneric Models and Levine's Linear Equating Procedures*. ACT research report series 90-12 Iowa City
- Brennan, R.L. (2006). *Educational Measurement*, 4th ed. Westport, CT: Praeger Publishers
- Brennan, R.L. (2001). *Generalizability Theory*. New York: Springer-Verlag.
- Brennan, R.L. (2011). Generalizability Theory and Classical Test Theory. *Applied Measurement in Education, 24*: 1–21. Iowa: Taylor & Francis Group, LLC. DOI: 10.1080/08957347.2011.532417.
- Brody, W. and Powell, N.J. (2003). A new approach to oral testing. *Educational and Psychological Measurement*
- Burrows, B. (2016). What is Standardized Test? Retrieved February 7th, 2019 from <https://www.studyusa.com/en/a/1284/what-is-a-standardized-test>
- Cao, Y. (2008). *Mixed-format test equating: Effects of test dimensionality and common-item sets*. Unpublished doctoral dissertation. University of Maryland.

- Casselman, B. L., Ohlsen, B. R. & Atwood, C. H. (2017). How we have used item response theory and classroom management to improve student success rates in large general chemistry classes. <http://dx.doi.org/10.21577/0100-4042.20170006>
- Chen, F., Huang, X. & MacGregor, D. (2009). Equating or Linking: Basic Concepts and a Case Study. Originally presented at CAL, Washington, 2009
- Chen, H. & Holland, P. W. (2009). *Nonlinear Levine observed score equating. Or is it?* Paper presented at the annual meeting of the National Council of Measurement in education, San Diego, CA.
- Chen, H., Cui, Z., Zhu, R., & Gao, X. (2010). *Evaluating the effects of differences in group abilities on the Tucker and Levine observed-score methods for common-item nonequivalent groups equating.* ACT Research report series (2010-(1). Iowa City, IA: American College Testing.
- Chen, H., Livingston, S. A., & Holland, P. W. (2011). Generalized equating functions for NEAT designs. In A. A. von Davier (Ed.), *Statistical models for test equating, scaling and linking* (pp. 185–200). New York, NY: Springer-Verlag.
- Chong, H. Y. (2007). *A simple guide to the item Response theory (IRT) and Rash Modelling:* Retrieved February 27th, 2018 from <http://www.creative-wisdom.com>
- Chong, H. Y. & Sharon E. O. P. (2005) Test Equating by Common Items and Common Subjects: Concepts and Applications. *Practical Assessment, Research & Evaluation* 10 (4), 1-19. Retrieved July 20th, 2017 from <http://pareonline.net/getvn.asp?v=10&n=4>
- Demir, S., & Güler, N. (2014). Common tests for the non-material equivalent groups working in sync pattern. *International Journal of Human Sciences*, 11(2), 190-208.
- Donlon, T. (1984). *The College Board technical handbook for the scholastic aptitude test and achievement tests.* New York: College Entrance Examination Board.
- Dorans, N. J. (2000). Recentering and realigning the SAT score distributions: How and why. *Journal of Educational Measurement* 39(1), 59-84.
- Dorans, N. J., Moses, T. P., & Eignor, D. R. (2010). *Principles and practices of test score equating* (Research Report No. RR-10-29). Princeton, NJ: Educational Testing Service.
- Dorans, N.J. & Holland, P.W. (2000). Population invariance and the equitability of tests: Basic theory and the linear case. *Journal of Educational Measurement*, 37(4), 281-306.

- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists* (Vol. 4): Lawrence Erlbaum.
- Enu, V.O. (2014). *The use of IRT in the Validation and Calibration of Mathematics and Geography items of Joint Command Schools Promotion Examination in Nigeria*. A Ph.D Post-Field Seminar presented to International Centre for Educational Evaluation (ICEE), Institute Of Education. University Of Ibadan
- Eya, N.M. (2015). Investigating the Contents of the Senior Secondary School Chemistry Curriculum that can Inculcate Entrepreneurial Skills among Students in Nigeria. *A journal of International Journal for Cross-Disciplinary Subjects in Education (IJCDSE), Volume 6, Issue 2, June 2015*
- Fahmy, A. F. M. (2000). Systemic Approach to teaching and learning chemistry for 21st century workshop on New Trend in chemistry. Bud a pest, Hungary, August 5-10,2000.
- Fan, X. (1998). Item response theory and classical test theory: A Empirical Comparison of their Item/Person Statistics. *Educational and Psychological Measurement*, 58(3), 357-374.
- Federal Ministry of Education (2007). Senior Secondary Education Chemistry Curriculum for SS 1 – 3, Abuja, Nigeria educational Research and Development Council (NERDC) Press.
- Federal Ministry of Education (2004). National Policy on Education. Lagos: Federal Ministry of Information, Printing Division.
- Felan, G. D. (2002). *Test Equating: Mean, linear, Equipercentile and item Response theory*. Paper Presented at the Annual Meeting of the Southwest Educational Research Association. Austin TRXAS.
- Gao, H. (2004). *The effect of Different Anchor Tests on the Accuracy of Test Equating for Test Adaptation*. Unpublished Ph.D Dissertation. Ohoi University.
- Giginna, L. I. & Nweze, B. N. (2014). Creativity in Chemistry Teaching: Effects of e – Learning on students' Achievement in Acids, Bases and Salts. In Z.C. Njoku (Ed). *55th Annual Conference Proceedings of Science Teachers Association of Nigeria*, 253 – 263. Ibadan, HEBN Publisher PLC.
- Gill, J., Johnson, P. & Clark, M. (2010). *Research Methods for Managers*, SAGE Publications.

- Haertel, E. H. (2004). *The Behaviour of Linking Items in Test Equating*. (Case Report. 630). Educational Testing Service, Princeton
- Hambleton, R., Swaminathan, H., & Rogers, H. (1991). *Fundamentals of item response theory*. Newberry Park, CA: Sage.
- Hambleton, R.K. & Jones, R.W. (1993). Comparison of Classical Test Theory and Item Response Theory and their Applications to Test Development. *Educational Measurement. Issues and Practice* 12 (3): 38-47.
- Harbor-Peters, V. F. A. (1999). *Noteworthy points on Measurement and Evaluation*. Enugu: Snaap Pres Ltd.
- Heh, V. K. (2007). Equating accuracy using small samples in the random groups design. PhD dissertation, Ohio University.
- Holland, P. W., von Davier, A. A., Sinharay, S., & Han, N. (2006). *Testing the untestable assumptions of the chain and poststratification equating methods for the NEAT design*. ETS research report series 2006 (1), 2006.
- Hou, J. (2007). *Effectiveness of the hybrid Levine equipercentile and modified frequency estimation equating methods under the common-item nonequivalent groups design*. PhD (Doctor of Philosophy) thesis, University of Iowa, 2007. <http://ir.uiowa.edu/etd/339>.
- Jimoh, A. T. (2001). The place of women education in the development of science, mathematics and technology. *42nd Annual conference Proceeding of Science Teacher Association of Niger*. pp. 170-173.
- Keeves, J.P & Watanabe, R. (2003). International handbook of educational research in the Asia- Pacific Region. Springer science and business media. Copyright <https://books.google.com.ng/books?isbn=1402010079>
- Kolawole E.B. (2001). *Test and Measurement*. Yemi prints and publishing services, Ado-Ekiti.
- Kolen, M. J. & Brennan, R. L. (2004). *Test Equating, Scaling and Linking Methods and Practices*. New York: Springer-Verlag.
- Kolen, M. J. & Brennan, R. L. (2004). *Test equating: methods and practices*. (2nd ed.) New York, NY: Springer –Verlag
- Kolen, M. J. & Brennan, R. L. (2014). *Test Equating, scaling and linking methods and practices*. Statistics for Social Behavioral Sciences. DOI: 10.1007/978-1-4939-0317-7_2, Springer Science+Business Media New York 2014

- Kolen, M.J. & Brennan, R.L. (1995). *Test equating methods and practices*. New York: Springer.
- Kolen, M.J. & Brennan, R.L. (2014). *Test equating: methods and practices*. Springer science and business media. Copyright <https://books.google.com.ng/books?isbn=1493903179>
- Kpolovie, P. J., Ololube, N. P. & Ekwebelem, A. B. I.(2011). Appraising the performance of secondary school students on the WAEC and NECO SSCE from 2004 – 2006. *International Journal of Scientific Research in Education* 4(2), 105 – 114.
- Korau, Y.K. (2006). *Educational Crises Facing Nigerian Secondary Schools and Possible Solutions being a paper presented at Faculty of Education, University of Ibadan.*
- Lakes, K.D. and Hoyt, W.T. (2009). Applications of Generalizability Theory to Clinical Child and Adolescent Psychology Research. *Journal of Clinical Child and Adolescent Psychology*, 38(1): 144–165.
- Lamprianou, L. (2007). *An investigation into the test equating methods used during 2006, and the potential for strengthening their validity and reliability*. Cyprus Testing Service, Manchester.
- Lawson, S. (1991). One parameter latent trait measurement: Do the results justify the effort? In B. Thompson (Ed.), *Advances in educational research: Substantive findings, methodological developments* (Vol. 1, pp. 159-168). Greenwich
- Lee, O. K. (2003). Rasch simultaneous vertical Equating for measuring Reading Growth. *Journal of Applied Measurement*, 4(1), 10-23.
- Lissitz, R. W. & Huynh, H. (2003). Vertical Equating for State Assessments: Issues and Solutions in Determination of Adequate yearly Progress and School Accountability. *Practical Assessment, Research & Evaluation* 8(10): 2003.
- Liu, J., Zu, J., Curley, E. & Carey, J. (2014). *Test score equating using discrete anchor items versus passage-based anchor items* (Research Report No. RR-14-14). Princeton, NJ: Educational Testing Service. DOI: 10.1002/ets2.12015
- Livingston, S. A. (2004). *Equating Test Scores (without IRT)*. Princeton: Educational Testing Service.
- Livingston, S. A., & Kim, S. (2009). The circle-arc method for equating in small samples. *Journal of Educational Measurement*, 46(3), 330–343.

- Livingston, S. A., & Kim, S. (2010). *An Empirical Comparison of Methods for Equating With Randomly Equivalent Groups of 50 to 400 Test Takers*. (ETS Research Rep. No. RR-10-05). Princeton, NJ: ETS.
- Magno, C. (2009). Demonstrating the difference between Classical Test Theory and Item Response Theory Using Derived Test Data. *The International Journal of Educational and Psychological Assessment*, 1 (1), 1-11
- Mbella, K.K. (2012). *Data Collection Design for Equivalent Groups Equating: Using a Matrix Stratification Framework for Mixed-Format Assessment*. PhD dissertation, University of North Carolina, Greensboro
- Minor, L.L. (2013). *Generalizability and Dependability of Brief Behavior Rating Scales for Social Skills*. An Unpublished PhD Dissertation, Louisiana State University and Agricultural and Mechanical College.
- Morrison, S. & Free, K. (2001). Writing multiple-choice test items that promote and measure critical thinking. *Journal of Nursing Education* 40(1): 17-24.
- Moyinoluwa T. D. (2015). Analyzing the psychometric properties of Mathematics in public examinations in Nigeria. *Research on Humanities and social Sciences vol. 5, No. 7, 2015*.
- Mroch, A. A., Suh, Y., Kane, M. T., & Ripkey, D. R. (2009). An evaluation of five linear equating methods for the NEAT design. *Measurement*, 7, 174–193.
- Mulemwa, (2002). *The Challenge of Providing Quality Scientist in Africa*. Proceeding 43rd Annual Conference of Science Teachers Association of Nigeria.
- Muraki, E., Hombo, C. M. & Lee, Y. W. (2000). Equating and linking of performance assessments. *Applied Psychological Measurement, Vol. 24 No. 4, December 2000, 325–337 Sage Publications, Inc.*
- Ndifon, O. C. and Ndifon, A.R. (2012). Public examination in Nigeria and punishing malpractice: human rights perspective in *British journal of arts and social sciences ISSN: 2046-9578,1.5 (2) (2012) ©Britishjournal Publishing, Inc.2012*
- Nenty, H. J. (2004). *From Classical Test Theory (CTT) to Item Response Theory (IRT): An introduction to a desirable transition*. In: OA Afemikhe, JG Adewale (Eds.): *Issues in Educational Measurement and Evaluation in Nigeria*. Institute of Education, University of Ibadan, Ibadan, Nigeria, pp.372- 384.
- Nkpono, H.L. (2001). *Application of latent trait models in the development and standardization of a physics achievement test for senior secondary schools*. Unpublished Ph.D Thesis, University of Nigeria.

- Nwagwu, C.C. (2002). Students' academic performance and the crisis of confidence in Nigerian education system. *Ilorin journal of Education*, 21, 10.
- Obinne, A. D. E. (2011). A psychometric analysis of two major examinations in Nigeria: Standard error of measurement. *International Journal of Educational Sciences*, 3(2), 137-144.
- Obinne, A.D.E., Nworgu, B.G., & Umobong, M.E. (2013). An investigation into differential item functioning of tests conducted by the two major examination bodies in Nigeria. *Advances in Educational Research* 2 (1), 1-8.
- Obioma, G. and Salau, M., (2007). A Predictive Validity of Public Examinations: A Case Study of Nigeria. *A Paper Presented at the 33rd Annual Conference of International Association for Educational Assessment (IAEA), held in Baku, Azerbaijan.*
- Ojerinde D.P. and Onyeneho P. (2012). A Comparison between Classical Test Theory and Item Response Theory: Experience from 2011 Pre-Test in the Use of English Paper of the Unified Tertiary Matriculation Examination (UTME). *Journal of Educational Assessment in Africa* 29 (7) 173 – 191.
- Okeke, O. J. (2011). *Effect of mindmapping teaching strategy on students' achievement, interest and retention in senior secondary school chemistry*. An unpublished PhD thesis, University of Nigeria, Nsukka.
- Okoro, D.U. (2004). *'Quality in education with particular reference to learning achievement'*. A paper presented at the national conference on education for all. Abuja, 20th – 23rd August.
- Okoye, R. O. and Nwafor, V. C. (2009). A Comparison of SSCE Questions Set by the West African Examinations Council and the National Examinations Council, Nigeria. *An International Multi-Disciplinary Journal, Ethiopia Vol. 3 (5)*
- Olatunji, D. S. (2015). *Analysis of linear and equipercentile equating of senior school certificate examination economics multiple-choice papers In Kwara State, Nigeria*. Unpublished Ph.D. Thesis, University of Ilorin.
- Olatunji, D.S. (2007). *Effects of number of options on psychometric properties of multiple choice tests in economics*. Unpublished M.Ed Thesis, University of Ilorin, Ilorin.
- Olen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices* (2nd ed.). New York, NY: Springer.
- Olorundare, A.S. (1998). Problem Solving in Mathematics and Science Education. *The Nigeria Experience, Journal of Educational Policy* 5(1), 1-19.

- Olutola, A.T. (2011). *Analysis of item parameters of senior school certificate multiple choice in biology in Ekiti State, Nigeria*. Unpublished Ph.D Thesis, University of Ilorin, Ilorin.
- Omo-Egbekuse, J. and Afemikhe, O. A. (2014). Equating for quality assurance in innovative examining system: Application of Kernel and Equipercentile Methods. *A paper presented at the International Conference of the Institute of Education, University of Ibadan*
- Oni, J.A. (2001). *'Resources and resources utilization as correlates of school academic performance in the secondary Pre-vocational education in Ogun State, Nigeria'*. Unpublished PhD thesis, University of Ibadan, Ibadan.
- Osho, A. H. (2019). *Comparison of Separate and Concurrent Equating of 2016 Basic Educaion Certificate Examination Mathematics Multiple choice Test in Lagos State*. Unpublished Ph.D. Thesis, University of Ilorin.
- Owolabi, H.O. (2004). Assessment in the Classroom. In Ogunsakin Ed. Teaching in Tertiary Institutions. Paper presented at the workshop on teaching for newly recruited university lecturers. Faculty of education, University of Ilorin.
- Oyekan, S.O. (2000). *Foundations of Teachers Education*. Ondo: Egunola Printers Nig Ltd.
- Ozdemir, B. (2017). Equating TIMSS Mathematics Subtests with Nonlinear Equating Methods Using NEAT Design: Circle-Arc Equating Approaches. *International Journal of Progressive Education, Volume 13 Number 2, 2017*
- Pandeya, S. (2015). Steps for Construction of a Standardized Test. Education Ke Concept. Retrieved on February 8, 2019 from <http://sudhake.blogspot.com/2015/03/steps-for-construction-of-standardised.html>
- Peter, J. K., Nwachukwu, P. O. and Anthony, B. I. E. (2011). Appraising the Performance of Secondary School Students on the WAEC and NECO SSCE from 2004 to 2006. *International Journal of Scientific Research in Education (IJSRE) 4(2), 105-114 ISSN:1117-3259*
- Pido S. (2010). Comparison of Item Analysis of Uganda Certificate of Education results obtained using IRT and CTT approaches. *Journal of Educational Assessment in Africa. (30) 29 (7) 59-67*
- Ponocny, I. (2002). The Applicability of some IRT Models for Repeated Measurement Designs; Conditions, Consequences and Goodness of fit test. *Methods of psychological Research online 7, 1-12. Retrieved July 9, 2017, from http://www.nicd.edu.na*

- Poulsen, J. and Hewson, K. (2013). *Standardized Testing: Fair or Not?* University of Lethbridge Teaching Centre. Retrieved February 8, 2019 from <https://www.uleth.ca/teachingcentre/standardized-testing-fair-or-not>
- Powers, S. (2010). *Impact of matched samples equating on equating accuracy and the adequacy of equating assumptions*. Unpublished doctoral dissertation: University of Iowa.
- Ricker, K. L. & von Davier, A. A. (2007). *The impact of anchor test length on equating results in a nonequivalent groups design* (ETS Research Report RR-07-44). Princeton, NJ: Educational Testing Service.
- Ryan, J. (2011). *A Practitioner's Introduction to Equating with Primers on Classical Test Theory and Item Response Theory*. Arizona State University Frank Brockmann, Center Point Assessment Solutions
- Saage, O. (2009). *Causes of Mass Failures in Mathematics Examination among Students*. A Commissioned Paper presented at Government Secondary School, Science Day Karu Abuja. 1st March 2009.
- Salako, R. J., Adegoke, B. O. & Ogundipe, L. O. (2017). Performance Appraisal of NECO and WAEC SSCE: An Empirical Evidences from Mathematics and Physics. *International Journal of Innovative Social & Science Education Research* 5(3):1-10, July-Sept., 2017
- Schumacker, R. E. (2005). *Test Equating*. Retrieved Oct 2nd 2017 from <http://www.appliedmeasurementassociates.com/white%20papers/TEST%EQUATING.Pdf>
- Schumacker, R. E. (2010). Classical test analysis. *Applied Measurement Associates LLC*
- Suen, H.K., & McClellan, S. (2003). Item construction principles and techniques. *Encyclopedia of vocational and technological education* (Vol. 1, pp. 777-798). Taipei: ROC Ministry of Education.
- Shin, M. (2015). *An investigation of subtest score equating methods under classical test theory and item response theory framework*. Doctoral Dissertations May 2014 - current. 403. http://scholarworks.umass.edu/dissertations_2/403
- Silvestre-tipay, J. L. (2009). Item Response Theory and Classical Test Theory: An Empirical comparison of Item/Person Statistics in a Biological Science Test. *International Journal of Educational and Psychological Assessment* 1 19-31.

- Sinharay, S., & Holland, P. W. (2009). *The missing data assumptions of the NEAT design and their implications for test equating* (ETS Research Rep. No. RR-09-16). Princeton, NJ: ETS.
- Skaggs, G. (2005). Accuracy of Random Groups Equating with very small Samples. *Journal of Educational Measurement* 42 (4): 309-330.
- Stiggins, R. J. (2008). *An introduction to student-involved assessment for learning* (5th ed.). Columbus, OH: Pearson Merrill Prentice Hall
- Sunnassee, D. (2011). *Conditions affecting the accuracy of classical equating methods for small samples under the NEAT design: A simulation study*. ProQuest LLC, Ph.D Dissertation, The University of North Carolina at Greensboro
- Svend, K. & Christensen K. B. (2002). Analysis of local Independence multidimensional in graphical loglinear Rasch Models. *Education and Psychological Measurement* 45 (6), 856-865.
- Tan, X. & Michel, R. (2011). Why do standardized testing programmes report scaled scores? R & D connections No. 16, Sept, 2011. Educational Testing Service.
- Tang. W, Cui, Y. & Babenko, O. (2014). Internal Consistency: Do We Really Know What It Is and How to Assess It? *Journal of Psychology and Behavioral Science* June 2014, Vol. 2, No. 2, pp. 205-220. Published by American Research Institute for Policy Development retrieved http://jpbsnet.com/journals/jpbs/Vol_2_No_2_June_2014/13.pdf
- Tanguma, J. (2000). *Equating test scores using the linear method: A primer*. Paper presented at the annual meeting of the Southwest Educational Research Association. Dallas, TX.
- Topczewski, A., Cui, Z., Woodruff, D., Chen, H. & Fang, F. (2013). *A Comparison of Four Linear Equating Methods for the Common-Item Nonequivalent Groups Design Using Simulation Methods*. ACT Research Report Series, Iowa City, IA 52243-0168
- Udo, N.N. (2008). *Effect of advanced organizer on students' achievement and retention in senior secondary physics*: Unpublished M.Sc. (Ed) Thesis, Nigeria University, Abakaliki
- Udofia, N. A & Udoh, K. I. (2017). Comparative Analysis of WAEC and NECO Senior Secondary School Mathematics Examination. *Mathematical Theory and Modeling* Vol.7, No.6, 2017

- Ugwu, A. N. (2008). Current Issues on Implementation of Senior Secondary Science Curricular in Nigeria. In U. Nzewi (Ed). *49th Annual Conference proceedings of Science Teachers Association of Nigeria*, 23 – 26, Ibadan, HEBN Publisher PLC.
- Ukonze, J.A., Onuoha, J., Isah, E. and Ogurzor, A.I. (2013). Determination of the functionalities of public examination for achieving national transformation in schools. *Journal of research on humanities and social sciences*. [www.iiste.org/journals/index.php/JEP/index3\(9\)](http://www.iiste.org/journals/index.php/JEP/index3(9)) p 21- 26.
- Van der Linden W. J. (2006). Equating scores from Adaptive to Linear Tests. *Applied Psychological Measurement*, 30(6), 493-508.
- Von Davier A.A. & Han, N. (2004). *Population Invariance and Linear Equating for the Non- equivalent groups design*. Educational Testing Service, Princeton, NJ
- Von Davier, A. A. & Kong, N. (2003). *A unified Approach to Linear Equating for the Nonequivalent groups Designs*. Educational Testing Service, Princeton, NJ
- Von Davier, A. A. & Kong, N. (2005). A unified Approach to Linear Equating for the Nonequivalent groups Designs. *Journal of Educational and Behavioural Statistics*, 30 : 313-342.
- Von Davier, A. A., Holland, P.W. & Thayer, D.T. (2004). The chain and post-stratification method for observed-score equating: Their Relationship to population-invariance. *Journal of Educational Measurement*, 41 (1), 15-32.
- Von Davier, A. A., Holland, P.W. & Thayer, D.T. (2004). *The Kernel Method of Test Equating*. New York. Springer-Verlag.
- Von davier, A.A. & Chen, H. (2013). *The Kernel Levine equipercentile observed-score equating function*. Educational Testing Service, Princeton, NJ
- Wang, T., Lee, W., Brennan, R. L., & Kolen, M. J. (2008). A comparison of the frequency estimation and chained equipercentile methods under the common-item nonequivalent groups design. *Applied Psychological Measurement*, 32(8), 632- 651.
- Wang, T., Won-Clan, L, Brennan, R.L. & Kolen, M.J. (2006). A Comparison of the Frequency Estimation and Chained Equipercentile Method under the Common-Item Non-Equivalent Groups Design: Retrieved August 3, 2017 from <http://www.education.ulowa.edu/casma/documents/17comparechreg.rpt.pdf>.
- Wang, W. (2013). *Mixed-format test score equating: Effect of item-type multidimensional, length and composition of common-item set and group ability difference* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No.3608502)

- West African Examination Council (2002). Chief Examiners Report. Lagos: WAEC
- Wilberg, M. (2004). *Classical Test Theory VS. Item Response Theory – An Evaluation of the Theory test in Swedish Driving – License Test (Em, No.50)*, Sweden: UMED University.
- Wolf, R. (2013). *Assessing the Impact of Characteristics of the Test, Common-items, and Examinees on the Preservation of Equity Properties in Mixed-format Test Equating*. PhD dissertation, University of Pittsburgh
- Wolkowitz, A. & Davis-Becker, S. (2015). Evaluating Common Item Block Options when Faced with Practical Constraints. *Practical Assessment, Research & Evaluation*, 20(19). Available online: <http://pareonline.net/getvn.asp?v=20&n=19>
- Wragg, E.C. (2001). *Assessment and Learning in Secondary School*, London: Routledge Falmer
- Wright, B. & Stone, M. (1999). *Measurement Essentials* (2nd Ed.). Wilmington, Delaware, WIDE RANGE INC.
- Zimowski, M. F., Muraki, E., Mislevy, R. J., & Bock, R. D. (2003). BILOG-MG 3 [computer programme]. Chicago: Scientific Software.
- Zu, J. & Liu, J. (2010). Observed score equating using discrete and passage-based anchor items. *Journal of Educational Measurement*, 47(4), 395–412.
- Zuru, A. A. (2009). An address by ICCON president on the occasion of the fourth induction ceremony of the institutes of chartered chemists of Nigeria at Oranmiyan Hall, Lagos Airport Hotel, Ikeja, Lagos