

**LEVEL OF EXAMINERS' SATISFACTION WITH
WAEC EXAMINATION PROCESS AND MARKING
CONDITIONS IN NIGERIA:
IMPLICATIONS FOR QUALITY ASSURANCE AND
RELIABLE ASSESSMENT OUTCOMES**

By

Dr. Lydia B. Akande

Lecturer I,

Department of Islamic, Christian and Comparative Religious Studies

Kwara State University, Malete

Email: Lydia.akande@kwasu.edu.ng

Tel. (Mobile): 08135448801

&

Dr. Mary O. Esere

Senior Lecturer,

Department of Early Childhood/Primary and Special Education.

Ag. Director, Counselling and Career Services Centre

Kwara State University, Malete

E-mail: mary.esere@kwasu.edu.ng

Mobile: 08033907176

2014

**LEVEL OF EXAMINERS' SATISFACTION WITH WAEC
EXAMINATION PROCESS AND MARKING CONDITIONS IN
NIGERIA: IMPLICATIONS FOR QUALITY ASSURANCE AND
RELIABLE ASSESSMENT OUTCOMES**

FINAL REPORT

By

Dr. Lydia B. Akande

Lecturer I, Department of Islamic, Christian and Comparative Religious Studies
Kwara State University, Malete
Email: Lydia.akande@kwasu.edu.ng
Tel. (Mobile): 08135448801

&

Dr. Mary O. Esere

Senior Lecturer, Department of Early Childhood/Primary and Special Education.
Ag. Director, Counselling and Career Services Centre
Kwara State University, Malete
E.mail: mary.esere@kwasu.edu.ng
Mobile: 08033907176

Prepared Under Grant from Tertiary Education Trust Fund (TETFUND)

Prepared for
TETFUND
A6 Zambesi Crescent
Off Aguiyi Ironsi Street, Maitama
Abuja, Nigeria

Submitted By
The University Research Council
Kwara State University, Malete
PMB 1530, Ilorin
Kwara State, Nigeria

EXECUTIVE SUMMARY

Globalisation has propelled West African Examination Council (WAEC) to benchmark its examination practices and procedures against other examination boards in Sub-Saharan Africa. This, notwithstanding, examinations across nations take place under different socio-economic and socio-political contexts. Local circumstances play a significant role on how valid and reliable assessment outcomes would be. Level of examiner satisfaction is one of the local circumstances that may affect validity and reliability of examinations outcomes. The study is a survey employing a descriptive correlational approach that informed data collection, analysis and interpretation. Level of examiners' satisfaction with marking condition and their overall level of satisfaction with the marking exercise were described in relation to the participants' demographic characteristics. The study targeted all WAEC examiners who were involved in marking WAEC final examination scripts. Data was collected by use of a survey questionnaire. The questionnaire was divided into four sections. Section 1 was on demographic information items. The items solicited personal and background information that would most likely have a bearing on ones level of satisfaction with marking. Section 2 comprised items that asked about specific contextual elements of marking. Such elements were categorised into six groups, namely pre-marking WAEC procedures, marking facilities/services, catering services, security services, remuneration and timing.. Data analysis included conducting descriptive statistics in the form of frequencies, means and standard deviations to describe the population demographically. Further analysis included multiple regression analysis to determine the extent of association between examiners' level of satisfaction with marking conditions and examiners' overall satisfaction with marking. Results showed that examiners reported a fairly high level of satisfaction with the examination processes. However, of all the different categories of examination processes, examiners were most satisfied with the laid out procedures of the examination process. The examiners were least satisfied with remuneration, security, timing and catering during marking. WAEC services and examination procedures seemed to influence the examiners' overall level of satisfaction more than the other categories of examination processes. On the basis of this result, it was recommended that WAEC should consider improving on remuneration and catering, but also keep improving on their services to keep examiners satisfied with the examination processes in Nigeria. This will go a long way in maintaining quality assurance and reliable assessment outcomes.

Key words: Examiners' level of Satisfaction, WAEC, Nigeria, quality assurance, reliability, examination outcomes.

· INTRODUCTION AND BACKGROUND TO THE STUDY

The West African Examinations Council is the foremost indigenous public examination body in West African sub-region. At different stages of the educational process, the assessment movement is always based on standards and outcomes, holding educational institutions accountable for student learning (Esere & Idowu, 2011). Examinations are an integral part of any education and training system. They play a crucial role in driving the teaching and learning processes; feedback obtained from examination outcomes provides yardsticks for measuring educational standards and quality among other uses (Nenty, Odili & Munene-Kabanya, 2009).

Reliability and validity are important throughout the process of examination from paper production, through administration, to scoring. In this paper, the researcher is particularly interested in examiners' satisfaction (Mertens, 2010) which is concerned with how consistent the judgments of the scorers are. (McIntire & Miller, 2007) regards the examinations processes, as critical for any country if the country must reach its objective of developing an educated and informed nation. According to Thobega, Mosalakgotia and Gabalebatse (2010), an educational system with no laid out assessment and examination procedures cannot lay claim to quality in education. The quality of examinations processes and procedures have a direct bearing on the quality of the examinations outcome hence reliability of performance measures. One major activity that has a bearing on the quality of performance measures obtained from examinations is the marking of the examination. Marking is a very demanding task when viewed from the perspective of the examiner (Masole, 2010). There may be several..... Scoring is a strenuous exercise requiring dedicated examiners who can: work in a structured environment; follow directions; recognize and monitor error in their own judgment; work long hours; maintain motivation; and accuracy (Johnson, et al, 2009). As a consequence, it is desirable to

screen would-be examiners prior to training. Tests scores need to be reliable so that they represent students' true level of knowledge and inform proper decision-making about students' learning and placement or selection for further studies and/or employment (Denscombe, 2000; Frisbie, 1988 ; Suto & Nádas, 2008; Masole, 2010). Goldstein (2001) posit that they are also useful as indicators of teachers and especially school performance at national and international levels as well as monitoring examination standards over time. Although we do not expect tests scores to be perfect measurements, their error should be tolerable. We need to determine this level of tolerable error by conducting a reliability study.

Studies have shown that inter-rater reliability is difficult to achieve. For example, Rennert-Ariev (2005) writes that written essays were scored differently by different raters and that even the same rater scored responses differently at different times. . Valentine, 1932; Murphy, 1978, 1982; Newton, 1996; Pinot de Moira, Massey, Baird and Morrissey, 2002; in Suto and Nádas (2008b) posit that inter-marker agreement in public examination marking in the UK have been found to be imperfect, varying significantly among examination subjects as well as among teams of markers.

There may be several contextual factors that affect the way an examiners score candidates papers. Eklof (2010) has asserted that assessment scores should take into account the motivation of the test-taker and by the same line of argument, one could assert that examiners' satisfaction with contextual factors within which they operate could also explain some variability in the examination scores. Such contextual factors are thus important in enhancing the quality and reliability of the examination scores. Quality of the marking process depends to a large extent on who the examiners are. Many advanced examination systems contract subject matter specialists, be they teachers or officers from other fields, some even use retired personnel to do the marking, while in some situations the

examination is a preserve of school-based teachers. In England, the General Certificate of Secondary Education (GCSE) is a product of both internal and external assessment systems; internal assessment is generally given a lower weight than external assessment when computing final grades for GCSE (Harlen, 2007).

WAEC in Nigeria engages teachers to mark the external examination so as to generate the final examination scores. Teachers from schools around the country are contracted to mark the subjects they teach at school. WAEC contracts teachers to mark external examinations because teachers are the most convenient subject matter specialists to contract. However the fact that the quality of examination scores that teachers produce may be affected by contextual factors (Leighton, Gokiart, Cor and Hefferman, 2010) holds true for Nigerian teachers as it does for teachers elsewhere. Satisfaction with marking conditions may be one of the contextual factors that may impact on the performance of the examiners, and by extension, on the reliability of examination scores produced.

Satisfaction with marking condition would entail an examiners' overall feeling about marking in terms of specific facets of the process that include, facilities for marking, catering security, remuneration, general services during marking, and general procedures (Gini, 2000). It is also worth noting that the extent to which the listed factors impact examiners satisfaction may vary with other population variables such as gender, training in marking, teaching experience, qualification, position, the subject being marked, and experience in marking. The Nigerian arm of West African Examination Council may have done everything right to ensure validity and reliability of the examinations and even the marking process, but if examiners are not satisfied with the marking, then the examination results would be at risk of being unreliable. This study investigated the level of examiners' satisfaction with conditions they work under during WAEC marking exercise.

Problem of the Study

Scoring is a very demanding task conducted continuously over a short period of time by WAEC. During this time, examiners are expected to mark a lot of scripts, putting them under pressure. This has the potential to increase error. Some extraneous variables may also interfere with the reliability of the assessment outcome. This study is aimed at investigating such variables with a view to preferring some valid solutions. The objectives are;

1. Determine the level of examiners' satisfaction with factors that make-up marking conditions of the examiners.
2. Identify marking conditions that have a direct relationship with examiners' overall satisfaction with the marking exercise.
3. Identify factors that WAEC should improve to provide suitable marking conditions.

Purpose of the study

The grades awarded to candidates have serious implications for the candidate's fate: admission to tertiary institutions and entry into formal employment hence methodological processes for assuring quality during scoring should be built into all stages of examination processing. The main purpose of this study was therefore to investigate the level of examiners' satisfaction with WAEC examination process and marking conditions.

METHODOLOGY

Design: Dunn (2001) defined a research design as an organised collection of procedures used by researchers to collect behavioural data. The main purpose of this study was to investigate level of examiners' satisfaction with WAEC examination process and marking conditions in Nigeria. Therefore the research design that was adopted for the study was a descriptive correlational approach that informed data collection, analysis and interpretation. Level of examiners' satisfaction with marking condition and their overall level of satisfaction with the marking exercise were described in relation to the participants' demographic characteristics. Descriptive survey is concerned with the collection and summarising of numerical data. However, the major function of descriptive survey is the presentation of information in a convenient, applicable and understandable manner. It also focuses on the people, their beliefs, opinions, perceptions, motivations and attitudes as well as behaviours. This method is preferred because it captures the complexity of everyday behaviour (Stangor, 2004). It also describes variables in given situations and establishes relationships between variables (Hassan, 1995).

Population, Sample and Sampling Procedure: The population for the study comprised all WAEC examiners in Nigeria while the target population comprised all WAEC examiners in Kwara State, Nigeria. Daramola (1995) , defined a sample as the selected group which is a fair representation of the population of interest. He defined a sampling procedure as a systematic process employed to select a required proportion of a target population. Purposive random sampling technique was used to select 1000 examiners who were involved in marking final examination scripts during the 2013 examination cycle. But only 600 correctly filled and returned questionnaire forms were used for the study. Purposive sampling technique is a method whereby respondents are purposively selected for a study because they possess the desired traits that the researcher is searching for.

Instrumentation: Data were collected by use of a survey questionnaire. The questionnaire was divided into four sections. Section 1 was demographic information items. The items solicited personal and background information that would most likely have a bearing on ones level of satisfaction with marking. A few items from the section included gender, *qualification*, teaching experience, position at work, marking experience etc. Section 2 comprised items that asked about specific contextual elements of marking. Such elements were categorised into six groups, namely facilities, catering services, security services, remuneration procedure, WAEC services and procedures, and general marking procedures. Under each of the marking conditions, specific questions were listed with Likert-type response categories. Examiners were to indicate their level of satisfaction with the question by picking a category on the scale. The categories were; very dissatisfied, dissatisfied, satisfied, and very satisfied. The survey was supplemented by in-depth interview.

Pilot Testing

According to Daramola (1995), a try out procedure is essential for the development of a sound plan. Mustapha (2006) stated that pilot study is a trial-run of the actual investigation and it involves pre-testing of a research instrument in order to remove any form of ambiguity and irrelevant items. Therefore, in order to determine the usability of the instrument, a pilot testing was conducted. The instrument was pilot tested on a representative sample of twenty (20) respondents who did not form part of the final sample for the study. Difficult and ambiguous items that the respondents of the pilot test could not answer were restructured. The purpose was to assess the usability of the questionnaire from their responses.

Psychometric Properties of the Instrument

The psychometric properties of the instrument lie in its validity and reliability. The properties are presented as follows:

Validity: Validity is a vital property of a measuring instrument because it answers the basic question of whether the items of an instrument reflect the significant aspects of the purpose of a study. Ogunlade, (1982) described validity as the extent to which a test measures what it claims to measure. He further said that, it is vital for a test to be valid in order for the results to be accurately applied and interpreted. Therefore, the content validity of the instrument was determined by giving the questionnaire to measurement and evaluation experts in the Departments of Counsellor Education and Social Science Education so as to offer suggestions for the modification of the instrument. The experts' suggestions and corrections were incorporated in the final draft of the questionnaire. On a second look, the experts unanimously agreed that the instrument has content validity.

Reliability: Reliability of an instrument is the degree to which an instrument yields consistent results when it is administered over a number of times. Ogunlade stated that a test can be said to be reliable if it consistently and under varying conditions measures what it ought to measure. To ensure the reliability of the instrument, the test re-test method was adopted. The researcher administered the same instrument on twenty (20) respondents within an interval of three weeks. The two sets of scores, obtained were computed using the Pearson Product Moment Correlation Co-efficient Formula. A correlation coefficient of 0.73 was obtained. This was considered appropriate as reliability measure for an instrument of this nature.

Procedure for Administration

The questionnaire forms were personally administered by the researchers with the help of 4 trained research assistants. The respondents were informed of the importance of the study and were guided on how to complete the questionnaire. The researchers established adequate rapport with the respondents in order to ensure maximum cooperation and returns. The respondents were encouraged to respond accurately and promptly to the questionnaire.

Table 1 shows examiners' level of satisfaction with each of the marking conditions on the scale of highly satisfied, satisfied, unsatisfied and highly unsatisfied. The table shows that the respondents were highly satisfied/satisfied with pre-marking WAEC procedures, and marking facilities/services but highly unsatisfied/unsatisfied with security services, remuneration procedures, catering services as well as timing.

Furthermore, examiners were compared on their level of satisfaction according to selected moderating variables as shown in Table 2.

Table 2: Level of Satisfaction by Selected Moderating Variables

Marking Condition	Variable	Diff. Between	More Satisfied
Catering	Gender	Males & Females	Males
Remuneration	Gender	Males & Females	Females
Facilities	Gender	Males & Females	Males
Security	Ethnic Group	Yoruba & Non Yoruba	Yoruba
Procedures	Experience	1-5yrs & 16yrs & above	1 – 5 yrs
Timing	Qualification	1 st Degree & Masters/Ph.D	1 st Degree

Table 2 shows differences in levels of satisfaction with certain factors of marking, according to population attributes like gender, ethnic group, qualification, and marking experience. Notice from the table that females were more satisfied with remuneration than males, but males were more satisfied with facilities and catering than females were. The Yorubas were more satisfied with security than their Non Yoruba counterparts while in terms of WAEC pre-marking procedures, those with 1 to 5 years of experience were more satisfied than those with 16 years and above marking experience. In terms of timing, the first degree holders were more satisfied than holders of Masters and Ph.Ds.

Table 3: Multiple Regression Output between Marking Conditions and Examiners' Satisfaction with marking

Variable	Coefficient	Std Error	Beta(β)	t	p-value
(Constant)	65.88	1.39		47.42	.00
Procedures	-1.56	.40	-.09	-3.85*	.00
Catering	.68	.38	.04	1.82	2.00
Remuneration	-.08	.42	-.00	-.18	1.86
Facilities	-.53	.13	-.09	-4.01*	.00
Security	.85	.47	.04	1.82	2.00
Timing	-.54	.38	-.03	-1.40	1.16
Model Summary					
F(6,594) = 7.08, $p \leq 0.00$					
$R^2 = 0.02$					

*significant, $p \leq 0.05$

Table 3 shows Multiple Regression Analysis result which indicated that WAEC pre-marking procedures and facilities significantly influenced level of examiners' satisfaction whereas the remaining variables did not.

DISCUSSION

This study investigated Level of Examiners' Satisfaction with WAEC Examination Process and Marking Conditions in Nigeria. The target sample was all examiners who were involved in marking final WAEC examination scripts. However, only 600 properly filled and returned scripts were used for data analysis. The small sample size was informed by high rates of examiners opting to commit all their time to marking examination scripts because each script marked is paid for while completing the survey questionnaire was not paid for. Figures 1 to 5 showed the demographic characteristics of the respondents. Figure 1 shows that there were more females than males at 65% and 35% respectively; the highest qualification for most examiners was Bachelor's degree, followed by those with Post Graduate Diploma in Education (PGDE) degrees and those with Masters and Ph.D degrees. Interestingly, 100% of the examiners are trained teachers. Most of the examiners are Yorubas (68.6%), Majority of the examiners have 6 to 15 years WAEC examination marking experience. Although marking of the candidates scripts is done in a relatively

short period of time, most of the examiners were satisfied with pre-marking WAEC procedures. Examiners' marking experience was found not to affect their satisfaction with marking conditions. New examiners were more satisfied than their older counterparts.

Examiners reported high levels of satisfaction with general marking procedures, and WAEC facilities/services, and low levels of satisfaction were reported with remuneration, security services, timing and catering services. The West African Examinations Councils' mandate is to conduct school examinations and any other examinations for the Ministry of Education and issue certificates in respect of the examinations. It is therefore plausible to expect the council to handle examinations processes and procedures in a satisfactory manner because such procedures fall within the core functions of the Council. It is also noteworthy that WAEC works hand in hand with the examiners throughout the test development process and the examination cycle. During the test development process, they are involved as content experts and during the examination cycle, they are involved as examiners, invigilators and examination centre administrators. It is therefore logical that the examiners were satisfied with WAEC services and procedures since they participated in formulating such procedures. Some examiners reported low levels of satisfaction with remuneration (especially the male examiners), marking facilities and catering services (especially the female examiners). Those are the marking conditions that WAEC must work to improve, however as long as the current marking exercise does not change; it may not be easy to satisfy every examiner on these marking conditions. The examiners are diverse by level of operation, marking venue and even individually, so harmonising everything across the diversity of attributes may not be easy to achieve.

To determine the exact contribution of each of the marking conditions to examiners' overall satisfaction with the marking exercise, a multiple regression analysis was conducted.

The analysis revealed that WAEC services and WAEC procedures explained a higher percentage of variability and were significant in the examiners' overall satisfaction with the marking exercise. Other factors like security services, remuneration, catering, and timing did not make any significant contribution to the examiners' satisfaction with marking. It is encouraging to realise that the satisfaction of examiners is not hinged on remuneration, but rather on factors that WAEC has total control over. Nevertheless, WAEC should still make efforts to improve on those conditions that examiners were minimally satisfied with, including remuneration and catering. It is also recommended that WAEC should improving on examination services, processes and procedures so as to keep examiners satisfied with the marking exercise.

RECOMMENDATIONS

The grades awarded to candidates have serious implications for the candidates' fate in terms of admission to tertiary institutions and entry into formal employment hence methodological processes for assuring quality during scoring should be built into all stages of examination processing. Screening for quality examiners during hiring stage is very important to improve the reliability of scores.

It is normally hypothesized that examiners' concentration is highest in the morning whilst they are fresh and as the day progresses, their accuracy in scoring diminishes. Examiners in this report would rather have the timeline for marking extended to give room for more relaxed marking instead of the existing situation where they have to rush so as to meet up with the timeline. This, the examiners believe will go a long way in entrenching quality assurance and reliable assessment outcome.

SUGGESTIONS FOR FURTHER RESEARCH

The focus of this study was limited to the moderating variables of gender, ethnicity, qualification and teaching experience; hence there is need for further investigation into other variables that could affect the examiners level of satisfaction with marking conditions.

The validity and reliability of the instrument used in this study need to be researched into on national level.

STATEMENT OF BUDGET EXPENDITURE

ACTIVITY	DETAILS	SUB-TOTAL N : K	TOTAL N : K
Ground work/ Archival and Library Search and Literature Review: Every meaningful research work should be based on existing works that preceded it. With the completion of such research, it will add more to the existing knowledge in the field of study. Literature review provided the basis for this study. It situated the current study in the context of the previous related studies.	i) Logistics for Internet search (N8,000 x 3 months)	24,000.00	70,000.00
	ii) Logistics for visits to various libraries of institutions	46,000.00	
Survey: The results of the survey served as needs assessment for the field study	i) Logistics for instrument development and production	40,000.00	70,000.00
	ii) Logistics for establishing psychometric properties of the instrument	5,000.00	
	iii) Logistics for field testing	25,000.00	
Intervention: Questionnaire distribution was the key intervention of	i) Logistics for questionnaire distribution and		

the study and was carefully planned and executed	ii)	collection Logistics for training of 4 research assistants	100,000.00	
	iii)	Logistics for purchase of appropriate gadgets and materials	10,000.00	
			10,000.00	120,000.00
Post-test Analysis: Having collated the required data, there was the need for the investigators to stay in a conducive environment for analysis and interpreting of the data and report writing	i)	Logistics for data analysis	50,000.00	
	ii)	Logistics for report writing	40,000.00	
	iii)	Logistics for evaluation and presentation of findings to the school authority. (Both hard and soft copies)	10,000.00	100,000.00
Project Outcome Dissemination: To ensure that the project outcome is given widest publicity, it will be published in an international journal of repute.		Processing and Publication of the project outcome (N55,000)		55,000.00
Allowances: A token fee was paid to the investigators and research assistants to encourage and motivate them	i)	Principal Investigator (N15,000 x 3 months)	45,000.00	
	ii)	Assistant Investigator (N10,000 x 3 months)	30,000	
	iii)	Four Research Assistants (N10,000 each)	40,000	115,000
Miscellaneous: This took care of unforeseen expenditures like incentives to respondents, incentives to support staff		N30,000.00		30,000.00
		TOTAL =		N560,000.00

REFERENCES

- Anderson, D. J., Major, R. L., & Mitchell, R. R. (1992). *Teacher supervision that works: A guide for university supervisors*. New York. Praeger Publishers.
- Airasian, P. W., & Russell, M. K. (2008). *Classroom Assessment Concepts and Applications*. Boston: McGraw Hill.
- Ary, D., Jacobs. L. C., & Razavieh, A. (1985). *Introduction to research in education (3rd ed.)*. New York. Holt, Rinehart and Winston.
- Chamberlain, S. (2008). *Do marking reliability studies have validity?* IAEA conference. Cambridge, 7-12 September.
- Daramola, S.O. (1995). *Research and statistical methods in education for students and researchers in tertiary institutions*. Ilorin: Bamitex Printing and Publishing.
- Denscombe, M. (2000). Social conditions for stress: Young people's experience for doing GCSEs. *British Educational Research Journal*, 26(3), 359-374.
- Dunn, D.S. (2001). *Statistics and data analysis for the behavioural sciences*. New York: McGraw-Hill Co.
- Esere, M.O. & Idowu, A.I. (2011). Continuous assessment practices in Kwara State secondary secondary schools. *Nigerian Journal of Research in Education and Society: International Perspective*, 2(1), 71-77.
- Eye, A. V., & Schuster, C. (1998). *Regression analysis for social sciences*. New York, Academic Press.
- Fowles, D. (2008). *Does marking Images of essays on screen retain marker confidence and reliability?* A paper presented at IAEA Conference. Cambridge
- Frisbie, D. A. (1988). Reliability of scores from teacher-made tests. *Educational Measurement- Issues and Practice*, 9, 55-65.
- Gini, A. (2000). *My job myself: Work and the creation of the modern individual*. New York. Routledge. .
- Goldstein, H. (2001). Using pupil performance data for judging schools and teachers: scope

- and limitations. *British Educational Research Journal*, 27(4), 433-442.
- Green, S. B., Salkind, N. J., Akey, T. M. (2000). *Using SPSS for windows*. (2nd Ed.). New Jersey: Prentice Hall.
- Harlen, W. (2007). *Assessment of learning*. Sage Publications Company: London.
- Hassan, T. (1995). *Understanding research in education*. Lagos: Merrifield Publishing Company.
- Hinkle, D. E., Wiersma, W., Jurs, S. G. (1998). *Applied statistics for the behavioral sciences*, (4th Ed.). New York: Houghton Mifflin.
- Kerlinger, F. N. (1979). *Behavioral research: A conceptual approach*. New York. Holt, Rinehart and Winston.
- Johnson, R., Penny, J. A., & Gordon, B. (2009). *Assessing performance designing scoring and validating performance tasks*. New York: The Guilford Press.
- Leighton, J. P., Gokiart, R. J., Cor, M. K., & Hefferman, C. (2010). Teacher beliefs about the cognitive diagnostic information of classroom versus large-scale tests: Implications for assessment literacy. *Assessment in Education; Principles, Policy and Practice*, 17(1), 7-21.
- Masole, M. (2010). *How reliable are examiners in scoring scoring candidates' work?* Being text of a paper presented at the AEAA conference held in Abuja.
- McIntire, S. A., & Miller, L. A. (2007). *Foundations of Psychological Testing: A practical Approach*. (2nd ed.). London: Sage Publications, Inc.
- Maxwell, G. S. (2004). *Progressive assessment for learning and certification: Some lessons from School-based assessment in Queensland*. A paper presented at the third Conference of the Association of Commonwealth Examination and Association Boards. Nadi, Fiji.
- Mertens, D. M. (2010). *Research and evaluation in education and psychology*. California: SAGE Publication, Inc.
- Mustapha, A (2006). *Foundation in business research methods*. Ilorin: Bamitex printing and

Publishing.

- Nenty, H. J., Odili, J.N., & Munene-Kabanya, A. N. (2009). Assessment training among secondary school teachers in Delta State of Nigeria: Implications for sustaining standards in educational assessment. *Journal of Educational Assessment in Africa*, 3, 110- 123,
- Nitko, A. J. (2004). *Educational Assessment of Students*. New Jersey: Merrill Prentice Hall.
- Suto, W. M. I., & Nádas, R. (2007). The Marking Expertise projects: Empirical investigations of some popular assumptions. *Research Matters: A Cambridge Assessment Publication* 4, 2 -5.
- Suto, W. M. I., & Nádas, R. (2008a). *Towards a new model of marking accuracy: An investigation of IGCSE Biology*. A paper presented at IAEA Conference
- Suto, W. M. I., & Nádas, R. (2008b). What determines GCSE marking accuracy? An exploration of expertise among maths and physics markers, *Research Papers in Education*.
- Suto, W. M. I., & Nádas, R. (2008c). *Towards a new model of marking accuracy: An Investigation of IGCSE Biology*. A paper presented at IAEA Conference. Cambridge.
- Thobega, M. Mosalakgotla, B. & Gabalebatse, M. (2010). *Level of examiners satisfaction with JCE and BGCSE marking conditions*. Being text of paper presented at the AEAA conference held in Abuja
- Thorndike, R. M., & Thorndike-Christ, T. (2010). *Measurement and Evaluation in Psychology and education*. New York: Pearson Education, Inc.

ACKNOWLEDGEMENTS

The investigators are grateful to TETFUND for the grant that facilitated this research work.



.....
Dr. Lydia B. Akande
(Principal Investigator)



.....
Dr. Mary O. Esere
(Assistant Investigator)