

Multiple-choice Construction Competencies and Items Quality: Evidence from selected Senior High School Subject Teachers in Kosofe Local Government Area, Lagos State, Nigeria.

Oluwatomi Modupeola ALADE & Oluwatobi Newton ODUNSI

Department of Education Foundations, Faculty of Education, University of Lagos, Nigeria.

Abstract

This study investigated the relationship between between senior high school subject teachers' competencies in constructing multiple-choice tests and the resultant quality of test items in Kosofe Local Government Area of Lagos State. The research examined the characteristics of each test item, focusing on the difficulty index and the discrimination index. Employing a descriptive-correlative research design, the study sampled five senior secondary schools in Kosofe Local Government Area using simple random sampling technique. Data collection involved a questionnaire named Teachers' Multiple-Choice Test Construction Competence Questionnaire (TTCCQ-MC) which comprised twenty statements. The study sample were randomly selected among the teachers in the twenty-five senior secondary schools that participated. The statistical tools used for analyses included mean, standard deviation, correlation, and multiple regression to address the research questions and hypotheses. Findings indicated a lack of statistically significant correlation between teachers' test construction competencies and the quality of multiple-choice test items. This suggested that there were potential inadequacies in the attention given by teachers to ensure item quality, leading to the presence of problematic items that may compromise overall assessment validity. The study concludes by emphasizing the critical role of test construction competencies in upholding assessment quality. It recommends collaborative efforts among teachers, adherence to standardized guidelines, solicitation of student feedback, and ongoing research and innovation to enhance multiple-choice test construction practices and elevate educational assessments.

Keywords: Multiple-choice test, test construction competencies, item quality, difficulty index, discrimination index.

Introduction

In education, the effectiveness of teaching and learning hinges upon the ability to gauge how much and how well students have absorbed the essential knowledge, skills, and abilities imparted during instruction. Assessment stands as a cornerstone of this process, furnishing educators, students, and stakeholders with invaluable insights. Yet, despite its

critical role, studies have consistently highlighted the challenges educators face in accurately assessing student learning outcomes. The repercussions of flawed assessments extend beyond the confines of the classroom, impacting schools, teachers, students, and society at large.

In Nigeria, while strides have been made to address the assessment of students' achievements within certain contexts, there remains a pressing need to delve deeper into the intricacies of classroom-based tests, particularly in underexplored domains. Imagine a community where individuals lack fundamental abilities to tackle problems or make informed decisions, where numeracy and literacy skills are scarce commodities. In such environments, fostering development becomes an uphill battle. Education emerges as the linchpin for progress, intertwined with human capital growth and individuals' overall well-being (Battle & Lewis, 2002).

Aligned with societal aspirations, educational objectives are crafted to nurture human resources, laying the groundwork for holistic development (Nitko, 2001). Learning outcomes are derived from these objectives, guiding the establishment of instructional goals aimed at fulfilling diverse educational aspirations (Kubiszyn & Borich, 2013; Nitko, 2001). However, a pivotal question arises: how can we ascertain whether students possess the requisite knowledge, skills, and abilities to thrive in society? Academic achievement and performance serve as barometers of individuals' societal utility, underlining the pivotal role of assessment in the educational sphere.

Assessment, encompassing the formal process of gathering, analyzing, and disseminating standardized information about students' competencies, transcends mere data reporting (Bunch, 2012). It serves as a compass, steering decisions regarding students, curricula, programs, and educational policies. Yet, while standardized tests offer structured assessment mechanisms, non-standardized examinations, often crafted by classroom teachers, provide a more flexible approach. These teacher-made tests play a vital role in pre-assessing, formatively evaluating, and summative gauging student performance, shaping crucial educational decisions.

Integral to effective assessment is teachers' competency in test construction, a facet often overlooked yet profoundly influential. Research underscores the relationship between teachers' test construction skills and the quality of assessment instruments, emphasizing the need for reliable and valid evaluations. However, studies in Nigerian secondary schools have revealed pervasive shortcomings in the reliability and validity of classroom assessments, attributing these deficiencies to teachers' inadequate test-building abilities (Ali, 1999).

Given the pivotal role of assessments in educational systems, the imperative for teachers to possess test construction competency and uphold assessment quality cannot be overstated. Thus, this study aims to unravel the correlation between teachers' abilities to craft high-quality multiple-choice items and their assessment efficacy. By scrutinizing the relationship between test construction competencies, test item quality, and teaching experience among senior high school teachers in Kosofe Local Government, this research endeavors to shed light on a critical yet understudied aspect of educational assessment.

Educational assessments serve as vital tools for gauging students' learning outcomes and instructional effectiveness. UNESCO's standards for educational assessment outline the competencies essential for teachers in facilitating effective teaching and learning processes across various grade levels (UNESCO, n.d.). These competencies encompass a spectrum of abilities crucial for crafting assessments that accurately measure student performance. Indeed, the quality of assessments hinges significantly on teachers' competence (Chau, 2003).

A well-constructed test aligns closely with instructional objectives and enables teachers to gauge students' mastery of specific content (Frey, 2007). Conversely, poorly designed test items can yield misleading results, obscuring true student performance and instructional efficacy (Frey, 2007). Koksals (2004) identified several inherent flaws in poorly designed tests, emphasizing the importance of addressing these issues to enhance the quality of classroom-based assessments.

Key factors contributing to poorly constructed tests include vague task specifications, unclear instructions, and inadequate consideration of students' proficiency levels (Koksals, 2004). Moreover, failure to align test items with instructional objectives and the inadvertent production of ambiguous items can compromise the validity and reliability of assessments (Koksals, 2004).

Teachers play a pivotal role in ensuring the quality of assessments by meticulously addressing these concerns during test construction. Clear task specifications, unambiguous instructions, and alignment with instructional objectives are paramount in fostering assessment validity and reliability. By scrutinizing these aspects and querying the clarity, correctness, and relevance of test items, teachers can mitigate the risk of producing flawed assessments (Koksals, 2004).

Agu et al. (2013) addressed concerns regarding the quality of classroom achievement tests in Nigeria, emphasizing the importance of teachers' test construction skills. Their study, "Measuring Teachers' Competencies in Constructing Classroom-Based Tests in Nigerian

Secondary Schools: Requirement for a Test Construction Skill Inventory," introduced a Test Construction Skill Inventory (TCSI) to assess secondary school teachers' proficiency in test creation. The TCSI demonstrated high reliability (coefficient of 0.73) and was recommended as a valuable tool for evaluating teachers' test construction abilities in Anambra State, Nigeria.

Similarly, Hamafyelto et al. (2015) investigated the relationship between commerce teachers' test construction proficiency and test quality in Borno State, Nigeria. Their study revealed deficiencies in test questions, particularly in content validity and cognitive domain representation. Most items focused on lower cognitive levels, indicating a lack of requisite test construction skills among senior secondary school teachers. They advocated for seminars and workshops to enhance teachers' competence in test development.

The effectiveness of educational assessment methods is crucial for comprehensive evaluation of students' knowledge and skills. However, despite the prevalent use of multiple-choice tests by teachers, there is a lack of assurance regarding adherence to established principles in test construction due to inadequate training. Research indicates that many secondary school teachers lack the necessary skills for constructing tests, which directly impacts the quality and validity of assessments. Consequently, low-quality test items hinder educators' ability to provide targeted support to students, potentially misclassifying their level of mastery and impeding educational opportunities. Therefore, this study aims to investigate the correlation between multiple-choice test construction competencies and the quality of test items among senior high school teachers in Kosofe Local Government.

Hypotheses

1. There is no statistically significant relationship between multiple-choice test construction competencies of teachers and the quality of multiple-choice test items in Kosofe Local Government.
2. There is no statistically significant relationship between test length and the number of good items produced by senior high school teachers in Kosofe Local Government.

Methodology

A descriptive-correlative research design was employed to test the hypothesis raised for this study. The design was found suitable for the study because it focuses on describing the relationships between variables without attempting to establish causation. The population for the study was 459 teachers. Simple random sampling technique was also used to select the five (5) senior secondary schools in Kosofe Local Government area of Lagos State. The total sample size of the research was made up of 25 teachers (respondents) in the five

(5) senior secondary schools selected in Kosofe Local Government.

A 20-item instrument titled Teachers' Multiple-Choice Test Construction Competence Questionnaire (TTCCQ-MC) was used to collect relevant data. The response mode was designed after a 4-point Likert scale of Strongly Agree as SA, Agree as A, Disagree as D and Strongly Disagree as SD. The validated using content validity and the reliability determined during the trial testing of the TTCCQ-MQ using test-retest reliability. A correlation coefficient of 0.79 was obtained and was adjudged suitable for the study. Copies of the instrument were administered on the respondents on the spot and collected and analysis. The collated data were analyzed using Pearson Product Moment Correlation Coefficient. The hypotheses were tested at 0.05 level of significance.

Results

Hypothesis 1: There is no statistically significant relationship between multiple-choice test construction competencies of teachers and the quality of multiple-choice test items in Kosofe Local Government.

Table 1: Correlation between Multiple-Choice Test Construction Competencies and Proportions of Good Items

		TC_Competence	Item_Quality
TC_Competence	Pearson Correlation	1	-.21
	Sig. (2-tailed)		.32
	N	25	25
Item_Quality	Pearson Correlation	-.21	1
	Sig. (2-tailed)	.32	
	N	25	25

The result in Table 1 shows that the Pearson correlation coefficient (r) was -0.21. This negative value indicates a weak negative correlation between multiple-choice test construction competencies and the number of good items. As the level of multiple-choice test construction competencies increases, the number of good items tends to decrease slightly, and vice versa. However, the correlation is close to zero, suggesting a weak relationship. Since this p-value (0.32) is greater than the significance level (0.05), the null hypothesis (H_0) was upheld. This means that there is no statistically significant relationship between multiple-choice test construction competencies and the number of good items at the 0.05 significance level.

Hypothesis 2: There is no statistically significant relationship between test length and the number of good items produced by senior high school teachers in Kosofe Local Government.

Table 2: Correlation between Test Length and the Number of Good Items

		Test length	No of Good Item
Test length	Pearson Correlation	1	.08
	Sig. (2-tailed)		.69
	N	25	25
No of Good Item	Pearson Correlation	.08	1
	Sig. (2-tailed)	.69	
	N	25	25

**Correlation is significant at the 0.01 level (2 -tailed)

Evidence from Table 2 above shows the Pearson correlation coefficient (r) between test length and the number of good items was 0.08. This coefficient quantifies the strength and direction of the linear relationship between the two variables. The p -value associated with the correlation coefficient was 0.69. This p -value is greater than the typical significance level of 0.05. Since the p -value (0.69) is greater than the chosen significance level (α), the null hypothesis (H_0) was accepted. This indicates that there is no statistically significant relationship between test length and the number of good items produced by senior high school teachers at the 0.05 significance level.

Based on the analysis, there is no statistically significant relationship between test length and the number of good items produced by senior high school teachers in Kosofe Local Government. The correlation coefficient (0.08) indicates a very weak positive correlation, which is not statistically significant.

Discussion of Findings

Hypothesis 1 examined the relationship between multiple-choice test construction competencies and the quality of the multiple-choice items they construct. The findings regarding research hypothesis 1 revealed that there was no significant relationship between the teachers' test construction competencies and the quality of the multiple-choice test items they produced. In other words, regardless of whether teachers scored high or low on test construction competencies, it did not significantly impact the quality of the multiple-choice items they constructed. This finding as presented contradicts Chau's (2015) perspective that a teacher's competence in test construction is directly related to ensuring the quality of the test he or she constructs. If the teachers' test construction competencies were not significantly related to the items' quality in multiple-choice tests, then it appears that though they possessed appreciable or high levels of competencies, they paid inadequate attention to ensure the quality of the multiple-choice test items. Therefore, the problem items influenced the items' quality in such a manner that it did not correlate to the teachers' self-report on their ability to construct multiple-choice test items. Kubiszyn

and Borich (2013) and Amedahe and Asamoah-Gyimah (2016) have pointed out that problem items, due to factors related to the test, such as difficult-to-read font size, unclear instructions, and ambiguous items with clues to correct answers, can compromise the validity of assessment results for making relevant educational decisions about students and classroom teachers. This highlights the importance of addressing item quality and avoiding the presence of problem items in test construction processes to ensure accurate and meaningful assessment outcomes.

Findings based on the examination of research hypothesis 2 showed that there was a significant positive relationship between test length and the number of good items produced by the research participants. The respective correlation coefficients for research hypothesis 2 ($r = .083$) mean that the probability of producing good items when increasing test length is low. The findings in relation to research hypotheses 2 suggest that while there may be a slight positive relationship between increasing test length and the number of good items in the multiple-choice test, it is crucial to pay close attention to the quality of the test items being produced. This means ensuring that the multiple-choice items are well-written and effectively assess the intended content. It is important to recognize that test length alone cannot guarantee the presence of a significant number of good items. The relationship between test length and the number of good items was not strong enough to rely solely on test length as a predictor of item quality. Other factors may come into play and influence the number of good items produced. According to Crocker and Algina (2008), increasing test length will improve assessment results only when the test constructor pays critical attention to producing well-written items free from technical flaws. Besides, the items should have appropriate difficulty and discrimination indices (Allen et al., 2008).

Conclusion

In this study, it became evident that despite teachers' self-reported competencies, the quality of multiple-choice test items did not meet expected standards. This discrepancy suggests a potential gap between perceived abilities and actual practices among senior high school subject teachers in Kosofe Local Government Area, Lagos State, Nigeria. The presence of problematic items highlights the importance of prioritizing the quality of test construction to ensure the validity of assessment outcomes, aligning with previous research emphasizing the significance of valid assessment tools.

Additionally, while a positive correlation is observed between test length and the number of good items produced, it is crucial to acknowledge that test length alone does not

guarantee assessment quality. Attention to item quality, including clarity, effectiveness in assessing content, and absence of technical flaws, is paramount. Teachers should also consider factors such as item difficulty and discrimination indices to enhance assessment accuracy and effectiveness.

Recommendations

Based on the findings, the following recommendations are suggested: Firstly, fostering collaboration and peer review among teachers can enhance the quality of multiple-choice items by leveraging collective expertise. Secondly, standardized guidelines for multiple-choice test construction should be developed and disseminated within the education district to ensure consistency and fairness. Thirdly, implementing a system for continuous monitoring and evaluation of assessment practices, with a focus on multiple-choice test quality, is essential for identifying and addressing persisting issues. Lastly, establishing feedback mechanisms where students can provide input on test item quality and clarity can help teachers make necessary improvements in their construction practices.

References

- Adane, L. O. (2013). Factors Affecting Low Academic Achievement of Pupils in Kemp Methodist Junior High School in Aburi, Eastern region. Aburi: University of Ghan.
- Adegoke, B. A. (2013). Comparison of Item Statistics of Physics Achievement Test using Classical Test and Item Response Theory Frameworks. *Journal of Education and Practice*, 4(22), 87-96.
- Adeola, O. A., & Fajonyomi, A. A. (1999). Adult Learners' Performance on Standardized and Non-Standardized Test as a Function of Sex and Location. 4(1), 32-36.
- Adodo, S. O. (2013). Effects of Two-tier Multiple Choice Diagnostic Assessment Items on Students' Learning Outcome in Basic Science Technology. *Academic Journal of Interdisciplinary Studies*, 2(2), 201-210.
- Agu, N. N., Onyekuba, C., & Anyichie, C. A. (2013). Measuring Teachers' Competencies in constructing Classroom-based Tests in Nigerian Secondary Schools: Need for a test construction skill inventory. *Educational Research and Reviews*, 8(8), 431-439.
- Alele-Williams, G. (2002). Measurement and evaluation in mathematics. The way forward, 1(1), 1-7.
- Ali, A. A. (1999). *Basic Research Skills in Education*. Enugu: Orient Printing and Publishing.

- Allen, M. J., & Yen, W. M. (2002). *Introduction to Measurement Theory*. Grove, IL: Waveland Press, Inc.
- Amedahe, F. K. (2002). *Fundamentals of educational research methods*. Cape Coast: University of Cape Coast Press.
- Amedahe, F. K., & Asamoah-Gyimah, K. (2014). *Introduction to Educational Research*. Cape Coast: University Printing Press.
- Amedahe, F. K., & Asamoah-Gyimah, K. (2016). *Introduction to Measurement and Evaluation* (7th ed.). Cape Coast: Hampton Press.
- Baker, J. O. (2003). *Testing in modern classrooms*. George Allen and Urwin Ltd.
- Battle, J., & Lewis, M. (2002). The increasing significance of class: The relative effects of race and socio-economic status on academic achievement. *Journal of poverty*, 6(2), 21-35.
- Bhattacharjee, A. (2012). *Social science research: Principles, methods, and practices* (2nd ed.). Retrieved from <http://scholarcommons.usf.edu/oa-tesxtbook/3>.
- Bunch, M. B. (2012). *Aligning curriculum, instruction, and assessment*. Retrieved from <http://measurementinc.com/sites/default/files/2017-08/Aligning%20Curriculum%2C%20Assessment%2C%20and%20Instruction%20%28M.%20Bunch%2C%202012%29.pdf>.
- Callahan, C. M. (2006). *Assessment in the classroom: The Key to Good Instruction* (The Practical Strategies Series in Gifted Education). Waco, TX: Prufrock Press, Inc.
- Chau, T. N. (2003). *The Quality of Primary Schools in different Development Contexts*. Paris: UNESCO Publishing International Institute for Educational Planning.
- Chloi, K., Nam, J. H., & Lee, H. (2001). The Effects of Formative Assessment with Detailed Feedback on Students' Science Learning Achievement and Attitude Regarding Formative Assessment. *Science Education International*, 12(2), 28-34.
- Clark, P. V., & Creswell, J. W. (2015). *Understanding research: A consumer's guide* (2nd ed.). New York: Pearson.
- Cohen, L., & Swerdik, M. J. (2010). *Psychological Testing and Assessment: An introduction to tests and measurement* (7th ed.). New York: McGraw-Hill.
- Cohen, L., Manion, L., & Morriison, K. (2000). *Research Methods in Education* (5th ed.). New York: Routledge.
- Cook, D. A., & Beckham, T. J. (2006). Current concepts in validity and reliability for (Vol. 119). *The American Journal of Medicine*.

- Crocker, L., & Algina, J. (2008). *Introduction to classical & modern test theory*. Mason: Cengage Learning.
- D'Agostino, J. V. (2007). *Quantitative Research, Evaluation and Management Section*. Columbus: Ohio State University.
- Domino, G., & Domino, M. L. (2006). *Psychological testing: An introduction* (2nd ed.). Cambridge: Cambridge University Press.
- Dosumu, C. T. (2002). *Issues in teacher-made tests*. Ibadan: Olatunji and Sons Publishers.
- Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of Educational Measurement* (5th ed.). New Jersey: Prentice-Hall.
- Education Up Close. (2005). *Writing effective tests: A guide for teachers*. Retrieved from <http://www.glencoe.com/sec/teachingtoday/educationup.phtml/print/40>.
- Erguven, M. (2014). Two Approaches to Psychometric Process: Classical Test Theory and Item Response Theory. *Journal of Education*, 2(2), 23-30.
- Esemonu, N. P. (2002). Assessment of non-cognitive behaviours in secondary schools for national development. *Education for National Development and Integration*, 88-93.
- Frey, B. B. (2017, October 19). Introduction to quality test construction. Retrieved from Special Connection: <http://www.specislconnection.ku.edu/cgi-bin/cgiwrap/speccom/mainphp?cat=assessment§ion=main§ion=qualitytest/main>.
- Furr, R. M., & Bacharach, V. R. (2014). *Psychometrics: An Introduction*. London: Sage.
- Gareis, R. C., & Grant, W. L. (2015). *Teacher-made assessments: How to Connect Curriculum, Instruction, and Student Learning* (2nd ed.). New York: Routledge.
- Gattullo, F. (2000). Formative Assessment in Primary (Elementary) ELT Classes: An Italian case study. *Language Testing*, 17(2), 278-288.
- Gronlund, N. E., & Linn, R. L. (1990). *Measurement and Evaluation in teaching* (6th ed.). New York: McMillan.
- Guskey, T. R. (2003). How Classroom Assessments Improve Learning. *Educational Leadership*, 60(5), 6-11.
- Guskey, T. R., & Jung, L. A. (2013). *Answer to Essential Questions about Standards, Assessments, Grading, & Reporting*. Thousand Oaks, CA: Corwin.

- Haeley, J. F. (2012). *Statistics: A Tool for Social Research* (9th ed.). Belmont: Wadsworth, Cengage Learning.
- Hamafyelto, R. S., Hamman-Tukur, A., & Hamafyelto, S. S. (2015). Assessing Teacher Competence in Test Construction and Content Validity of Teacher-made Examination Questions in Commerce in Borno State, Nigeria. *Journal of Education*, 5(5), 123-128.
- Hambleton, R. K., & Jones, R. W. (2003). Comparison of Classical Test Theory and Item Response Theory and their Applications to Test Development. *Educational Measurement: Issues and Practice*, 12(3), 38-47.
- Hamman, T., & Kamis, A. B. (2010). Content Analysis of B.Sc Bio-chemistry Examination QUESTIONS Implication for Testing, teaching and development African. *Journal Research in Education.*, 1, 59-62.
- Haris, R. J. (2002). *What every parent needs to know about standardised tests: How to understand the tests and help kids score high!*. New York: McGraw-Hill.
- Joshua, R. B., & Christensen, L. B. (2005). *Fundamentals of Test and Measurement in Education*. Calabar: University of Calabar Press.
- Kaplan, R. M., & Saccuzo, D. P. (2007). *Psychological Testing: Principles, Applications and Issues*. Pacific Grove: Brooks Cole Pub. Company.
- Kern, D. E., Thomas, P. A., & Hughes, M. T. (2009). *Curriculum Development For Medical*. Baltimore: The Johns Hopkins University Press.
- Kinyua, K., & Okunaya, L. O. (2014). Validity and Reliability of Teacher-made Test: Case study of Year 11 Physics in Nyahururu District of Kenya. *African Educational Research Journal*, 2(2), 61-71.
- Koksal, D. (2004). Assessing Teachers' Testing skills in Elt and Enhancing their Professional Development through Distance Learning on the Net. *Turkish Onlines J. Distance Education*, 5(1).
- Kubiszyn, T., & Borich, G. D. (2013). *Educational Testing and Measurement: Classroom AApplication and Practice* (10th ed.). Hoboken, NJ: Wiley.
- Lagos State Government Mistry of Education. (2019). *Annual School Census Report*. Ikeja.

- Leighton, J. P., & Gokiert, R. J. (2005, October 17). The cognitive effect of test item features. Retrieved from University of Alberta: <http://www.education.ualberta.ca/educ/psych/crame/files/NCMEfinalLeighandGokiert.pdf>
- Leighton, J. P., & Gokiert, R. J. (2005, October 17). The cognitive effect of test item features. Retrieved from University of Alberta: <http://www.education.ualberta.ca/educ/psych/crame/files/NCMEfinalLeighandGokiert.pdf>
- 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.
- Mayers, A. (2013). *Introduction to Statistics and SPSS in Psychology*. Boston: Pearson Education, Inc.
- McBride, N. L. (2001). *An Item Response Theory Analysis of the Scales From The International Personality Item Pool and the Neo Personality Inventory-Revised*. Master of Sciences Thesis submitted to the Faculty of Virginia Polytechnic Institute and State University.
- McMillan, H. J. (2013). *Sage handbook of research on classroom assessment*. Thousands Oaks: Sage.
- Morrow, J. R., Jackson, A. W., Disch, J. G., & Mood, D. P. (2000). *Measurement and evaluation in human performance* (2nd ed.). Champaign IL: Human Kinetics.
- Nitko, A. J. (2001). *Educational Assessment of Students* (3rd ed.). New Jersey: Prentice-Hall.
- NTI (National Teachers' Institute). (2006). *Manual for re-training of primary school teachers*. In *School-based assessment*. Kaduna: National Teachers Institute.
- Oescher, J., & Kirby, P. C. (2001, August 8). *Assessing teacher-made test in secondary school math and science classroom*. Retrieved from <http://eric.ed.gov/ERICweb%20porlets/%20records/detailmini.js>
- Okoli, C. E. (2014). *Introduction to Educational and Psychological Measurement*. Shomolu: Behenu Press & Publishers.
- Ololube, N. P. (2008). Evaluation Competencies of Professional and Non-professional Teachers in Nigeria. *Studies in Educational Evaluation*, 34(1), 44-51.

- Osuala, E. O. (2005). *Introduction to Research Methodology* (3rd ed.). Onitsha: African-Fep Publishers Limited.
- Polit, D. F., & Beck, C. T. (2004). *Nursing Research: Principles and Methods*. Philadelphia, PA: Lippincott Williams & Wilkins.
- Quaigrain, A. K. (1992). *Teacher–competence in the use of Essay Tests: A Study of Secondary Schools in the Western Region of Ghana*. Cape Coast: University of Cape Coast Press.
- Quaigrain, K., & Arhin, A. K. (2017). Using Reliability and Item Analysis to Evaluate a Teacher-developed Test in Educational Measurement and Evaluation. *Cogent Education*, 4, 1-11.
- Rivera, J. E. (2007). *Test Item Construction and Validation: Developing a Statewide Assessment for Agricultural Science Education*. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download>
- Rudner, L. M., & Schafer, W. D. (2002). *What Teachers Need to Know about Assessment*. Washington, DC: National Education Association.
- Sasu, E. O. (2017). *Testing Practices of Junior High School Teachers in the Cape Coast Metropolis*. Cape Coast: University of Ghana Press.
- Shillingburg, W. (2016). *Understanding Validity and Reliability in Classroom, School-wide, or District-wide Assessments to be used in Teacher/Principal Evaluations*. Retrieved from <https://cms.azed.gov/home/GetDocumentFile?id=57f6d9b3aadebf0a04b2691a>.
- Teacher Assessment Practices and Perceptions: The use of Alternative Assessments within the Quebec Educational Reform. (n.d.). Retrieved from <http://www.archipel.uqam.ca/9490/1/D1368.pdf>
- The increasing significance of class. The Relative Effects of Race and Socioeconomic Status on Academic Achievement. (2002). *Journal of Poverty*, 6(2), 21-35.
- Tshabalala, T., Mapolisa, T., Gazimbe, P., & Ncube, A. (2015). Establishing the Effectiveness of Teacher-made Tests in Nkayi District Primary Schools. *Nova Journal of Humanities and Social Sciences*, 4(1), 1-6.
- UNESCO. (2009). *UNESCO Guidelines on Teachers' Competence Procedures*. 4..