

Adherence to Testing Principles: Exploring the Congruence between what Teachers Preach and Practice

***Aliu Nandzo**

ORCID: <https://orcid.org/0000-0001-5228-7208>

Department of Education Studies, Tumu College of Education, Ghana

Email: aliunandzo10@gmail.com

Eric Anane, PhD

ORCID: <https://orcid.org/0000-0001-8505-8903>

Institute of Education, University of Cape Coast, Ghana

Email: eanane@ucc.edu.gh

Andrews Cobbinah, PhD

ORCID: <https://orcid.org/0000-0002-5315-659X>

Department of Education and Psychology, University of Cape Coast, Ghana

Email: andrews.cobbinah@ucc.edu.gh

Ruth Keziah Annan-Brew, PhD

ORCID: <https://orcid.org/0000-0002-6434-510X>

Department of Education and Psychology, University of Cape Coast, Ghana

Email: ruth.Keziah@ucc.edu.gh

***Corresponding Author:** aliunandzo10@gmail.com

Copyright resides with the author(s) in terms of the Creative Commons Attribution CC BY-NC 4.0.

The users may copy, distribute, transmit and adapt the work, but must recognize the author(s) and the

East African Journal of Education and Social Sciences

Abstract: This study explored the congruence between teachers' self-reported classroom assessment practices and their actual practice in the Sissala East Municipality, Ghana, using the sequential explanatory mixed method design. The study focused on teachers at the basic school level, whereby 220 teachers participated in the study through some sampling procedures. Data was gathered through a questionnaire and teaching documents. Data analysis involved descriptive statistics and thematic approach. Based on the findings, the study concluded that although the investigated junior high school teachers possessed adequate understanding of test construction procedures, they did not implement some of the procedures in their classroom assessments. Therefore, several recommendations were made.

Keywords: Assessment practices; test construction; objective-test; multiple-choice; adherence; testing.

How to cite: Nandzo, A., Anane, E., Cobbinah, A. and Annan-Brew, R. K. (2024). Adherence to Testing Principles: Exploring the Congruence between what Teachers Preach and Practice. East African Journal of Education and Social Sciences 5(6), 1-7. Doi: <https://doi.org/10.46606/eajess2024v05i06.0409>.

Introduction

The use of traditional assessment tools (conventional testing methods, such as quizzes, exams and standardized tests, used by teachers to measure students' learning) continues to play a critical role within Ghana's curricular frameworks/educational system. According to Reynolds et al (2009), "assessment" is a broad term

encompassing a well-organized body of data essential for determining the quality of people, things and events. Apart from gathering data, assessment is viewed as a process by which test administrators establish students' knowledge, understanding and skills per a demarcated set of predetermined instructional objectives or study guidelines (Marriot & Lau, 2008). Opara et al.

(2015) further argued that assessment encapsulates teachers' actions to help students learn and determine students' improvement. Educational systems recognize the impact of assessment because it is a crucial aspect of educational practice, facilitating teaching and learning activities and mediating classroom interactions between teachers and students (Amua-Sekyi, 2016). Based on how teachers utilize outcomes of assessment, assessment is a strong tool that can be used to boost or limit the learning process. Therefore, the assessment process is usually seen as the educational enterprise's cornerstone of teaching and learning.

Depending on the purpose, two forms of assessment are often utilized by teachers to gather data on their learners. These are formative assessment and summative assessment. According to Hortigüela Alcalá et al. (2019), the foremost objective of formative assessment is to enhance the instructional process. The process helps teachers to better appreciate their job (to enhance their teaching approach and allows students to learn more (improve upon their weaknesses). Therefore, its primary goal is to inform teachers how to support student growth and learning rather than to grade or qualify the student (Martos-Garca et al., 2017). However, Summative assessment is an evaluative process conducted after an instructional process, such as at the end of the academic year. It assesses students' mastery of learning objectives to provide insight into their knowledge and areas of deficiency (Guskey, 2003). This form of assessment aims at evaluating the performance of students against established criteria aligned with specific learning standards.

In the Ghanaian school system, a test is one of the commonly used assessment procedures adopted by teachers for continuous evaluation of students' learning and grading purposes. This implies that testing plays a vital role when it comes to educational assessment in the country. Tamakloe et al. (1996) as cited in Quansah et al. (2019) defined a test as "a device or procedure for measuring a sample of an individual's behaviour in a specific learned activity or discipline" (p.2). Considering the critical role testing plays in the educational assessment process in Ghana, every teacher must study educational measurement and evaluations as part of their initial training to become a professional teachers. The essence of studying the course in educational assessment is to equip them with the

relevant skills and competencies to develop tests that will produce valid and reliable results.

Studies intimate that the main challenge with the use of classroom assessment organized by teachers concerns making valid and reliable interpretations of the results (American Educational Research Association et al., 1999; Mertler, 1999). Hence, instructors need to implement classroom assessments that would provide a realistic picture of students' performance, accurately reflecting what the assessment claims to measure and reliably providing replicable information that can be used for decision-making.

Various nations exert strenuous effort in preparing teachers to become competent assessment literates, capable of developing and using results of tests to make valid and reliable instructional and learning-related decisions. Notwithstanding, teachers are often found wanting when it comes to the actual implementation of testing principles (Frey & Schmitt, 2010; Ololube, 2008; Quansah & Amoako, 2019).

Stakeholders constantly criticize the way teachers implement classroom assessments. For example, studies have shown that teachers poorly develop items or mostly make use of previously used test items (Frey & Schmitt, 2010; Ololube, 2008; Onyechere, 2000). Onyechere, (2000) for example revealed that teachers failed to follow the principles of constructing questions that would meet the intended purpose. Similarly, in Ghana, teachers failed to follow the principles guiding testing principles (Anhwere, 2009; Oduro-Okyireh, 2008; Osman et al., 2021; Quansah & Amoaka, 2018; Quansah et al., 2019). It is worth noting that although the issue of teachers' assessment practices has been widely researched and documented, the studies mostly only self-reported measures including the use of questionnaires (Anhwere, 2009; Oduro-Okyireh, 2008; Osman et al., 2021) and questionnaire together with interviews in gathering data. The use of these self-reported measures could offer respondents the opportunity to provide socially desirable responses. This means that teachers could easily provide responses that are in sharp variance with their actual practice.

Other studies that sought to examine the gap between theory and practice (e.g. Quansah et al., 2019), resorted to the use of document analysis only. While the use of this data collection procedure could provide an avenue for understanding

teachers' actual practice of testing, it did not offer the opportunity to relate teachers' claims of their testing practices to their actual practice in the same context. That is to say, the use of document analysis only will not provide a comprehensive understanding of the extent to which teachers' claim of their testing practices aligns with their actual practice. Hence, it is essential to carry out an explorative study to establish the congruence between teacher's claim of their competences in test construction and their actual practice of test construction in a single study. Therefore, this study sought to fill the gap.

Methodology

This section focused on the research methods employed. Specifically, it includes the research design, the study population and sampling, statistical treatment of data, data collection instruments and ethical considerations.

Design

This study employed the sequential explanatory mixed method design. According to Creswell and Plano Clark (2018), this type of mixed method is adopted when the objective of the study is to use both quantitative and data collection procedures to complement the results from the quantitative procedures with the results of the qualitative data. This method demands that the researcher gathers data for the quantitative phase first before carrying out the qualitative data collection. Creswell (2011) stressed that the essence of the qualitative phase is to aid in expatiating the results obtained from the quantitative phase obtained earlier.

Population and Sampling

The study focused on teachers at the basic school level in the Sissala East Municipality of the Upper West Region in Ghana. The municipality had 78 Junior High Schools clustered into 9 educational circuits with 483 teachers. However, five circuits, comprising 61 schools were selected from the nine circuits using simple random sampling. Out of this number, 55 schools, comprising 220 teachers, were considered sufficient for the study, based on a formula by Krejcie and Morgan (1970).

Instruments

Quantitative data was gathered by adapting a questionnaire developed by Oduro-Okyireh (2008), with a composite reliability of .70. The questionnaire measured three dimensions of testing; test construction, test administration and test scoring.

However, this paper focused on only the test construction dimension with nineteen (19) items. For qualitative data, samples of teaching documents were assessed. These included syllabi for the four core subjects (English Language, Mathematics, Integrated Science and Social Studies), lesson plans and sample test papers.

Statistical Treatment of Data

Data analysis was done in two phases: quantitative data analysis and qualitative data analysis. Quantitative data, gathered using the questionnaire was analysed using descriptive statistics, such as mean scores and standard deviations. Qualitative data, gathered based on sample test items was analysed using the thematic approach in the sense that similar themes were assembled together.

Ethical Considerations

In conducting this study, several ethical issues were addressed. First, before administering the questionnaire, the researchers explained the objective of the study to the respondents. It was indicated to them that participation was voluntary. Hence, only those who offered their consent were engaged in the study. To ensure the anonymity of the participants, the respondents were told to not write their names on the questionnaire. Besides, the respondents were assured that data emanating from them would strictly be kept confidential.

Results and Discussion

The results of the study were presented based on research questions.

Research Question 1: What do Junior High School Teachers in the Sissala East Municipality believe about specific assessment practices?

To answer this research question, data for the quantitative phase was analyzed. The research question sought to establish the belief of teachers regarding various aspects of test construction. The principles of test construction were measured using 19 statements structured on a five-point Likert-type scale. The scale had the following response categories: never = 1; rarely = 2; sometimes = 3; very often = 4 and always = 5. For proper interpretation, an average of 3.0 was used as the rule of thumb. This means an average score less than 3.0 implies that the teachers' beliefs were negative just as the average score of 3.0 or higher indicates positive beliefs about test construction.

It can be seen from the results in Table 1 that the overall mean of 3.89 with the standard deviation of 1.03 is far above the threshold of 3.0. This is an indication that generally, respondents possessed positive attitude toward test construction procedures. All except one item recorded the mean score of greater than the threshold of 3.0. For

example, a statement like “*it is important to relate the instructional objectives of the subject matter to the test*” recorded the mean score of 4.73 and the standard deviation of .58. Another statement about “*It is important to select test format suitable for testing stated objectives*” had the mean score of 4.47 and a standard deviation of .79.

Table 1– Results of Responses on the Quantitative Phase

SN	Statement in the Questionnaire	Mean	SD
1	I relate the instructional objectives of the subject matter to the test.	4.73	.58
2	It is important to inform students in advance before the test.	4.44	.84
3	It is important to inform students about the uses of the test scores	3.62	1.01
4	It is important to select a test format suitable for testing the stated objectives.	4.47	.79
5	It is important to select a test format convenient for scoring the test.	4.22	.99
6	It is important to prepare and use the table of specifications.	3.48	1.14
7	It is important to inform students what chapters the test will cover.	4.08	1.13
8	It is important to prepare more questions than needed.	2.90	1.29
9	It is important to write questions in advance (at least two weeks) for review.	3.47	1.26
10	It is important to write questions a day before the test to avoid leakage.	3.40	1.38
11	It is important to review questions (at least one week) after they have been written.	3.48	1.29
12	It is important to ensure questions are challenging to high achievers.	3.23	1.35
13	It is important to prepare a marking scheme immediately after the test is written.	3.82	1.34
14	It is important to write clear information for entire and in each section of the test	4.58	.73
15	It is important to ensure the questions are simple and clear.	4.17	.94
16	It is important to ensure the test is a representative sample of the content covered.	4.67	.66
17	It is important to ensure that students have enough time to complete the test.	4.20	1.01
18	It is important to ensure that the test assesses the knowledge or skills taught	4.09	1.04
19	It is important to prepare students adequately for a test	4.51	.80
Overall Mean Score		3.98	1.03

These results as seen in Table 1 are consistent with previous studies by Ololube (2008) and Agu et al. (2012) in Nigeria, Benzehaf (2017) in Morocco and Owusu-Mensah (2019) in Ghana. For instance, in Nigeria, Agu et al. (2012) revealed that teachers considered listed principles while constructing test items. In Ghana, Owusu-Mensah (2019) revealed that teachers considered the purpose of the test before developing test items and preparing marking schemes. They also matched instructional objectives with tests. This implies that the results of the teachers' assessments could be used to make sound decisions about students learning. Effective assessment procedures offer a dependable estimate of student achievement and connect with instructional objectives, allowing teachers to make sound decisions on methods for instruction (Brookhart, 2011). In addition, adopting rigorous assessment procedures enhances students' motivation by providing precise and attainable goals (Stiggins, 2014). Additionally, effective assessment processes foster equity by ensuring that all students are assessed according to standard and objective criteria (Popham, 2018).

Research Question 2: What are the actual assessment practices of Junior High School Teachers in the Sisasala East Municipality?

The objective of this this research question was to ascertain whether or not the respondents, aside from their self-reported adherence to the basic principles of test construction, actually put their knowledge into practice. To achieve this objective, document analysis was done based on samples of test items (teacher-made tests) constructed and used in four subject areas (English Language, Mathematics, Integrated Science and Social Studies). The analysis of the sample papers was done using six principles of test construction as parameters (indicators). These included item1, item6, item14, item15, item17 and item 18 (see Table 1).

Table of Specification

Based on a critical examination of sample papers, it was observed that in 29 out of 32 test items that were examined (90.6%), the teachers did not develop or use a table of specifications for the tests. Therefore, such tests did not cover all the relevant

domains of learning in the four subject areas. This finding implies that decisions made about the achievement of learners based on the results of such tests would be weak in terms of validity and reliability. There is a need to develop and use the table of specifications because in the view of Cantwell and Scevak (2010), the paramount consideration when planning classroom assessments is to ensure that the course material or examination materials are balanced appropriately to cover all the important skills, knowledge and understanding. Assessing some material more than once at the expense of other important content lowers the validity of the assessment. Thus, the blueprint serves as a basis for setting the number of assessment tasks and for ensuring that the assessment will have the desired emphasis and balance (Nitko & Brookhart, 2014). Etsey (2012) stressed that the table of specifications ensures that justice is done to all the topics covered in the course, and helps the teacher to determine the content validity of the test.

Higher Level of Thinking

Further scrutiny of the papers revealed that although the items were distributed across the various cognitive levels spelled out in the revised Bloom's taxonomy, most of the items were merely measuring lower-level thinking processes (thinking processes that focusing on the recall of basic facts). For example, an item in Integrated Science was crafted as 'The width of a rectangle is 4 centimetres and the length is 3 centimetres. What is the area?' This type of item only requires the learner to recall a simple algorithm or mathematical operation involving multiplication, where even a learner with superficial knowledge could get it correct. This finding confirms Amua-Sekyi's (2016) observation that teachers do not nurture evaluative thinking skills in their students in their assessment and teaching practices. Designing assessments that only focus on lower-level thinking processes implies that learners might find it difficult to apply the knowledge and understanding in real situations. In line with this trend, Nitko (2001) advised that assessment should help learners to think critically and to apply the material to their daily life problems, rather than merely recalling facts or describing concepts. A further examination of the content of test items concerning subject matters outlined on the syllabi and in the lesson plans across the various subject areas showed that the items measured important concepts. This is commendable

because according to Cantwell and Scevak (2010), assessment should align well with valuable content being taught, but not focus on irrelevant or trivial concepts.

Clarity of Instructions

It was further revealed that teachers failed to write very clear questions. Several of the papers examined across the four subject areas revealed some errors in the items. For instance, in Social Studies, one of the items was crafted as "Lanes drawn on the maps to correct places of equal height above sea level is known as" In this item, two keywords, Lines and connect, were wrongly worded as Lanes and correct, respectively. Similarly, some of the items on the sample test items were double-barreled questions or ambiguous. These glaring typographical errors could potentially affect the learning target the item was meant to assess. These results confirm several previous investigations into teachers' assessment practices and found that teachers have poor test construction skills (Ololube, 2008; Onyechere, 2000; Frey & Schmitt, 2010). It has been observed that such teachers either construct poor items. Similarly, in Ghana, the classroom assessment practices of teachers have been studied and the results showed that Ghanaian teachers have limited skills in constructing objective and essay-type tests (Oduro-Okyireh, 2008; Anhwere, 2009; Quansah & Amoako, 2018; Quansah et al., 2019). Teachers need to ensure clarity and correctness when writing items to enhance the validity of the results. Literature has pointed out that questions, tasks or items that are not well written, are ambiguous or lack clarity, often yield assessments that are low in validity (Amedahe, 2014; Cantwell & Scevak, 2010).

Reflection of Knowledge and Skills Taught

Finally, papers that constituted the instrument were assessed to determine their appropriateness in measuring knowledge or skills taught in class. To achieve this objective, the items were examined against topics taught in their lesson plans across the various subject areas. It was revealed that the use of the test was appropriate for assessing the knowledge and skills taught in the lesson plan. This finding implies that the teachers were aware that different assessment tools were appropriate for assessing different learning targets. Through the alignment of assessment processes with the skills that are being taught, it is ensured that students are evaluated based on what they are expected to learn and apply, which promotes equity and relevance in

the educational system. As stated by Wiggins and McTighe (2005), assessments that are directly connected to the content of the instruction provide teachers with the ability to properly measure the level of comprehension that their students have and to direct adjustments in the teaching process. Brookhart (2011) highlights the fact that evaluations that are connected with the subject that is being taught not only increase students' performance but also assist in the process of defining goals that are both reasonable and attainable for learners.

Conclusions and Recommendations

Based on the evidence garnered from the study, it can be concluded that the although junior high school teachers in the Sissala East Municipality possessed adequate understanding of test construction procedures, they did not implement some of the procedures in their classroom assessments. Therefore, it is recommended that stakeholders in education, especially, the Management of Ghana Education Service in the Sissala East Municipality should enhance teachers' test construction skills through workshops and seminars and periodically evaluate tests developed and used for classroom assessment to ensure that they are aligned with test construction procedures.

References

Agu, N., Onyekuba, C. and Anyichie, A. C. (2012). Measuring teachers' competencies in constructing classroom-based tests in Nigerian secondary schools: Need for a test construction skill inventory. *Academic Journal*, 8(8), 431-439.

Amedahe, F. K. (2014). Test construction practices in secondary schools in the Central Region of Ghana. *The Oguua Educator*, 2, 52-63.

American Educational Research Association, American Psychological Association and National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association (AERA).

Amua-Sekyi, E. T. (2016). Assessment, Student Learning, and Classroom Practice: A Review. *Journal of Education and Practice*, 7(21), 1-6.

Anhwere, Y. M. (2009). Assessment practices of teacher training colleges tutors in Ghana. Master of Philosophy Unpublished Thesis. University of Cape Coast, Cape Coast, Ghana.

Benzehaf, B. (2017). Exploring teachers' assessment practices and skills. *International Journal of Assessment Tools in Education*, 4(1), 1-18.

Brookhart, S. M. (2011). *Formative assessment strategies for every classroom: An ASCD action tool*. ASCD.

Cantwell, R. H. and Scevak, J. J. (Eds.). (2010). *An academic life: A handbook for new academics*. Aust Council for Ed Research.

Creswell, J. (2011). *Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. New York: Pearson.

Creswell, J. W. and Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Thousand Oaks: Sage Publications.

Etsey, Y. K. A. (2012). *Assessment in Education*. Cape Coast: UCC Printing Press.

Frey, B. B. and Schmitt, V. L. (2010). Teachers' classroom assessment practices. *Middle Grades Research Journal*, 5(3).

Guskey, T. R. (2003). How classroom assessments improve learning. On *Formative Assessment: Readings from Educational Leadership* (EL Essentials).

Hortigüela Alcalá, D., Palacios Picos, A., & López Pastor, V. (2019). The impact of formative and shared or co-assessment on the acquisition of transversal competencies in higher education. *Assessment & Evaluation in Higher Education*, 44(6), 933-945.

Krejcie, R. V. and Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, 30 (3), 607-610.

Marriot, P., & Lau, A. (2008). The use of online summative assessment in undergraduate financial accounting course. *Journal of Accounting Education*, 26, 773-90.

Mertler, C. A. (1999). Teachers' (mis)conceptions of classroom test validity and reliability. Paper presented at the annual meeting of the Mid-Western Educational Research Association, Chicago, IL.

Nitko, A. J. (2001). *Educational assessment of students* (3rd. ed.). Upper Saddle River, NJ: Merrill.

- Nitko, A. J. and Brookhart, S. M. (2014). Educational assessment of students (6th ed.). Pearson.
- Oduro-Okyireh, G. (2008). Testing practices of senior secondary school teachers in the Ashanti region of Ghana. Doctoral Unpublished Dissertation, University of Cape Coast, Cape Coast, Ghana.
- Ololube, N. P. (2008). Evaluation competencies of professional and non-professional teachers in Nigeria. *Studies in Educational Evaluation*, 34(1), 44-51.
- Onyechere, I. (2000). New face of examination malpractice among Nigerian youths. *The Guardian Newspaper* July, 16.
- Opara, I. M., Onyekuru, B. U., & Njoku, J. U. (2015). Predictive Power of School-Based Assessment Scores on Students' Achievement in Junior Secondary Certificate Examination (JSCE) in English and Mathematics. *Journal of Education and Practice*, 6(9), 112-116.
- Osman, S., Bordoh, A. and Eshun, I. (2021). Basic school teachers' conceptions of assessment in the Sissala East Municipality. *International Journal of Research and Innovation in Social Science*, 5(3), 311-324.
- Owusu-Mensah, J. (2019). Teacher characteristics as correlates of classroom assessment practices at Atwima Nwanbiagya South District. Master of Philosophy Unpublished Thesis, University of Cape Coast, Cape Coast, Ghana.
- Popham, W. J. (2018). Classroom assessment: What teachers need to know (8th ed.). Pearson.
- Quansah, F. and Amoako, I. (2018). Attitude of senior high school (SHS) teachers towards test construction: Developing and validating a standardized instrument. *Research on Humanities and Social Sciences*, 8(1), 25-30.
- Quansah, F., Amoako, I. and Ankomah, F. (2019). Teachers' test construction skills in Senior High Schools in Ghana: Document analysis. *International Journal of Assessment Tools in Education*, 6(1), 1-8.]
- Reynolds, C., R., Livingston, R. B., & Willson, V. (2009). *Measurement and assessment in education* (2nd. ed.). Ohio: Pearson.
- Stiggins, R. J. (2014). *Assessment literacy for teachers: Fostering learning-centered assessments*. Pearson.
- Wiggins, G. and McTighe, J. (2005). *Understanding by design* (2e). Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).