

# **TESTING IN CHRISTIAN RELIGIOUS STUDIES IN GHANAIAN SENIOR HIGH SCHOOLS**

**MARTIN YAW ADJEI<sup>1</sup>, ANTHONY MENSAH<sup>2\*</sup>, GABRIEL KUMAH<sup>3</sup>, SAMUEL  
AYETEY CHAITEY<sup>3</sup>,**

<sup>1</sup>St. Joseph's College of Education, Post Office Box 15, Bechem-Ahafo, Ghana

<sup>2,3,3</sup>Abetifi Presbyterian College of Education, P. O. Box 19, Abetifi-Kwahu, Eastern  
Region, Ghana

\*Corresponding author: [tonymenss22@gmail.com](mailto:tonymenss22@gmail.com)

## **Abstract**

The study sought to examine testing in Christian Religious Studies (CRS) in Ghanaian Senior High Schools. The cross-sectional descriptive survey research design was adopted for the study. Using the census method, 96 CRS teachers in the Senior High schools in the Bono, Bono East and Ahafo regions of Ghana were involved in the study. Data gathered from the use of the questionnaire was analysed using frequencies, percentages, means and standard deviations. It was found out that, to a high extent, teachers adhered to the principles of test construction, administration and scoring. The study recommended that, teachers should write test items at least two weeks before time for test administration so that they can have enough time to read over and correct wrongly constructed items before test is administered. Also, teachers should prepare classroom a day before test is taken and in the case of lack of infrastructure, then classroom should be made ready and prepared hours before test is taken. Teachers should also use the "DO NOT DISTURB SIGN" at the entrance of classroom during test administration. Again, teachers should shuffle scripts before scoring, and should score a particular item on all papers at a sitting. Again, the teachers should periodically rescore previously scored items.

**Keywords: Assessment, Testing, Christian Religious Studies,**

## **Introduction**

A test is an essential tool that helps to quantify constructs which helps one to make a value judgment about the degree to which such constructs might probably exist in an individual. Nobody can study in an entire educational system without being exposed to a wide range of educational testing as a means of assessment. This is because constantly in an educational system, decisions have to be made about students, curricula and programmes, and educational policies. According to Nitko [1], decisions about students include managing classroom instruction, placing students into different types of programmes, assigning them to appropriate

categories, guiding and counseling them, selecting them for educational opportunities, and credentialing and certifying their competence. Decisions about curricula and programmes include decisions about their effectiveness and about ways to improve them. It is worth noting, nevertheless, that educational assessments, the most prevalent of which are tests in Ghana's educational system, provide some of the information needed to make these kinds of decisions.

According to the National Council on Measurement in Education's (NCME) standard for Educational and Psychological testing, "a test is a device or procedure in which a sample of an examinee's behavior in a specific domain is gathered and then evaluated and scored using a standardized approach" [2]. However, it must be noted that the psychological attributes of an individual cannot be measured directly as can height or weight. It will never be possible to prove the existence of such a psychological construct. Only observation of an individual's behaviour can the degree to which any trait characterises him or her. It becomes more prudent if one can quantitatively relate the subjective judgments of individuals about the estimated amount of construct or trait that exists in a person by establishing standards for such measurement. To gather information about students, a variety of assessment procedures might be used. These include formal and informal observation of students, paper-and-pencil tests, a student's performance on homework, laboratory work, projects, and oral questioning and analysis of students' records.

Standardised achievement, aptitude, and intelligence tests, such as those used in developed countries such as the United States of America (USA), Canada, and the United Kingdom, are almost non-existent in Ghana's educational system. The tests that are conducted by the West African Examinations Council (WAEC) at the terminal points of the educational system cannot be said to be standardised since they not meet all of the requirements for standardised achievement tests. Examples of the WAEC conducted tests are the Basic Schools Certificate Examination (BECE) and the West African Senior School Certificate Examination (WASSCE). According to Linn and Gronlund [3], the characteristics of a carefully constructed standardised achievement test include the following: The test items are of high technical quality. They have been developed by educational and test specialists, tried out experimentally (pretested), and selected based on difficulty, discriminating power, and relationship to a clearly defined and rigid set of specifications; Directions for administering and scoring are so precisely stated that the procedures are standard for different users of the test; Norms based on national samples of students in the grades where the test is intended for use are provided as aids in interpreting the scores; as well as Equivalent and comparable forms of the tests are usually provided as well as information concerning the degree to which the tests are comparable.

Given the extent of prevalence of classroom achievement tests in Ghanaian schools and the variety of uses to which the results from these tests are put, there is the need to research the achievement assessment practices of CRS teachers. However, from personal experiences as teachers in the study area, our interaction with some CRS teachers in the Senior High schools in the area gave a very worrying situation. This is because, their assessment practices were with a lot of flaws which included double-barreled questions which were ambiguous, poor construction of test items with poor wording/grammatical errors, poor administration, poor scoring, etc., and these raises a lot of questions. Did these teachers receive enough training on assessment practices? or does training in assessment in education contribute to competence in assessment?

Could it also be that CRS teachers who received instruction in assessment in education are refusing to put the knowledge they acquired into practice? Based on these challenges, coupled with the literature gap on testing so far as CRS is concerned, the researcher had the motivation to conduct this study and find out testing in Christian Religious Studies in Ghanaian Senior High Schools. The research attempts to find solutions to following research questions:

1. “How do CRS teachers adhere to the principles of test construction in Senior High schools?”
2. How do CRS teachers adhere to the principles of test administration in Senior High schools?”
3. “How do CRS teachers adhere to the principles of test scoring in Senior High schools?”

The study involved CRS teachers from the Senior High schools in the Bono, Bono East, and Ahafo regions of Ghana. The results of this study would help stakeholders such as the National Council for Curriculum and Assessment (NaCCA) and the Ministry of Education (MoE) to determine the state of affairs with respect to testing in the Ghanaian educational system. Also, the findings of this study reveal the challenges that CRS teachers face in test construction, test administration, and test scoring in Senior High schools so that such needs can be addressed. Positive suggestions would be offered as a means of addressing these flaws in the assessment. Finally, the study adds to the literature on the subject and serves as a guide to future researchers on testing in Ghanaian schools. The subsequent paragraphs consider the Methodology, Results and Discussions, Conclusions, and Recommendations.

## **Methodology**

### **Research Design**

The design used for the study was the cross-sectional descriptive survey. According to Johnson [4], a cross-sectional survey collects data to make inferences about a population of interest at one specific point in time. Levin [5] explains that a cross-sectional study is employed for studies covering relatively short periods. The method was used for this study because it offers the researcher the opportunity to assess, observe and describe the assessment practices of CRS teachers in Senior High schools in the Bono, Bono East, and Ahafo regions at a specific point in time.

### **Population**

Mugenda and Mugenda [6] defined a population as an entire group of individuals, events, or objects with some observable characteristics. The population for the study comprised all CRS teachers in the Senior High schools in the Bono, Bono East, and Ahafo Regions. There was a target population of 124 CRS teachers in the Bono, Bono East, and Ahafo regions of Ghana [7]. Out of this, there was an accessible population of 96 CRS teachers in the 82 Senior High schools in the Bono, Bono East, and Ahafo regions. CRS teachers were used because they could provide information about how CRS teachers adhere to the principles of test construction, administration, and scoring in Senior High schools in the Bono, Bono East, and Ahafo regions.

**Sample and Sampling Procedure**

The quality of any research not only stands or falls by the appropriateness of methodology and instrumentation but also by the suitability of the sampling strategy that is adopted [8]. All 96 accessible CRS teachers in the 82 Senior High schools in the Bono, Bono East, and Ahafo regions were involved in the study. This constituted the sample size of CRS teachers for the study. The census study was employed to involve all the CRS teachers in the Bono, Bono East, and Ahafo regions of Ghana due to their limited number.

**Research Instruments**

A questionnaire was the main instrument for data collection for this study. A self-designed questionnaire (for CRS teachers) was employed in this study. Reasons for the choice of the instrument are that the questionnaire affords greater assurance of confidentiality and anonymity to respondents [9]. Again, a questionnaire is described as a structured instrument for gathering data from a potentially large number of respondents, within a shorter possible time when especially the population is easily accessible to make it uneconomical for reasons of time or funds to interview every subject in the study [10]; [11]. The questionnaire is also appropriate when the respondents are literates and since the CRS teachers could read and write, the study adopted the questionnaire.

In a reference to the above, this study employed a structured questionnaire to gather data from the respondents. To ensure a quick and easy response to the items, the questionnaire was the closed-ended type and was drafted on a five-point Likert scale (1=Strongly Disagree (SD); 2= Disagree (D); 3= Uncertain (U); 4= Agree (A) and 5= Strongly Agree (SA). This made it possible for the analysis of data using both descriptive and inferential statistics.

The questionnaire consisted of 43 items in four sections (A, B, C, and D). The 'A' part entailed five (5) items geared towards obtaining information about the demographic characteristics of the respondents. Section B consisted of sixteen (16) items that aimed at obtaining information on how CRS teachers adhere to the principles of test construction. Section C had eleven (11) items that looked at how CRS teachers adhere to the principles of test administration. Finally, Section D was made up of eleven (11) items that considered how CRS teachers adhere to the principles of test scoring.

**Validity and Reliability of the Instrument**

Validity and reliability test was performed on the research instrument. Validity refers to the degree to which evidence and theory support the interpretation of test scores entailed by the use of tests. The validity of an instrument is the extent to which it does measure what it is supposed to measure. Face and content validity tests were executed on the instrument. In ensuring face validity, the instrument was given to experts to review and find out whether the items in the questionnaire measure the characteristic or trait of interest. In terms of content validity, the experts who were also familiar with the construct of interest of the research reviewed the questionnaire items for readability, clarity, and comprehensiveness and finally concluded on the items which should be part of the final questionnaire.

A pilot test of the instruments was conducted in three Senior High schools in the Ashanti Region to check the reliability of the instrument. Apart from proximity reasons, this area was chosen for the pilot testing because the curriculum and CRS syllabus implemented in Senior High schools in the Ashanti Region are similar to that of Senior High schools in the Bono, Bono East, and Ahafo regions, and hence CRS teachers may adopt similar practices in test

construction, test administration and test scoring in the Ashanti Region as it pertains in the study area. The data gathered was analysed and the Cronbach’s alpha was established for each of the items that fall under the three research questions formulated to guide the study. The Cronbach's alpha of .72 was obtained for the CRS teachers’ questionnaire. Fraenkel and Wallen [12] posited that “A good rule of thumb for research purposes is that reliability should be at least .70 and ideally higher” [12]. With this, the instruments could be said to be of good quality capable of collecting useful data for the study.

**Data Analysis**

This study sought to assess testing in Christian Religious Studies in Ghanaian Senior High Schools. To answer the research questions formulated to guide the study, the type of statistics that was employed in the analysis of the data was descriptive statistics. Specifically, the data from the questionnaire was used to analyse research questions 1, 2 & 3 through the computation of frequencies, percentages, as well as mean of means distributions. This was done using Statistical Product and Service Solutions.

**Results and Discussion**

Data was analysed and presented systematically beginning with the background information of the respondents, followed by the research questions that guided the study. Table 1 show the characteristics of CRS teachers from the Senior High schools in the Bono, Bono East, and Ahafo regions of Ghana, who served as respondents for the study.

**Table 1: Characteristics of CRS Teachers (n=96)**

Variable	Subscale	No.	%
Gender	Male	69	71.9
	Female	27	28.1
Age	20-29 years	6	6.2
	30-39 years	45	46.9
	40-49 years	21	21.9
	50-59 years	24	25.0
Years in Teaching Service	Under 5 years	9	9.4
	6-10 years	54	56.2
	Above 10 years	33	34.4
Academic Qualification	Bachelor's Degree	75	78.1
	Masters of Arts Master of Philosophy	15	15.6
Major Subject(s) Areas	Religious Studies	6	6.2
	Other subject areas	52	54.2
Professional Teaching Qualification	Teacher's Cert “A”	44	45.8
	PGDE	3	3.1
	Bachelor of Education	33	34.4
		39	40.6
	Master's in Education	21	21.9

**Source: Field data, 2021**

From Table 1, all the 96CRS teachers were involved in the study. This represents a return rate of 100.0%. Concerning the gender of the teachers involved in the study, 77.9% were males, while 28.1% were females. So a greater number of respondents in the study area were males. Also, the majority of the respondents were between 30-39 years. This is because, 6.2% were between 20-29 years, 46.9% were between 30-39 years, 21.9% were between 40-49 years, and 25.0% were between 50-59 years. In line with years in the teaching service, 9.4% had worked under 5 years, 56.2% had 6-10 years of working experience, and 34.4% had more than 10 years of working experience. Therefore, the majority of the teachers had taught between 6-10 years. This means that the majority of the teachers had taught CRS for quite a good number of years and may have a great deal of experience when it comes to assessment practices of Christian Religious Studies in Senior High Schools. Regarding academic qualification, 78.1% had Bachelor's Degree, 15.6% Master of Arts, and 6.2% had Master of Philosophy. As a result, the majority of those sampled had a Bachelor's Degree as their highest academic achievement. In terms of the major subject(s) areas offered, 54.2% majored in Religious Studies whereas 45.8% majored in other subject areas which included: political science, social studies, guidance and counseling, theology, social work, sociology, history, African studies, and history. Therefore, it is evident that the majority of the CRS teachers majored in Religious Studies and this is important for the study because CRS teachers who majored in Religious Studies are better prepared to teach the subject compared with their counterparts who did not major in Religious Studies. Although the majority of the CRS teachers majored in Religious Studies, it is very worrying that a considerable number of the CRS teachers (44, 45.8%) did not major in Religious Studies and this presupposes that, they did not do methods of teaching Religious Studies as well as assessment practices in Christian Religious Studies. In line with a professional teaching qualification, 3.1% had Teacher's Cert "A", 34.4% had Post Graduate Diploma in Education (PGDE), 40.6% had Bachelor of Education, and 21.9% had Masters in Education. This means that the majority of the respondents were professional teachers who had a Bachelor of Education.

This section presents the results and discussions of data collected to answer the five research questions formulated to guide the study. It comprised data from the questionnaire.

**How CRS Teachers Adhere to the Principle of Test Construction**

Research Question 1: How do CRS teachers adhere to the principles of test construction in Senior High Schools in the Bono, Bono East, and Ahafo regions of Ghana?

This research objective sought to find out how CRS teachers adhere to the principles of test construction in Senior High Schools in the Bono, Bono East, and Ahafo regions of Ghana. Table 2 presents the extent to which CRS teachers adhere to the principles of test construction.

**Table 2: Extent to which CRS Teachers Adhere to the Principles of Test Construction**

Statements: As a CRS teacher, I do the following when constructing test items:	M	SD
I state the purpose of the test.	4.56	.56
I specify the construct to be measured.	4.22	.74

I use a test specification table.	4.38	.74
I match learning outcomes to the items.	4.38	.55
I construct test items when it is time to assess.	3.94	1.20
I ask colleagues in subject area to review test items.	4.19	.89
I consider the meaning of wording against different ethnic background.	4.00	.87
I consider students' language proficiency.	4.16	.67
I consider variation of students with respect to physical disability.	4.00	.83
I consider the time individual will spend on a question.	4.22	.86
I try solving the questions myself to determine the time required.	4.34	.82
I provide clear and simple instructions on how test is to be answered.	4.38	.65
I evaluate test items given to the students.	4.25	.50
I write test items at least two weeks before time for test administration.	4.13	.82
I write more test items than needed.	4.25	.75
I follow the principles of test construction for each format.	4.50	.56

**Source: Field Data, 2021**

Scale:

1 = Strongly Disagree;                      2 = Disagree;                      3 = Uncertain;  
 4 = Agree;                                      5 =Strongly Agree

Mean of means = 4.24

Mean of standard deviation = .75

Table 2 sought to find out the extent to which CRS teachers adhere to the principles of test construction. A mean of means of 4.24 and a mean of the standard deviation of .75 was realized. This means that the majority of the teachers agreed that they adopted the principles of test construction to a high extent. Further discussions of individual items are presented in the paragraphs below.

From Table 2, a mean of 4.56 and a standard deviation of .56 were achieved for the statement: “I state the purpose of the test”. This means that the teachers strongly agreed with the statement. Gronlund [13] asserts that “the key to effective achievement testing is careful planning” (p. 15). It is during the planning stage that the purpose of the test must be determined. Again, when the teachers were asked whether they specify the construct to be measured, the respondents agreed to the statement. Here, a mean of 4.22 and a standard deviation of .74 were obtained for this item showing the respondents agreed that they specify the construct to be measured. Also, from Table 2, the teachers agreed that they use a test specification table. This is evidenced by the mean score of 4.38 and a standard deviation of .74 for this item. The mean is approximately 4, showing that the respondents agreed to the statement. This finding confirms that of [14] that a table of specifications must be kept before the teacher and continually referred to as the items are written.

Regarding the statement: “I match learning outcomes to the items”, the majority of the teachers agreed to the statement. This can be seen from the mean of 4.38 and a standard deviation of .55 that were realized. Mehrens and Lehmann [15] assert that test items must be related to and match the instructional objectives. Also, a mean of 3.94 and a standard deviation of 1.20 were recorded for the item “I construct test items when it is time to assess”. This means that the majority of the teachers agreed to the statement. This is because the mean falls on scale 4 (agree) looking at the scale under Table 2. However, the high standard deviation obtained which is higher than the mean of the standard deviation of .75 indicates that there were variations and that not all the respondents agreed to the statement. However, it remains that the majority of the teachers agreed to the statement.

The finding depicts that, most of the teachers agreed that they ask colleagues in the subject area to review test items. With a mean of 4.19 and a standard deviation of .89, it could be concluded that the mean falls into the scale of 4 (agree). Thus, the majority of the teachers ask colleagues in the subject area to review test items. Etsey [14] has suggested that the items must be critically examined at least a week after writing them. He has emphasised that where possible, fellow teachers or colleagues in the same subject area should review the test items. Again, when the respondents were asked whether they consider the meaning of wording against the different ethnic backgrounds, they agreed to the statement. Here, a mean of 4.00 and a standard deviation of .87 were obtained for this item showing that the respondents consider the meaning of wording against the different ethnic background. Also, from Table 2, the teachers consider students’ language proficiency. This is evidenced by the mean score of 4.16 and a standard deviation of .67 for this item. The mean is approximately 4, showing that the respondents agreed to the statement. Regarding the statement: “I consider the variation of students concerning physical disability” the majority of the teachers agreed to the statement. This can be seen from the mean of 4.00 and a standard deviation of .83 that were realized. This means that the teachers consider variations of students concerning physical disability. Again, when the teachers were asked whether they consider the time individual will spend on a question, they agreed to the statement. Here, a mean of 4.22 and a standard deviation of .86 were obtained for this item showing the respondents consider the time individual will spend on a question.

Also, from Table 2, the teachers indicated that they try solving the questions themselves to determine the time required. This is evidenced by the mean score of 4.34 and a standard deviation of .82 for this item. The mean falls on scale 4 (agree) showing that the respondents try solving the questions themselves to determine the time required. The finding also depicts that, most of the teachers provide clear and simple instructions on how the test is to be answered. With a mean of 4.38 and a standard deviation of .65, it could be concluded that the mean falls into the scale of 4 (agree). Regarding the statement “I evaluate test items given to the students”, a mean of 4.25 and a standard deviation of .50 were obtained for the statement. Thus, the majority of the teachers evaluate test items given to the students. From Table 2, a mean of 4.13 and a standard deviation of .82 were achieved for the statement: “I write test items at least two weeks before time for test administration”. This means that the teachers agreed to the statement. Etsey [14] explains that test items must be written in advance (at least two weeks) of the testing date to permit reviews and editing. Again, when the teachers were asked whether they write more test items than needed, the respondents agreed to the statement. Here, a mean of 4.25 and a standard deviation of .75 were obtained for this item showing that the respondents agreed that they write more test items than needed. This finding is in agreement with [15] who suggested that the initial number of items should be 25% more while [16] has suggested 10% more items

than are needed in the test. Also, from Table 2, the teachers strongly agreed that they follow the principles of test construction for each format. This is evidenced by the mean score of 4.50 and a standard deviation of .56 for this item. The mean is approximately 5, showing that the respondents strongly agreed with the statement.

From the foregoing, it can be concluded that CRS teachers adhered to the principles of test construction to a high extent. In adhering to the principles of test construction, the CRS teachers state the purpose of the test, specify the construct to be measured, use a test specification table, and match learning outcomes to the items. Also, the teachers construct test items when it is time to assess, ask colleagues in the subject area to review test items, consider the meaning of wording against the different ethnic backgrounds, and consider students' language proficiency. The teachers indicated that they consider the variation of students concerning physical disability, consider the time individual will spend on a question, try solving the questions themselves to determine the time required and provide clear and simple instructions on how the test is to be answered. Again, the teachers evaluate test items given to the students, write test items at least two weeks before time for test administration, write more test items than needed, and follow the principles of test construction for each format.

**How CRS Teachers Adhere to the Principle of Test Administration**

Research Question 2: How do CRS teachers adhere to the principles of test administration in Senior High Schools in the Bono, Bono East, and Ahafo regions of Ghana?

This research objective sought to find out how CRS teachers adhere to the principles of test administration in Senior High Schools in the Bono, Bono East, and Ahafo regions of Ghana. Table 3 presents the extent to which CRS teachers adhere to the principles of test administration.

**Table 3: Extent to which CRS Teachers Adhere to the Principles of Test Administration**

Statements: As a CRS teacher, I do the following when administering test items...	M	SD
I make students aware of the rules and regulations covering the test.	4.50	.71
I make room for adequate ventilation and lighting.	4.28	.80
I make provision for extra sheets and writing materials.	4.44	.71
I allow students to start and stop test on time.	4.47	.71
I give more instructions during the time the students are taking the test.	4.06	1.09
I inform students in advance about areas for the test.	4.47	.75
I prepare classroom a day before test is taken.	3.38	.82
I inform students about the test format.	4.38	.78
I make provision for emergencies during the time the test is taken.	4.38	.65
I proof read all the items.	4.65	.60
I use "DO NOT DISTURB SIGN" at the entrance of classroom.	2.13	1.22

**Source: Field Data, 2021**

Scale:

1 = Strongly Disagree;                      2 = Disagree;                      3 = Uncertain;  
4 = Agree;                                      5 = Strongly Agree

Mean of means = 4.10

Mean of standard deviation = .80

A mean of means of 4.10 and a mean of the standard deviation of .80 were realized. This means that the majority of the teachers agreed that they adopted the principles of test administration to a high extent. Details of the individual items are presented in the subsequent paragraphs.

The finding depicts that, most of the teachers strongly agreed that they make students aware of the rules and regulations covering the test. With a mean of 4.50 and a standard deviation of .71 it could be concluded that the mean falls into the scale of 5 (strongly agree). Thus, the majority of the teachers make students aware of the rules and regulations covering the test. This finding supports that of [14] that, in administering test items, students must be made aware of the rules and regulations covering the conduct of the test. Penalties for malpractice such as cheating should be spelled out and adhered to. Again, when the respondents were asked whether they make room for adequate ventilation and lighting, they agreed to the statement. Here, a mean of 4.28 and a standard deviation of .80 were obtained for this item showing that the respondents make room for adequate ventilation and lighting. The physical conditions that need to be in place to ensure maximum performance on the part of students include adequate workspace, quietness in the vicinity, good lighting and ventilation, and comfortable temperature [14]; [13]; [19]. Good lighting is important in effective test administration. This facilitates students' reading of instructions and test items without straining their eyes, thereby working faster [13]. Also, from Table 3, the teachers make provision for extra sheets and writing materials. This is evidenced by the mean score of 4.44 and a standard deviation of .71 for this item. The mean is approximately 4, showing that the respondents agreed to the statement. Regarding the statement: "I allow students to start and stop test on time" the majority of the teachers agreed to the statement. This can be seen from the mean of 4.47 and a standard deviation of .71 that were realized. This means that the teachers allow students to start and stop tests on time. Nunnally [17] and [18] assert that a good working rule is to try to set a time limit such that about 90 percent of the students will feel that they have enough time to complete the test. Again, when the teachers were asked whether they give more instructions during the time the students are taking the test, they agreed to the statement. Here, a mean of 4.06 and a standard deviation of 1.09 were obtained for this item showing the respondents give more instructions during the time the students are taking the test. There is the need to consider the psychological conditions in test administration which include, threatening behaviour of invigilators, and interruption to give instructions and announcements [14]; [13]; [19].

Also, from Table 3, the teachers indicated that they inform students in advance about areas for the test. This is evidenced by the mean score of 4.47 and a standard deviation of .75 for this item. The mean falls on scale 4 (agree) showing that the respondents inform students in advance about areas for the test. On the fairness criterion, consideration is given to whether students have been given advance notice of the test, whether students have been adequately prepared for the test and whether students understand the testing procedures [20]; [14]; [19]. The teacher is to prepare his students in advance for the test [14]. Etsey [14] has emphasised that for students' maximum performance, they should be made aware of when (date and time) the test will be given, the conditions (number of items, place of test, open or closed book) under which

the test will be given, the content areas (study questions or list of learning targets) that the test will cover, the emphasis or weighting of content areas, the kinds of items (objective-types or essay-types) on the test, how the test will be scored and graded, and the importance of the results of the test.

The finding also depicts that, most of the teachers were uncertain as to whether they prepare classroom a day before test is taken. With a mean of 3.38 and a standard deviation of .82 it could be concluded that the mean falls into the scale of 3 (uncertain). Regarding the statement, "I inform students about the test format", a mean of 4.38 and a standard deviation of .78 were obtained for the statement. Thus, the majority of the teachers inform students about the test format. From Table 3, a mean of 4.38 and a standard deviation of .65 were achieved for the statement: "I make provision for emergencies during the time the test is taken". This means that, the teachers agreed to the statement. This finding is in agreement with [19] who explains that emergencies during test administration must be expected and well catered for. Again, when the teachers were asked whether they proof read all the items, the respondents strongly agreed to the statement. Here, a mean of 4.65 and a standard deviation of .60 were obtained for this item showing that the respondents strongly agreed that they proof read all the items. Also, from Table 3, the teachers disagreed that they use "DO NOT DISTURB SIGN" at the entrance of classroom. This is evidenced by the mean score of 2.13 and a standard deviation of 1.22 for this item. The mean is approximately 2, showing that the respondents disagreed to the statement. Noise and distraction in the testing environment should be kept at the barest minimum if not eliminated completely. Interruptions within and outside the testing room has the tendency of affecting student's performance [15]; [19]. Etsey [14] has pointed out that it is helpful to hang a -Do Not Disturb. "People should stay away from the testing room because there is a "Testing in Progress" notice on the door. It can be concluded that CRS teachers adhered to the principles of test administration to a high extent. In adhering to the principles of test administration, the CRS teachers make students aware of the rules and regulations covering the test, make room for adequate ventilation and lighting, make provision for extra sheets and writing materials, and allow students to start and stop tests on time. Also, the teachers give more instructions during the time the students are taking the test, inform students in advance about areas for the test, inform students about the test format, make provisions for emergencies during the time the test is taken, and proofread all test items. However, the teachers did not use "DO NOT DISTURB SIGN" at the entrance of the classroom. Also, teachers were uncertain as to whether they prepared the classroom a day before the test is taken.

### **How CRS Teachers Adhere to the Principle of Test Scoring**

Research Question 3: How do CRS teachers adhere to the principles of test scoring in Senior High Schools in the Bono, Bono East, and Ahafo regions of Ghana?

This research objective sought to find out how CRS teachers adhere to the principles of test scoring in Senior High Schools in the Bono, Bono East, and Ahafo regions of Ghana. Table 4 presents the extent to which CRS teachers adhere to the principles of test scoring.

**Table 4: Extent to which CRS Teachers Adhere to the Principles of Test Scoring**

Statements: As a CRS teacher, I do the following when scoring test items...	M	SD
I mark papers just after the test is taken.	3.63	.75
I prepare a scoring guide.	4.40	.70
I make sure test takers are kept anonymous.	3.65	1.02
I grade the responses item by item.	4.21	.78
I keep scores of previous items out of sight.	4.06	.90
I periodically rescore previously scored items.	2.00	.94
I shuffle scripts before scoring.	2.13	.86
I score essay test when I am physically sound and mentally alert in a sound environment.	4.38	.74
I constantly follow the scoring guide.	4.34	.54
I score a particular item on all papers at a sitting.	2.09	.88
I provide comments and correct errors on scripts.	4.65	.48

**Source: Field Data, 2021**

Scale:

1 = Strongly Disagree;                      2 = Disagree;                      3 = Uncertain;  
 4 = Agree;                                      5 =Strongly Agree

Mean of means = 3.60

Mean of standard deviation = .78

Table 4 sought to find out the extent to which CRS teachers adhere to the principles of test scoring. A mean of means of 3.60 and a mean of the standard deviation of .78 were realized. This means that the majority of the teachers agreed that they adopted the principles of test scoring to a high extent. Further discussions of individual items are presented in the paragraphs below.

From Table 4, a mean of 3.63 and a standard deviation of .75 were achieved for the statement: “I mark papers just after the test is taken”. This indicates that the teachers agreed to the statement. Again, when the teachers were asked whether they prepare a scoring guide, the respondents agreed to the statement. Here, a mean of 4.40 and a standard deviation of .70 were obtained for this item showing the respondents agreed that they prepare a scoring guide. According to [14] and [16], the marking scheme must be prepared when the items are still fresh in the teacher’s mind and always before the administration of the test. This way, defective items that do not match their expected responses would be recognised and reviewed. Also, from Table 4, the teachers agreed that they make sure test takers are kept anonymous. This is evidenced by the mean score of 3.65 and a standard deviation of 1.02 for this item. The mean is approximately 4, showing that the respondents agreed to the statement. However, the high standard deviation obtained which is higher than the mean standard deviation of .75 indicates that there were variations and that not all the respondents agreed to the statement. However, it still remains that the majority of the teachers agreed to the statement. Mehrens and Lehmann [15], [16] and [14] suggest that there is the need to score the scripts anonymously. Instead of using students' names, scripts should be recognized by code numbers or some other method. The goal of this idea is to lessen the halo effect. This occurs when a scorer's general opinion about a person has an impact on how the paper is graded. Regarding the statement: “I grade the responses item by item”, the majority of the teachers agreed to the statement. This can be seen from the mean of 4.21 and a

standard deviation of .78 that were realised.

Also, a mean of 4.06 and a standard deviation .90 were recorded for the item “I keep scores of previous items out of sight”. This means that, the majority of the teachers agreed to the statement. This is because the mean falls on scale 4 (agree) looking at the scale under Table 4. The finding depicts that, most of the teachers disagreed that they periodically rescore previously scored items. With a mean of 2.00 and a standard deviation of .94 it could be concluded that the mean falls into the scale of 2 (disagree). Thus, the majority of the teachers do not periodically rescore previously scored items. Again, when the respondents were asked whether they shuffle scripts before scoring, they disagreed to the statement. Here, a mean of 2.13 and a standard deviation of .86 were obtained for this item showing that the respondents do not shuffle scripts before scoring. Research by [15] has shown that a student's essay grade will be influenced by the position of the paper, especially if the preceding answers were either very good or very poor. Mehrens and Lehmann [15] have pointed out that randomly reshuffling of scripts is especially significant when teachers are working with high- and low level classes and read the best scripts first or last. Also, from Table 4, the teachers score essay test when they are physically sound and mentally alert in a sound environment. This is evidenced by the mean score of 4.38 and a standard deviation of .74 for this item. The mean is approximately 4, showing that the respondents agreed to the statement. According to Karpicke and Roediger [21], over excitement, depression, and any type of psychological or mental disequilibrium will affect the consistency of the scores of essay-type tests.

Table 4 sought to find out the extent to which CRS teachers adhere to the principles of test scoring. A mean of means of 3.60 and a mean of the standard deviation of .78 were realized. This means that the majority of the teachers agreed that they adopted the principles of test scoring to a high extent. Further discussions of individual items are presented in the paragraphs below.

From Table 4, a mean of 3.63 and a standard deviation of .75 were achieved for the statement: “I mark papers just after the test is taken”. This means that the teachers agreed to the statement. Again, when the teachers were asked whether they prepare a scoring guide, the respondents agreed to the statement. Here, a mean of 4.40 and a standard deviation of .70 were obtained for this item showing the respondents agreed that they prepare a scoring guide. According to [14] and [16], the marking scheme must be prepared when the items are still fresh in the teacher's mind and always before the administration of the test. This way, defective items that do not match their expected responses would be recognised and reviewed. Also, from Table 4, the teachers agreed that they make sure test takers are kept anonymous. This is evidenced by the mean score of 3.65 and a standard deviation of 1.02 for this item. The mean is approximately 4, showing that the respondents agreed to the statement. However, the high standard deviation obtained which is higher than the mean standard deviation of .75 indicates that there were variations and that not all the respondents agreed to the statement. However, it still remains that the majority of the teachers agreed to the statement. Mehrens and Lehmann [15], [16], and [14] suggest that there is the need to score the scripts anonymously. Scripts should be identified by code numbers or any other means instead of the names of students. This principle is to reduce the halo effect. This happens when a scorer's general impression of a person influences how the paper is scored. Regarding the statement: “I grade the responses item by item”, the majority of the teachers agreed to the statement. This can be seen from the mean of 4.21 and a standard deviation of .78 that were realised.

**Conclusions**

These conclusions were drawn as a result of the findings realized for the study. The CRS teachers adhered to the principles of test construction, administration and scoring to a high extent. In adhering to the principles of test construction, the CRS teachers indicated that they construct test items when it is time to assess and at the same time write test items at least two weeks before time for test administration. This is a contradiction that raises a lot of questions. How can teachers claim to write test items at least two weeks before time for test administration and at the same time construct test items when it is time to assess? Perhaps, the teachers were not frank in responding to the items and may not have realized that they contradicted themselves in ticking the “agreed” items. Also, although the CRS teachers adhered to the principles of test administration to a high extent, they did not prepare the classroom a day before the test is taken, and they did not use the “DO NOT DISTURB SIGN” at the entrance of the classroom. It could be that, perhaps, the teachers do not know that there is the need for them to put the “DO NOT DISTURB SIGN” during test administration. Also, teachers’ inability to prepare the classroom a day before the test is taken may be due to a lack of infrastructure. Again, in adhering to the principles of test scoring, teachers did not shuffle scripts before scoring, and they did not score a particular item on all papers at a sitting. Again, the teachers did not periodically rescore previously scored items.

**Study Recommendations**

The following recommendations were made for policymakers:

1. It is suggested that teachers should write test items at least two weeks before the time for test administration so that they can have enough time to read over and correct wrongly constructed items before the test is administered. This would also ensure that teachers have enough time to ensure that the test covers all the major areas that need to be examined.
2. Also, teachers should prepare the classroom a day before the test is taken and in the case of lack of infrastructure, the classroom should be made ready and prepared hours before the test is taken. Teachers should also use the “DO NOT DISTURB SIGN” at the entrance of the classroom during test administration.
3. Again, teachers should shuffle scripts before scoring, and should score a particular item on all papers at a sitting. Again, the teachers should periodically rescore previously scored items.

**References**

- [1] Nitko, A.J. (1996). Educational Assessment of Students (2nd ed.) Englewood Cliffs, NJ: Prentice-Hall/ Merrill Education
- [2] NMCE(2014). Technical Recommendations for psychological test and diagnostic techniques: Washington DC: American Psychological Association
- [3] Linn, R. L. & Gronlund, N. E. (1995). Measurement and assessment in teaching (7th ed.). Englewood Cliffs, New Jersey; Prentice Hall.
- [4] Johnson, B. (2001). Toward a new classification of non-experimental quantitative research. Educational Researcher, 30, 3-13. doi:10.3102/0013189X030002003
- [5] Levin, K. A. (2006). Study Design III: Cross-Sectional Studies. Evidence-Based Dentistry, 7, 24-25.

- [6] Ugenda, O.M. and Mugenda, A.G. (2003) *Research Methods, Quantitative and Qualitative Approaches*. ACT, Nairobi.
- [7] *Education Information Management System (2020)*. Brong Ahafo Education Office: Sunyani, Ghana.
- [8] Cohen, L, Manion, C and Morrison, K. (2011). *Research Methods in Education*. London; Routledge.
- [9] Sarantakos, S. (2005). *Social research* (3<sup>rd</sup> ed.). Melbourne: MacMillan Education.
- [10] Osuala, C. S. (2005). *Introduction to research methodology African*. Fep Publishers Onisha Nigeria.
- [11] Deng, P. (2010). The internationalization of Chinese firms: A critical review and future research. *International Journal of Management Reviews*, 14(4), 408-427.
- [12] Fraenkel, J. R, & Wallen, N. E. (2003). *How to design and evaluate research in education*. New York: McGraw Hill.
- [13] Gronlund, G., M. (2013). *Focused observations: How to observe young children for assessment and curriculum planning*. Redleaf Press.
- [14] Etsey, Y. K. A (2004). *Assessing performance in schools: Issues and practice*. Ife
- [15] Mehrens G. & Lehmann, Y. (2014). Improving Students' Teaching through authentic assessment. Paper presented at the first International conference on enhancing Teaching and learning through assessment. Hong Kong Polytechnic University, Hong Kong.
- [16] Amedahe, F. K. (2004). *Testing practices in secondary schools in the Central Region of Ghana*. Unpublished master's thesis. University of Cape Coast.
- [17] Nunnally, H. (2015). *Psychological Testing*. NY: Macmillan Publishing Company
- [18] Abel, L. R & Frisbie A.D. (2014). *Essentials of Educational Measurement* (5th ed). Englewood Cliffs, New Jersey: Prentice Hall
- [19] Tamakloe, E. K., & Amedahe, F. K. (2013). *Principles and methods of teaching*. Cantoment. Accra: Black Mask.
- [20] Amedahe, F. K. & Gyimah, K. A. (2003). *Measurement and evaluation*, Cape Coast.
- [21] Roediger, H. L., Karpicke, J D. (2006). *Psychol. Sci.* 17, 249