UNIVERSITY OF CAPE COAST

TEACHER CHARACTERISTICS AS CORRELATES OF CLASSROOM ASSESSMENT PRACTICES AT ATWIMA NWABIAGYA SOUTH DISTRICT

OWUSU-MENSAH JAWARD

2019

UNIVERSITY OF CAPE COAST

TEACHER CHARACTERISTICS AS CORRELATES OF CLASSROOM ASSESSMENT PRACTICES AT ATWIMA NWABIAGYA SOUTH DISTRICT

BY

OWUSU-MENSAH JAWARD

Thesis submitted to the Department of Education and Psychology of the Faculty of Educational Foundations, College of Education Studies, University of Cape Coast, in partial fulfilment of the requirements for the award of Master of Philosophy Degree in Educational Measurement and Evaluation.

AUGUST 2019

DECLARATION

Candidate's Declaration

I hereby declare that this thesis is the result of my own original research and that no part of it has been presented for another degree in this university or elsewhere.

Candidate's Signature: …… Date: ……..……

NAME: ……

Supervisors' Declaration

We hereby declare that the preparation and presentation of the thesis were supervised in accordance with the guidelines on supervision of thesis laid down by the University of Cape Coast.

Principal Supervisor's Signature: …… Date: ……..……

Namer: ……

Co-Supervisor's Signature: ……. Date: ……..…….

Name: ……..

ABSTRACT

The purpose of the study was to investigate teacher characteristics as correlate of classroom assessment practices at Atwima Nwabiagya South District J.H.Ss. Descriptive survey design with a quantitative approach was used in the conduct of the study. A sample of 219 respondents was randomly selected from a population of 366 teachers. A 42-item Teacher Self-Perceived Assessment practices questionnaire was used to gather data from the respondents. Frequencies and percentages were used to analyse the data to answer research questions 1 to 5. Data on research question 6 was answered using hierarchical multiple regression. Hypothesis 1 was tested using independent samples t-test. It was found that teachers used a number of assessment formats in assessing students. Prevalent among them were essay items to assess students' understanding, giving class exercises to students immediately after teaching, use of oral questions to assess students, as well as the use of exams to assess students at the end of the term. It was also found that teachers considered the purpose of the test before developing test items, prepared a marking scheme immediately after constructing test items, matched instructional objectives with test items, teachers however, did not use test specification table while writing test items. None of the demographic variables (experience, age and qualification) had a significant effect on teachers' assessment practices. No significant difference was found in teachers' assessment practices with regard to gender. It was recommended that the Ministry of Education and school administrators should train and encourage teachers regarding the importance of using of test specification tables when constructing test items.

ACKNOWLEDGEMENTS

In the quest to undertake this study, I relied on the services of many people without whose assistance the work would not have been successful. I wish to recognize and acknowledge my indebtedness to Dr. Eric Anane, Examination Officer, Institute of Education, University of Cape Coast, who as my principal supervisor, pain-stakingly read through the original manuscript and offered valuable suggestions. I would like to register my sincere thanks to Dr. Kenneth Asamoah- Gyimah, Head of Department of Education and Psychology of the university of Cape Coast, who as my co-supervisor, also read through the manuscript and offered very useful suggestions. I wish to express my sincere apprecation to Headteachers and teachers of the forty-eight selected public Junior High Schools of Atwima Nwabiagya South District who took part in the study, accepted me and cooperated with me during the data gathering process. Colleagues and friends especially Gifty Aba Essel also deserve my sincere gratitude. However, any errors or shortfalls are mine.

DEDICATION

To my late father, Shieck Zacharia Duad Mensah

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

Background to the Study

Globally, students' achievement goals have been regarded as important achievement-related outcomes to be promoted in the classroom learning process (Harlen & Crick, 2003). Classroom assessment encompasses a range of activities from construction of assessment tasks, administration, marking and grading the tasks to interpreting the results. Information generated through assessment can help teachers to evaluate the effectiveness of their teaching strategies. It is essential to use assessment feedback to make decisions about teaching and learning so as to ensure that meaningful learning takes place. It is important to state that classroom assessment is a global exercise or practice that has caught the attention of many researchers. For example, in a synthesis of over 250 studies, Black and William (1998) concluded that improvement of learning occurs when teachers use classroom assessment information to establish knowledge, skills and attitudes possessed by their students and incorporate that information in planning for lessons. The major purpose of conducting classroom assessment was to obtain information about students' progress in learning and the achievement attained (Airasian, 2001; Gronlund, 2003; McMillan, 2008; Popham, 2008).

Again, research evidence showed that classroom assessment is an essential ingredient for effective teaching and learning (Gipps, 1990; Black &

Wiliam, 1998; Shephard, 2000; Stiggins, 2002). Crooks (1998) reviewed the impact of classroom assessment practices on students and observed that the choice of classroom assessment approach has a vital effect on the extent to which teaching and learning can be enhanced. Proper choice of classroom assessment method allows teachers to diagnose problems faced by students in attaining desirable learning outcomes and in devising appropriate remedial measures to redress the situation. In a nutshell, classroom assessment can be viewed as a totality of all the processes and procedures used to gather useful information about the progress in teaching and learning which facilitates in regulating the pace and strategies of teaching.

The frequency of classroom assessment has also been noticed as an important exercise. In a study, Shirvani (2009) observed that the frequency of assessment has a mediating effect on student engagement in learning. Research by Marcell (2008) showed that when the frequency of testing is increased, there is increased student involvement in responding to questions and in discussing the subject matter. Other scholars maintained that frequent testing helps students to monitor their learning and reinforces their engagement with the course as a result of immediate feedback provided (Haigh, 2007; Leeming, 2002). It has also been established that frequent testing has positive impact on future retention of material learnt (Roediger & Karpicke, 2006). Since retention of material is one of the important components of mastery learning (Wolf, 2007), it can be inferred that frequent testing contributes to mastery learning.

As a matter of principle, assessment practices must be appropriate in order to produce results that are appropriate. Tanner and Jones (2006) argued

that appropriateness of an assessment procedure will always be influenced by the original purpose of the assessment, and the intended audience of the results. Teachers, therefore, should always ask themselves why they should assess their learners. According to Broadfoot, Winter and Weeden (2000) assessment is classified according to its purposes. These scholars further enlighten that the assessment terms are not descriptions of kinds of assessment, but rather, of the uses to which information arising from the assessment is put. In other words, each component of assessment should fit the purpose.

Undeniably, some studies have indicated that some practitioners (teachers) in the field do face some challenges in their daily assessment in the classroom. Broadfoot, et al. (2002) admit that some teachers are faced with challenges as they develop formative assessments. The initial challenge being that these teachers cannot clarify for themselves what they understand by formative assessment, they cannot also decide how they can make initial changes to their practice that will help learners actively engage in their own learning and be clear about their current performance and decide what they need to do next. In addition, Broadfoot et al. (2000) bring forward another challenge as being that teachers find formative assessment as time consuming. It may be true that the introduction into the classroom of formative assessment strategies, such as sharing objectives and self-assessment may appear to consume time for teaching. Teachers on the other hand, need to spend more time marking and giving feed back to the learners.

Collecting and interpreting data formatively are another challenge for teachers (Jones & Tanner, 2006). These scholars claim that any assessment

can be formative if it is used not just to categorize a child as an 'A' or 'F', but to identify how performance can be improved. An assessment becomes formative if the information collected is used by the teacher or by the pupil in the learning processes to promote the learning. To curb the situation, teachers need to internalise the function of classroom assessment, and not view it as something in competition with teaching. Rather, they should view assessment as an integral part of teaching and learning. They should desist from using assessment as a tool solely for ranking learners and schools (Heritage, 2009).

According to Stiggins (1991), "teachers spend much of their instructional time ('a third to a half') in assessment related activities. While this would suggest the need for teachers to be knowledgeable on assessment practices, that is not the case." (p. 85). Ward (1980) reported that only about half of the teachers in a national survey had received pre- service instruction in tests and measurement. As more and more emphasis is being placed on student performance and teacher accountability, measurement and assessment are becoming increasingly important to all educators. It is critical therefore that tutors possess not only comprehensive knowledge of subject matter but also the ability to assess the learning of the subject matter. A concerted effort to use valid measurement and assessment practices would strengthen the effectiveness of teaching and learning.

In Ghana, classroom assessment also plays a critical role in our education system. According to Asamoah-Gyimah (2002), classroom or teacher-made tests are frequently used as a major evaluating device of students' progress in Ghana schools. Testing practices are, therefore, indispensable tools in the educational enterprise. Ghanaian educators as well

as others concerned with the progress of the leaner need information on the learner for decision making. There is the need for the teacher to describe the nature and extent of the learner's learning in terms of how far the aims and objectives of teaching have been achieved and what is left to be covered. This need calls for assessment of the learner (Tamakloe, Atta, & Amedahe, 1996).

Again, both the trained and the untrained teachers in the classroom, from the basic level to the university level, construct, administer and score classroom achievement tests irrespective of whether they have had training in measurement and evaluation or not. Decisions mostly taken on the students have far reaching consequences that affect the student. Policy makers have ignored the training and equipping teachers with the skill in test construction, test administration and test scoring for the fact that results have always been achieved out of tests, even in the private institutions.

Amedahe (1989) stated that:

> Although teachers may succeed in their teaching to some extent without following prescribed principles in testing their students, more could be achieved if scientific principles and practices deemed useful in measurement are followed. This is rather crucial in the Ghanaian educational system where we lack, to a large extent, the availability of standardized achievement and intelligence tests found in developed countries like the United States of America (U. S. A.) and Britain. What could be termed standardized achievement tests (these cannot strictly meet the definition of standardized tests) are the tests conducted by the West African Examinations Council (WAEC), which are taken

nationally and internationally in the West African sub-region at the terminal points of the educational system. Examples are the Middle School, the Junior Secondary School and the present Secondary School Leaving Certificate Examinations (pp. 3-4).

Not all teachers in the schools in Ghana have undergone professional training in testing techniques (Amedahe, 1989). It is, therefore, unascertainable whether these teachers and for this study, teachers in the Basic schools in Ghana with little or no skills in measurement and evaluation are coping with the task of testing with specific regard to construction, administration and scoring of classroom or teacher-made tests in the schools. To what extent are the characteristics (gender, age, experience and qualification) of these teachers affecting their test practices in the classroom? In other words, to what extent does teacher charateristics correlate assessment practices in the classroom?

Statement of the Problem

The practice of assessment in the Basic Schools in Ghana is based on the premise that teachers have had a course or training in "testing". This premise assumes that Basic School teachers construct, administer and score classroom or teacher-made tests based on the basic principles on testing in measurement and evaluation (Oduro, 2015). This premise is not, in most cases, necessarily true. In addition to this, studies by Amedahe (1989) and Akyeampong (1997) reported that most Ghanaian teachers had limited skills for constructing the objective and essay type tests, which are the most frequently used instruments in our schools. This is because most initial teacher

training programmes do not make adequate provision for a course in educational testing until recently that assessment as a course, has been introduced in colleges of education but it is combined with research methods and is only two credit hours which is inadequate for the course.

Amedahe (2000) further, stated that "teacher-made tests may be influenced by a number of factors, notably among them are, training in assessment techniques, class size and a particular school's policy in assessment standards with implications on validity and reliability of the assessment results" (p. 112-113). Most studies conducted in Ghana, in the area of basic school teachers' assessment practices focused attention on teachers' assessment practices, the influence of their knowledge (Asare, 2015; Hattori & Saba, 2008; Konadu, 2015) and assessment impact on classroom learning (Amua-Sekyi, 2016; Oduro, 2015). One may ask, are the factors as stated by the Ghanaian test expert exhaustive? Could it also be possible that the characteristics of basic school teachers would have any relation in the way they assess students? There are no readily available answers for these questions because in Ghana, not much information regarding this issue has been documented hence, it is unclear whether certain teacher characteristics such as gender, age, experience and qualification would have a relation with assessment practices of Junior High School teachers as well.

Empirically, much is known about assessment practices among teachers in Ghana and most of the studies (e.g., Anhwere, 2009; Akyeampong, 1997; Kankam et al., 2014) have found limited skills in assessment and testing practices among Senior High School (SHS) teachers and tutors of Colleges of education. Little is however know about the relationship that exist between

teacher characteristics (such as gender, age, experience, qualification) and classroom assessment practices in the Junior High Schools. The few studies that focused on teacher characteristics and assessment practices were conducted among teachers in the western world. For example, Alkharusi (2011a) conducted a study in Sultanate of Oman (a Middle East Arab country) on teachers' self-perceived assessment skills as a function of gender, subject area, grade level, teaching experience, and in-service assessment training. Alkharusi found significant differences on the self-perceived assessment skills with respect to teachers' gender, subject area, grade level, teaching experience, and in-service assessment training.

A critical examination of the literature appears to suggest that studies done in the area of assessment practices (eg. Alkharusi, 2011a) focused much attention on teachers' self-perceived assessment skills. This implies that the previous studies did not measure the actual assessment practices of teachers. Additionally, most of the studies conducted in the area of assessment practices in Ghana were also conducted among teachers in the Senior High Schools and Colleges of education (e.g., Anhwere, 2009; Akyeampong, 1997; Kankam et al., 2014). It is therefore important to investigate the extent to which teacher characteristics such as gender, age, experience and qualification would relate to teachers' classroom assessment practices. Based on this, the present study seeks to investigate the relationship that exist between teachers' characteristics such as age, experience, and qualification and teachers' classroom assessment practices. The study also seeks to examine the testing practices of teachers in Junior High Schools in Ghana in terms of the development or construction, grading, analysis and feedback of teacher-made tests.

Purpose of the Study

The study sought to investigate teacher characteristics as correlate of classroom assessment practices at Atwima Nwabiagya South District J.H.Ss. The focus was on the relationships that exist between teacher characteristics such as gender, age, experience and qualification and their classroom assessment practices. Specifically, the study sought to find out the;

1. prevalent classroom assessment practices among teachers;
2. test construction practices of teachers;
3. test administration practices of teachers;
4. test scoring practices of teachers;
5. grading practices of teachers;
6. contribution of teachers' age, qualification and experience to classroom assessment practices;
7. difference that exist among teachers in terms of how they practice assessment in relation to gender;

Research Questions

The following research questions were posed to guide the study:

1. What are the prevalent classroom assessment practices among teachers in Atwima Nwabiagya South District?
2. What is the test construction practice of teachers in the Atwima Nwabiagya South District?
3. What is the test administration practices of teachers in the Atwima Nwabiagya South District?
4. What is the test scoring practices of teachers in the Atwima Nwabiagya South District?

5. What is the grading practice of teachers in the Atwima Nwabiagya South District?
6. What demographic variables (age, experience, qualification) predict teacher classroom assessment practices?

Research Hypotheses

1. H_0: There is no significant difference among teachers in terms of their assessment practices with reference to gender.

   H_1: There is significant difference among teachers in terms of their assessment practices with reference to gender.

Significance of the Study

Gosling (1967) (as cited in Amedahe, 1989) stated "that the role of teachers in testing is too important to be left to chance" (p. 23). Since teachers in the Basic Schools of Ghana are homogeneous to some extent, in terms of their training qualification and duties in the classroom, the conclusions and recommendations that come out of the study could be quite relevant and a guide to all teachers in the Basic School.

Moreover, the study could serve as an important reference source for researchers, school administrators, National Teaching Council (NTC) of the Ghana Education Service (GES) and the Institute of Education, (IOE) of the University of Cape Coast, (UCC) in their effort to improve the management of testing in the teacher training colleges with adequate information about what is actually involved in classroom assessment practices. Additionally, examamination bodies, curriculum experts as well as students would benefit from the findings of this study. It is hoped that the study would complement studies already undertaken in this subject matter. Besides, the study could

contribute to knowledge regarding the improvement of testing practices, specifically, construction, administration and scoring of teacher-made tests in the Basic schools in Ghana.

Delimitations

Issues concerning assessment practices are so broad that it would not be feasible for one study to identify all. This study focused on assessment practices, construction, administering, scoring and grading of assessment result and how teacher characteristics such as age, gender, experience and qualification affect it. The research was confined to Junior High School teachers both trained and untrained at the Atwima Nwabiagya South District in the Ashanti Region of Ghana.

Limitations

A study of this nature revealed challenges which might affect the validity of the responses. First, the nature of the design resulted in issues of validity and reliability. Survey designs captured brief moments in time just like taking a photograph of an on-going activity. This suggests that the validity of the findings in survey studies were time-bound and therefore, the validity of such studies reduces with time passage. In line with this study, the result of this present study might not hold in ten years since teachers' characteristics as correlates of assessment practices might change over time. Caution should be taken when reporting findings of this study in some time to come.

Definition of Terms

Assessment Practices: It involves a wide range of activities from designing paper-pencil tests and performance measures to grading, communicating assessment results, and using them in decision making.

Teacher Characteristics: This encompasses teachers' age, gender, qualification as well as number of years (experience) in the teaching field.

Organisation of the Study

The study has been organised into five chapters. The first chapter discusses the introduction, which highlights the background to the study, the statement of the problem, and the purpose for the study. Also indicated in this chapter are the research questions, the significance and delimitation of the study. Chapter Two reviews the literature related to the study. The review involves empirical studies and conceptual framework. The Third chapter describes the methodology used for the study. This involves the research design, population and sampling procedure, the research instrument, the pre-testing procedure, the procedure for data collection and the data analysis. In Chapter Four, the results are discussed while the final chapter summarises the study and provides conclusions. Recommendations are given in the last section of the chapter based on the findings of the study.

CHAPTER TWO

LITERATURE REVIEW

The purpose of the study was to investigate teacher characteristics as correlate of classroom assessment practices at Atwima Nwabiagya South District J.H.Ss. This chapter reviews the earlier studies related to the purposes of the study. Among the issues to be discussed include the conceptual framework, theoretical and empirical evidences surrounding the topic. The areas reviewed include the following:

1. The concept of classroom assessment pratices.
2. Historical development of testing.
3. Assessment practices issues in Ghanaian Basic Schools.
4. Theoretical framework of assessment practices.
    i. Classroom assessment environment model
    ii. Achievement goal theory
5. Empirical review on:
    i. Types of classroom assessment practices
    ii. Construction of teacher made test
    iii. Administering of teacher made test
    iv. Scoring of teacher made test
    v. Characteristics of teachers and their classroom assessment practices.

The Concept of Classroom Assessment Practices

Classroom assessment aims at improving student learning and motivation to learn (Gronlund, 2006). It has become a tool for improving classroom teaching and learning (Shavelson et al. 2008). In this regard, Gronlund (2006) suggests that a sound classroom assessment requires a clear conception of all intended learning outcomes of the instruction and a variety of assessment procedures that are relevant to the instruction, adequately sample student performance, and fair to everyone. In addition, a sound assessment requires the specifications of criteria for judging successful performance and timely and detailed feedback to students emphasizing strengths of their performance and weaknesses to be corrected (Gronlund, 2006).

Teachers are required to develop classroom assessment that aligns with practices recommended by experts of educational assessment. For example, assessment experts have recommended that students should clearly be informed about the grading procedure in advance and involved in the assessment process (Stiggins & Chapuis 2005; Stiggins, Frisbie & Griswold, 1989). Student personal characteristics such as ability, effort, motivation, interest, and neatness of work should not be incorporated into grading due to the lack of objective measurement (Stiggins et al. 1989); a final grade for borderline cases should be determined using additional academic achievement data rather than non-achievement data (Stiggins et al. 1989); and students should be given continuous and formative assessment feedback rather than judgmental feedback about their academic performance (Brookhart, 1994).

Moreover, teachers are encouraged to use more than one assessment method in order to have enough, accurate evidence of student learning (Nitko

2001; Stiggins et al. 1989). It is emphasized that the assessment should match the learning targets and provides meaningful feedback to students (Nitko, 2001). In addition, the American Federation of Teachers (AFT), the National Council on Measurement in Education (NCME), and the National Education Association (NEA) (1990) have jointly defined seven Standards for Teacher Competence in Educational Assessment of Students (Brookhart, 1994).

The standards emphasized that teachers should competently be able to choose and develop assessment methods appropriate for instructional decisions; administer, score, and interpret results of externally produced and teacher-made assessment; use assessment results when making educational decisions; develop valid assessment-based grading procedures; communicate assessment results; and recognize unethical, illegal, and inappropriate methods and uses of assessment (AFT, NCME & NEA 1990). Like in the western world, West African countries have embraced continuous assessment of students, with so much emphasis laid on the classroom assessment being comprehensive (Stiggins et al. 1989).

For example, before the creation of West African Examinations Council in 1952, pre-tertiary institutions in Ghana were taking examinations of British accredited bodies. These were one-shot, theory biased examinations which were often criticized for not fully assessing ability to apply the knowledge and skills acquired. From this experience many West African States started making some form of changes in their educational reforms that gave birth to continuous assessment usage. Nigeria was the first to start the implementation of the reform. The first batch of students took their first Junior Secondary School in 1987. The reforms in Ghana started in 1987 and the first

batch of Ghana's Basic Education Certificate Examination candidates took their examination in 1990. Sierra Leone and the Gambia have also followed suit. The use of continuous assessment brought a lot of stake holders in education attention to the importance of classroom assessment practices especially in the first and second cycle institutions (Akplu, 1989).

Unfortunately, findings from past and recent studies of classroom assessment practices have consistently expressed a concern about the adequacy of teachers' skills in assessment. For example, in an earlier survey of statistical analyses of test results for 336 elementary and secondary school teachers, Gullickson (1982) found that a substantial proportion of teachers reported using relatively little statistical information such as means, medians, and standard deviations to describe assessment results. Also, these same teachers did not have an adequate understanding of basic testing concepts such as item difficulty and reliability. Parallel to Gullickson's (1982) study, Mertler (1998, 1999) found in two studies of 625 K-12 Ohio state teachers that teachers did not spend much time conducting statistical analyses of the assessment data with no significant differences based on teacher's gender and years of teaching experience. Further, Hills (1991) identified four misuses of classroom assessment in schools including using grades for controlling students' behavior, assigning grades that are contingent on improvement, using tests that are technically inadequate, and deviation from established standardized-test administration procedures.

In Ghana some test experts have also raised similar concerns on the issue of assessment practices. For example, Amedehe (1989) stated that not all teachers in the Ghanaian schools have undergone professional training in

testing techniques. Quagrain (1992) also indicated that some teachers in Ghana does not follow professional testing procedures. All these issues surrounding testing practices raise concern for in-depth study into testing practices of teachers in Ghana continually to address all challenges in the area.

Historical Development of Testing

Rudman, Mehrens and Wanous (1983) noted that few studies of teacher testing practices existed. Amedahe (1989) indicated, that "research work on classroom testing in Ghana was probably non-existent" (p. 15). According to Amedahe (1989), "teaching and learning can be likened to two sides of a coin. Without testing, the teacher would not be in a position to know whether the objectives set out to be achieved at the outset have been attained" (p. 1).

Tests constitute an integral part of the assessment process in education. The historical development of the field of measurement in education is closely related to similar developments in psychology. The Chinese set the pace for what is known today as "assessment under examination conditions". The practices of assessment had been established by the Chinese long before 1115 B.C. (DuBois, 1966; DuBois, 1970; Ebel, 1972; Lindon & Lindon, 1968; Kuo, 1915).

According to DuBois (1970), the first systematic programme of testing was initiated in China as far back as 2200 B. C. At that time, China had no hereditary ruling class, and initial appointments as well as continuance in employment were based on examinations. The tests covered the examinee's knowledge of civil law, military affairs, agriculture, revenue and geography. Civil servants were tested every three years. The Indian civil service adopted

the Chinese method of selecting civil servants. The United States of America also adopted the Chinese method for selecting civil servants in 1883.

Tests were developed for many occupations in the United States. As a means of validating the tests, attempts were made to show that there was a statistical relationship between scores on the test and later performance on the job. Formal testing in schools did not commence until in the 12th century by the Arabs using the Chinese method. Assessment by means of written tests was first used by Jesuits at St. Ignatio (DuBois, 1970). The development of academic tests was pioneered in Britain, particularly at the University of London. Under its initial stage, testing and awarding degrees were recognized as a legitimate basis for decision making (Ebel, 1972).

The birth of psychology as a science in testing took place in the laboratory of Wilhelm Wundt in Leipzig, Germany. However, Galton (1967) developed psychological testing by studying individual differences. Many of the methods in psychological testing can be traced to Alfred Binet who focused on the assessment of human abilities. Thorndike and Alfred Binet contributed immensely to the development of the current principles in testing. Thorndike (1903) published the first book in educational measurement. In this book, he indicated that whatever exists at all exists in some amount.

The historical antecedent of testing from ancient Chinese first developed test by Alfred Binet, introduction of test in the school system tells that tests have specific purposes and uses. Historically, tests were first used for making decisions regarding recruitment and selection of civil servants. Later it was introduced into the academic circles with the same mandate of basing decisions on it regarding awards of degrees and other certificates. Today

teachers in all levels of education base important decisions such as selection, promotion, certification just but to mention a few on tests result. This presupposes that assessment practices are so important and therefore call for special attention. The present study will therefore go into the assessment practices in the country, specifically assessment practices at the basic school level.

Teachers' Assessment Practices

Educational assessment is an important aspect of the teaching profession. It refers to the process used in the classroom by the teacher to obtain information about students' performances on assessment tasks, either as a group or individually, using a variety of assessment methods, to determine the extent to which students achieve the target instructional outcomes (Gronlund, 2006). Information gathered from the educational assessment is used for making various educational decisions including planning classroom instruction, placing students into learning sequences, monitoring students' progress, diagnosing students' learning difficulties, providing students and parents with feedback about achievements, evaluating effectiveness of teaching, and assigning grades (Nitko, 2001). A variety of methods are used by the teachers in their daily classroom assessment including traditional assessments such as multiple-choice, true-false, matching, completion, and short-answers and alternative assessments such as portfolios, student self-assessment, observations, and other performance-based assessments (Gronlund, 2006). The quality of these assessments and their consequences on teaching and learning depends on teachers' competence and knowledge in the

educational assessment (Alkharusi, 2011b, 2011d; Alkharusi, Kazem & Al-Musawai, 2011).

However, in a survey of assessment skills of 69 pre-service teachers, Volante and Fazio (2007) found that the self-described levels of assessment skills remained relatively low for the pre-service teachers across the four years of the teacher education program, thereby suggesting the need for in-service assessment training to ensure an acceptable level of assessment skills. Along similar lines, Wolfe, Viger, Jarvinen and Linkman (2007) proposed that teachers' self-perceived confidence in assessment should be a vital component in the professional development of in-service teachers. Further, Alkharusi (2009) found that assessment knowledge of pre-service teachers tended to vary as a function of gender. Specifically, in a survey of 211 pre-service teachers, Alkharusi (2009) found that males tended to have on average a higher level of knowledge in educational assessment than females. In a two-week classroom assessment workshop for seven in-service teachers, Mertler (2009) pre and post-tested teachers' assessment skills. The results showed that the assessment training had a positive impact on teachers' assessment skills as well as on their feelings regarding assessment and confidence in using assessment.

Moreover, when comparing assessment skills of 62 pre-service teachers and 71 in-service teachers, Johns and Davis (1991) found that both groups need a better understanding of standardized testing procedures and ethical considerations in test-wiseness. Green (1992) examined differences in opinions about testing between 236 pre-service teachers and 553 in-service teachers. Green (1992) found that pre-service teachers had less favorable attitudes toward classroom testing and more favorable attitudes toward

standardized testing than in-service teachers, thereby drawing teacher educators' attention toward the relevance of assessment training to the realities of classroom assessment. Likewise, in a study of 67 pre-service teachers and 10 in-service teachers, Mertler (2004) found that pre-service teachers demonstrated lower levels of assessment literacy in choosing and developing appropriate assessment methods; administering, scoring, and interpreting assessment results; and recognizing unethical assessment practices. Recently, Alkharusi, Kazem, and Al-Musawai (2011) compared assessment skills of 279 pre-service teachers and 233 in-service teachers from Oman. Results indicated that in-service teachers tended to have a higher level of perceived skillfulness in educational assessment than pre-service teachers, thereby testifying the value of including teaching experience when preparing teachers in educational assessment. In conclusion, the classroom assessment literature reveals some contradictions between teachers' practices and recommendations of educational assessment experts regarding issues of classroom assessment and grading, thereby necessitating the need for a better understanding of teachers' skills in educational assessment.

Furthermore, in a survey of 893 teachers in 34 schools, Bol, Stephenson, O'Connell, and Nunnery (1998) investigated teachers' frequent uses of traditional and alternative assessment methods in relation to teaching experience, grade level, and subject area. The traditional methods of assessment included close-ended examinations, quizzes, and other written assignments. The alternative methods of assessment included performance and observation-based assessment methods. Results showed that the most experienced teachers indicated the use of alternative assessment more often

than the least experienced teachers. Also, elementary school teachers reported using alternative assessment more often than did middle and high school teachers. Mathematics teachers reported using traditional assessment methods much less frequently than did teachers in other subject areas. In a related study, Snow-Renner (1998) examined teachers' assessment practices in Colorado classrooms relative to students' opportunities to learn. Survey responses were received from 737 mathematics and science teachers in grades 4, 8, and 10 as well as from 116 elementary school and 223 secondary school students. Elementary school teachers reported a greater emphasis on alternative assessments than did secondary school teachers. Students in different classrooms reported experiencing differential opportunities to learn relative to reform-oriented assessments, suggesting that teachers' assessment practices may affect students' perceptions of the classroom environment and that this effect may vary across classrooms. Snow-Renner (1998) attributed such results to fluctuations in teacher's capacity and knowledge about assessment and to ambiguous policy definitions of assessment reforms in Colorado.

Also, an investigation of classroom assessment practices of 246 third preparatory science teachers from 112 schools in Oman, Alsarimi (2000) found that teachers indicated using short answer, completion, oral exams, extended answer, and multiple-choice item formats with no significant differences based on teacher's gender and years of teaching experience. Also, Alsarimi (2000) found that the teachers indicated using four main sources of information when assigning grades to students: final exams, midterm exams, class participation, and oral questioning. Also, these same teachers tended to

incorporate some non-achievement factors such as student's effort in grading. The teachers commented that the grades reflect student improvement, effort, and knowledge of the subject matter. Recently, Zhang and Burry-Stock (2003) surveyed 297 teachers across grade levels and content areas about their classroom assessment practices. They found that mathematics and science teachers reported grading on non-achievement factors more frequently than did teachers in social studies and non-academic subjects.

The aforementioned studies tend to confirm that classroom assessment practices may be unique from one grade level, teaching experience, and subject area to another. It is also evident from the classroom assessment literature that there seems to be some contradictions between teachers' practices and recommendations of educational assessment experts. Therefore, teachers' assessment skills need a considerable scrutiny.

Assessment practice issues in Ghana

In Ghana, Amedahe (1989) in a study of the assessment practices of secondary school teachers in 18 secondary schools in the Central Region found that teachers lacked the skills and principles of test construction. The limited skills in testing practices and in the management of assessment practices were due to their inability to receive training in assessment practices. In another study of student assessment procedures in junior secondary schools in 11 districts in the country, it was also found out that teachers did not have adequate training in the management of assessment practices (Curriculum Research & Development Division [CRDD], 1999). It was reported that 55% of the teachers interviewed felt they were not confident in the testing and measurement practices because they did not have any training in testing and

measurement. Moreover, Etsey (2003) in a study of teacher trainees in 24 teacher training colleges, recommended that Teacher Education Division of the Ghana Education Service mandates the teacher education curriculum planners in the country to make courses in the management of assessment practices a priority in the first-two years in the training of teacher trainees in the teacher training colleges in the country.

Aidoo-Taylor (1993) pointed out that:

> Testing practices require a great deal of record keeping and frequent measurement of student performance. It means more work for teachers. Teachers would have to construct tests, and other forms of testing instruments, mostly on their own. It demands more dedication and professionalism from teachers, and the adjustment may be painful for some teachers (p. 35).

According to Akyeampong (1997) (as cited in Ghana Muster Research Report, 2000), the system of assessment practices in the teacher training colleges has virtually remained the same throughout teacher training reforms. No statements of standards have ever been developed to guide the teaching, learning and assessment practices in the teacher training colleges, until the introduction of the current Diploma (Basic Education) programme. As a matter of policy, the Institute of Education of the University of Cape Coast, expects each tutor in the teacher training college to give at least one class assignments/exercises and class test using different test formats, depending on

the number of credits of the course of study in a semester. Principals, Vice-Principals, Assessment Officers and Heads of Departments in the colleges are required to assist in management of assessment practices with regards to the effective management of college assessment committee, handling of continuous assessment scores of the students and putting the teacher -made tests scores into good use.

The above studies all conducted in Ghana actually suggest that teachers from the basic level, secondary school level even tertiary (specific to teacher training colleges) do face some challenges in their testing practices. Among some of the causes of this challenge were attributed to the fact that, the teachers' in question lacked the skill and training in the area of measurement and evaluation as well as policy direction in the training of teachers in the country. The question that readily comes to mind is that, is it always the case that teachers' assessment practices are been influenced by their knowledge in testing practices alone? Is it possible that in our local context, gender, age, subject area, teaching experience and qualification also influence assessment efficiency? This current study is specifically interested in the influenced of teacher characteristics such as gender, age, teaching experience and teacher qualification on assessment practices in Ghana.

Theoretical Framework

Achievement Goal Theory

Achievement goal theorists have traditionally identified two types of achievement goals: mastery and performance goals (Ames, 1992; Dweck, 1986; Nicholls, 1984). Mastery goals center on the development of competence whereas performance goals center on the outward showing of

competence (Ames, 1992; Dweck, 1986; Nicholls, 1984). Teachers who adopt mastery goals are expected to persist in the face of difficulty following high standard assessment practices, and have high intrinsic motivation (Ames, 1992; Dweck, 1986; Nicholls, 1984). In contrast, teachers' who adopt performance goals are expected to minimally persist in the face of difficulty in following approved standard of assessment practices, and have low intrinsic motivation (Ames, 1992; Dweck, 1986; Nicholls, 1984). Since its origin in the late 1970s and early 1980s, achievement goal theory has undergone a number of theoretical advances. For example, Elliot and Church (1997) and Middleton and Midgley (1997) have developed a trichotomous framework of achievement goals that further differentiates performance goals into approach and avoidance goals. In particular, three achievement goals have been identified: (a) mastery goals that focus on improving competence, (b) performance-approach goals that focus on displaying competence, and (c) performance-avoidance goals that focus on avoiding a display of incompetence (Elliot & Church, 1997; Middleton & Midgley, 1997). This model is crucial when scrutinizing competencies of teachers' assessment practices.

Classroom Assessment Environment Model

Brookhart (1997) has developed a theoretical model based on a synthesis of classroom assessment literature and social cognitive theories of learning and motivation. In this model, the classroom assessment environment is conceived as a classroom context experienced by students as the teacher establishes assessment purposes, assigns assessment tasks, sets performance criteria and standards, gives feedback, and monitors outcomes. Based on this

model, students' perceptions of the classroom assessment environment have been thought to influence students' motivational beliefs and achievement-related outcomes. This postulation has been examined for elementary, middle, and high school students as well as for college-level students (e.g., Brookhart & Bronowicz, 2003; Church, Elliot, & Gable, 2001; Rodriguez, 2004; Wang, 2004).

However, most research on the assessment environment has used individual student scores as the unit of analysis rather than the average score of students at the classroom level (e.g., Church et al., 2001; Wang, 2004). Proponents of this approach argued that students within the same classroom differ in how they interpret and perceive the various practices in the classroom as a result of differential treatment and their different prior experiences brought to the classroom (Ames, 1992). Nevertheless, Church et al. asserted that "composite measures of perceived classroom [assessment environment] have been shown to be internally consistent, and composite indicators yield a more comprehensive assessment of the perceived classroom environment than do individual indicators" (p. 51). Likewise, Brookhart (2004, p. 444) maintained that "classes have an assessment 'character' or environment" that originates from the teacher's assessment practices, and that "students construct their own meaning of the classroom assessment environment based in part on their group experiences" (p. 445).

The concept of classroom assessment environment was first introduced by Stiggins and Conklin (1992) as a result of their observations of the assessment practices of four teachers in three sixth grade classrooms. According to Stiggins and Conklin (1992), the classroom assessment

environment included eight key elements. These were assessment purposes, assessment methods, criteria for selecting the assessment methods, quality of assessment, feedback on assessment results, teacher's assessment background and preparation, teacher's perception of students, and assessment policy (Stiggins & Conklin, 1992). Given that of interest in the classroom environment research are "students' perceptions of the meaning" of the classroom assessment practices (Ames, 1992b, p. 264), Stiggins and Conklin's (1992) conceptualization of the classroom assessment environment centered more on teachers' practices than on students' perceptions of these practices (Brookhart & Durkin, 2003). As such, based on a synthesis of classroom assessment and motivation literature, Brookhart (1997) developed a theoretical framework for the role of classroom assessment in student motivation and achievement. In this framework, classroom assessment environment was construed as a classroom context experienced by students as the teacher establishes assessment purposes, assigns assessment tasks, sets performance criteria and standards, gives feedback, and monitors outcomes (Brookhart, 1997).

Classroom assessment environment model has to do with classroom context experienced by students as the teacher establishes assessment purposes, assigns assessment task, sets performance criteria and standards, gives feedback, and monitors outcome. Based on this model, students' perceptions of the classroom assessment environment have been thought to influence students' motivational beliefs and achievement related outcomes. Proponents of this model argued that students within the same classroom differ in how they interpret and perceive the various practices in the classroom

as a result of differential treatment and their different prior experiences brought to the classroom (Ames, 1992). With this background this present study is interested in finding out the kind of classroom assessment environment that are mostly created by teachers who have received in-service training on educational measurement and those who have not received any of such training. Again, it is of interest to this study, the kind of assessment environment employed by teachers with respect to gender.

Empirical Review

Types of Classroom Teacher-Made Tests

Assessment made by tutors of students' attainment, knowledge and understanding is called variously as teacher-made or classroom made test and school -based assessment (Amedahe, 1989). The rationale of teacher-made tests is linked with the constructivist model of learning. In this model, it is important to understand what the student knows and how he/she articulates it in order to develop his/her knowledge of understanding. In this model, it is learning with understanding which counts and to this end, information about existing ideas and skills is essential. Work in psychology and learning tells us similarly that for effective learning, the task must be matched to the student's current level of understanding (Gipps, 1992a), and either pitched at the level to provide practice or slightly higher in order to extend and develop the student's skills. If the new task is much too easy, the students can become bored, and if much too difficult, the student can become de-motivated (Gipps, 1992b).

Teacher or classroom- made test is essentially both an informal and formal activity. Tutors may pose questions, observe activities, and evaluate students' work in a planned and systematic or ad hoc way (McCallum,

McAlister, Brown & Gipps, 1993). Classroom tests are in most cases teacher – made tests.

Teachers have an obligation to provide their students with the best instruction possible. This implies that they must have some procedure(s) whereby they can reliably and validly evaluate how effectively their students have learnt what has been taught them (Mehrens & Lehmann, 1991). The classroom or teacher-made test is one such tool. Moreover, they are more likely to reflect today's curriculum. This is especially true in subject-matter areas as science and social studies.

Classroom or teacher-made tests can also be tailored to fit a tutor's particular instructional objectives in providing optimal learning to the student (Bejar, 1984). Without classroom or teacher-made tests, those objectives that are unique to a particular topic in the college might never be evaluated. Emphasis is always on the desirability and importance of the tutors being able to construct their test items based on the basic principles in measurement and evaluation.

A survey by Stiggins and Bridgeford (1985) on the uses of various types of tests [teacher-made objectives; standardized objectives; and structured (planned and systematically designed to include pre-specified purposes, exercise, observations and scoring procedures) and spontaneous (arises naturally in the classroom upon which the tutor makes a judgment of the student's level of development] reported that the tests are (a) for assigning grades and evaluating the effectiveness of an instructional treatment, (b) for diagnosis, (c) for remedial teaching, (d) to motivate students to learn to

improve in their work, (e) to provide the basis for guidance in selection and placement in the world of work and (f) for certification.

In a study of teachers in the United States of America, it was observed that in the pre-service education programmes, teachers were only trained to teach but not to assess (test) their students (Gullickson, 1986; Gullickson & Ellwein, 1985; and Marso & Pigge, 1989). Some of the major deficiencies commonly associated with teacher or classroom - made tests are; (a) ambiguous questions, (b) excessive wording, (c) lack of appropriate emphasis and (d) use of inappropriate item formats.

Classroom or teacher-made tests can take various forms. There are several ways in which items have been classified by format – supply and selection types; free answer and structured answer; essay and objective (Ebel & Frisbie, 1986; Gronlund & Linn, 1990; Thorndike & Hagen, 1977). Item types are grouped into two major categories – essay and objectives. Some educators argue that essay tests are more susceptible in scoring than the objective tests. However, classroom teachers exclusively use both since one cannot be used exclusively to measure all learning outcomes. According to Gronlund (1985) and Nitko (1983), essay forms are of two main types; extended response type and the restricted response types. With regard to the objective type tests, the multiple choice, short- answer/fill-in- the blanks, matching and true or false types are the most commonly used by tutors in the teacher training colleges in Ghana (Bartels, 2003).

Essay-Type Tests

According to Amedahe & Etsey (2003), an essay test item is a test that gives freedom to respondents to compose their own responses using their own

words. The essay test items consist of relatively few items, but each demands an extended response. There are two types of essay tests items. These are the restricted response type and the extended response type. The restricted response type limits the respondents to a specified length of the response while the extended response type does not limit respondents in the form and the scope of the answer.

Discussing some of the differences between the essay and the objective tests, Ebel & Frisbie (1986) pointed out the following:

1. Essay tests require an individual to organize and express his/her answers in his/her own words. In the essay or "free response" item, the student is not restricted to a list of responses from which he/she is to select the answer. Objective tests, on the other hand, require that the student either supplies a brief answer or choose the correct answer from among several alternatives.
2. An essay test consists of fewer questions but calls for lengthy answers. Efficiency and reliability are therefore likely to be superior in objective tests.
3. Different skills and processes are involved in taking the tests (Ward, Kline & Fluagher, 1986).
4. The quality of essay test is dependent largely on the skill of the person grading the answer and that of an objective test, on the skill of the test constructor.
5. Essay tests are relatively easy to prepare but more difficult to grade accurately since they are graded by humans (who may be subjective) rather than by impersonal machines.

6. Essay tests afford both the students and the tutors the opportunity to be individualistic. Objective tests afford this freedom of expression (item writing) only to the test maker.

Objective-Type Tests

An objective test requires a respondent to provide a briefly response which is usually not more than a sentence long. The objective test-items normally consist of a large number of items and the responses are scored objectively, to the extent that competent observers can agree on how responses should be scored (Amedahe & Etsey, 2003).

There are two major types of objective tests. These are the selection type and the supply type. The selection type consists of the multiple-choice type, true or false type and matching type. The supply type has variations as completion, fill-in-the blanks and short answer. Objective type test items are most useful when class sizes are very large and when there is limited time to submit the results of the test (Amedahe & Etsey, 2003). Objective tests are more susceptible to guessing; and the distribution is determined almost completely by the test (Mehrens & Lehmann, 1991). A true or false test item consists of a statement marked true or false. A respondent is expected to demonstrate his command of the material by indicating whether the given statement is true or false. One of the limitations in constructing this type of objective test is the probability of getting right answer by guessing is 0.5. It can be used to assess only a limited number of educational objectives, and can be used to evaluate definitions, facts, meaning of expressions, recognition, and interpretation of charts/graphs. One advantage of the true-false format is that it is suitable for classroom short–time evaluation. One format of the objective

test formats that mostly eliminates guessing is the completion type. The short answer is the type of objective test item that is known as the supply, completion, and fill-in-the blank. It consists of a statement or question and the respondent is required to complete it with a short answer usually not more than one line. It is useful for testing knowledge of facts or recall of specific facts (e.g. "knowledge objective" in Bloom's taxonomy of educational objectives). One demerit of the format is that it can be such that there is more than one answer, and that situation makes the scoring of the test subjective. A matching-type test is a special form of the multiple-choice format. It is not very flexible, though it is useful for testing variables that are compatible. The matching type of objective consists of two columns. The respondent is expected to associate an item in column A with a choice in column B on the basis of a well-defined relationship. Column A contains the premises and column B the responses or options.

A multiple-choice test item is a type of objective test in which the respondent is given a stem and then is to select from among three or more alternatives (options or responses) the one that best completes the stem. The incorrect options are called foils or distracters.

There are two types of multiple-choice tests. These are the single "best response" type and the "multiple responses" type. The single "best response" type consists of a stem followed by three or more responses and the respondent is to select only one option to complete the stem. The "multiple responses" type consists of a stem followed by several true or false statements or words. The respondent is to select which statement(s) could complete the stem.

The multiple-choice format can be used to evaluate educational objectives in the cognitive domain of learning (Bloom, 1956). This is widely used in the schools, including the teacher training colleges, and in national or public examinations. It is susceptible to guessing, but the probability of guessing correct responses decreases as the number of options increases. To minimize the probability of guessing, it is recommended that generally, items be made to have about five options. Thorndike and Hagen (1969) identified six functions served by teacher-made tests. These are (1) motivation; (2) diagnosis and instruction; (3) differentiation and certification; (4) guidance and counseling; (5) curriculum development; and (6) selection and placement.

Construction of Classroom or Teacher-Made Tests

Tests constitute an integral part of the assessment process in education. There is a general consensus among testing experts, (for example, Cunningham, 1986; Ebel, 1972; Mehrens & Lehman, 1991; Nitko, 1983) concerning what makes an item a good test item. It needs to be emphasized, however, that this knowledge is necessary but not sufficient to ensure good test construction. The test constructor needs to have commitment to construct a good classroom or teacher–made test because it takes a great deal of time and effort to do so. The specific instructions for constructing test items are divided into two categories; supply and selection items. The supply items include essay and short answer formats while the selection formats include matching, true or false, and multiple-choice item formats (Cunningham, 1986).

Good tests do not just happen. They require adequate and extensive planning so that the instructional objectives, the teaching strategy to be

employed, the textural material, and the evaluative procedures are all related. The steps or stages in any test construction have been delineated in most works. Mehrens and Lehmann (1991) identified the following stages and steps as being relevant to the construction of the classroom or teacher-made tests:

1. Specify the course or unit content
2. List the major course or unit objectives
3. Define each objective in terms of students behaviour
4. Discard unrealistic objectives
5. Prepare a table of specifications
6. Decide on the type of item format to be used.
7. Prepare test items that match the instructional objectives
8. Evaluate the degree to which the students have learned the objectives.
9. Revise the objectives and/or teaching material and/or test on the basis of the results.

Contributing to the stages and steps involved in the construction of a classroom or teacher-made test, Chase (1979) indicated five distinct stages. These stages are:

1. Designating the purpose of testing;
2. Choosing between norm-referenced and criterion –referenced approaches based on competencies;
3. Developing a table of specification. Here, two major principles are always involved when the classroom teacher decides to use the table of specification; i.e., preparing a framework of what the essential objectives of the learning instruction are to be assessed and the

appropriate test content that reflects the various topics treated to be assessed in making appropriate decisions.

It is very important to prepare the table of specifications before the beginning of instruction since the table would assist the tutor in organizing his/her teaching -all the resources needed to plan on using in teaching a course (Mehrens, 1984a; Mehrens& Kaminski, 1989). The use of a table of specification will ensure that (a) only those objectives actually involved in the instructional process are assessed, (b) each objective would receive a proportional emphasis on the test in relation to the emphasis placed on the objective by the tutor, and (c) no important objective or content area would be inadvertently omitted.

The table of specifications aids immensely in the preparation of test items, in the production of a valid and well-balanced test, in the clarification of objectives to both the tutor and the students, and in assisting the tutor to select the most appropriate teaching strategy.

Constructing appropriate test items formats. These are mostly of two main types, i.e., essay or objective test format. Among the objective test item formats are completion test items, multiple-choice test items, true or false test items, and matching test items preparing the test for administration

The following factors are also to be considered in the writing of teacher-made test items. The test constructor ought to:

1. Carefully define the instructional objectives,
2. Prepare a table of specifications, keep it before him/her, and continually refer to it as test items are written,
3. Formulate well-defined questions.

4. Avoid excess verbiage,
5. Have the test items based on information that the students know,
6. Try to avoid race and sex bias,
7. Write each test item on a separate sheet,
8. Prepare more items than actually needed by creating item banks,
9. Write and key the test items as soon as possible after the material has been taught,
10. Prepare the test items well in advance to permit review and editing;
11. Avoid specific determiners,
12. Be careful when rewording a faulty test items,
13. Insert some novelty into the test,
14. Prepare a scoring key or marking scheme while the items are fresh in your mind.
15. Give clear and concise directions for the entire test as well as sections of the test.
16. Avoid textbook or stereotyped language and not to concentrate on past and examinations questions (Amedahe & Etsey, 2003).

Thorndike and Hagen (1977) have listed some of the common faults of the objective tests. They are:

1. The content tends towards the trivial. Most of the time the test constructor limit themselves to the specifics and the factual.
2. The type of test exercises that the test constructor uses is poorly adapted to appraising the mental processes or content that the test is supposed to measure. Most of the time, the test constructor includes a number of different item types not because each is especially effective

for assessing certain objectives, but just to achieve a diversity of format.

3. Most of the questions test constructors write for tests are ambiguous. The wordings usually are either not clear, or the task to be performed is not specified. Students had to guess or use extrasensory perception to determine what the test constructor is asking and wants for an answer.
4. The characteristics of the tests are inappropriate for the purpose for which the test constructor wants to use the test results. As a matter of fact, seldom the test constructor explicitly states the purpose for which he/she wants to use the results.
5. Usually, most test constructors suddenly realize that it is time to give a test and sit down to write items with the idea of producing a certain number of items and never does plan a test to truly reflect their teaching emphasis.

Contributing to the discussion on the basic principles of test construction, Adamolekun, (1985) noted that in the writing of any classroom or teacher-made tests for whatever purpose, the classroom teacher should consider the following:

1. Identify the purpose of the test i.e. what the teacher wants to achieve by the test.
2. Select the test item type that will best measure the learning outcome.
3. Obtain a representative sample of student behaviour which the teacher would want to evaluate (e. g. in the cognitive domain; does the teacher want to test knowledge of facts, comprehension, application, analysis, synthesis or evaluation?)

4. Construct test items of the proper level of difficulty.
5. Try to eliminate factors that are extraneous.
6. Make a test which will contribute to improved teaching/ learning practices i.e. what will enable teachers to teach better and students to learn better.

For objective tests, whatever the format, the same general principles guide the tutor in the construction of the test. However, additional guidelines are:

1. Keep reading difficulty level to the appropriate level of the students.
2. Guard against one item providing the clue to another item.
3. Avoid interdependent items i.e. items which are deliberately based on another item.

Administration of Teacher –Made Tests

According to the Standards for Educational and Psychological testing National Council on Measurement in Education (NCME, 1999), the test administrator must establish conditions so that each examinee can do his/her best. The physical conditions should be as comfortable as possible, and the examinees should be as relaxed as possible, even though the evidence is inconclusive regarding the effects of physical and environmental conditions on test performance. Whereas distractions during testing have little effect on the scores of high schools and college students, young children may be affected (Anastasi, 1981; Gronlund & Linn, 1990; Thorndike & Hagen, 1977; Trentham, 1975) Lindquist (1951) further pointed that some types of improper administration of tests cause bias in the results for entire groups, and render

inter-group or intra-group comparisons of little value. Test administration is therefore, equally important as the preparation of the test.

It should be noted that individuals perform better at any endeavor, including test taking, if they approach the experience with positive attitude, and that students cannot perform better when they are excessively tense and nervous. It is known that test anxiety affects optimum performance (Clawson, Firment & Trower, 1981; Culler & Hollohan, 1980; Sarason, Hill, & Zimbrado, 1964). Things that create anxiety are (1) warning students to do their best "because the test is important"; (2) telling students that they must work fast in order to finish on time; (3) threatening dire consequences if they fail; and (4) threatening students with tests if they do not behave. When administering the classroom or teacher-made test, the test constructor should make sure that the students understand the directions of the test. In addition, the teacher should keep the students informed of time remaining.

Careful proctoring should take place so that cheating is eliminated. Many a time test administrators are seen in the examination rooms, but are found to be reading novels, marking scripts and doing other things that contribute to cheating during the administering of the test (Lindquist, 1951). Frary (1981) and Roberts (1987) pointed out that the best way to detect cheating is to observe students during the examination and not by being preoccupied at one's desk. Moreover, cheating during the administration of a test would occur when a test is poorly planned in which test items do not cover what was taught or focuses on straight memorization of trivial rather than upon reasoning skills. Discussing the basic principles related to the administration of classroom or teacher-made tests, Tyler (1950) pointed out

that certain rules have been useful in actual practice in the administration of any testing programme and for that matter, could be used for administering any classroom or teacher-made test.

These are:

1. Select the tests carefully, preferably in co-operation with a college assessment committee.
2. Order the tests well in advance of the date on which they are to be used.
3. Plan in detail for the administration of the tests. Choose examiners and proctors with great care.
4. Avoid overemphasis on the tests.
5. Seat the students in alternate chairs, if possible.
6. Make announcements slowly and clearly in a voice that is loud enough to be heard throughout the room.
7. Have the blanks on the front of the booklets filled out. Be sure to announce the date, how names are to be written, and other items that may need clarification.
8. Time the examination with extreme care, using a watch or a wall clock
9. Move about the room occasionally to see that students are working on the right part of the test, but do not stand gazing over a student's shoulder and do not constantly move nervously from student to student.
10. Stop the test immediately when the time is up and collect the answer booklets.

11. As soon as a certain test has been given, have all the students taking the test turn in their booklets promptly. Alphabetize and check the test papers against the class list.
12. Except in cases of illness, see that all absentees make up the test. This necessity will cause the tutor much trouble and worry, but it is unavoidable, for complete data are essential if the results of the test are to be used successfully in either teaching or guidance.
13. Endeavour to have the tests promptly scored.
14. Have the scores of each student entered on an individual cumulative – record card, and make the card available to both the students. The card may be shown to parents and guardians when necessary (pp. 61 – 63).

In addition to other studies, Amedahe & Etsey (2003) indicated that in ensuring maximum performance of students in the administration of teacher-made tests, teachers should follow certain guidelines. These are as follows:

1. Prepare students for the test. Provide information such as
    a. when the test will be given (date and time),
    b. under what conditions the test will be given, for example number of items, open book or closed book, place of test,
    c. the content of the areas the test will cover, for example, study questions or list of learning targets,
    d. emphasis or weighting of content areas,
    e. the kinds of items of the test (objective ─types or essay ─type tests),
    f. how the test will be scored and graded, and
    g. the importance of the results of the test.

2. Students must be made aware of the rules and regulations covering the conduct of the test. Penalties for malpractice such as cheating should be clearly spelt out.
3. Re-usable tests must be kept secure. Efforts must be made to keep students from taking out copies of the test.
4. Avoid giving tests immediately before or after a long vacation, holidays or other important events where all students are actively involved physically/emotionally.
5. Avoid giving tests when students would normally be doing something pleasant, for example having lunch.
6. The sitting arrangement must allow enough space so that students will not copy each other's work.
7. Adequate ventilation and lighting is expected in the examination room.
8. Provision must be made for extra answer sheets and writing materials.
9. Students should start the test promptly and stop on time.
10. Invigilators are expected to stand at a point where they could view all students. They should once a while move among the pupils to check on malpractices. Such movements should not disturb the students. The teacher/invigilator must be vigilant. Reading novels, newspapers, grading papers are not allowed in the examination room.
11. Threatening behaviors should be avoided by the invigilators, for example, speeches like 'if you do not write fast, you will fail'. Students should be made to feel at ease.
12. The testing environment should be free from distractions. Noise should be kept to a very low level if it cannot be eliminated or removed.

Interruptions within and outside the classroom should be reduced. It is helpful to hang a "Do not DISTURB – EXAMINATION IN PROGRESS" sign at the door.

Scoring Essay Tests

Because essay tests are less objective, there is a far greater risk that the marks that students receive on these tests, and the decisions made based on the marks, will not be valid or reliable. Therefore, it is very important that technical care be taken in developing, marking, and using essay tests. There are several ways that essay tests can be improved to contribute to better learning and teaching and to be more valid and reliable, including writing detailed specifications, writing prompts that clearly define the student's task, preparing test marking guidelines, and training markers and marking in groups (Capper, 2007).

In essay tests, the measurement of a student's ability does not end with his/her answer, but depends to a large extent on the person who reads his/her answer, as well as on the grading method used. The effectiveness of an essay test depends to a large degree on how well it is graded. Unreliable grading has been one of the major and most valid criticisms leveled against the use of grading (Charmey, 1984).

In grading essay tests, the test constructor must (1) use appropriate methods to minimize biases, (2) pay attention only to the significance of the answer, (3) be careful not to let personal idiosyncrasies affect the grading, and (4) apply uniform standards to all papers (Freedman, 1981). Undoubtedly, the uniformity of grading standards is a crucial aspect of grading essay tests. For without uniformity, there is no fair way of comparing students. Moreover,

without uniformity, one cannot be certain that the scores represent a valid measure of the students' achievement (Spandel, 1981).

Two commonly used methods that have been developed for grading essay tests are the analytical method and the global (holistic) method (Spandel, 1981; Stiggins, 1991; Frisbie & Griswold, 1989). In the analytical method, (also known as "point-score" method), the student's score is broken down into specific points contained in the answer. Component parts such as (1) effectiveness of expression, (2) logical organization, and (3) "support of statements" are specified and assigned points or values. The major advantages of the analytical method of grading essay tests are as follows:

1. It can yield very reliable scores when used by a conscientious tutor (Ward, 1980).
2. The very process of preparing the detailed answer may frequently bring to the test constructor's attention such errors as faulty wording, extreme difficulty and complexity of the question, and unrealistic time limits.
3. Hence, if the model answer had been prepared before the test was administered, the question could have been reworded or the time extended.
4. The fine subdivision of the model answer can make it easier to discuss the grade given to the student (Cunnigham, 1986).

Two major limitations of analytical method of scoring are that: (1) it is very laborious and time consuming, and (2) in attempting to identify the elements, undue attention may be given to superficial aspects of the answer (Diederich, 1967).

In holistic (or sometimes known as rating method or global method), the ideal answer is not subdivided into specific points or component parts. It simply serves as a standard. The scorer makes a single overall judgment of the quality of the writing sample. The rater reads the student's work, forms a general impression, and using some standard, assigns a rating to the response. No single factor is given undue weight. Rather, all factors are taken into account in forming the judgment of the response (Stanes, 1992). The crux of this method is to select papers that vary in quality to serve as anchor points to be awarded to the student's response. (Vacc, 1998).

According to Charmey (1984), the validity of holistic scoring has not been convincingly demonstrated. In addition, the process that test constructors follow when marking quick evaluations of a writing sample's quality (i.e. which features of the writing sample contribute most to the holistic score) is unclear (Freedman & Calfee, 1983). The most important limitation of holistic scoring is that it gives no meaningful diagnostic information beyond the comparative ranking it represents. Even if it is assumed that the score is reliable, the test constructor cannot tell much that he/she might want to know about the student. For example, a low score might represent an inability to control sentence structure, a major spelling incapacity, a total misreading of the question, or misguided attempt to be whimsical or creative. A high score might mean a correct but boring response or a genuinely creative piece of prose. All we have is a single score, where we might wish to have *a* profile (Vacc, 1998).

The second limitation on the value of holistic score emerges from its connection to its particular test group. It cannot represent an absolute value in

itself. A second-stage operation is needed to give meaning to the ranking. This means that every time a holistic scoring is completed, those responsible for reporting scores need to make a fresh decision about where cutting levels should be (Kirsch & Guthrie, 1980).

The third important limitation of holistic scores is also common to all test scores: reliabilities are customarily overestimated and inescapable inaccuracy of scores tends to be ignored. All tests yield approximations of ability levels, and even the most highly developed multiple-choice tests (which have almost perfect scoring reliability) would report a wide band of possible error (Tiedt, 1983). In view of the objectivity in scoring the essay tests, test developers have developed a number of principles to be followed in scoring essay tests to ensure to a greater degree, consistency of the scores. Most test specialists have agreed on these principles as useful in scoring essay tests (Ebel, 1972; Lindquist, 1951; Mehrens & Kaminski, 1989; Spandel, 1981; Tyler, 1950). Based on the literature, the principles for scoring essay tests can be summarized as cited from the works above. They are as follows:

1. Check scoring rubric (or marking scheme) against actual responses.
2. Be consistent in grading. Graders are human and therefore fallible, and graders should not be influenced by the first few papers they read. (Hales &Tokar, 1975).
3. Randomly shuffle the papers before grading.
4. Grade only one question at a time for all papers to reduce “the hallo effect” (the quality of the response to one question influence the reader’s evaluation of quality of the response to subsequent questions).
5. Try to grade all responses to a particular question without interruption.

6. Grade the responses without knowing identity of the students.
7. The mechanics of expression should be judged from the content. [Research has shown that when teachers are told to disregard spellings, punctuations, and grammatical errors in some papers which do not require them, they still assign lower grades to papers containing such errors in such papers (Chase, 1979; and Scanell & Marshall, 1966). Where they are relevant, they should be scored for their own sake].
8. If possible, have two independent readings of the test and use the average as the final score.
9. Provide comments and correct errors.
10. Set realistic objectives (Chase, 1979).

According to Amedahe and Etsey (2003), the following principles should be followed by teachers when scoring essay test items:

1. Prepare a form of scoring guide, either an analytic scoring rubric or holistic scoring rubric.
2. Grade the responses item by item and not script by script. The teacher should also score all responses to each item before going to the next item. This reduces carryover effect. The carryover effect occurs when the mark for a question is influenced by the performance on the previous question.
3. Keep scores of previously graded items out of sight when evaluating the rest of the items.
4. Score the essay test when you are physically sound, mentally alert and in an environment with very little or no distraction.

5. Constantly follow the scoring guide as you score. This reduces the rater drift which is the tendency to either not paying attention to the scoring guide over time or interpreting it differently as time passes. Also avoid being influenced by first few papers read. [These could make the teacher either too harsh or too lenient].

Grading Practices among Teachers

In the simplest sense, a grade is a type of assessment judgment and 'grading' is the process of assigning a grade. Newton (2007) proposes that grading should not be discussed as a specific assessment purpose at the decision level; rather, it should be considered at the judgment level as purely a standards-referenced technical process. However, as McMillan (2008) pointed out, even when teachers use the same grading scale and the same grading guidelines, there is little consistency in teachers' grading across schools. Teachers vary considerably in terms of weighting different factors in determining grades and unforeseen unique situations constantly arise in the classroom setting, which require teachers to make professional decisions. There is ample evidence derived from studies on classroom teachers' assessment and grading practices that grading is a complex decision-making process influenced by various internal and external factors. McMillan and Nash (2000) proposed a model of teachers' classroom assessment and grading decision-making including both internal and external influencing factors. The most salient internal factor was the teacher's philosophy of teaching and learning. The major external factors were identified as mandated statewide learning factors and high-stakes tests, district grading policies and parents.

McMillan and Nash's (2000) model has been supported by studies conducted in other contexts. For example, Cheng and colleagues (Cheng, Rogers, & Hu, 2004; Cheng, Rogers, & Wang, 2008; Cheng & Wang, 2007) compared classroom assessment practices including grading practices by teachers of English as a second or foreign language (ESL/EFL) in three tertiary institutional contexts: Canada, Hong Kong and China. Their studies showed that these teachers' preferences for different methods of assessment and grading were influenced by the beliefs they held about assessment, their assessment purposes, their teaching experiences and educational training, the nature of their instructional contexts such as the goal of the programme, class size as well as the dominance of external large-scale high-stakes standardised testing. Zoeckler (2007) examined how American high school English language teachers attempted to arrive at a fair grade while weighting both achievement and non-achievement factors and the role of teachers' expectations. Results of this study indicated that grading was influenced by the local grading systems, teachers' perceptions of student effort, and their concerns for moral development.

McMillan (2008) argued that one of the most difficult issues in grading is how to deal with non-achievement factors such as effort, work habits and motivation. He refers to these factors as academic enablers. Teachers tend to consider these non-achievement factors in grading because they are traits that teachers cultivate and regard as important for students' ultimate achievement. In a questionnaire survey with elementary, middle, and high school teachers in the USA, Randall and Engelhard (2010) found that under most circumstances, these teachers abided by the official grading policy of the participating school

district, assigning grades based primarily on achievement. However, in some borderline cases, they rely more heavily on other student characteristics such as motivation, behaviour and effort. Simon, Chitpin, and Yahya (2010) investigated teacher candidates enrolled in the Bachelor of Education programme in a Canadian university. Their study found that student effort, participation, and late or missed assignments were identified as key concerns in the participants' thinking about grading.

Teachers' values and beliefs about teaching/learning and their considerations of the purposes and consequences of grading provide a rationale for their grading decision- making. Brookhart's (2004b) systematic review of literature in classroom assessment found that the practice of educational assessment occurred at the intersection of three practical bases: instruction, classroom management and classroom assessment, and at the intersection of three theoretical bases: psychology, sociology and measurement. She recommended that in order to evaluate the meaning, value, accuracy, and consistency of classroom assessment information, the intersection nature of classroom assessment should be acknowledged. Using Brookhart's (2004b) framework, Simon et al. (2010) found that the pre-service teachers in their study attached greater importance to assessment for classroom management, student motivation, and social justice purposes, than to support learning. In an earlier paper, Bishop (1992) argued that teachers cannot act as judges and coaches at the same time and suggested that teachers should give up the judging role to external assessment and focus on developing mentoring relationships with their students to fully function as coaches.

Grading policies have a direct effect on the grades that students receive, it is extremely important that schools carefully consider what practices best measures students' performance (Reeves, 2008). Grading practices have long be a controversial issue among educators and academics. Through grades are accepts as a standard and inherent part of education system, there is some disagreement as to what exactly is the function of grades. There has been much debate over whether grades should be designed to communicate a student's performance in variety of areas, including behaviour and participation or whether they should just represent a student's proficiency in a given subject. Some educators have even questioned the value of using grades at all, claiming that using extrinsic rewards to reinforce learning teaches students to care more about their performance on assessment than on what they learn (Edwards, 1999). The grading practices used by many teachers are designed to communicate student's performance in a number of areas, including both academic achievement and behavioral factors such as student effort, conduct and attitude (Allen, 2005). When teachers assign grades, especially final grades, they are communicating a number of messages to students with single mark. According to Zoeckler (2007), teachers often attempt to communicate message that include; level of expectations, level of academic achievement, encouragement and disappointment. Educators often use grades as both a punishment for bad behaviour and a motivational tool for good behaviour (Wormeli, 2006).

How ever. Some educators now recommend that grades should not be based on behaviour and other non-academic factor, but only on student's mastery of the material in a given subject. Grading students on what they

know and can do, and not on other factors, will help teachers provide students and parents with specific feedback on what learning areas need improvement (Andy, 2011). Grading only on achievement is a key elements of standards-based grading, a practice that is gaining in popularity that focuses solely on student's proficiency on well-defined course objectives (Scriffiny, 2008). Scriffiny (2008), summarized seven common characteristics of a standards-based grading which include;

1. Students are graded either entirely or almost entirely on how well they progress towards learning objectives.
2. Standards based grading system measures only a student's most recent level of mastery over the course materials.
3. In order to avoid distorting student's grades away from their actual level of proficiency, standards-based grading only incorporates summative assessments such as tests or essay, not assessments like homework.
4. Information from formative assessments can be used to provide valuable feedback to both the student and their parents.
5. Students can redo summative assessment until they have demonstrated proficiency.
6. Many standards-based grading system uses rubrics. Rubric define the specific learning criteria against which teachers will compare a student's proficiency level.
7. Standards-based grading systems often use a scale different from A, B, C, D and F to record student's grades on report cards. One common

scale is 4, 3, 2 and 1. The score provided in a standards-based system correspond to performance standards.

The most fundamental measurement principle related to meaningful assessment and grading is the principle of validity (Linn &Gronlund, 2000; Stiggins, 2001). Validity is important because the sole purpose of grades is to accurately communicate to others the level of academic achievement that a student has obtained (Snowman & Biehler, 2003). If the grades are not accurate measures of the student's academic achievement, then they do not communicate the truth about the level of the students' academic achievement. Unfortunately, as stated by Cizek (1996), grades continue to be relied upon to communicate important information about academic performance and progress. The major reason for assigning grades is to create a public record of a student's academic achievement that can accurately and effectively communicate to others the level of mastery of a subject a student has demonstrated (Airasian, 2001; Linn & Gronlund, 2000; Oosterhof, 2001; Stiggins, 2001). Nitko (2001), points out that grade are used by students, parents, other teachers, guidance counselors, school officials, post-secondary educational institutions and employers. Therefore teachers must assign grades with utmost care and maintain their validity.

A study by Baron (2000) shows that there is lack of coherence in the beliefs about grades held by parents and students and those held by education community. Even in the same schools, teachers often hold very different views about the purposes of grades and fail to communicate with their colleagues about their grading practice. Friedman and Frisbie (1995), made a particularly strong argument for making sure that report card grades accurately report

information to parents about a student's academic progress and that teachers and administrators shares a common understanding of what information a grade should communicate. They suggested that since grades becomes part of students' permanent record, the purpose of these grades must be to communicate a valid summary of a student's academic achievement in the subjects. Grading systems used by teachers vary widely and unpredictably and often have low levels of validity due to the inclusion of nonacademic criteria used in the giving of grades (Allen & Lambating, 2001; Brookhart, 2004; Frary, Cross & Weber, 1993; Olson, 1989). Teachers have been found to make decisions about grades related to student efforts in attempts to be fair in their grading practices (Barnes, 1985). It has been shown that grades are used as a motivational tool as well as to develop good study habits and desirable classroom management behaviour (Oosterhof, 2001).

Nevertheless, non-academic factors are often used as criteria for assigning grades because some teachers consider the consequences of grades more important than the value of clear communication of information and the interpretability of the grades (Brookhart, 1993). It follows then that instead of the grade being a function of what a student has learned, it has become a function of many variables. Since, important decisions are often based on a student's grade. Grades can open up or close down important learning opportunities for students (Jasmine, 1999). With high grades students get admitted to colleges and universities of their choice and receive scholarships and tuition assistance, since grades are a major selection criterion in the college and university admission process. More so, it is very difficult for students to get admitted to some schools, if their grades are not sufficiently

high. Therefore, invalid grades that understate the students' knowledge may prevent a student with ability to pursue certain educational or career opportunities. Based on principles of attribution and social cognitive theories, when students receive grades lower than ones that accurately show their true level of academic knowledge, it may lead students to believe they lack the ability to succeed academically and lower their sense of self-efficacy as well as their motivation to learn (Pintrich & Schunk, 2002).

In an attempt to explore grading practices, issues of judgment, communication, and character development in grading through a framework which exposes the underlying moral issues in grading, (Zoeckler, 2007) examined how teachers arrived at a fair grade while weighing both achievement and non-achievement factors. The role of teacher expectations were also examined using a theoretical framework which considers grading processes in terms of truth, worthwhileness, trust, and intellectual and moral attentiveness. Zoeckler (2007) collected data from rural high school teachers in upstate New York through interviews. What emerged in this study was that teachers continue to struggle with issues of fairness as they grade students' work. The main argument that Zoeckler made was that teachers' grading and feedback to students is influenced by teachers' values and beliefs. Zoeckler argued that even though teachers' moral issues in assessment often go unexplained, they play a major role in the assessment practices they adopt.

McMillan (2008) conducted a study to document the differences in actual assessment and grading practices conducted for a specific class taught by teachers across a range of subjects. Results of the study indicated that secondary school teachers use a multitude of factors when grading students

work. A mixture of factors to determine grades were organized into four clearly distinct components: academic achievement, academic enablers (such as effort, ability, improvement, and participation), use of external benchmarks, and use of extra credit and borderline cases). Academic achievement was considered to be the most important process in grading students' work. Two thirds of teachers who participated in this study were reported to have agreed with the use of academic enablers (effort, ability, and improvement) when grading students' work. Teachers defend their choice to use non-achievement factors, such as effort, because they see them as some form of borderline to determine grades; and that such factors are good proxy for student achievement. Both primary and secondary school teachers have been found to use non-achievement factors when they award grades to their students to raise or lower grades except in borderline cases, they were also found to reward hard work by raising borderline grades and some would lower borderline grades for lack of effort.

Lekoko and Koloi (2007) conducted a survey with pre-service teachers enrolled in education classes at the University of Botswana. The purpose of this study was to explore students' perceptions regarding the correlation of teacher's feedback and the grades that teachers award to students. Students reported some experiences regarding how their work was graded and the nature of feedback they receive from their lecturers. This study showed that when lecturers grade students' work they did not provide adequate comments that could help students understand where they went wrong, teachers gave low marks that are not accounted for in terms of what and how the teacher arrived at the marks, there was no reconciliation of marks and comments

accompanying them, and teachers made ticks that were incompatible with the marks given. The main argument that Lekoko and Koloi (2007) made in this study was that when there is a discrepancy between teachers' comments and grades that students receive, students are left frustrated as this robs them of the potential to improve in their learning. For this reason, it is essential that teachers should be given sufficient assessment training that would enhance their grading practices and equip them with skills of giving effective, efficient, and useful feedback to students.

In sharing a narrative perspective on views about grading and giving students feedback Wormeli (2006) emhasises that "Assessment and feedback, particularly during the course of learning, are the most effective ways for students to learn accountability in their work and in their personal lives" (p. 14). Wormeli stressed the need for teachers to use grading and feedback practices that can best serve the interests of students. Some of the recommendations that Wormeli made were that when grading and giving students feedback, teachers should clearly show what students did, what they were supposed to do, and then help them compare and contrast the two.

The Role of Teacher-Made Classroom Assessment

Assessment is a systematic process for collecting information that can be used to make inferences about characteristics of people or objects (Reynolds, Livingstone, & Wilson, 2009). Assessment is not just about collecting data, but is also a processes used to appraise students' knowledge, understanding, abilities or skills and it is inextricably linked to a course or program's intended learning outcomes (Marriot & Lau, 2008). The overall scope of assessment can be viewed within five main dimensions:

> *(1)Why assess*? Deciding why assessment is to be carried out and what outcomes the assessment is expected to produce. (*2)What to assess?* Deciding, realizing or otherwise coming to an awareness of what one is looking for in people being assessed. *(3)How to assess*? Selecting from among available means, those assessments we regard as being most truthful and fair for various sorts of valued knowledge. (*4) How to interpret*? Making sense of the outcomes of the observations or measurement or impressions we gather through whatever means we employ; explaining, appreciating, and attaching meaning to the raw 'events' of assessments. *(5) How to respond*? Finding appropriate ways of expressing our response to whatever has been assessed and communicating it to those concerned. (Rowntree, 1977, p. 11).

These dimensions make an important contribution to the framework in which classroom assessment practices should be viewed. All these are a clear indication that classroom assessments play an integral part of the teaching and learning process. Teachers have a wide range of classroom assessment methods to employ. These methods provide teachers with access to powerful assessment of students' learning. Through classroom assessments, teachers collect various forms of information in order to make informed, consistent, and appropriate judgments regarding students' learning outcomes. Teachers and school administrators are the main decision makers on the forms of assessment and specific assessment tasks employed in schools (Cavanagh, Waldrip, Romanoski, Dorman, & Fisher, 2005). Teachers control classroom assessment environments by choosing how they assess their students, the frequency of these assessments, how and when they give students feedback.

McMillan (2008) found that "Assessment of students at classroom level is very critical because effective decision making is based to some extent on the ability of teachers to understand their students and to match actions with accurate assessments" (p. 5).

Perceptions about Subjective and Objective Items

Debates regarding students' assessment methods have always been top agendas in many educational forums. Educators have had divided opinions on the best methods of assessing student learning outcomes. Some educators advocate for the use of traditional forms of assessments such as multiple choice tests and other forms of objective tests, others advocate for contemporary views of assessments such as portfolios, journal critiques, and research essays. Those who support traditional forms of assessment believe that such tests are more focused on improving the cognitive side of instruction, i.e. the skills and knowledge that students are expected to develop within a short period of time, are based on fair grading practices, have higher content coverage (Segers & Dochy, 2001; Linn, Baker & Dunbar, 1991).

Those who support the use of alternative assessments argued that traditional forms of assessment (multiple-choice and true-false) test facts and skills in isolation, seldom requiring students to apply what they know and can do in real-life situations.

They also indicate that traditional forms of assessments do not match the emerging content standards. They argued that over-reliance on this type of assessment often leads to instruction that stresses basic knowledge and skills. They argued that if teachers want to measure students' ability to engage in some form of a debate, write a poem, tune an engine, use a microspore, or

prepare a meal, authentic assessments are the only ones that can measure such skills rather than traditional forms of assessments (O'Day & Smith, 1993; Reynolds, Livingston & Willson, 2009).

The road from theory to practice seems to be a rocky one, as both assessment forms have acceptable qualities and in particular traditional forms of assessments cover an acceptable range of curricular content, indicating high levels of content and construct validity. Just like other forms of assessments, traditional tests are also focused on improving the cognitive side of instruction, i.e. the skills and knowledge that students are expected to develop within a short period of time (Segersand & Dochy, 2001). Even though alternative assessments are still at their infancy stage; there is an emerging body of literature that indicates their benefits in teaching and learning (Hargreaves, Earl, & Schmidt, 2002; Kleinert, 2001).

A number of studies have been conducted to ascertain teachers' perceptions about the use of alternative assessments and there is evidence that though these assessments are seen as valuable, teachers hold negative perceptions about them. Caeter (1984) conducted a study with secondary school teachers to find out: (a) Teachers' accuracy in identifying particular skills being tested in a widely used criterion-referenced reading test, (b) Teachers perception of personal ability to write items similar to those written for formal tests, (c) Teachers practice in developing items for classroom tests. Results in this study showed that many teachers were unable to recognize particular skills being tested by individual items that they used for testing students, and that teachers had more problems in constructing test items that

test higher order skills such as essay items but felt more skilled in constructing items that test basic skills such as objective items.

A study conducted by Kleinert, Kennedy, and Kearns (1999) indicated that teachers expressed levels of frustration in the use of alternative assessments. Some major issues that teachers have against the use of alternative assessments are that they require more time for students to complete, and for teachers to supervise and assess. Teachers were also concerned about competencies they have in reliably grading these forms of assessments and that such assessments are more teacher based than student based. Findings from these studies open more points of view regarding whether teachers possess necessary skills needed for using different types of tests, and the perceptions they hold regarding the use of different types of tests in classroom assessment practices.

Teacher Beliefs about Classroom Assessment Practices

"Teacher beliefs" is a term that has been viewed with differing perspectives because of its complexity. For instance, Oliver and Koballa, (1992) conducted a study where they asked teachers to give a definition of "teacher beliefs". Teachers gave differing perspectives; some associated "belief" with other psychological constructs such as knowledge, values, and attitudes. Others viewed "belief" as a process that influences their behaviors, attitudes, and practices. Some researchers simply concluded that because of its complexity, the term "teacher beliefs" cannot be defined easily (Cantu, 2001). McMillan and Nash (2000) held a discussion with teachers regarding their beliefs, values and purpose of classroom assessments as well as their grading practices. This discussion showed that teacher beliefs and values were not

directly linked with measurement principles. Rubie-Davies, Flint, and McDonald (2011) argued that even though teacher beliefs have been found to play a major role in influencing their thoughts and behaviors that contribute to student learning outcomes; they are less studied, compared to students' beliefs. A practice that McMillan (2005) and Popham (2008) argued against as they assert that research on teacher beliefs must be intensified and conducted on continuous basis, particularly because understanding teacher beliefs can lead to better ways of understanding their classroom practices.

Regardless of these mixed perceptions and misconceptions surrounding the concept of teacher beliefs, some researchers hold the view that "teacher beliefs" form an important process as they are associated with what they know or perceive as important. This may influence how teachers conceptualize their work and contribute to decisions, behaviors and practices they display in their daily classroom activities (Boog, 2003; Mansour, 2009). Teacher "beliefs" play an integral role in teaching and learning that include assessment practices that teachers adopt (Fang, 1996). Based on their beliefs, teachers adopt various assessment practices, indicating that such practices are not constant, but keep on changing, making it pertinent to study them on continuous basis (McMillan, 2008; Popham, 2008).

The complexity of "teacher beliefs" has led to ways of understanding assessment practices adopted by different groups of teachers. Based on their belief about classroom assessments, teachers can be classified into three main sub-categories. The first group is made up of realists. Realist teachers believe in the use of paper and pencil types of assessments where learners are expected to recognize rather than generate their own answers (Segers &

Dochy, 2001; Windschitl, 1999, Nitko, 1994). These types of assessments are focused on improving the cognitive side of instruction, i.e. the skills and knowledge that students are expected to develop within a short period of time (Segers & Dochy, 2001). Realist teachers by nature believe in norm-referenced testing. In norm-referenced testing students' mastery of core knowledge and skills of the curriculum and is evaluated relative to the performance of others (Nitko, 1994). Realists' teachers tend to rely more on paper and pencil objective tests that can be scored easily and be used to compare students.

The second group of teachers is made up of contextual teachers, who are more likely to use alternative assessments such as student portfolios, group-work assessments and performance based assessments. Contextual teachers believe that there is a changing perspective in classroom assessments. The changing perspective is driven by the need to use classroom assessments that recognize, teach and assess knowledge, skills, and abilities that students need beyond class environments. Contextual teachers believe that overreliance on the use of traditional methods of assessments such as multiple choice tests, true or false and other related types of tests only measure the recall of knowledge instead of higher level learning skills. They advocate for the increased use of performance testing that seem better suited for testing complex mental abilities like extended writing and problem solving skills (Haladyna, Downing & Rodriguez, 2002).

Contextual teachers are more likely to use criterion based testing evaluation to determine what students know and don't know based on a set criterion (Tzuriel, 2000; Nitko, 1994).

The third group of teachers is made up of relativists. They base their assessment practices on the developmental theory. They believe that children learn best in classrooms or environments where instruction is developmentally appropriate. They take into account that students developmental levels vary. Relativists' teachers believe that children have opportunities to learn and be assessed in different ways to address the learning mode that is most appropriate for each child's unique developmental level (Schunk, 2008; Siegeler, Deloache, & Eisenberg, 2003; Steinberg, 2008). As they believe that students have different learning needs, relativist teachers therefore use multiple assessment practices that accommodate students' diverse needs such as written tests, oral presentations, visual, technological presentations, drama, media and so on (Hargreaves, Earl, Moore, & Manning, 2001).

Classroom Assessment Preferences, Practices and Professional Development

Assessment preference refers to an imagined choice between alternatives in assessment and the possibility of practicing these assessment alternatives (Van de Watering, Gjibels, Docky, & Van de Rijt, 2008). While assessment practices refers the activities that teachers do in relation to conducting classroom assessment- from test planning to reporting to utilization of test results (Gonzales & Fuggan, 2012; Volante & Fazio, 2007; Zhang, 1995; Zhang & Burry-Stock, 2003).Studies indicate that teachers with very minimal participation in professional development programs may not be adequately prepared to meet the demands of classroom assessment (Galluzo, 2005; Mertler, 2009; Shaffer, 1991; Volante & Fazio, 2007; Zhang, 1995). Teachers are expected to be skillful in collecting information about the

performance of their students in school (Ames, 1992; Bennet & Gitomer, 2009; Nitko & Brookart, 2007; Harlen, 2007; Musial et al., 2009). However, assessment problems are particularly prominent in performance assessment, interpretation of standardized test results, and grading procedures (Zhang & Burry-Stock, 2003). Teachers also are less aware of the pedagogical, managerial and communicative aims of classroom assessment.

To be effective, teachers must be aware that it is not enough to present a lesson to their students and hope that they understand it (Sato et al., 2008). They should realize that learning occurs when there is interplay between the teaching process and the outcomes (Bond, 1995; Mory, 1992). When teachers assess learning, they should be able to identify specific goals and objectives for each subject or lesson (Segers, Dochy, & Cascallar, 2003; Jones & Tanner, 2008), systematically gauge the extent to which these anticipated outcomes actually occur and determine to what degree learning takes place (Raty, Kasanen, & Honkalampi, 2006; Stiggins & Chappuis, 2005). In addition, when teachers do assessment in the classrooms, they also are required to define the role of assessment in making instructional and educational decisions (Danielson, 2008; Stake, 2004; Stiggins, 2008). Teachers more often become immersed in their jobs and lose sight of the exact purpose of assessment.

Consequently, teachers may not achieve their purpose for assessment or they overlook another form of assessment that may be more appropriate (Rust, 2002). Hence, it is very important that teachers conduct assessments with a clear purpose in mind and believe that their assessment promotes excellence in students (Astin, 1991; Earl & Katz, 2006; Hill, 2002; Murray, 2006; Sanchez & Brisk, 2004). On the other hand, effective schools need to

rethink the roles of assessment. Here, "effective" means maximizing learning and well-being for students. Hence, two questions need to be answered. First, what uses of assessment are most likely to maximize student learning and well-being? Second, how can assessment be used best to facilitate student learning and well-being? The usual response to these questions would be to provide professional development and training to teachers on classroom assessment and how to maximize the information gathered from assessment. Research studies report that teachers' assessment and evaluation practices are still incongruent with recommended best practices (Galluzzo, 2005; Mertler, 2003, 2009; Volante & Fazio, 2007; Zhang & Burry-Stock, 2003).

Assessment Practices among Teachers

Sofo, Ocansey, Nabie, and Asola (2013) conducted a study on assessment practices among secondary physical education teachers in Ghana. Participants included a purposive sample of 63 secondary physical education teachers (43 males and 20 females). An open-ended questionnaire was used in data collection. Data were analyzed using qualitative content analysis comprising of both inductive and deductive analyses. The findings of the study indicated that teachers used observation, skill test, knowledge test, demonstration, peer observation, and oral report as assessment practices. The deductive analysis of the study further indicated that the highest percentage of assessment practice utilised by the teachers in practical lessons was teacher observation (70.11%), while that with the lowest percentage was oral report (1.15%). Teachers also reported knowledge test (81.43%) and individual report (7.14%) as the highest and lowest percentage of assessment practices used in their theory lessons respectively. The findings of the study also

indicated that most of the teachers who participated in the study used assessment for documenting learning, rather than for accountability purposes.

Benzehaf (2017) conducted a study by exploring teachers' assessment practices and skills. The study focused on High schools English teachers in the town of El Jadida, Morocco. Both quantitative and qualitative data were gathered for the purpose of the study. Forty (40) respondents which consisted of 24 male teachers and 16 female teachers were randomly sampled for the study. The respondents were English teachers in high school in the town of El Jadida. Both questionnaires and interview guides were used in data collection. The findings of the study indicated that teachers used varied number of assessment strategies ranging from home assignments to in-class written tests but mainly for summative purposes. With reference to frequency of assessment, 72.5% of the respondents indicated that they assessed their students once a month; whereas 27.5% of the teachers stated they did so twice a semester. The findings of the study further showed that respondents also used a number of assessment practices. Among the assessment practices used by teachers included essay questions, true/false questions, and fill-in-the-blanks.

Zhang, and Burry-Stock (2003) investigated teachers' assessment practices and self-perceived assessment skills within the framework of classroom assessment literature and the Standards for Teacher Competence in Educational Assessment of Students. Factor analytical technique was applied to study the relationship between the constructs of assessment practices and self-perceived assessment skills. Teachers' assessment practices and self-perceived assessment skills were examined in a MANOVA design to

determine how they may vary as a function of teaching level, content area, teaching experience, and measurement training. The constructs of assessment practices and self-perceived assessment skills overlap to some extent in the underlying dimensions they measure, yet each contains a certain degree of uniqueness. The similarity between assessment practices and self-perceived assessment skills was supported by a strong correlation coefficient of .71 and by similar patterns of item loadings on four of the underlying dimensions they measure (paper-pencil test; standardized testing, test revision, and instructional improvement; performance assessment; and non-achievement-based grading). Where the two factor structures differ, the construct of assessment practices does a better job of subsuming inherently related activities under the same dimension than does that of self-perceived assessment skills. The findings of the study indicated that teachers differ in their assessment practices due to the nature of classroom assessment delineated by teaching levels. A general difference emerges between elementary and secondary teachers in terms of the assessment methods used and teachers' concerns for assessment quality. While secondary teachers rely mostly on paper–pencil tests and were concerned about the quality of assessment, elementary teachers often use performance assessment as an alternative.

Ogunkola (2013) investigated the instructional assessment practices pattern, techniques and challenges of science teachers in Barbados. The study employed a cross-sectional survey design. The population for the study included all science teachers in secondary schools in Barbados. Science teachers included teachers of Biology, Chemistry, Physics, Integrated Science, Agricultural Science, and Human and Social Biology. Twelve out of 22

secondary schools were randomly selected. All the science teachers at the twelve schools formed the sample for the study. Data was collected from fifty five secondary school science teachers. Descriptive statistics (frequency and percentages), one way Analysis of Variance (ANOVA), as well as one-sample t-test were used in analyzing the research questions. The findings of the study revealed that all the teachers used collaborative assessment practices. Such practices include discussing assessments with peers before administering and learning from other teachers through sharing of ideas about assessment practices that work. Ninety-six percent of the teachers reported that they assess students' skills and follow guidelines of the assessment development process. The findings of the study further revealed that there was no significant difference in the instructional assessment practices of the male teachers (M = 105, SD = 7.084) and the female teachers (M= 104.50, SD = 6.584); (t = 0.269; df = 53; p = 0.789). That is to say, both males and females did not differ in terms of the use of assessment practices. The findings of the study also revealed a non-significant difference in the teachers' instructional assessment practices and their experience, professional qualification as well as their academic qualification.

Test Construction Practices among Teachers

Agu, Onyekub and Anyichie (2012) conducted a study that was geared towards the development and validation of an instrument for assessing test construction skills of secondary school teachers in Nigeria. The study developed and validated a Test Construction Skill Inventory (TCSI) for assessing the secondary school teachers' competencies in constructing classroom-based tests. Factor analysis was done on the 30- item instrument

developed by the researchers. 25 items were found to be factorially valid. Five hundred and forty-three (543) secondary school teachers in Onitsha education zone, Anambra state, Nigeria were used for the study.

Proportionate stratified random sampling was used to ensure that teachers from different local government areas that make up Onitsha education zone were selected in the same proportion as they existed in the population. To do this, 40% of teachers in each of the three local government areas in Onitsha education zone were sampled (292 from Onitsha North LGA; 107 from Onitsha South LGA; and 268 from Ogbaru LGA). The sample comprised of 120 male and 423 female teachers. Teachers sampled for the study with less than ten years teaching experience were taken as 'less experienced' (n = 200) while those that have been in the teaching profession from ten years upwards were taken as 'experienced teachers' (n = 323). Cronbach alpha was used to analyze data for research questions 1 and 2 while mean and standard deviation statistical procedures were used to analyze data for research question. The Test Construction Scale Inventory was found to be reliable with a coefficient of 0.73 and the secondary school teachers found almost all the 25 items important skills for quality classroom-based test construction. The finding of the study revealed that almost all the teachers took the following skills into consideration while constructing test items: outlined the content covered for the term before setting test items form the content covered, prepared a test blueprint as a guide in the test construction process, ensured that the items are measuring the determined objectives , prepared a marking guide while constructing the test items, consulted standard text books in the subject for guide, gave clear instructions to guide the test takers, submitted tests meant for

promotional examinations for expert editing on time. The findings of the study however revealed that quite a number of the respondents did not consider the following skills while constructing test items: avoid the use of interlocking items, avoid items that measure opinion.

In their study, Quansah and Amoako (2018), developed and validated a standardised instrument in measuring teachers' attitude towards test construction. The study further explored the attitude of teachers towards test construction. The instrument was developed based on literature as well as personal experiences of the researchers. The developed instrument was administered to 432 Senior High School teachers in the Cape Coast Metropolis. Through an exploratory factor analysis, four dimensions were obtained which include: planning, item construction, item review and assembling. A confirmatory factor analysis was then conducted to examine the factor loadings of the items. After critical evaluation, the items on the instrument remained 32 which was on a four point Likert scale. Further analysis revealed an overall negative attitude of SHS teachers towards test construction.

Ololube (2008) also evaluated test construction skills of professional and non-professional teachers in Nigeria and reported that professional teachers tend to construct effective evaluative instruments more than the non-professional teachers. Ololubeemphasised that professional teachers have the propensity to employ the various assessment techniques correctly, which is unlikely to happen in the case of non-professional teachers.

In another study, Onyechere (2000) found that some teachers craft poor tests while others continue to use replica of test items because they seem to have

inadequate skills in test construction. Hamafyelto et al. (2015) also discovered that Senior High School (SHS) teachers in Borno State, Nigeria, constructed items which focused on lower cognitive operations. In Ebinye's (2001) view, test construction has been found to be a major source of anxiety among many teachers in Nigerian schools, especially, less experienced ones. This anxiety stems greatly from lack of skills in test construction among these teachers.

In another study, Anhwere (2009) conducted a study to investigate whether tutors in the teacher training colleges follow the basic laid down principles of testing practices, especially, test construction, administration and scoring of classroom or teacher-made tests in the teacher training colleges in Ghana. The study adopted the descriptive design and covered 20 public teacher training colleges in Ghana comprising 230 male and 80 female teacher training college tutors. The main instrument used for the study was a questionnaire. The instrument was developed and pre-tested by the researcher. The data was analyzed using means and standard deviations, frequency and percentages, and independent samples-t test. For all tests, the level of significance was at 0.05. The study showed that, teacher training college tutors did not follow the basic principles of testing in the construction of teacher-made or classroom tests, and that they perceived the management of assessment practices in the colleges as extra load to their teaching activities.

In a similar studies, Quansah, Amoako, and Ankomah (2018) explored the test construction skills of Senior High School (SHS) teachers in the Cape Coast Metropolis. Using a qualitative document analysis, samples of end-of-term examination papers in Integrated Science, Core Mathematics and Social Studies in three selected SHS in the Cape Coast Metropolis were randomly

(Lottery method) selected. The assessment tasks on the sampled instruments were critically examined by experts in the area of Educational Measurement and Evaluation. The results revealed that the teachers have limited skills in the construction of end-of-term examinations. This was evident as issues were found with the content representativeness and relevance of the test, reliability, and fairness of the assessment task which were evaluated. It was recommended that head teachers take up the challenge of inviting resource persons from recognised academic institutions to organise workshops for teachers on regular basis to sharpen teachers' skills on effective test construction practices.

Test Administration and Scoring Practices among Teachers

According to Rukundo and Magambo (2010), the paramount guiding principle in administering any classroom test is that all examinees should be given a fair chance to demonstrate their achievement of the learning outcomes intended or planned. Rukundo and Magambo expatiated that physical and psychological environment in which the examination is taking place has to be conducive for the examinee to facilitate the achievement of the testing outcome. The authors further explained that the factors that might interfere with validity of the measurement also have to be controlled. Linn and Miller (2005) postulated that even though the evidence regarding the effects of physical and environmental conditions on test performance is inconclusive, examinees should be as relaxed as possible and distractions should be eliminated or minimised. Linn and Miller further stated that whereas distractions during testing are known to have little effect on the scores of students, they may have profound effect on especially young children

According to Mehrens and Lehmann (1999), students are likely to perform better at any endeavour, including test taking, when they approach the experience with a positive attitude. Unfortunately, teachers frequently fail to help students develop positive attitudes toward tests. Students are not likely to perform at their best when they are excessively tense. Thus, the experience of test anxiety among some students.

Ogwang (2007) also stipulated that the process of examination administration is a difficult task, as it is sometimes marred by irregularities. Ogwang iterated that examination administration requires a lot of additional resources, both human and monetary ones to curb examination malpractices.

In his view, Cottrell, (2001) advanced that it is imperative that test administrators are qualified enough and trusted persons. Cottrell, emphasised that this is to ensure that tests are properly managed to obtain valid and reliable results. In Cottrell's view, test administrators need to have the opportunity to learn their responsibilities as a prerequisite to accurate test results. Similarly, Fontana (1995) underscored that a well prepared test is easy to administer, as compared with a poorly prepared test. Fontana further explained that it is equally important to realise that a successful test administration exercise is a product of test planning. Fontana concluded that cheating is most likely to occur in a poorly planned test, thus, proving a challenge to test administration.

In their study, Rukundo and Magambo (2010) advised that instructors should avoid giving hints to students who ask about individual items. The authors explained that if the item is ambiguous, it should be clarified for the entire group; instructors should however refrain from helping a pupil to

answer an item if it is not ambiguous. Rukundo and Magambo further iterated that the challenge is that at times, refraining from giving hints to examinees who ask for help may be difficult especially for new comers in the field of testing. Nevertheless, giving unfair aid to some students decreases the validity of the test results and lowers class morale.

Vickers and Nichols' (2005) study involved 63 raters scoring 35,534 essays written in response to a seventh-grade reading item with a four-point rubric for scoring. A group of 32 raters was randomly assigned to the online training condition and a group of 31 raters was randomly assigned to the standup training condition. Each group then scored the papers assigned to them until all papers were scored. Results indicated that the online and standup groups achieved comparable results in terms of reliability (measured by agreement of first and second raters) and validity (measured by agreement with expert raters on validity papers seeded into the scoring activity). In addition, the group trained online was able to score about 10 percent more responses than the standup group in the same period of time, suggesting that online training may have promoted more efficient scoring as compared to the standup training.

Differences among Teachers in Assessment Practices based on Gender

This aspect of the review focused the differences that exist among male and female respondents with respect to how they practice classroom assessment. Studies conducted in this area of gender and classroom assessment practices is not conclusive, some suggest significant differences while other give counter findings. For example, Alkharusi (2011) aimed at investigating teachers' self-perceived assessment practices and skills as a

function of gender. Participants were 213 Omani teachers from Muscat public schools. A 25-item Self-Perceived Assessment Skills Scale was developed and used in the study. Results indicated statistically significant differences on the self-perceived assessment skills with respect to teachers' gender.

Again, Ndalichako (2015) explored perceptions of secondary school teachers on assessment. A total of 4160 Tanzanian teachers were involved in the study. A questionnaire comprising questions that sought to establish demographic information of participants as well as their perceptions on classroom assessment. In order to establish whether the observed differences in the mean perception score for male and female was statistically significant, an independent sample t-test was performed. Result shows that there is a statistically significant difference in teachers' perception of assessment by gender. The *t*-test indicates that female teachers had more favorable perceptions of assessment than male teachers. In terms of classroom practices, the finding suggests that female teachers tend to use assessment more often to facilitate and support teaching and learning than their male counterparts.

Similarly, when examining educational assessment knowledge of 259 pre-service teachers who completed an educational assessment course, Alkharusi (2011b) found that male teachers tended to have on average a higher level of educational assessment knowledge than female teachers. Moreover, Alkahrusi et al. (2012) surveyed 165 in-service teachers from Muscat governorate about their attitudes, competence, knowledge, and practices in educational assessment. They found that although teachers held a favorable attitude towards and perceived themselves as being competent in educational assessment, teachers used a variety of assessments in the classroom primarily

for assigning grades and motivating students to learn, with some variations by gender. However, Alsarimi (2000) investigated classroom assessment practices of 246 third preparatory science teachers from 112 schools in Oman. The study found out that teachers indicated using short answer, completion, oral exams, extended answer, and multiple-choice item formats with no significant differences based on teacher's gender and years of teaching experience.

A number of studies have been conducted to examine educational assessment attitudes, competence, knowledge and practices. For example, Alsarimi (2000) investigation of classroom assessment practices of 246 third preparatory science teachers from 112 schools in Oman, found that teachers indicated using short answer, completion, oral exams, extended answer, and multiple-choice item formats with no significant differences based on teacher's gender and years of teaching experience.

Alkharusi (2011c) examined self-perceived assessment skills of 213 Omani teachers. He found that female teachers perceived themselves more skillful than male teachers in writing test items and communicating assessment results. Also, science teachers perceived themselves more skillful than English language teachers and fine arts teachers in developing performance assessment and analyzing assessment results. Further, sixth grade teachers indicated higher levels of self-perceived skills in developing performance assessment than eighth and tenth grade teachers. Furthermore, teaching experience correlated positively with self-perceived assessment skills, and that teachers with in-service assessment training showed a higher level of assessment skills than those without in-service assessment training. Similarly, when examining

educational assessment knowledge of 259 pre-service teachers who completed an educational assessment course, Alkharusi (2011b) found that male teachers tended to have on average a higher level of educational assessment knowledge than female teachers.

Differences among Teachers in Assessment Practices based on Experience

On the issue of teacher experience and assessment practice several studies have reported significant differences among teachers. Koloi-Keaikitse (2012) surveyed 691 primary and secondary school teachers in Botswana about their classroom assessment practices. Results indicated factors related to teachers' educational level, teaching experience, and assessment training contributed positively to beliefs, skills, and uses of desirable classroom assessment practices. Again, Zhang and Burry-Stock (2003) surveyed 297 teachers across grade levels and content areas about their classroom assessment practices. The study confirm that teachers are unique in terms of classroom assessment practices with reference to their teaching experience.

Also, in an investigation of 516 in-service teachers, Alkharusi (2011a) found that in-service assessment training and teaching experience correlated positively with educational assessment knowledge. Moreover, Alkharusi (2011) aimed at investigating teachers' self-perceived assessment practices and skills as a function of gender. Participants were 213 Omani teachers from Muscat public schools. A 25-item Self-Perceived Assessment Skills Scale was developed and used in the study. Results indicated statistically significant differences on the self-perceived assessment practices and skills with respect to teaching experience. Furthermore, examined the effects of classroom assessment practices on students' achievement goals. The study included

1,636 ninth grade students and 83 science teachers from Muscat public schools in Oman. Results from hierarchical linear modeling techniques showed that teachers differed significant per their experience level.

Alufohai and Akinlosotu (2016) of secondary school teachers' continuous assessment (CA) practices in Edo Central Senatorial District, Nigeria. The study was undertaken to determine the influence of gender, age, years of experience and area of educational specialization on teachers' attitude towards CA practices in secondary schools in the district. 543 teachers were drawn from the population of 1084 teachers across the district. However, 512 questionnaires were recovered and used for analysis. The t-test statistics was used to test the hypotheses. Result shows that there was a statistically significant differences between teachers in their practice of continuous assessment with respect to teaching experience. Moreover, Alkharusi (2011a) investigation of 516 in-service teachers, the study found that in-service assessment training and teaching experience correlated positively with educational assessment knowledge.

Recently, Alkharusi et al. (2012) surveyed 165 in-service teachers from Muscat governorate about their attitudes, competence, knowledge, and practices in educational assessment. They found out that although teachers held a favorable attitude toward the practice of assessment and perceived themselves as being competent in educational assessment. They however, demonstrated a low level of knowledge in educational assessment. Teachers used a variety of assessments in the classroom primarily for assigning grades and motivating students to learn, with some variations by gender, grade level,

and subject area. Teaching load and teaching experience accounted for some of the variations in teachers' educational assessment practices.

Differences in Assessment Practices Based on Qualification

Arulappen (2013) conducted a study to find out the perception of teachers about classroom assessment and self- perceive of teachers' competency in classroom assessment in Malaysia. The study further sought to find out the significant mean differences between teachers' competency in classroom assessment based on academic qualification, professional qualification, experience in teaching, experience in constructing classroom assessment and experience in attending assessment related courses. A total 367 teachers were randomly drawn from Johor Bahru District secondary schools using sampling frame provided by Johor Bahru District education Office to be the respondents of the study. The study used a questionnaire which had a Cronbach Alpha value of 0.95 to establish teachers' perceptions about classroom assessment and Cronbach Alpha value of 0.85 for self-evaluation of teachers' competency in constructing classroom assessment. Research question 1 and 2 were answered using descriptive analysis using statistical test frequency, mean and percentage whereby research questions 3 to 7 were answered using inferential analysis using statistical test one-way ANOVA. Hypotheses 1 to 5 were tested using One Way ANOVA. The findings of the study indicated that there was a significant mean difference between teachers' competency in classroom assessment based academic qualification ($F = 2.772$, $p < .041$). Post Hoc test revealed that teachers who have bachelor degree and master's degree (sig. = .041) as their academic qualification had the highest competency in classroom assessment compared

to teachers with diploma and certificate (sig. = .998) as their academic qualification.

Similarly, Zhang and Burry-Stock (2003), investigated teachers' assessment practices across teaching levels and content areas, as well as teachers' self-perceived assessment skills as a function of teaching experience and measurement training in the southeastern state of Alabama. Data from 297 teachers on the Assessment Practices Inventory were analyzed in a MANOVA design. The findings of the study indicated that as grade level increases, teachers rely more on objective tests in classroom assessment and show an increased concern for assessment quality ($p < .001$). The findings of the study also revealed that teachers' involvement in assessment activities reflects the nature and importance of the subjects they teach ($p < .001$). In furtherance, regardless of teaching experience, teachers with measurement training reported a higher level of self-perceived assessment skills in using performance measures; in standardized testing, test revision, and instructional improvement; as well as in communicating assessment results ($p < .05$) than those without measurement training.

In a study, Alkharusi et al. (2011) purposed to find assessment knowledge, skills, and attitudes of 217 in-service teachers in Oman. The study found that teachers who had a pre-service course in educational assessment demonstrated on average a higher level of educational assessment knowledge than those who did not have a pre-service assessment course.

It seems that results of the aforementioned studies in the Sultanate of Oman did not differ to some extent from those around the world. They generally point to a conclusion that classroom assessment might be unique

from one teacher to another depending on gender, teaching experience, teaching grade, qualification, and assessment training.

Chapter Summary

The chapter reviewed empirical studies on the following headings: types of classroom teacher-made tests, essay-type tests, objective-type tests, construction of classroom or teacher-made tests, administration of teacher-made tests, scoring essay tests, grading practices among teachers, test construction among teachers, as well as differences in test construction based on teacher characteristics (gender, experience and qualification). It was clear from the literature that most of the studies conducted on assessment practices in Ghana focused on the impact of assessment on classroom learning(Amua-Sekyi, 2016; Oduro, 2015), influence of teachers' knowledge on assessment practices (Asare, 2015; Hattori & Saba, 2008; Konadu, 2015). Additionally, majority of the studies that focused on teacher characteristics and assessment practices were conducted in other countries where the educational seems to differ from that of Ghana. The findings of these studies revealed that teachers reported using short answer, completion, oral exams, extended answer, and multiple-choice item formats with no significant differences based on teacher's gender and years of teaching experience. Similar studies also found that male teachers on average had a higher level of educational assessment knowledge than female teachers. It appears, however that there is no known study among basic school teachers in Ghana regarding teacher characteristics (such as gender, age, experience and qualification) as correlates of classroom assessment practices. In view of that, the present study examined teacher characteristics as correlates of classroom assessment practices at Atwima

Nwabiagya South District J.H.Ss. Specifically, the study assessed the prevalent of classroom assessment practices among teachers, test construction practices of teachers, test administration practices of teachers, test scoring practices of teachers, grading practices of teachers, contribution of teacher age, qualification and experience to classroom assessment practices, as well as difference that exist among teachers in terms of how they practice assessment in relation to gender. The next chapter presents the research methods employed in carrying out the study.

CHAPTER THREE

RESEARCH METHODS

The purpose of the study was to investigate teacher characteristics as correlate of classroom assessment practices at Atwima Nwabiagya South District J.H.Ss. The chapter discusses how the study was conducted. It is divided into five parts. The first part deals with research design, and the second; deals with the population and sample and sampling procedure. The third part covers the research instrument (including pre-testing that was used) while the fourth part deals with data collection procedure. The last part covers how data collected was analysed.

Research Design

According to Trochim (2000), a research design provides the glue that holds the research together. He explains that the design is used to structure the research, thereby showing how all the major parts of the research work together to address the central research problem. Nwadinigwe (2005, p.33) also emphasised the importance of design to research reiterating that "basically, research design, as an important aspect of research, must be the most appropriate to approximately measure what is being measured and obtain the data that will validly lead to a conclusion that is also valid".

The design that was used for this study was a descriptive survey. The descriptive survey design is a type of design that can be explained as the process of gathering data in order to answer research questions or test

hypothesis which concerns the existing status of a phenomenon. This type of survey attempts to provide an accurate and objective description of a picture of an on-going situation or real life situation (Quartey & Awoyemi, 2002). Osuala (2001) also noted that: "descriptive surveys are versatile and practical, especially to the researcher in that they identify present needs" (p. 35). Descriptive research involves collecting data in order to test hypothesis or answer questions concerning the current status of the subjects of the study. It determines and reports the way things are (Gay, 1992).

Koul (1997) indicated that survey studies are conducted to collect detailed description on existing phenomenon with the intent of employing data to justify current conditions, practices or make more intelligent plans for improving them. He further explained that, in addition to analysing, interpreting and reporting on the status of an organisation for future guidance, descriptive surveys can be used to determine the adequacy of an activity by comparing results to establish standards. It also has another advantage of producing a good amount of responses from a wide range of people (Frankel & Wallen (2000). This design is ideal because this study will provides a report on Basic School teachers' perception on classroom assessment practices.

This does not mean that the descriptive survey design is not without weakness. Marczyk, DeMatteon and Festinger (2005) observe that survey designs, like all non-experimental designs, no matter how convincing the data may be, cannot rule out extraneous variables as the cause of what is being observed. This is because descriptive survey designs do not have control over the variables and the environment that they study. This suggest that findings from survey are most often influenced by factors other than those attributed by

the researcher. Seifert and Hoffgung (1991) also identify problems with survey designs to include the possibility of producing untrustworthy result because they may delve into people's private matters. Again, since descriptive survey designs most often make use of questionnaires, it becomes limited to respondents who are literate.

However, attempts have been made to minimize the limitation(s) of survey design in this study. These include avoiding issues which respondents considered sensitive and personal. Also, all members of the target group were literates and the researcher used very simple language to make the items easy to understand and answer.

Population

Population refers to all individuals of interest to the researcher (Marczyk, DeMatteon & Festinger, 2005). Agyedu, Donkor and Obeng (1999) also explain population as a set of individuals (objects, subjects, events) that have common observable characteristics for which a researcher is interested. According to Fink (1995), the criteria for the inclusion of a unit in a survey are based on characteristics of respondents who are eligible for the participation in the survey. For this reason, the target population for the study was all Public School teachers in the Atwima Nwabiagya South District within the Ashanti Region of Ghana. The target population for the study was 434 teachers of the 62 public Junior High Schools in the Atwima Nwabiagya South District within the Ashanti Region of Ghana. The accessible population was made up of 336 teachers in 48 selected public Junior High Schools in the District (GES, Atwima Nwabiagya South District Metropolitan Directorate 2018).

Sampling Procedure

Breakwell, Harmond, Fife-Shaw and Smith (2006) underscore that population are often extremely large or infinite, thus making it impossible or too costly to study. They therefore opine that a sample be drawn from the population. Breakwell, et al (2006), noted, that in addition to the cost savings this entails, it is usually appropriate to make more and more detailed observations of each sampled element. There are sixty-two (62) public Junior High Schools in the district which are located in five circuits. There are four hundred and thirty-four (434) teachers in the 62 schools. Circuit one has fifteen schools, the second Circuit has eighteen schools, the third Circuit has twelve schools and the fourth and fifth Circuits have nine and eight schools respectively.

The selection of schools and respondents were done in a two-stage sampling procedure. The first stage involved proportionate stratified sampling of schools from the five circuits for the study. Due to the large number of schools in the district, forty-eight (48) public junior high schools out of sixty-two (62) was used for the research. A common ratio of .774 was used to determine the number of schools from each circuit. 12 schools were selected from the circuit one, 14 schools from circuit two, 9 schools from circuit three, 7 schools from circuit four and 6 schools from circuit five. This sampling procedure was deemed appropriate for the study because the researcher would compare results among the circuits. The researcher shook a receptacle and randomly picked up to determine the schools using number codes. That was done with replacement, in order to maintain the same probability for the schools in each circuit.

Three hundred and thirty-six teachers representing the accessible population were selected from forty-eight public junior high schools for the study. A sample size of two hundred and nineteen was used for the study. This sample size was chosen using table for determining sample size from a given population provided by Krejcie and Morgarn (1970) as cited in Sarantakos (2005) which shows that for a population of 336, a sample size of 219 should be adequate. However, to cater for non-response, a sample size of 240 was used. Disproportionate stratified sampling procedure was adopted at the second stage for the selection of the respondents (teachers) with each school selected with five teachers irrespective of the subject he or she teaches in the school. With this procedure, representative sample for both schools and teachers was obtained. The selected schools were serially coded.

In each school, the special code numbers for the teachers were then written on pieces of papers and were put in an urn. Simple random sampling technique was then used to select five teachers in each public junior high school. The slips of papers were picked one after the other without looking into the pool. Once a number code representing a teacher was selected, it was recorded as a sample with the urn shaken while the chosen piece of paper put back into the urn and reshuffled. The process was repeated in each school till the 240 teachers were obtained for the study.

Data Collection Instrument

The main instrument that was used in gathering data was a structured questionnaire. A 42-item Teacher Characteristics Assessment Practices questionnaire was adapted from Anhwere (2009) and used in this study. The 42 items were on three-point Likert scale ranging from 3 (more often) to 1 (not

used). Questionnaire was used due to its effectiveness in collecting data from a large number of people within short possible time (Koul, 2001).The instrument was made up of five sections. Section A is on background data of respondents which is made up of four items. Section B is on 'Teachers Assessment Modes and Formats' which consists of seven items. Section C has to do with teachers test construction practices and was made up of thirteen items. Section D and E also have to do with test administration and scoring practices and grading practices of teachers respectively. Section D has thirteen items and E has five items.

Pilot-Testing

The instrument was pilot tested on thirty (30) junior high school teachers from Manhyia M/A, Akropong M/A 'A' and 'B' from Atwima Nwabiagya north district in the Ashanti region of Ghana. The questionnaire was administered to the thirty teachers in the selected schools by the researcher himself. The responses given by respondents were used to refine the questionnaire (Amedahe, 2002). The pilot-testing was necessary because it enhanced the content validity and reliability of the instrument , and to improve questions, format and scales after careful analysis of the items based on the comments passed by respondents concerning the weaknesses, clarity and ambiguity on all aspects of the questionnaire (Leedy & Ormrod, 2005). For example items like "I use yes/no items to assess my students" in the Section B aspect of the questionnaire was changed to "I use true/fale items to assess my students." Similarly, items like "I draft a scoring rubrics immediately after constructing my items" in the Section C aspect of the questionnaire was

chaged to "I prepare a making scheme immediately after constructing my items." In all two items were reworded after the pilot testing.

In addition, the pilot-testing provided the opportunity in assessing the appropriateness and practicality of the data collection instrument. Also, it tested the adequacy of the procedures that were used for the study. On the whole, the pre-testing helped to fine-tune the instrument (Leedy & Ormrod, 2005).

Validity and Reliability of the Instrument

According to Fraenkel, Wallen, and Hyun (2012), the content and face validity of research instrument must be determined by expert judgment. Therefore, to ascertain the content validity, the items on the questionnaire were shown to my supervisors for expert review. This was to examine whether: (a) the items were related to the research questions; (b) the items elicited the appropriate responses from the respondents; (c) the vocabulary structure was appropriate; (d) the items were properly arranged; (e) the items fitted into sections they had been placed; and (f) if any of the items were ambiguous. The suggestions given by the supervisors were used to improve the instrument and thereby helped to establish the content and face validity of the instrument.

To determine the reliability of the items on the questionnaires, Cronbach Alpha (α) was used to estimate the internal consistency of the sub-sections of the instruments. An alpha value of .70 or above was considered appropriate (Karagoz, 2016). Details of the reliability coefficients are presented in Tables 1 and 2.

Table 1- *Reliability Co-efficient of Piloted Instrument*

No.	Name of Scale	No. of Items	Cronbach's Alpha
1.	Teachers' assessment mode and formats	7	.76
2.	Teachers' test construction practices	13	.71
3.	Test administration practices	6	.70
4.	Test scoring practices	7	.72
5.	Grading practices of teachers	5	.73

From Table 1, the reliability coefficients of the scales of the instrument after the pilot testing ranged from .70 to .76.Generally, these coefficients are good indicators of internal consistency, since they were not below .70.

Table 2- *Reliability Co-efficient of Final Instrument*

No.	Name of Scale	No. of Items	Cronbach's Alpha
1.	Teachers' assessment mode and formats	7	.78
2.	Teachers' test construction practices	13	.74
3.	Test administration practices	6	.73
4.	Test scoring practices	7	.72
5.	Grading practices of teachers	5	.75

At the end of the final data collection, the reliability coefficients for students ranged from .72 to .78; an indication of good reliability.

Data Collection Procedure

Permission was sought from the heads in the schools, who allowed the researcher to administer the questionnaire, after an introductory letter had been delivered to the head teachers for their co-operation and assistance (see Appendix D).

On the 19thto 23rd of March 2018, the questionnaire was administered to circuits one, four and five by the researcher with the assistance of the head teachers in the schools. The questionnaire for circuits two and three was administered on the 26th to 30th of March 2018 by the researcher with the assistance of the head teachers in the schools

In each school, the researcher explained the purpose of the study to the head and the sampled teachers, and assured them of anonymity and confidentiality of their participation in the study. Questionnaires were then administered to the teachers with the assistance of the heads.

There was time for questions, during which respondents had the opportunity of asking questions on items that were not clear to them before responding to the questionnaire. This was because it helps to erase respondents' biases and prejudices (Trochim, 2000). This ensured good contact with the teachers to further explain the purpose of the study so that the researcher won the commitment of the teachers towards responding to items on the questionnaire and submitting them in good time. For the sake of anonymity and confidentiality of the results, participants were instructed not to write either their names or name of their school on the questionnaire. The researcher used two weeks in collecting the data for the study.

Data Processing and Analysis

The demographic variables of the respondents in the study as well as data on research questions one (1), two (2), three (3), four (4), and five (5) were answered using descriptive statistics of frequencies and percentages. Data on research question six (6) was answered using hierarchical multiple regression. This was because the researcher sought to find out which

demographic variable (i.e., experience, age and qualification) predicted more of classroom assessment practices. Data on hypothesis 1 was tested using independent samples t-test.

CHAPTER FOUR

RESULTS AND DISCUSSION

The study sought to investigate teacher characteristics as correlate of classroom assessment practices at Atwima Nwabiagya South District J.H.Ss. Out of 219 questionnaires administered, 202 of them were completely responded to and returned. This led to a response rate of 92% hence all the analysis in this chapter was based on 202 respondents. Additionally, before the analysis, items that were negatively worded were reversely score to enhance the quality of the analysis. This chapter has a description of the sample characteristics, followed by the analysis of the main data and then the discussion of the result.

Background Characteristics of Respondent

The study sought for the demographic characteristics of the respondents. These included the respondent's gender, years of teaching, educational qualifications and their ages. Table 1 presents detailed information on the distribution.

The results in Table 3 revealed that out of 202 respondents, 110 (54.5%) were female teachers while 92 (45.5%) were male teachers. This implies that teaching profession in Atwima Nwabiagya south district at the basic level is dominated by females. The results further indicated that majority of the respondents 71 (35%) had taught for 12 years and above indicating that, majority of the respondents were experienced. Regarding the educational

qualification of the respondents, the study revealed that majority of the respondents 133 (65.3%) had bachelor in education, 37 (18.3%) had diploma in education, and only 2(1.0%) of the respondents had MPhil.

Table 3-*Teacher Demographics*

Demographics	Frequency	Percentage (%)
Gender		
Male	92	45.5
Female	110	54.5
Years of Teaching		
From1-3yrs	25	12.4
4-7years	47	23.3
8-11years	59	29.2
12years and above	71	35.1
Educational Qualification		
Cert A	10	2
Diploma in education	37	18.3
Bachelor in education	133	65.8
BA/B.Sc without education	15	7.4
Master's in education	9	4.5
MA/M.Sc	4	2.0
M.Phil	2	1.0
Age range		
20-29	39	19.3
30-39	115	56.9
40-49	41	20.3
50-59	7	3.5

Source: Field survey (2018)

Information from Table 3 clearly shows that currently majority of teachers at the junior high school level at the Atwima Nwabiagya south district had qualification ranging from Diploma in Basic Education to Bachelor Degree which is an indication of inprovment in teacher quality. In furtherance,

majority of the respondents 115 (56.9%) were between the age range of 30-39 years, 41(20.3%) of the respondents were between the age range of 40-49 years, 39 (19.3%) of the respondents were between the age range of 20-29 years while only 7 (3.5%) of the respondents were between the ages of 50-59 years. This clearly shows that, the district has a lot of young teachers at the basic schools.

Research Question 1

What are the prevalent classroom assessment practices (formats and modes) among teachers in Atwima Nwabiagya South District?

This research question sought to find out prevalent classroom assessment practices among teachers. Seven (7) items were used in measuring this construct which were on a three-point Likert type scale (More Often, Often, Not used). Descriptive statistics of frequencies and percentages were used to analyse the responses of the respondents and the results are shown in Table 4.

The results in Table 4 indicated that majority of the respondents 145 (71.8%) used exams more often in assessing students at the end of the term. Again, majority of the respondents 139 (68.8%) reported that they gave exercises more often to assess students immediately after teaching. Interestingly, 138(68.3%) of the respondents indicated that they used oral questions more often to assess students during lesson periods. More so, 117(57.9%) representing the majority of the respondents reported that they gave essay items often to assess their students. For the use of multiple-choice items to assess students, majority 104(51.5%) often use to assess their students. With reference to using true/false items to assess students, majority

141(69.8%) of the respondents did not use true/false items in assessing their students.

Table 4- *Teachers Assessment Modes and Formats*

Statements	More often	Often	Not used
I give exercise to assess my students immediately after teaching	139(68.8)	50(24.8)	13(6.4)
I give homework to assess my students	79(39.1)	103(51.0)	20(9.9)
I ask oral questions to assess my students during lesson periods	138(68.3)	56(27.7)	8(4.0)
I give exams to assess my students at the end of the term	145(71.8)	50(24.8)	7(3.5)
I give essay items to assess my students	42(20.8)	117(57.9)	43(21.3)
I construct multiple–choice items to assess my students	58(28.7)	104(51.5)	40(19.8)
I use true/false items to assess my students	16(7.9)	45(22.3)	141(69.8)

Source: Field survey (2018), Percentages in parenthesis ( )

Information from Table 4 clearly indicated that the modes of assessment teachers at the junior high schools of Atwima Nwabiagya south district mostly used were the essay and the multiple-choice types.

Research Question 2

What is the test construction practice of teachers in the Atwima Nwabiagya South District?

This research question sought to find out the test construction practice of teachers in the Atwima Nwabiagya South District. Thirteen (13) items were used in measuring this construct which were on a three-point Likert type scale (More Often, Often, Not used). Descriptive statistics of frequencies and percentages were used to analyse the responses of the respondents and the results are shown in Table 5.

Table 5- *Teachers Test Construction Practices*

Statement	More often	Often	Not used
I write individual test items at least two weeks before the date	34(16.8)	92(45.5)	76(37.6)
I prepare a marking scheme immediately after constructing my items	113(55.9)	68(33.7)	21(10.4)
I copy questions from B.E.C.E. examinations questions	39(19.3)	91(45.0)	72(35.6)
I develop test items only when it is time to assess students.	22(10.9)	63(31.2)	117(57.9)
I use a test specification table when writing test items	35(17.3)	80(39.6)	87(43.1)
I copy test question from textbooks	46(22.8)	90(44.6)	66(32.7)
I match instructional objectives with test items	98(48.5)	83(41.6)	21(10.4)
I write items based on information that students know	71(35.1)	90(44.6)	41(20.3)
I arrange objectives test answers in a pattern to make scoring easy	36(17.8)	55(27.2)	111(55.0)
I consider the purpose of the test before developing test items	111(55.0)	71(35.1)	20(9.9)
My school conducts in-service training in test construction for teachers	21(10.4)	54(26.7)	127(62.9)
I prepare more items than needed before I review and select some for the test	42(20.8)	91(45.0)	69(34.2)
I evaluate the test as a whole before I make the final copy	98(48.5)	78(38.6)	26(12.9)

Source: Field survey (2018), Percentages in parenthesis ( )

In Table 5, majority of the respondents 111(55.0%) reported that they considered the purpose of the test more often before developing test items. 113(55.9) representing the majority indicated that they prepared a marking scheme more often immediately after constructing items but suprisngly 10.4%

of the respondents indicated that they did not prepare a marking scheme immediately after constructing test items. Regarding matching instructional objectives with test items, majority of the respondents 98 (48.5%) indicated that they matched instructional objectives with test items more often.

Interestingly, quiet a number of the respondents 87(43.1%) did not use test specification table when writing test items. Again, majority 136 (67.8%) of the respondents reported that they copied test questions from textbooks often, only 32.2% of the respondents reported that they did not copy test questions from textbooks. This information clearly indicated that there are lapses in items construction by teachers at Atwima Nwabiagya South district junior high schools.

Research Question 3

What is the test administration practice of teachers in the Atwima Nwabiagya South District?

This research question sought to find out the test administration practice of teachers in the Atwima Nwabiagya South District. Six (6) items were used in measuring this construct which were on a three-point Likert type scale (More Often, Often, Not used). Descriptive statistics of frequencies and percentages were used to analyse the responses of the respondents and the results are shown in Table 6.

In Table 6, majority of the respondents 152 (77.2%) reported that they ensured good seating arrangements more often to prevent students from coping from each other. Regarding giving hints to students on test items during examination, majority 175 (86.6%) of the respondents indicated that they did not gave hints to students about individual items during examination.

Table 6- *Test Administration Practices*

Statement	More often	Often	Not used
I ensure good seating arrangements to prevent students from copying from each other	152(75.2)	42(20.8)	8(4.0)
During examinations, I tell students that if they do not write fast, they will fail.	30(14.9)	38(18.8)	134(66.3)
During examinations, I prepare for and expect emergencies.	32(15.8)	87(43.1)	83(41.1)
I give hints to students when they ask about individual test items during examination.	10(5.0)	17(8.4)	175(86.6)
I read novels, newspaper or "whatsApp" friends when I invigilate tests/exams	11(5.4)	13(6.4)	178(88.1)
I inform students in advance about content/topics that the tests/examination cover	37(18.3)	64(31.7)	101(50.0)

Source: Field survey (2018), Percentages in parenthesis ( )

Also, with reference to writing fast during examination, majority 134 (66.3%) of the respondents did not inform students to write fast during examination in order to avoid failure but as many as 68 of the respondents told students to write fast during examination in order to avoid failure. The findings of the study also revealed that majority of the respondents 178 (88.1%) of the respondents did not read novels, newspapers or "whatsApped" friends during invigilation, only few of the respondents (24) read novels, newspapers or "whatsApped" friends during invigilation. Information from Table 6 indicates that teachers in junior high schools at Atwima Nwabiagya south district practice good test administration during examination though there were few of them on the other side.

Research Question 4

What is the test scoring practice of teachers in the Atwima Nwabiagya South District?

This research question sought to find out the test scoring practice of teachers in the Atwima Nwabiagya South District. Seven (7) items were used in measuring this construct which were on a three-point Likert type scale (More Often, Often, Not used). Descriptive statistics of frequencies and percentages were used to analyse the responses of the respondents and the results are shown in Table 7.

Table 7- *Test Scoring Practices*

Statement	More often	Often	Not used
I score essay tests, question by question	71(35.1)	79(39.1)	52(25.7)
I prepare marking schemes for essay tests after students have taken the test.	64(31.7)	54(26.7)	84(41.6)
I periodically re-score previously scored papers to check consistency in scoring	28(13.9)	69(34.2)	105(52.0)
I score answer scripts with the names of the students known to me	13(6.4)	12(5.9)	177(87.6)
I keep previously scored items out of sight when scoring the rest of the items	34(16.8)	61(30.2)	107(53.0)
The first few essays I score influence the rest of the scores I give	12(5.9)	26(12.9)	164(81.2)
I give a separate mark for the mechanics of writing such as correct grammar, flow of expression etc. when scoring essays	64(31.7)	64(31.7)	74(36.6)

Source: Field survey (2018), Percentages in parenthesis ( )

Regarding periodically re-scoring previously scored papers to check consistency in scoring, the findings of the present study indicated that majority 105 (52.0%) of the respondents did not periodically re-scored previously scored papers to check consistency in scoring. Ninety seven (48%) reported

that they engaged in periodically re-scoring previously scored papers to check consistency in scoring. The findings of the study further indicated that majority 177 (87.6%) of the respondents reported that they did not score the scripts with the names of the students known to them.

When asked whether the first few essays scored influence the rest of the scores they gave, majority 164 (81.2%) of the respondents reported that the first few essays they scored did not influence the rest of the scores they gave.

Again scoring as one of the assessment practices, teachers at the Atwima Nwabiagya south district in the junior high schools have good scoring skills.

Research Question 5

What is the grading practice of teachers in the Atwima Nwabiagya South District?

This research question sought to find out grading practice of teachers in the Atwima Nwabiagya South District. Five (5) items were used in measuring this construct which were on a three-point Likert type scale (More often, Often, Not used). For the purpose analysis, the results of most often and often were combined as one.

Descriptive statistics of frequencies and percentages were used to analyse the responses of the respondents and the results are shown in Table 8. Information from Table 8 shows that majority of the respondents 119 (58.9%) assigned grades to students work to reflect academic performance only while 83 (41.1%) of the respondents did not assign grades to students work to reflect academic performance only. Regarding using a standard base grading system for any school in grading, majority of the respondents 174(86.1%) reported

that they use standard base grading system for any school in grading whereas 28 (13.9%) of the respondents did not use standard base grading system for any school.

Table 8- *Grading Practices of Teachers*

Statements	More often	Often	Not used
I assigned grades to students work to reflect academic performance only	47(23.3)	72(35.6)	83(41.1)
I assign grades purposively to punish non-learning students	12(5.9)	20(9.9)	170(84.2)
I assign grades to motivate all students to learn harder	98(48.5)	65(32.2)	39(19.3)
I use a standard base grading system for any school in grading	112(55.4)	62(30.7)	28(13.9)
I grade students based on behavioural factors such as conduct, effort and attitude affects my grading	16(7.9)	46(22.8)	140(69.3)

Source: Field survey (2018), Percentages in parenthesis ( )

Also, when asked whether respondents assigned grades purposly to punish non-learning students, majority 170 (84.2%) of them reported that they did not assign grades purposly to punish non-learning students.

With reference to assigning grades to motivate all students to learn harder, majority 163(80.1%) of the respondents indicated that they assigned grades to motivate all students to learn harder. The results of the present study further revealed that majority 140 (69.3%) of the respondents reported that they did not grade the students based on behavioural factors such as conduct, effort and attitude.

Research Question 6

What demographical variables (experience, age and qualification) predict teacher classroom assessment practices?

This research question sought to find out which demographic variable (i.e., experience, age and qualification) predict more of classroom assessment practices. To test this research question, a hierarchical multiple regression was conducted to find out the contribution of each predictor to the model as well as which variable predicted more of assessment practice. The criterion variable was unidimensional and thus, the composite score was used for the analysis. Since each of the predictors (independent variables) was in categories, the predictor variables were dummy coded in order to facilitate the analysis of the data.

Before the analysis, normality of the data was checked; the results presented that the data did not violate the normality assumption. Results from the Durbin Watsons test, no autocorrelation was found in the data (d = 1.917). Again, there is a linear relationship between the predictors and assessment practice. (See Apendix C). The VIF and Tolerance values also revealed that there was no multicollinearity. With these conditions satisfied, a hierarchical multiple regression was conducted. Results on the analysis of data are presented in Table 9.

Table 9 provides results on hierarchical regression of demographic variables and assessment practice. The analysis examined the independent contribution of each demographic characteristic to the model. The overall model was found to be non-significant. In Model 1, years of experience was found as a non-significant predictor of assessment practice, F (3, 198) =.279, p

= .840. It was found that 0.4% of the variance in assessment practice was explained by years of teaching experience.

In Model 2, age was added to years of teaching experience. The overall model was still not significant, *F* (6, 195) =.318, *p* =.927. It was found that 10% of the variances in assessment practice were explained by years of teaching, experience and age of respondents. Age contributed 0.5% of the total variances in assessment practice and this was not significant *F* (3,195) = .360, *p* =.782.

Table 9- *Hierarchical Regression of Demographic Variables and Assessment Practices*

Model		Sum of Squares	df	Mean Square	F	Sig	R^2	R^2 Change
1	Regression	7.73	3	2.58	.279	.840	.004	.004
	Residual	1826.89	198	9.23				
2	Regression	17.79	6	2.97	.318	.927	.010	.005
	Residual	1816.83	195	9.32				
3	Regression	112.271	12	9.36	1.027	.426	.061	.051
	Residual	1722.35	189	9.11				

1. Predictor: Experience
2. Predictor: Experience, age,
3. Predictor: Experience, age, qualification

Criterion: Assessment Practice
*significant at .05 level
Model 1: F Change = .279, Sig. F Change = .840, df1=3, df2 =198
Model 2: F Change = .360, Sig. F Change = .782, df1=3, df2= 195
Model 3: F Change =1.728, Sig. F Change = .117, df1=6, df2=189

In Model 3, qualification was added to both age and years of teaching experience. There overall model was still not significant *F* (12, 189) =1.027, *p*=.426. It was however clear that years of experience, age of respondents as well the qualifications of the respondents accounted for 6.1% of the variances in assessment practice. The contribution of qualification to the model was 5.1%

of the total variances in assessment practice. This contribution was, however, not significant F (6,189) = 1.728, p =.117.

Table 10 presents the individual predictors of the categorical demographic variables.

Table 10- *Regression Coefficient of Categorical Demographic Variables*

Model		Unstandardized Coefficients		Standardized Coefficients			
		B	Std. Error	Beta	T	Sig.	VIF
1	(Constant)	21.36	.608		35.160	.000	
	4-7 years	-.232	.752	-.033	-.309	.758	2.210
	8-11 years	.284	.725	.043	.392	.696	2.379
	12 yrs & above	-.078	.706	-.012	-.111	.912	2.490
2	(Constant)	21.54	.645		33.380	.000	
	4-7 years	-.031	.791	-.004	-.040	.968	2.422
	8-11 years	.662	.835	.100	.793	.429	3.124
	12 yrs &above	.372	.868	.059	.429	.668	3.720
	30-39 years	-.557	.653	-.092	-.854	.394	2.266
	40-49 years	-.821	.858	-.110	-.958	.339	2.580
	50 to 59 years	-.282	1.351	-.017	-.209	.835	1.323
3	(Constant)	19.40	2.192		8.851	.000	
	4-7 years	-.048	.810	-.007	-.060	.952	2.598
	8-11 years	.893	.860	.135	1.038	.300	3.393
	12 yrs. & above	.342	.911	.054	.376	.708	4.193
	30-39 years	-.479	.661	-.079	-.725	.470	2.373
	40-49 years	-.693	.862	-.092	-.804	.423	2.666
	50 to 59 years	-.428	1.363	-.026	-.314	.754	1.378
	Diploma	2.395	2.226	.307	1.076	.283	16.430
	Bachelor of Edu	1.931	2.202	.304	.877	.382	24.165
	BA/B.Sc	2.484	2.369	.216	1.049	.296	8.550
	M.Ed	-.113	2.409	-.008	-.047	.963	5.476
	MA/MSC	3.679	2.680	.170	1.373	.172	3.091
	Mphil	5.800	3.062	.191	1.894	.060	2.037

*significant at .05 level

Hypothesis Testing

The study sought to test three hypotheses. These hypothesis was tested at a significance level of .05 and a confidence level of 95%. Statistical test tools such as independent samples t-test, was employed. It is important to

emphasise that some of the negative worded items on the standardised scale used in the study to measure "assessment practice" were coded reversely to enhance easy reporting and comprehension of the analysis.

Hypothesis 1

H_0: There is no statistically significant difference between teachers' in terms of assessment practices with reference to gender.

H_1: There is a statistically significant difference between teachers' in terms of assessment practices with reference to gender.

This hypothesis sought to examine whether a statistically significant difference exist in teachers' assessment practices in terms of gender. An independent samples t-test analysis was conducted to test this hypothesis. The dependent variable for the respondents was the composite score of respondents' assessment practices. Prior to the analysis, assumptions underlying the use of independent samples t-test were checked. Results from the normal Q-Q plot revealed that the data did not violate the normality assumption (Figure 1).

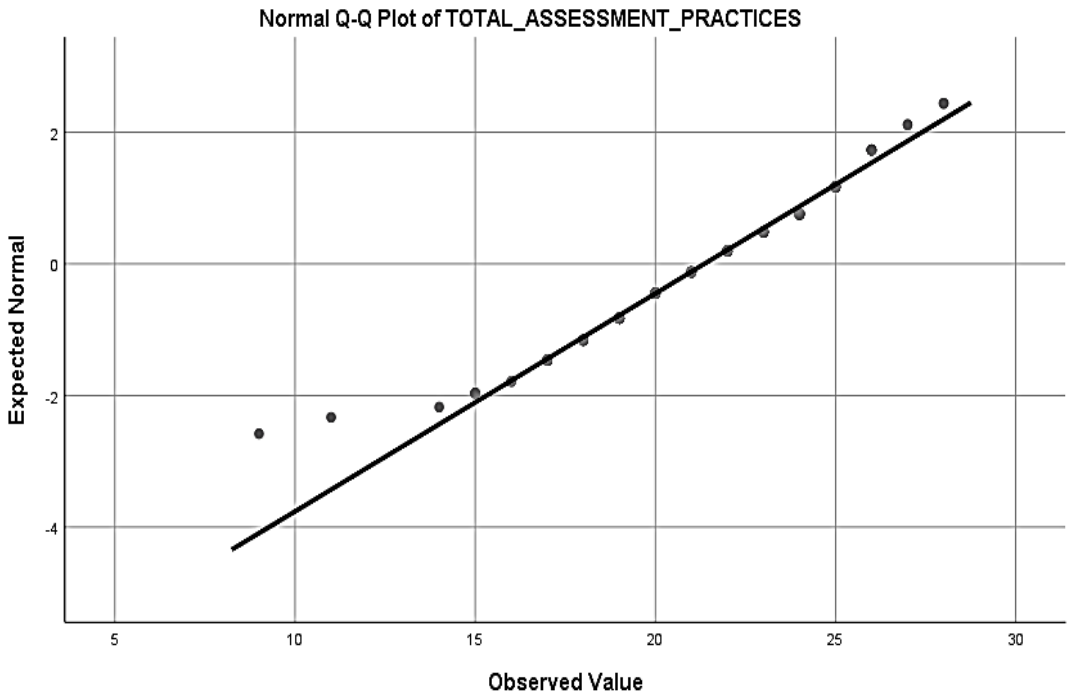


Figure 1 - The Normal Q-Q Plot

The result was supported by further analysis which showed that the mean (M = 21) is approximately the same as the median (Mdn = 21). The normality data was also confirmed by the histogram which showed that majority of the data score lied between ±2 standard deviation (See Appendix B). This give a "green light" for a parametric test tool to be used for the analysis. Hence the independent samples t-test.

The equal variance assumption was also tested to find out whether the variances between the groups are the same. Results from the Levene's test for equality of variances revealed that the equality of variance assumption was not violated (F = 1.415, p = .236). Table 11 further presented the actual analysis, testing the difference between the two groups with regards to the dependent variable.

Table 11- *Gender Difference in Teachers' Assessment Practices*

Gender	N	Mean	SD	df	t-value	p-value
Male	92	21.55	2.72			
				200	.830	.408
Female	110	21.20	3.26			

*significant at .05 level

Result from the analysis found a non-significant difference in teachers' assessment practices in terms of gender, t (200) = .830, p = .408. This indicates that male and female teachers in the Atwima Nwabiagya South District did not differ in terms of assessment practices. In other words, both male and female teachers in Atwima Nwabiagya South District engage in similar assessment practices.

Discussion

This section discusses the research findings in relation to teacher characteristics as correlate of classroom assessment practices on:

1. Prevalent classroom assessment practices.
2. Test construction.
3. Test administration and scoring.
4. Grading.
5. Contribution of teacher age, qualification and experience in predicting classroom assessment practices.
6. Difference among teachers in assessment practice in relation to gender, experience and qualification.

Prevalent Classroom Assessment Practices among Teachers in Atwima Nwabiagya South District

This research question sought to find out the prevalent classroom assessment practices among teachers in the Atwima Nwabiagya South District.

The results of the study indicated that teachers used a number of assessment practices in gathering valid information from student. Among some of the classroom assessment practices used by teachers in the classroom includes, essay items to assess students' understanding about a particular concept, multiple-choice items, giving class exercises to students immediately after teaching, the use of oral questions to assess students as well as the use of exams to assess students at the end of the term. Although the findings of the study revealed that teachers used a number of classroom assessment practices in assessing students, some assessment practices are more prevalent among the teachers in the classroom.

The findings of the study corroborated with the findings of Benzehaf (2017) who found that teachers used varied number of assessment strategies ranging from home assignments to in-class written tests but mainly for summative purposes. With reference to frequency of assessment, Benzehaf found that majority of the respondents assessed their students once a month; whereas quite a number of the teachers assessed their students twice a semester. The findings of Benzehaf further revealed that respondents used a number of assessment practices that was consistent with the findings of the current study. Benzehaf found out that teachers used essay questions, true/false questions as well as fill-in-the-blanks formats in assessing students in the classroom. This finding was at par with the findings of the current study. More so, the findings of this study buttressed the assertion of Gronlund, (2006) who iterated that a variety of methods are used by teachers in their daily classroom assessment including traditional assessments such as multiple-choice, true-false, matching, completion, and short-answers and alternative

assessments such as portfolios, student self-assessment, observations, and other performance-based assessments.

The findings of this study was however inconsistent with the findings of Sofo, Ocansey, Nabie, and Asola (2013) who found that teachers used observation, skill test, knowledge test, demonstration, peer observation, and oral report as assessment practices. The deductive analysis of Sofo, et al. study further indicated that the highest percentage of assessment practice utilised by teachers in practical lessons was teacher observation, while that of oral report had the lowest percentage. Teachers also reported knowledge test and individual report as the highest and lowest percentage of assessment practices used in their theory lessons respectively. The findings of Sofo, et al. also revealed that most of the teachers who participated in the study used assessment for documenting learning, rather than for accountability purposes.

The findings of the current study did not agree with the findings of Zhang, and Burry-Stock (2003) who found that teachers differ in their assessment practices due to the nature of classroom assessment delineated by teaching levels. According to the findings of Zhang, and Burry-Stock, a general difference emerged between elementary and secondary teachers in terms of the assessment methods used and teachers' concerns for assessment quality. While secondary teachers rely mostly on paper–pencil tests and were concerned about the quality of assessment, elementary teachers often use performance assessment as an alternative.

Test Construction practices of Teachers in the Atwima Nwabiagya South District

This research question sought to find out the test construction practices of teachers in the Atwima Nwabiagya South District. The findings of the study revealed that respondents engaged in the use of a number of test construction practices. It was evident in this study that most of the respondents more often considered the purpose of the test before developing test items. The idea of knowing the purpose for which a particular test is being administered is very crucial when constructing test items. That is to say it is incumbent on a test developer to know the purpose for which a particular test is being administered before going ahead to construct the individual items of the test. Knowing the purpose or the use to which the test results will be put to, will inform the test developer about the specific traits to measure. Without knowing the purpose of the test, one cannot construct valid items to measure the specific traits that one intends to measure. This will at the long run affect the final decision to which the results of the test will be put to. The findings of the current study further revealed that a vast number of the respondents more often prepared a marking scheme immediately after constructing test items.

Preparing a marking scheme immediately after constructing test items is an effective way of deciding on the length of time to give for students to answer the items of a particular test. That is to say, an instructor who prepares his/her marking scheme immediately after constructing test items will be in the position to tell how difficult, easy or demanding the various test items are. In that regard, such an instructor will be well informed on the basis of the length of time to give students in answering the various items of the test.

Again since the items on the test will be fresh on the mind of the test developer, preparing the marking scheme right after constructing the test items is a step in the right direction. This will help the test developer to easily provide the appropriate answers for each of the test items. In the same vein, the test developer will be in a good position to uncover any inconsistencies that might have been introduced in the options given in the case of multiple choice items. ‘

In furtherance, regarding matching instructional objectives with test items, the findings of the study found that almost half of the respondents matched instructional objectives more often with test items. Matching instructional objectives with individual items of a particular test reveals the true purpose of the test. Instructional objectives are the specific behaviours learners are able to demonstrate after an instruction session. An instructor who matches instructional objectives of a particular subject with the individual test items is able to measure the specific trait he wants to measure. This clearly helps to define the purpose of the test as well as the use to which the results of the test will be put to. Amasingly, the findings of the study revealed that, quite a number of the respondents did not use test specification table while writing test items. Again, the findings of the present study showed that respondents often copied test questions from textbooks.

The findings of this study was in line with the findings of Agu, Onyekub and Anyichie (2012) who found in their study that almost all the teachers took the following skills into consideration while constructing test items: teachers outlined the content covered for the term before setting test items form the content covered, teachers ensured that the items are measuring

the determined objectives, teachers prepared a marking guide while constructing the test items, teachers gave clear instructions to guide test takers. The findings of Agu, Onyekub and Anyichie also indicated that teachers submitted tests meant for promotional examinations for expert editing on time. This finding also agreed with the findings of the current study.

Interestingly, other findings of Agu, Onyekub and Anyichie (2012) contradicted the findings of the present-day study. For instance, their finding that, teachers prepared a test blueprint as a guide in the test construction process as well as the finding that teachers consulted standard text books in the subject for guide while constructing test items were not consistent with the findings of the current study.

Test Administration Practice of Teachers in the Atwima Nwabiagya South District

This research question sought to find out the test administration practices of teachers in the Atwima Nwabiagya South District. The findings of the study brought to bare that, majority of teachers at Atwima Nwabiagya south district in the junior high schools have good administration skills. This evidence is shown in Table 4.

The findings of this study agreed with the assertion of Rukundo and Magambo (2010), they hypothesised that the paramount guiding principle in administering any classroom test is that all examinees should be given a fair chance to demonstrate their achievement of the learning outcomes intended or planned. Rukundo and Magambo expantiated that physical and psychological environment in which the examination is taking place has to be conducive for the examinee to facilitate the achievement of the testing outcome. The authors

further explained that the factors that might interfere with validity of the measurement also have to be controlled. Linn and Miller (2005) postulated that even though the evidence regarding the effects of physical and environmental conditions on test performance is inconclusive, examinees should be as relaxed as possible and distractions should be eliminated or minimised. Linn and Miller further stated that whereas distractions during testing are known to have little effect on the scores of students, they may have profound effect on especially young children.

The findings of the present-day study was also consistent with the declaration of Rukundo and Magambo (2010) who advanced that instructors should avoid giving hints to students who ask about individual items. The authors explained that if the item is ambiguous, it should be clarified for the entire group; instructors should however refrain from helping a pupil to answer an item if it is not ambiguous. Rukundo and Magambo further iterated that the challenge is that at times, refraining from giving hints to examinees who ask for help may be difficult especially for new comers in the field of testing. Nevertheless, giving unfair aid to some students decreases the validity of the test results and lowers class morale.

This study was buttressed the assertion of Cottrell, (2001) who advanced that it is imperative that test administrators are qualified enough and trusted persons. Cottrell, emphasised that this is to ensure that tests are properly managed to obtain valid and reliable results. In Cottrell's view, test administrators need to have the opportunity to learn their responsibilities as a prerequisite to accurate test results. Similarly, the findings of the current study was in harmony with the proclamation of Fontana (1995) who underscored

that a well prepared test is easy to administer, as compared with a poorly prepared test. Fontana further explained that it is equally important to realise that a successful test administration exercise is a product of test planning. Fontana concluded that cheating is most likely to occur in a poorly planned test, thus, proving a challenge to test administration.

Test Scoring Practice of Teachers in the Atwima Nwabiagya South District

This research question sought to find out the test scoring practices of teachers in the Atwima Nwabiagya South District. The findings of the study brought to bare that, majority of teachers at Atwima Nwabiagya south district in the junior high schools have good test scoring skills. The findings of this study was consistent with the assertion of Amedahe and Etsey (2003) who outlined a number of princciples teachers should follow when scoring essay test items. Among some of these principles include: prepare a form of scoring guide, either an analytic scoring rubric or holistic scoring rubric, grade the responses item by item and not script by script, keep scores of previously graded items out of sight when evaluating the rest of the items. The authors emphasized that it is important for teachers to score essay test when they are physically sound, mentally alert and in an environment with very little or no distraction.

Grading Practice of Teachers in the Atwima Nwabiagya South District

This research question sought to investigate the grading practices of teachers in the Atwima Nwabiagya South District. The findings of the study showed good grading practice among teachers in junior high schools at the Atwima Nwabiagya south district which is in support of some past studies.

It was however evident in the findings of the current study that almost half of the respondents more often assigned grades to motivate all students to learn harder. Assessment results have a unique tendency of motivating students to study hard in order to realise their respect goals in life. It is therefore a step in the right direction for instructors to use the motivational tendency of assessment results in grading students to enhance teaching and learning in the classroom.

It is also important to emphasis that instructors should try as much as possible to mark scripts and class exercises of students on time and return these scripts to the student. This will enable such student to have a personal evaluation of their results and make amends where necessary after the entire items on the script have been thoroughly discussed with them by the instructor.

The findings of this study was in harmony with the findings of McMillan (2008) who found that secondary school teachers use a multitude of factors when grading students' work. McMillian postulated that, a mixture of factors in determining grades was organized into four clearly distinct components: academic achievement, academic enablers (such as effort, ability, improvement, and participation), use of external benchmarks, and use of extra credit and borderline cases. Academic achievement was considered to be the most important process in grading students' work.

The findings of McMillian's study indicated that two-thirds of teachers who participated in this study were reported to have agreed with the use of academic enablers (effort, ability, and improvement) when grading students'work. McMillian iterated that teachers defend their choice to use non-

achievement factors, such as effort, because they see them as some form of borderline to determine grades; and that such factors are good proxy for student achievement. The findings of McMillian's study revealed that both primary and secondary school teachers use non-achievement factors when they award grades to their students to raise or lower grades except in borderline cases, teachers were also found to reward hard work by raising borderline grades and some would lower borderline grades for lack of effort.

In furtherance, the findings of this study was in agreement with the findings of Zoeckler (2007) who examined how American high school English language teachers attempted to arrive at a fair grade while weighting both achievement and non-achievement factors and the role of teachers' expectations. Results of Zoeckler's study indicated that grading was influenced by the local grading systems, teachers' perceptions of student effort, and their concerns for moral development. The findings of this study was also consistent with the findings of (Allen, 2005) who recapitulated that grading practices used by many teachers are designed to communicate student's performance in a number of areas, including both academic achievement and behavioral factors such as student effort, conduct and attitude.

The findings of the present-day study was also consistent with the findings of Randall and Engelhard (2010) who found that under most circumstances, teachers abided by the official grading policy of the participating school district, assigning grades based primarily on achievement. However, in some borderline cases, teachers relied more heavily on other student characteristics such as motivation, behaviour and effort.

The findings of the current study was however in contrast with the findings of Lekoko and Koloi (2007) who found that lecturers grade students' work they did not provide adequate comments that could help students understand where they went wrong, teachers gave low marks that are not accounted for in terms of what and how the teacher arrived at the marks, there was no reconciliation of marks and comments accompanying them, and teachers made ticks that were incompatible with the marks given. The main argument that Lekoko and Koloi (2007) made in this study was that when there is a discrepancy between teachers' comments and grades that students receive, students are left frustrated as this robs them of the potential to improve in their learning. For this reason, the authors postulated that it is essential for teachers to be given sufficient assessment training that would enhance their grading practices and equip them with skills of giving effective, efficient, and useful feedback to students.

In sharing a narrative perspective on views about grading and giving students feedback Wormeli (2006) supported the premise of Lekoko and Koloi by emphasising that "Assessment and feedback, particularly during the course of learning, are the most effective ways for students to learn accountability in their work and in their personal lives" (p. 14). Wormeli stressed the need for teachers to use grading and feedback practices that can best serve the interests of students. Some of the recommendations that Wormeli made were that when grading and giving students feedback, teachers should clearly show what students did, what they were supposed to do, and then help them compare and contrast the two.

Demographic Variables (experience, age and qualification) and Teachers' Classroom Assessment Practices

This research question sought to find out which demographic variable (i.e.experience, age, and qualification) predicted more of teachers' classroom assessment. The findings of the study revealed that none of the demographical variables (experience, age and qualification) had a significant effect on teachers' assessment practices. That is to say, neither experience, age nor qualification significantly predicted teachers' assessment practices. Although none of the demographical variables had a significant prediction on teachers' assessment practice, each of the demographical variables (experience, age and qualification) contributed to the total variances in teachers' assessment practices in the classroom. For instance, the findings of the study revealed that variances in assessment practice were explained by the teaching experience of teachers. Although none of the demographic variables (experience, age and qualification) significantly predicted teachers' assessment practice, the total variances in assessment practice revealed that teachers' qualification predicted more of assessment practice.

The findings of the current study was however incongruent with the findings of Koloi-Keaikitse (2012) who found that factors related to teachers' educational level, teaching experience, and assessment training contributed positively to beliefs, skills, and uses of desirable classroom assessment practices. Similarly, the findings of this study was at variance with the findings of Zhang & Burry-Stock (2003) who found that teachers are unique in terms of classroom assessment practices with reference to their teaching experience.

Furthermore, the findings of the study contradicted the findings of Alkharusi (2011c) who found that teaching experience correlated positively with self-perceived assessment skills, and that teachers with in-service assessment training showed a higher level of assessment skills than those without in-service assessment training. Additionally, the findings of this study was not in line with the findings of Mertler (2004) who found that in-service teachers tended to have a higher level of perceived skilfulness in educational assessment than pre-service teachers, thereby testifying the value of including teaching experience when preparing teachers in educational assessment.

Difference in Teachers' Assessment Practices With Reference To Gender

This hypothesis sought to find out whether there was any significant in teachers' assessment practices in terms of gender. The findings of the study revealed a non-significant difference in teachers' assessment practices in terms of gender. This means that both male and female respondents did not significantly show any unique difference regarding assessment practices. In view of this finding, the null hypothesis (H_0) of Hypothesis 1 which states that there is no significant difference in assessment practice with regard to gender was retained.

The findings of this study concurred with the findings of Alsarimi (2000) who found that teachers indicated using short answer, completion, oral exams, extended answer, and multiple-choice item formats with no significant differences based on teacher's gender and years of teaching experience. Alsarimi opined that the teachers indicated using four main sources of information when assigning grades to students: final exams, midterm exams, class participation, and oral questioning.

The findings of this study was however not in line with the findings of Alkharusi (2011b) who found that male teachers tended to have on average a higher level of educational assessment knowledge than female teachers. Again the findings of the current study did not correspond with the findings of Ndalichako (2015) who found a statistically significant difference in teachers' perception of assessment by gender. According to Ndalichako, a t-test analysis indicated that female teachers had more favorable perceptions of assessment than male teachers. The findings of Ndalichako iterated that in terms of classroom practices, the finding suggested that female teachers tend to use assessment more often to facilitate and support teaching and learning than their male counterparts. In the same way, the findings of this study was at variance with the findings of Alkahrusi et al. (2012) who found that although teachers held a favorable attitude towards and perceived themselves as being competent in educational assessment, teachers used a variety of assessments in the classroom primarily for assigning grades and motivating students to learn, with some variations by gender.

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

The purpose of the study was to investigate teacher characteristics as correlate of classroom assessment practices at Atwima Nwabiagya South District J.H.Ss. Specifically, the study sought to find out the relationships that existed between teacher characteristics such as gender, age, experience and qualification on their classroom assessment practices. This chapter presented a summary of the study, key findings, the conclusions drawn as well as the recommendations made. The contribution of the study to knowledge, implications for counselling and suggestions for further study has also been presented.

Overview of the Study

Assessment is an integral aspect of every educational system. It is because assessment involves a systematic process of gathering educational relevant information for legal and instructional decisions. Since valid information is required for making informed decision about students in an educational system, it is important for instructors to employ the appropriate assessment practices in gathering such information. The purpose of this study therefore was to investigate the factors that are associated with assessment practices and skills of teachers at Atwima Nwabiagya South District J.H.Ss. specifically, the study sought to attain the following objectives:

1. The prevalent classroom assessment practices among teachers;

2. The test construction practices of teachers;
3. The test administration practices of teachers;
4. The test scoring practices of teachers;
5. The grading practices of teachers;
6. The contribution of teacher age, qualification and experience in predicting classroom assessment practices;
7. The difference that exist among teachers in terms of how they practice assessment in relation to gender;

To ascertain these purposes of the study, the following research questions and research hypotheses were analysed and tested respectively.

Research Questions

1. What are the prevalent classroom assessment practices among teachers in Atwima Nwabiagya South District?
2. What is the test construction practice of teachers in the Atwima Nwabiagya South District?
3. What is the test administration practice of teachers in the Atwima Nwabiagya South District?
4. What is the test scoring practice of teachers in the Atwima Nwabiagya South District?
5. What is the grading practice of teachers in the Atwima Nwabiagya South District?
6. What demographical variables (age, experience, qualification) predict more of teacher classroom assessment practices?

Research Hypotheses

1. H_0: There is no statistically significant difference between teachers in terms of their assessment practices with reference to gender.

   H_1: There is a statistically significant difference between teachers in terms of their assessment practices with reference to gender.

A descriptive survey design was used for the study. A sample size of two hundred and nineteen (219) respondents were randomly selected from a population of Three hundred and thirty-six teachers (366). A 42-item Teacher Self-Perceived Assessment practices questionnaire was used to gather data from the respondents. In order to ensure that the items on the questionnaires hang together, The instrument was pilot tested using thirty (30) junior high school teachers from Manhyia M/A, Koforidua M/A, Akropong M/A 'A' and 'B' from Atwima Nwabiagya north distrct in the Ashanti region of Ghana. The responses given by respondents were used to refine the questionnaire. The pilot-testing was necessary because it enhanced the content validity and reliability of the instrument. Also, it improved the questions format and scales after careful analysis of the items based on the comments passed by respondents concerning the weaknesses, clarity and ambiguity on all aspects of the questionnaire.

The data collected, was screened to ensure completeness and analysed using the Statistical Product for Service Solution (SPSS) Version 20 software. The demographic variables of the respondents in the study as well as data on research questions one (1), two (2), three (3) and four (4) were analyzed using descriptive statistics of frequencies and percentages. Data on research question five (5) was tested using hierarchical multiple regression. This was because

the researcher sought to find out which demographic variable (i.e., experience, age and qualification) predicted more of classroom assessment practices. Data on hypothesis 1 was tested using independent sample t-test. One way analysis of variance (ANOVA) was used to test data on hypothesis 2 and 3, this was because the researcher was interested in determining whether there are any significant differences exist among the three independent groups.

Key findings

The following were the findings of the study:

1. Teachers used a number of assessment modes and formats in assessing students. Prevalent among these assessment modes and formats used by teachers in the classroom included using essay items to assess students' understanding about a particular concept, multiple-choice items, giving class exercises to students immediately after teaching, the use of oral questions to assess students as well as the use of exams to assess students at the end of the term.
2. It was also found that teachers engaged in the use of some test construction practises. Dominant among these test construction practices included: teachers considered the purpose of the test before developing test items, teachers prepared a marking scheme immediately after constructing test items, instructors matched instructional objectives with test items, teachers however did not use test specification table while writing test items, and teachers copied test questions from test books.
3. In test administration, it was found in the study that teachers ensured good seating arrangements to prevent students from coping from each

other, teachers informed students to write fast during examination in order to avoid failure, teachers did not read novels, newspapers, neither did teacher engaged in "whatsApping" friends during invigilation.

4. The study also found that teachers did not periodically re-score previously scored papers to check consistency in scoring, teachers scored scripts of students without the names of such students known to them. It was also found in the study that the first few essay items scored by teachers, did not influence the rest of the scores teachers gave to students.
5. Regarding the grading practices of teachers, it was found that teachers did not assign grades to students' work to reflect academic performance only, teachers used standard base grading system for any school in grading, teachers did not assign grades purposively to punish non-learning students and teachers did not grade students based on behavioural factors such as conduct and attitude.
6. It was found in the study that none of the demographical variables (experience, age and qualification) had a significant effect on teachers' assessment practices. It was however found in the study that teachers' qualification predicted more of assessment practice.
7. The study found no significant difference in teachers' assessment practices in terms of gender.

Conclusions

It can be concluded from the study that the use of classroom assessment modes and formats such as, giving class exercises to students immediately after teaching, the use of oral questions to assess students, the use

of exams to assess students at the end of the term, the use of multiple-choice items in assessing students' understanding of a particular concept as well as the use of essay items in assessing students were prevalent among teachers in the Atwima Nwabiagya South District. This suggests that teachers used variety of assessment modes and formats to assessment students' understanding of what was thought in the classroom. The prevalence of different assessment formats among the teachers of Atwima Nwabiagya South District could also imply that each assessment format was used based on what the teachers intended to assess. This is to say, the use of a particular assessment format is dependent on the attribute or traits that one intends to measure.

It can also be concluded that teachers of Atwima Nwabiagya South District adhered to a number of test construction principles while crafting test items. Among some of these test construction principles included: considering the purpose of the test before developing test items, preparing a marking scheme immediately after constructing test items, matching instructional objectives with test items. This suggests that the items constructed by teacher measured a specific trait or attribute of interest to the teachers. This could also imply that most of the tests were well planned by the teachers. That's teachers defined what they wanted to assess and obtained an adequate sample of what has been taught in the classroom.

Most of the teachers however did not use test specification table while writing test items. This could imply that teachers did not ensure the representativeness of test items across topics covered and their corresponding learning outcomes. The use of a specification table assists an instructor to give

appropriate weighting to the different topics and the objectives of instruction to reflect their importance in the curriculum. It could also be concluded that although none of the demographical variables (experience, age and qualification) had a significant effect on teachers' assessment practice, qualification to some extent contributed a little more than half of the of the total variances in assessment practice. This could imply that a teachers' level of qualification could to some extent positively impact their knowledge in assessment practices. It can also be said that both male and female teachers in Atwima Nwabiagya South District did not significantly show any unique difference regarding assessment practices. Again, no significant difference was found in teachers' assessment practice with regard to their years of teaching experience. In furtherance, teachers in Atwima Nwabiagya South District did not differ in assessment practice with respect to the various qualifications the teachers had. Based on this, it could be concluded that assessment practices among teachers in Atwima Nwabiagya South District is the same for regardless of their gender, years of experience and qualification.

Recommendations for Policy and Practice

Based on the findings of the findings and the conclusions drawn, the following recommendations were made to guide the development of policy and practice:

1. Head teachers and school administrators should encourage teachers to continue the use of different assessment formats in assessing students. This will expose students the various assessment formats in examinations which will in turn help such students to acquire the

requisite skills in answering questions of different formats both in internal as well as external examinations.

2. Teachers should be trained and encouraged in the use of test specification tables when constructing test items. This would ensure the representativeness of test items across topics covered and their corresponding learning outcomes. The use of a specification table will also assist instructors to give appropriate weighting to the different topics and the objectives of instruction to reflect their importance in the curriculum.
3. Teachers should also be encouraged to desist from copying or lifting questions directly from test books. The services of assessment practitioners could be employed by school administrators to assist teachers in crafting items based on novel situations and problem creativity.
4. Head-teachers and school administrators should encourage and reward teachers who do not read novels, newspapers or engaged in "whatsApping" with friends during examination invigilation. This will perhaps serve as a deterrent to other teachers who occasionally engage in such negative behaviour.

Suggestions for Future Research

1. Future studies could be conducted to examine lecturers, assessment practices in specific courses within departments in tertiary institutions.
2. Similar studies should also be conducted to examine teachers' test construction competences and their attitudes towards crafting of the test items.

3. Future studies could also be conducted to examine the effect of classroom assessment practices on students' achievement goals.

REFERENCES

Adamolekun, J. (1985). Effective test construction: A case for the junior senior secondary school mathematics curricula. *Journal of the School of Pure Sciences, 1*(3), 197-224.

Agu, N., Onyekuba, C., & 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.

Agyedu, G. O., Donkor, F., & Obeng, S. (1999). *Research methods*. Kumasi: University of Education.

Airasian, P. (2001). *Classroom assessment: Concepts and applications* (4th ed.). Columbus, OH: McGraw-Hill.

Airasian, P.W. (2000). *Assessment in the classroom: A concise approach* (2nd ed.). Boston: McGraw-Hill

Akplu, H. (1989). *Implementation of Continuous Assessment in Teacher Training Colleges*. Unpublished Paper, University of Cape Coast

Akyeampong, A. K. (1997). *Continuous assessment in post-secondary teacher training colleges in Ghana: A case study evaluation*. Unpublished Ph.D. Thesis. University of Nottingham, Nottingham: United Kingdom.

Alkharusi, H. (2009). Correlate of teacher education students' academic performance in an educational measurement course. *International Journal of Learning, 16,* 1-15.

Alkharusi, H. (2011a). Teachers' classroom assessment skills: Influence of gender, subject area, grade level, teaching experience and in-service assessment training. *Journal of Turkish Science Education, 8*(2), 20-35.

Alkharusi, H. (2011b). An analysis of the internal and external structure of the teacher assessment literacy questionnaire. *International Journal of Learning, 18*(1), 515-528.

Alkharusi, H., Aldhafri, S., Alnabhani, H., & Alkalbani, M. (2012). Educational assessment attitudes, competence, knowledge, and practices: An exploratory study of Muscat teachers in the sultanate of Oman. *Journal of Education and Learning, 1*(2), 217-232.

Alkharusi, H., Kazem, A. M., & Al-Musawai, A. (2011). Knowledge, skills, and attitudes of pre-service and in-service teachers in educational measurement. *Asia-Pacific Journal of Teacher Education, 39*(2), 113-123.

Allen, J. D., & Lambating, J. (2001). *Validity and reliability in assessment and grading. Perspective of pre-service and in-service teachers and teachers' education.* Paper presented at the annual meeting of American Educational Research Association, Seattle.

Alsarimi, A. M. (2000). *Classroom assessment and grading practices in the Sultanate of Oman*. Unpublished dissertation, University of Pittsburgh, Pennsylvania.

Amedahe, F. K. (1989). *Testing practices in secondary schools in the Central Region of Ghana*. Unpublished Master's thesis, University of Cape Coast, Cape Coast.

Amedahe, F. K. (2000). Issues in combining continuous assessment scores with external examination scores for certification at the basic and secondary school levels. *Journal of Educational Management, 3*,112-123.

Amedahe, F. K. (2002). *Notes on educational research*. Unpublished paper, University of Cape Coast, Ghana.

Amedahe, F. K., & Etsey, Y. K. A. (2003). *Assessment criteria, feedback and evaluation*. Unpublished paper presented at the University of Cape Coast, Ghana.

Amua-Sekyi, E. T. (2016). Assessment, student learning and classroom practice: A review. *Journal of Education and Practice*, *7*(21), 1-6.

Anastasi, A. (1981). *Coaching, test sophistication, and developed abilities. American Psychologist, 36*, 1086-1093.

Andy, F. (2011). The grades game. *Principal Leadership, 11(6),* 48-52.

Anhwere, Y. M. (2009). *Assessment practices of teacher training college tutors in Ghana.* Unpublished Master's thesis, University of Cape Coast, Cape Coast.

Arulappen, D. R. (2013). *Perception of teachers about classroom assessment and self- perceive of teachers' competency in classroom assessment.* University Teknologi Malaysia, Malaysia.

Asamoah-Gyimah, K. (2002). *An evaluation of the practice of continuous assessment in SSS in Ashanti region of Ghana* (Doctoral dissertation, Master's Thesis. University of Cape Coast, Ghana).

Asare, K. (2015). Exploring the kindergarten teachers' assessment practices in Ghana. *Assessment*, *5*(8), 115-120.

Barnes, S. (1985). A study of classroom pupil evaluation: The missing link in teacher education. *Journal of Teacher Education*, *36*(4), 46-49.

Baron, P. A. B. (2000). *Consequential validity for high school grades: what is the meaning of grades for senders and receivers?* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

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

Bishop, J. H. (1992). Why US Students Need Incentives to Learn. *Educational leadership, 49*(6), 15-18.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, *5*(1), 7-74.

Breakwell, G. M., Hammond, S., Fife-Shaw, C., & Smith, J. A. (Eds.). (2006). *Research methods in psychology* (3rd ed.). London: SAGE Publications Ltd.

Broadfoot, P., Osborn, M., Planel, C. & Sharpe, K. (2000). *Promoting quality in education: does England have the answer?* London, Continuum.

Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education, 7*, 279-301.

Capper, J. (2007). *Testing to learn-Learning to test. International Reading Association*. Academy for Educational Development. New Washington: Delaware Publishing Company.

Charmey, D. (1984). The validity of using holistic scoring to evaluate writing: A critical overview. *Research in the Teaching of English, 18*, 65-81.

Chase, C. I. (1979). *Measurement for educational evaluation*. Philipines: Addison Wesley Publishing Company Inc.

Cheng, L., & Wang, X. (2007). Grading, feedback, and reporting in ESL/EFL classrooms. *Language Assessment Quarterly*, *4*(1), 85-107.

Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods, and procedures. *Language Testing*, *21*(3), 360-389.

Cheng, L., Rogers, W. T., & Wang, X. (2008). Assessment purposes and procedures in ESL/EFL classrooms. *Assessment & Evaluation in Higher Education*, *33*(1), 9-32.

Cizek, G. J. (1996). Grades: The final frontier in assessment reform. *Nassp Bulletin*, *80*(584), 103-110.

Clawson, T. W, Firment, C. K. & Trower, T. L. (1981). Test anxiety: Another origin for racial bias in standardized testing. *Measurement and Evaluation in Guidance, 13*(4), 210-215.

Cottrell, S., (2001). *Teaching Study Skills and Supporting Learning* (ed.). New York: Palgrave.

Crook, C. (1998). Children as computer users: The case of collaborative learning. *Computers & Education*, *30*(3-4), 237-247.

Culler, R. E., & Hollohan, C. J. (1980). Anxiety and academic performance. *Journal of Educational Measurement, 72*, 16-20.

Cunningham, G. K. (1986). *Educational and psychological measurement*. New York: Macmillan Publishing Company.

Diederich, P. B. (1967). Cooperative preparation and rating of essay tests. *English Journal, 56*, 145-150.

DuBois, P. H. (1966). *A test dominated society: China 115 B. C. – 1905 A. D. Proceedings of the 1964 Invitational Conference on Testing Problems. In A. Anastasi (Ed.)*. Testing Problems in Perspective, (p. 56). Washington D. C: American Council on Education

DuBois, P. H. (1970). *A history of psychological testing*. Boston: Allyn and Bacon Inc.

Ebel, R. L. (1972). *Essentials of Educational Measurement*. Englewood Cliffs. New Jersey: Prentice Hall Inc.

Fink, O. (1995). *How to sample in surveys. California*: Sage Publications Inc

Fontana, D., (1995). *Psychology for teachers* (3rd ed.). New York: Palgrave.

Fraenkel, J. R., Wallen, N. E., Hyun, H. H. (2012). *How to design and evaluate research in education.* The McGraw-Hill Companies, Inc.

Frankel, J. R., & Wallen, N. E. (2000). *How to design and evaluate research in education*. New York: McGraw Hill.

Frary, R. B. (1981). Cheating? *Teaching and Learning, 2*, 3-4.

Frary, R. B., Cross, L. H., & Weber, L. J. (1993). Testing and grading practices and opinions of secondary teachers of academic subjects: Implications for instruction in measurement. *Educational Measurement: Issues and Practice*, *12*(3), 23-30.

Freedman, S. (1979). How characteristics of student essays influence teachers' evaluations. *Journal of Educational Psychology, 71,* 328- 338.

Freedman, S. (1981). Influences on evaluators of expository essays: Beyond the text. *Research in the Teaching of English, 15*, 245-255.

Freedman, S. W, & Calfee, R. C. (1983). Holistic assessment: *Experimental design and cognitive theory*. In P. Mosenthal, L. Tamor & S. W.

Walmsley (Eds.). Research on writing: Principles and methods (pp. 75 98). New York: Longman.

Friedman, D.J., & Frisbie, D. A. (1995). The influence of report cards on the validity of grades reported to parents. *Educational and Psychological measurement, 55*(1), 5- 26.

Gay, R. L. (1992). *Educational research: Competencies for analysis and application*. New York: Macmillan Publishing Company

Ghana Education Service, Atwima Nwabiagya South District Metropolitan Directorate (2018). *Teachers' population*. Kumasi: Author.

Gipps, C. (1990). Begond testing: *Towards a theory of educational assessment*. London: Falmer Press.

Gronlund, G. (2003). *Focused early learning: A planning framework for teaching young children*. St. Paul, MN: Red Leaf Press.

Gullickson, A. R. (1982). The practice of testing in elementary and secondary schools.

Haigh, M. (2007). Sustaining learning through assessment: an evaluation of the value of a weekly class quiz. *Assessment & Evaluation in Higher Education*, *32*(4), 457-474.

Harlen, W., & Deakin Crick, R. (2003). Testing and motivation for learning. *Assessment in Education: Principles, Policy & Practice, 10*(2), 169-207.

Hattori, K., & Saba A. N. (2008). Comparison of classroom assessment practices: A case of Selected Ghanaian and Japanese mathematics lessons. *Journal of Educational Comparison 3*, 95-105.

Hills, J. R. (1991). Apathy concerning grading and testing. *Phi Delta Kappan, 72*, 540-545.

Kankam, B., Bordoh, A., Eshun, I., Bassaw, T. K., & Koranteng, F. Y. (2014). An investigation into authentic assessment practices of social studies teachers in the senior high schools (SHSs) in Ghana. *American Journal of Social Sciences, 2*(6), 166-172.

Karagöz, Y. (2016). *SPSS and Amos 23 applied Statistical Analysis.* Ankara, Turkey: Nobel Publishing House.

Kirsch, I. S., & Guthrie, J. T. (1980) Construct validity of functional reading tests. *Journal of Educational Measurement, 17*, 81-93.

Koul, L. (1997). *Methodology of Educational Research*. New Delhi: Vikas Publishing House Pvt. Ltd.

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

Leeming, F. C. (2002). The exam-a-day procedure improves performance in psychology classes. *Teaching of Psychology*, *29*(3), 210-212.

Lekoko, R. N., & Koloi, S. (2007). Qualms in marking university students' assignment by teaching staff: Does correlation of student's expectations and teacher's feedback matter? *Journal of Business, Management and Training, 4*, 34-45.

Lindquist, E. F. (1951). *Preliminary considerations in objective test construction*. In E. F. Lindquist, (Ed.). Educational Measurement (pp. 119-158). Washington DC: American Council on Education.

Linn, R. L., & Miller, M. D., (2005). *Measurement and Assessment in Teaching* (9th ed.). New Jersey: Pearson.

Linn, R.L., & Gronlund, N.E. (2000). *Measurement and assessment in teaching* (8th ed.). Englewood Cliffs, NJ: Merrill/Prentice Hall.

Marcell, M. (2008). Effectiveness of regular online quizzing in increasing class participation and preparation. *International Journal for the Scholarship of Teaching and Learning*, *2*(1), 7-10.

Marczyk, G., DeMatteo, D., & Festinger, D. (2005). *Essentials of research design and methodology*. New Jersey: John Wiley & Sons Inc.

McMillan, J. (2008). *Educational Research Fundamentals for the consumer*. Pearson Education: Boston.

McMillan, J. H., & Nash, S. (2000). Teacher Classroom Assessment and Grading Practices Decision Making.

McMillan, J. M. (2008). *Assessment essentials for student-based education* (2nd ed.). Thousand Oaks: Crown Press.

Mehrens, W. A. & Lehmann, I. J. (1991). *Measurement and evaluation in education and psychology*. New York: Harcourt Brace College Publishers.

Mehrens, W. A. (1984a). National tests and local curriculum: Match or mismatch? Educational Measurement: *Issues and Practice*, *3*(3), 9-15.

Mehrens, W. A., & Lehmann, K. (1999). *Measurement and evaluation in education and psychology* (4th ed.). New York: Wadsworth.

Mehrens, W., A. & Kaminski, J. (1989). Methods for improving standardized test scores: Fruitful, fruitless or fraudulent? *Educational Measurement: Issues and Practice, 8* (1), 14-22.

Mertler, C. A. (1998). Classroom assessment practices of Ohio teachers *Improving Schools, 12*(2), 101-113.

Mertler, C. A. (1999). Assessing student performance: A descriptive study of the classroom assessment practices of Ohio teachers. *Education*, *120*(2), 285-285.

National College Testing Association (2010). Professional standards and guidelines. Retrieved from http://www.ncta.testing.org/resources.

National Council on Measurement in Education (NCME). (1999). *Standards for educational and psychological testing*. Washington, D. C: Author

Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in education*, *14*(2), 149-170.

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

Nwadinigwe, I. P. (2005). *Fundamentals of research methods and statistics*. Lagos: Sibon Books Limited.

Oduro, E. O. (2015). *Assessment in mathematics classrooms in Ghana: a study of teachers' practices* (Doctoral dissertation, University of Sussex).

Ogunkola, B. J. (2013). Instructional assessment practices of science teachers in Barbados: pattern, techniques and challenges. *Academic Journal of Interdisciplinary Studies*, *2*(1), 57-65.

Ogwang, S. G., (2007). The persistence of malpractice in examinations conducted by Uganda National Examinations Board. *African Journal of Teacher Education*, *1*(1), 155-160.

Oosterhof, A. (2001). *Classroom application of educational measurement*. Upper Saddle River. NJ Prentice Hall.

Popham, J. 2008. *Transformative Assessment*. Alexandria, VA: Association for Supervision and Curriculum Instruction.

Quansah, F. & Amoako, I. (2018). Attitude of senior high school teachers toward test construction: developing and validating a standardised instrument. *Research on Humanities and Social Sciences, 8*(1), 2224-5766.

Quansah, F. A, Amoako, I. & Ankomah, F. (2018). Teachers' Test Construction Skills in Senior High Schools in Ghana: Document Analysis. *International Journal of Assessment Tools in Education, 6*(1), 1–8.

Randall, J., & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education, 26*(7), 1372-1380.

Roberts, D. M. (1987). Limitations of the score-difference method in detecting cheating in recognition test situation. *Journal of Educational Measurement, 24* (1), 77-81.

Roediger III, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*(3), 181-210.

Rukundo, A. & Magambo, J. (2010). Effective test administration in schools: principals and good practices for test administrators in Uganda. African Journal of Teacher Education, 1 (1), 166-173.

Sarantakos, S. (2005). *Social research* (4th ed.). New York: Pelgrave Macmillan.

Sarason, H., Hill, K., & Zimbrado, P. (1964). A longitudinal study of the relation of test anxiety to performance on intelligence and achievement

tests. *Monographs of the Society for Research in Child Development, 29*, 28-34

Scanell, D. P., & Marshall, J. C. (1966). Effect of selected composition errors on grades assigned to essay examinations. *American Educational Research Journal, 3*, 125-130.

Scriffiny, P. L. (2008). Seven reasons for standards-based grading. *Educational Leadership*, *66*(2), 70-74.

Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A., & Yin, Y. (2008). On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied measurement in education*, *21*(4), 295-314.

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.

Shirvani, H. (2009). Does your elementary mathematics methodology class correspond to constructivist epistemology? *Journal of instructional Psychology*, *36*(3), 245-259.

Simon, M., Chitpin, S., & Yahya, R. (2010). Pre-service teachers' thinking about student assessment issues. *International Journal of Education, 2*(2), 1-12.

Snowman, J. & Biehler, R.F (2003). *Psychology applied to teaching* (10th ed.). Boston: Houghton Mifflin.

Sofo, S., Ocansey, R. T., Nabie, M. J., & Asola, E. F. (2013). Assessment practices among Secondary Physical Education Teachers in Ghana. *International Online Journal of Educational Sciences, 5*(2), 274-281.

Spandel, V. (1981). *Classroom application of writing assessment: A teacher's handbook.* Portland, OR: North–west Regional Education Laboratory.

Spandel, V., & Stiggins, R. J. (1988). *Writing assessment in the classroom.* White Plains, New York: Longman.

Stanes, D. L. (1992). Analytic responses to conceptual style test as a function of instructions. *Child Development, 44*, 389-391.

Stiggins, R. (1991). *Facing the challenge of the new era of assessment*. Portland, OR: Northwest Evaluation Association.

Stiggins, R. J. (1991). Relevant classroom assessment training for teachers. *Educational Measurement: Issues and Practice*, *10*(1), 7-12.

Stiggins, R. J. (2001). *Student-involved classroom assessment* (3rd ed.) .Upper Saddle Rivers, NJ: Merrill/Prentice Hall.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, *83*(10), 758-765.

Stiggins, R. J., Frisbie, D. A., & Griswold, P. A. (1989). Inside high school: Building a research agenda. Educational Measurement: *Issues and Practice, 8*, 5-14.

Stiggins, R., & Chappuis, S. (2005). Putting Testing in Perspective: It's for Learning. *Principal Leadership*, *6*(2), 16-20.

Tamakloe, E. K., Atta, E. T., & Amedahe, F. K. (1996). *Principles and methods of teaching*. Accra: Black Mask Ltd.

Tanner, H., & Jones, S. (2007). Using video- stimulated reflective dialogue to learn from children about their learning with and without ICT. *Technology, Pedagogy and Education*, *16*(3), 321-335.

Thorndike, E. L. (1903). *An introduction to the theory of mental and social measurements*. New York: The Science Press.

Thorndike, R. L., & Hagen, E. (1977). *Measurement and evaluation in psychology*. New York: Wiley.

Tiedt, I. M. (1983). *The language arts handbook*. Englewood Cliffs, NJ: Prentice-Hall.

Trentham, L. L. (1975). The effect of distractions on sixth grade students in a testing situation. *Journal of Educational Measurement, 12*, 13-17.

Trochim, W. (2000). *The research methods knowledge base* (2nd ed.). Cincinnati, O. H.: Atomic Dog Publishing.

Tyler, R. W. (1950). *Basic principles of curriculum and instruction*. Chicago: University of Chicago Press.

Vacc, N. B. (1998, September). Writing evaluation: Examining four teachers' holistic and analytic scores. *The Elementary School Journal, 90*(1), 87-95.

Vickers, D., & Nichols, P. (2005). The comparability of online vs. standup training. *Journal of Educational Measurement, 31*(3), 220-233.

Ward, J. G. (1980). Teachers and Testing; *A Survey of Knowledge and Attitudes*.

Weeden, P., Winter, J., & Broadfoot, P. (2002). *Assessment*. Psychology Press.

Wolf, C. (2007). *Construction of an adaptive e learning environment to address learning styles and an investigation of the effect of media choice*. Dissertation, RMIT University, Melbourne.

Wormeli, R. (2006). Accountability: teaching through assessment and feedback, not grading. *American Secondary Education, 34,* 14-27.

Zhang, Z., & Burry-Stock J. A. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education, 16*(4), 323-342.

Zoeckler, L. G. (2007). Moral aspects of grading: A study of high school English teachers' perceptions. *American Secondary Education, 35,* 83-102.

APPENDICES

APPENDIX A

QUESTIONNAIRE

This questionnaire aims to collect data that will help the researcher to explore teacher characteristics as correlates of classroom assessment practices which is a chosen area of study. The exercise is for academic purpose only. Whatever information you give will be kept confidential. The questionnaire measures your perception and experience as a teacher. Instructions to fill out the questionnaire are given at the top of each section. Thank you very much for your cooperation.

SECTION A

BACKGROUND DATA OF RESPONDENT.

Please tick (√) or provide the appropriate response.

1. Gender

   Male [ ] Female [ ]

2. Number of years you have been teaching

   1 – 3 Years [ ]

   4 – 7 Years [ ]

   8 – 11 Years

   12 years and above [ ]

3. Highest Academic Qualification

   Cert 'A' [ ]

   Diploma in Education [ ]

   Bachelor of Education [ ]

   BA/B.Sc (without education) [ ]

   Master of Education (M.Ed) [ ]

MA/MSc [ ]

M.Phil in Education [ ]

PHD [ ]

4. Age-range in years

20 – 29 [ ]

30 -39 [ ]

40 – 49 [ ]

50-59 [ ]

SECTION B

TEACHERS ASSESSMENT MODES AND FORMATS

Indicate with a tick [√] your level of practice on the following activities regarding test modes and formats.

SN.	Statements	More often	Often	Not often	Not used
5.	I give exercises to assess my students immediately after teaching.				
6.	I give homework to assess my students.				
7.	I ask oral questions to assess my students during lesson periods.				
8.	I give exams to assess my students at end of the term.				
9.	I give essay items to assess my students.				
10.	I construct multiple –choice items to assess my students.				
11.	I use true/false items to assess my students				

SECTION C

TEACHERS TEST CONSTRUCTION PRACTICES

Indicate with a tick [√] your level of practice on the following activities regarding test construction

SN.	Statements	More often	Often	Not used
12.	I write individual test items two weeks before the date of testing.			
13.	I prepare a marking scheme immediately after constructing my items.			
14.	I copy questions from past BECE examinations questions.			
15.	I develop test items only when it is time to assess students.			
16.	I use a test specification table when writing test items.			
17.	I copy test questions from textbooks.			
18.	I match instructional objectives with test items.			
19.	I write items based on information that students know.			
20.	I arrange objective test answers in a pattern to make scoring easy.			
21.	I consider the purpose of the test before developing test items			
22.	My school conducts in-service training in test construction for teachers.			
23.	I prepare more items than needed before I review and select some for the test.			
24.	I evaluate the test as a whole before I make the final copy.			

SECTION D

TEST ADMINISTRATION AND SCORING PRACTICES

Indicate your level of practice with a tick [√] on the following activities regarding test administration and scoring

SN.	Statements	More often	Often	Not used
25.	I ensure good seating arrangements to prevent students from copying from each other.			
26.	During examinations, I tell students that if they do not write fast, they will fail.			
27.	During examinations, I prepare for and expect emergencies.			
28.	I give hints to students when they ask about individual test items during examinations.			
29	I read novels, newspapers or WhatsApp friends when I invigilate tests/examinations.			
30.	I inform students in advance about content/topics that the tests/examinations cover.			
31	I score essay tests, question by question.			
32.	I prepare marking schemes for essay tests after students have taken the test.			
33.	I periodically re-score previously scored papers to check consistency in scoring.			
34.	I score answer scripts with the names of the students known to me.			
35.	I keep previously scored items out of sight when scoring the rest of the items.			
36.	The first few essays I score influence the rest of the scores I give.			
37.	I give a separate mark for the mechanics of writing such as correct grammar, flow of expression etc. when scoring essays.			

SECTION E

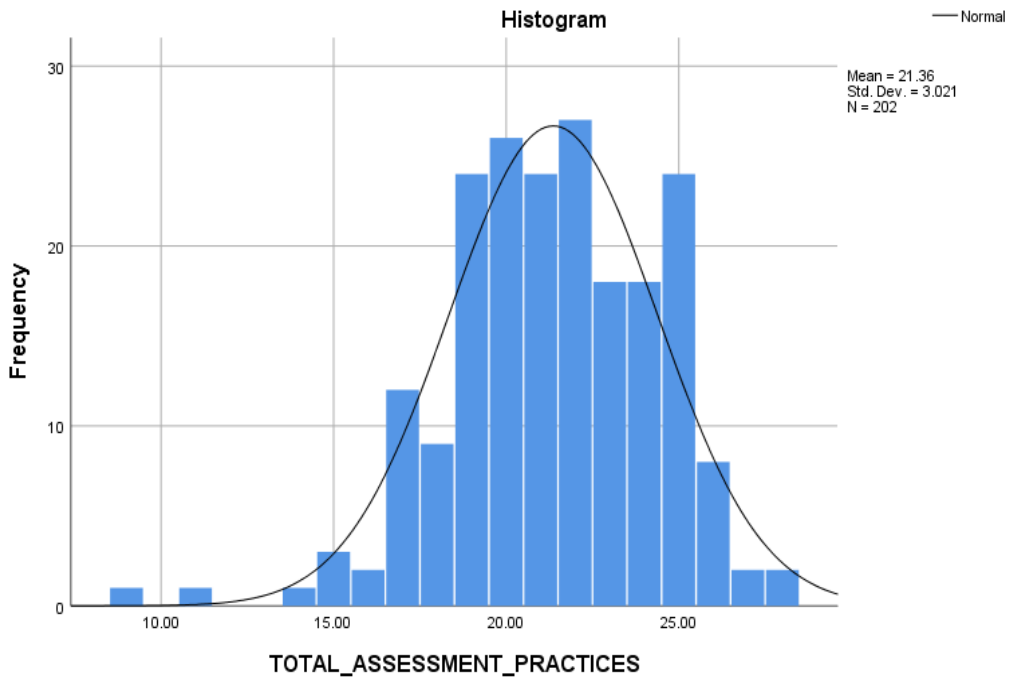
GRADING PRACTICES OF TEACHERS

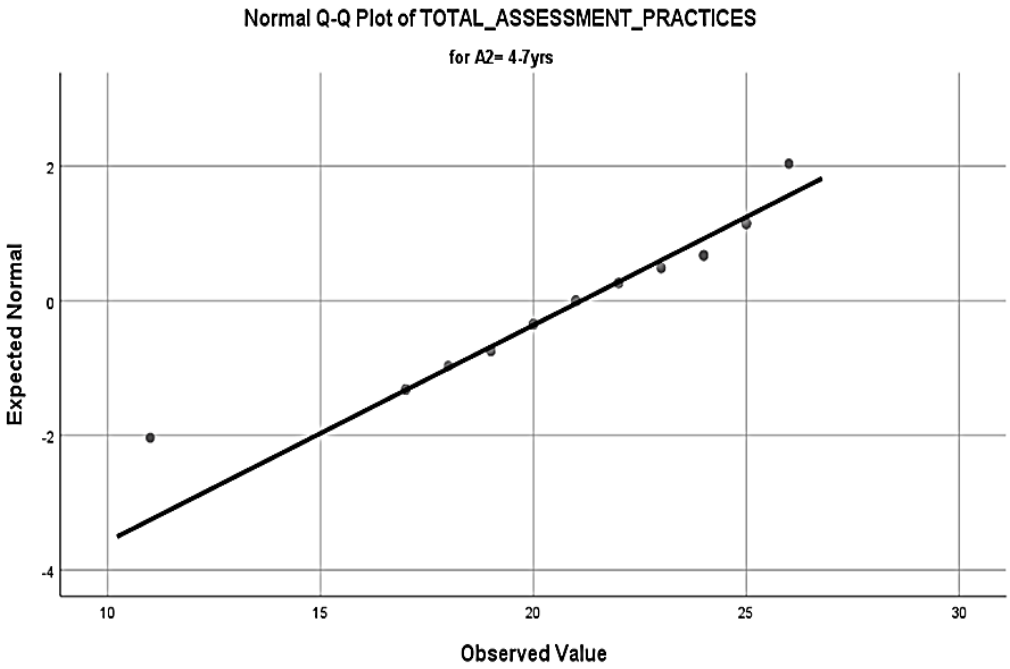
Indicate your level of practice with a tick [√] on the following activities regarding the test.

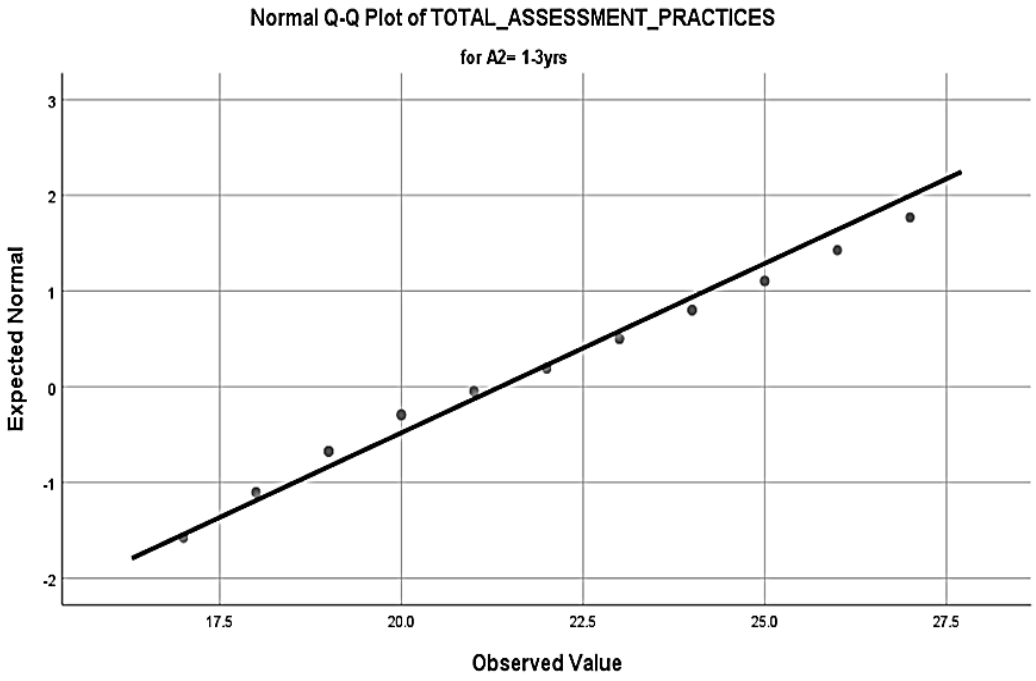
SN.	Statements	More often	Often	Not used
38.	The grade that I assigned to students work communicate academic performance only.			
39.	I assign grades purposely to punish non-learning students.			
40.	I assign grades to motivate all students to learn harder.			
41.	I use a standard base grading system for my school in grading.			
42.	Behavioural factors of students such as conduct, effort and attitude affect my grading.			

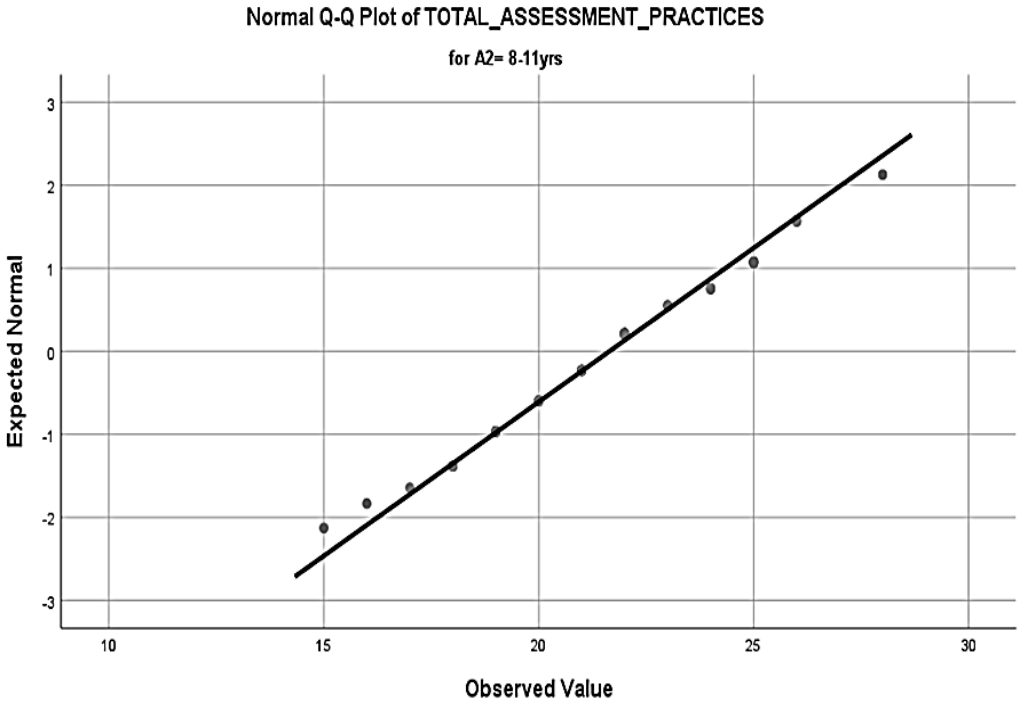
APPENDIX B

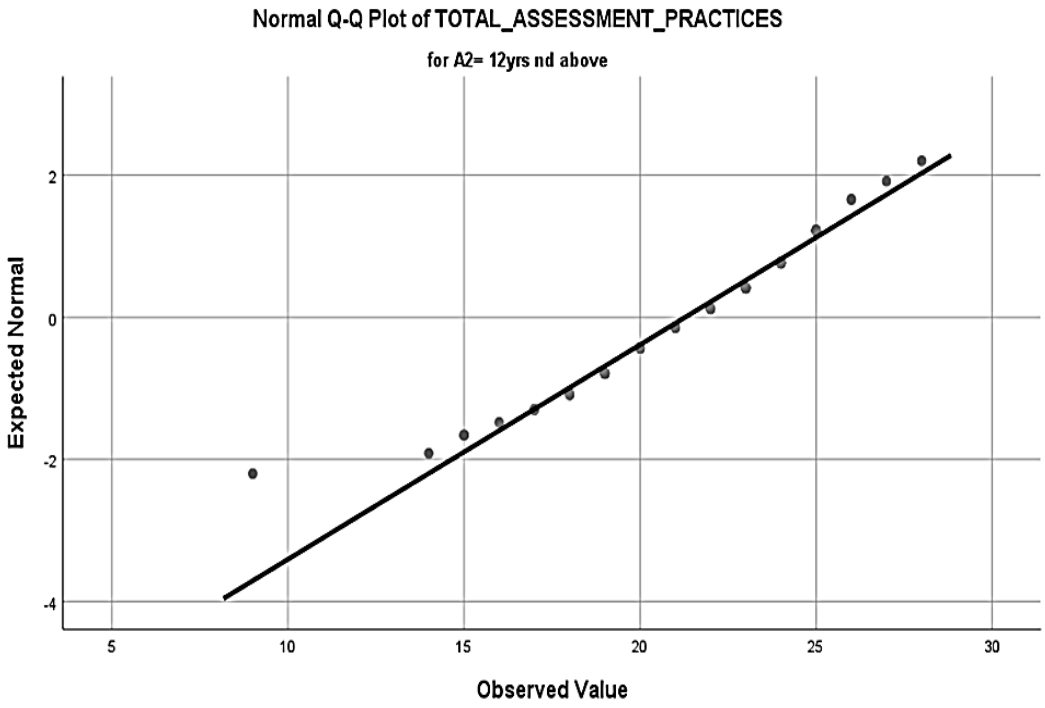
NORMALITY TEST

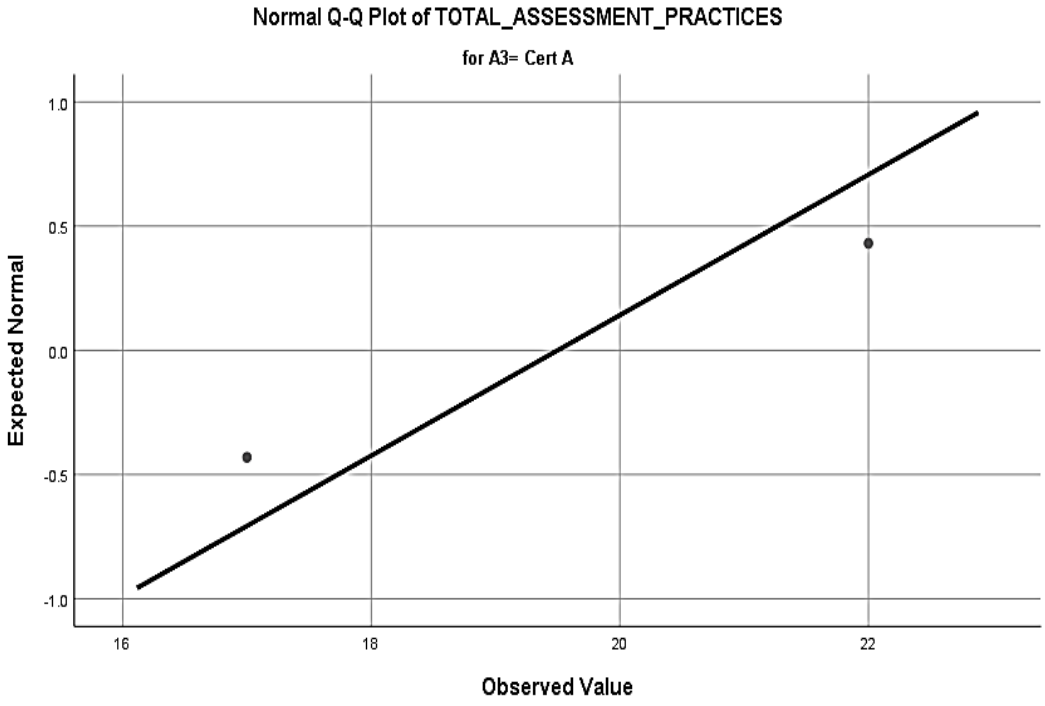


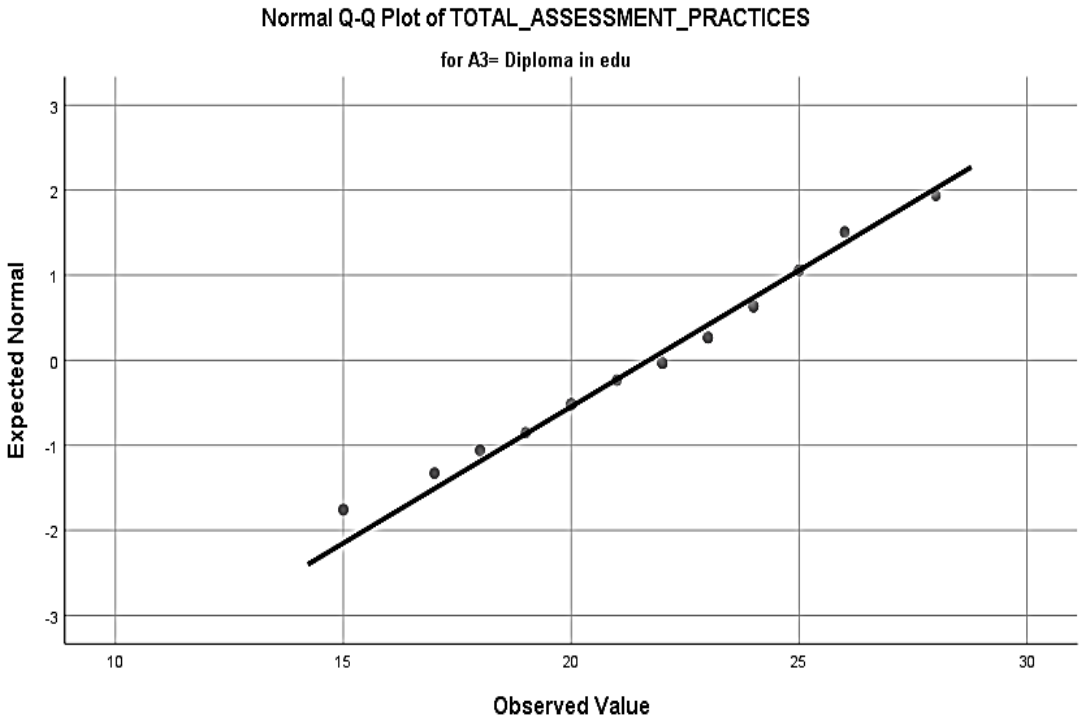


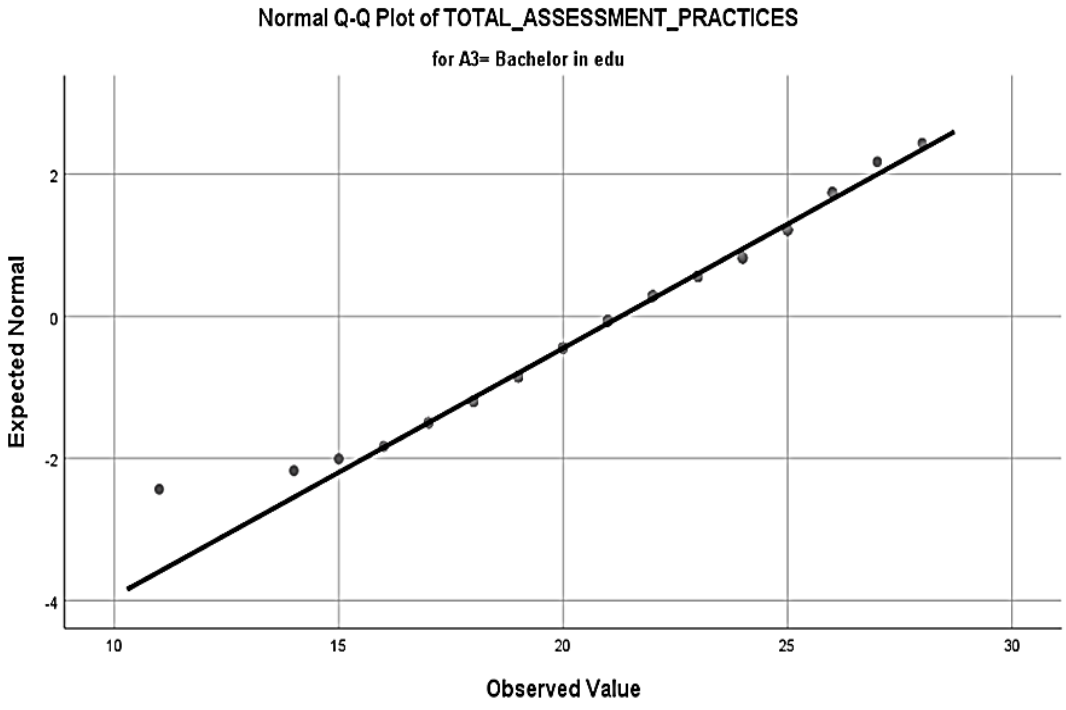


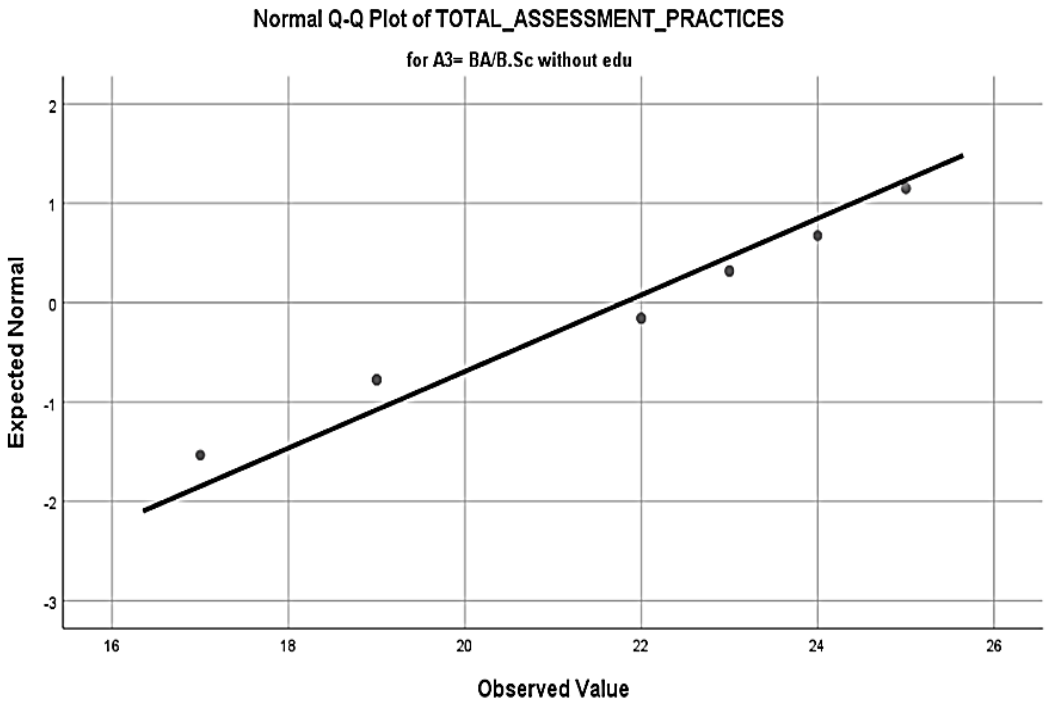


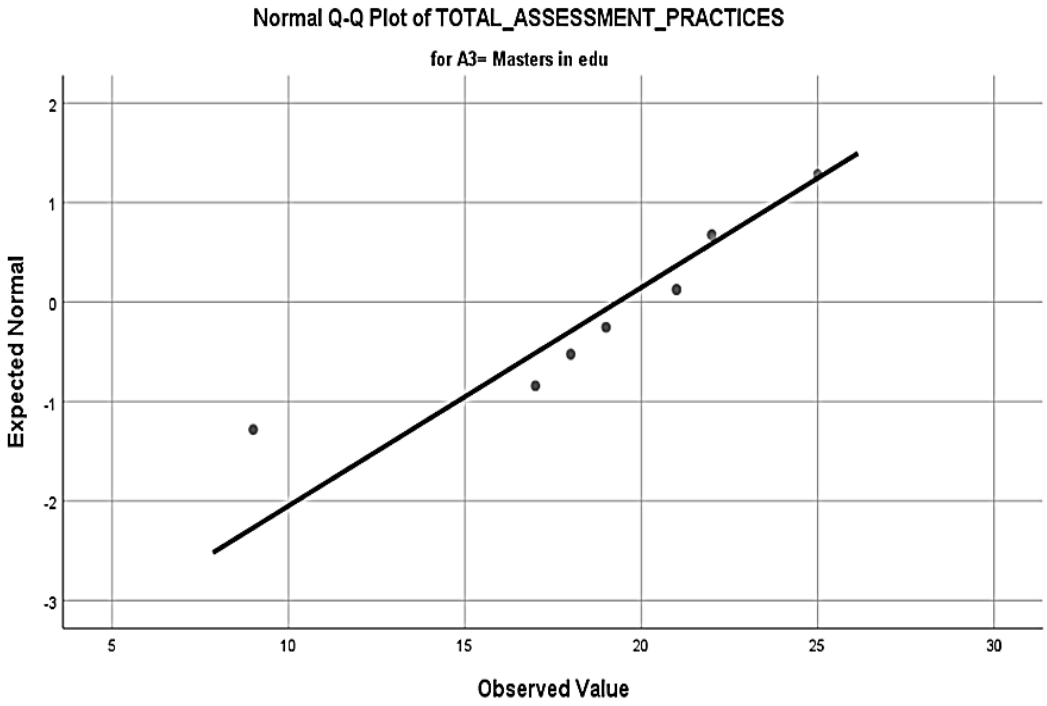


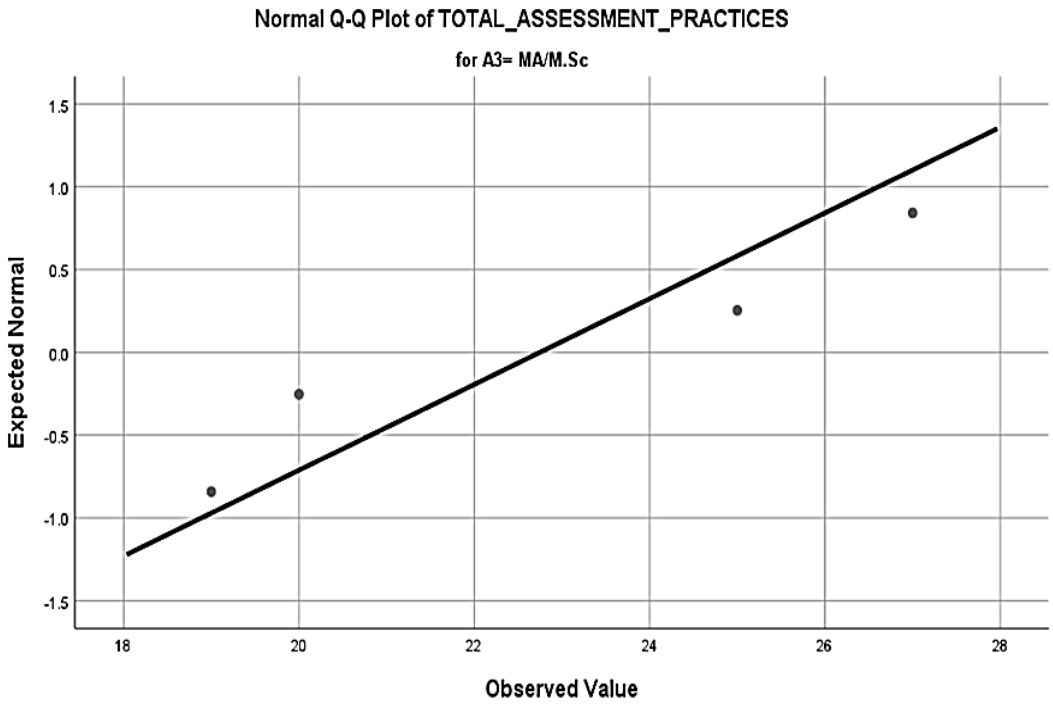






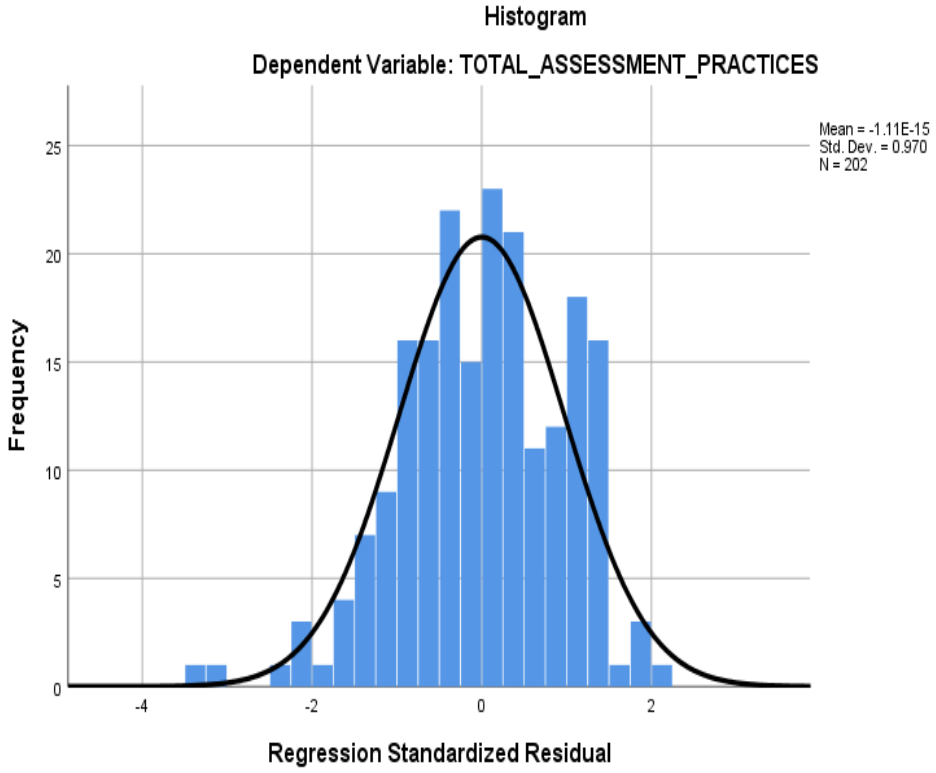


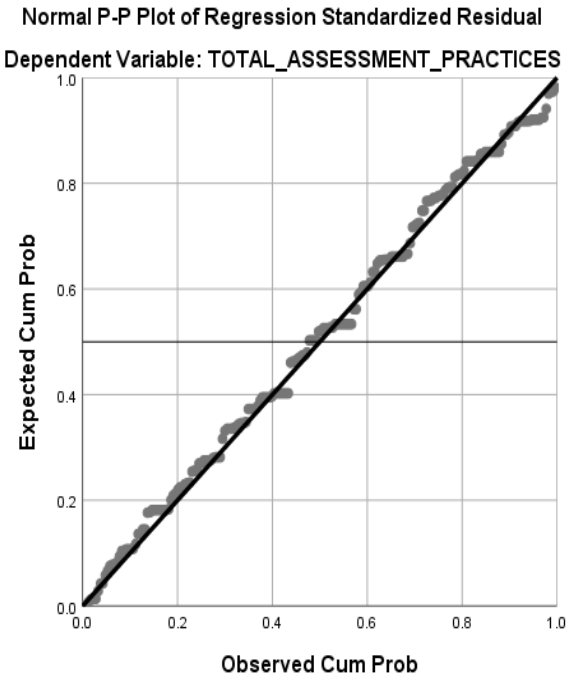


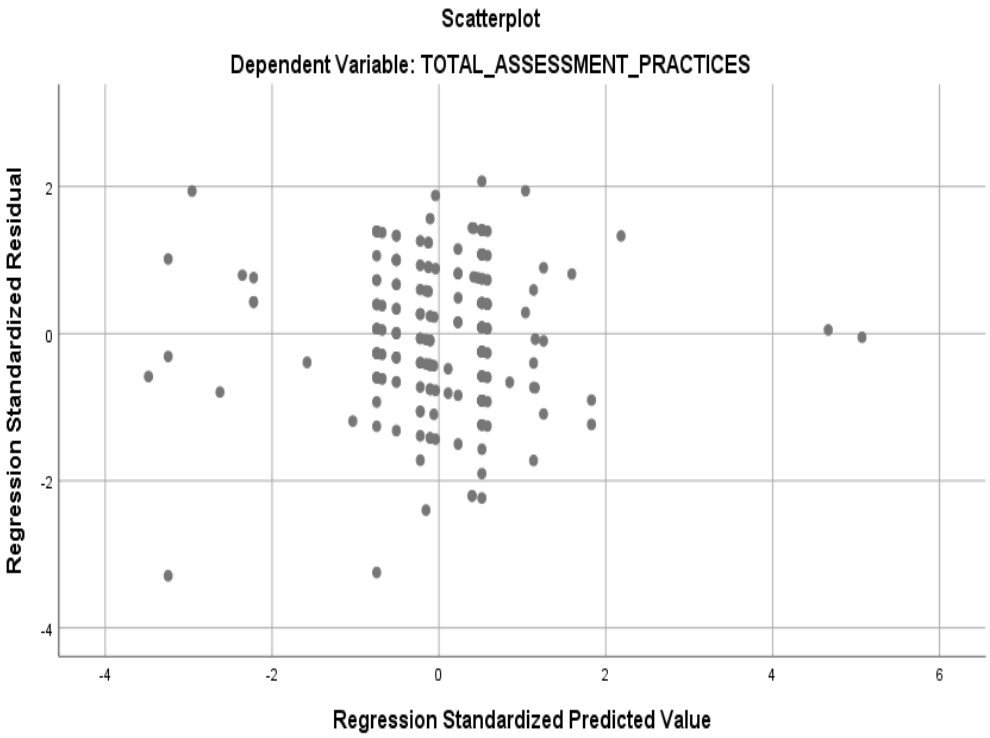


APPENDIX C

STANDARDISED REGRESSION PLOTS

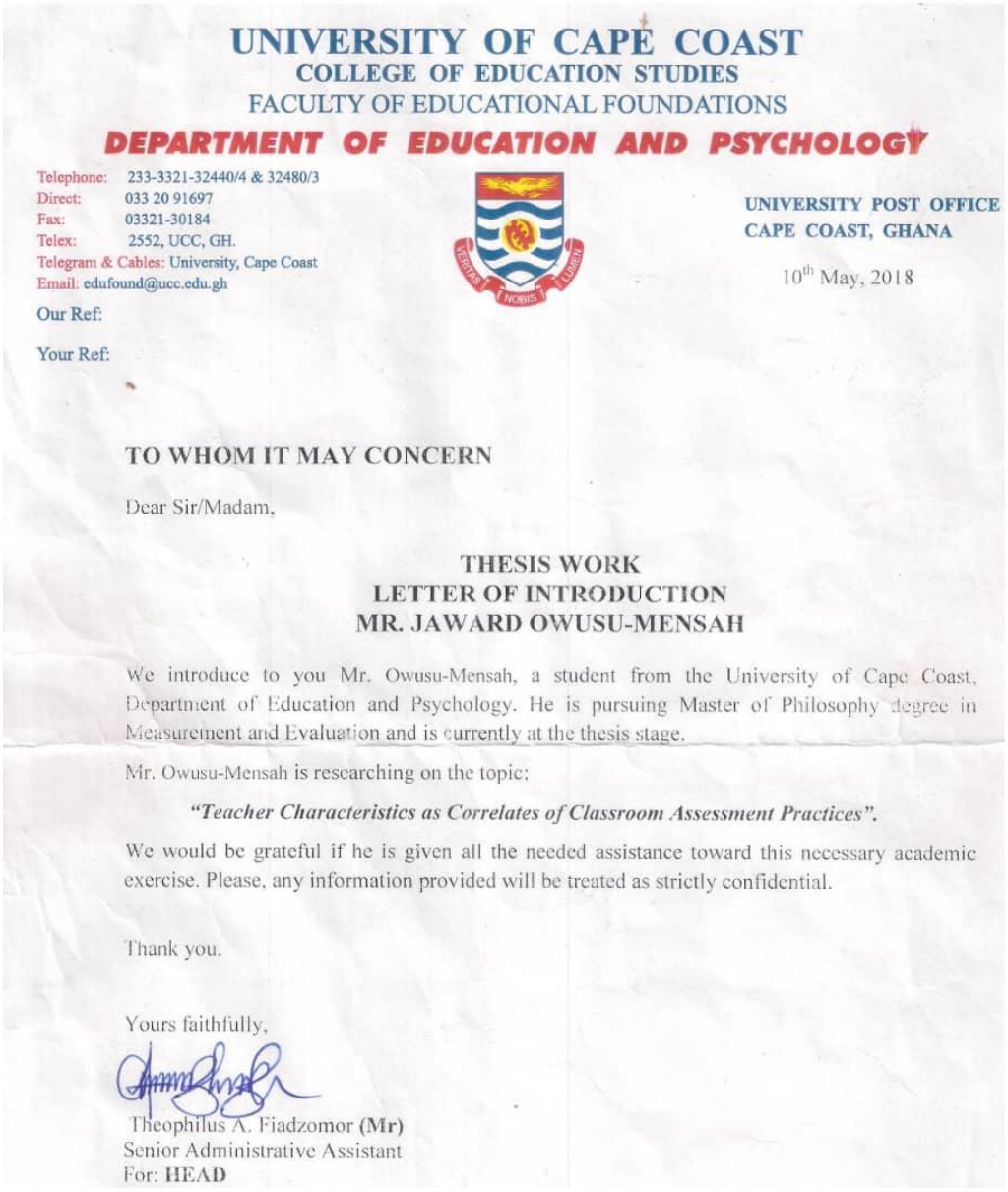






APPENDIX D

INTRODUCTION LETTER



UNIVERSITY OF CAPE COAST
COLLEGE OF EDUCATION STUDIES
FACULTY OF EDUCATIONAL FOUNDATIONS
DEPARTMENT OF EDUCATION AND PSYCHOLOGY

Telephone: 233-3321-32440/4 & 32480/3
Direct: 033 20 91697
Fax: 03321-30184
Telex: 2552, UCC, GH.
Telegram & Cables: University, Cape Coast
Email: edufound@ucc.edu.gh

UNIVERSITY POST OFFICE
CAPE COAST, GHANA

Our Ref:

Your Ref:

10th May, 2018

TO WHOM IT MAY CONCERN

Dear Sir/Madam,

THESIS WORK
LETTER OF INTRODUCTION
MR. JAWARD OWUSU-MENSAH

We introduce to you Mr. Owusu-Mensah, a student from the University of Cape Coast, Department of Education and Psychology. He is pursuing Master of Philosophy degree in Measurement and Evaluation and is currently at the thesis stage.

Mr. Owusu-Mensah is researching on the topic:

"Teacher Characteristics as Correlates of Classroom Assessment Practices".

We would be grateful if he is given all the needed assistance toward this necessary academic exercise. Please, any information provided will be treated as strictly confidential.

Thank you.

Yours faithfully,

Theophilus A. Fiadzomor **(Mr)**
Senior Administrative Assistant
For: **HEAD**