UNIVERSITY OF CAPE COAST

PREDICTORS OF ADHERENCE TO TEST CONSTRUCTION

PRINCIPLES: THE CASE OF SENIOR HIGH SCHOOL TEACHERS IN

SEKONDI-TAKORADI METROPOLIS

FRANCIS ANKOMAH

2020

UNIVERSITY OF CAPE COAST

PREDICTORS OF ADHERENCE TO TEST CONSTRUCTION

PRINCIPLES: THE CASE OF SENIOR HIGH SCHOOL TEACHERS IN

SEKONDI-TAKORADI METROPOLIS

BY

FRANCIS ANKOMAH

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 Measurement and Evaluation

FEBRUARY 2020

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 thesis laid down by the University of Cape Coast.

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

Name: ……………………………………………………………………….

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

Name: ………………………………………………………………………

ii

ABSTRACT

The study examined teachers' test construction skills, commitment to work, and attitude towards test construction as predictors of adherence to test construction principles among SHS teachers in the Sekondi/Takoradi Metropolis. The study was carried out employing the descriptive cross-sectional survey design. The study targeted teachers who teach mathematics, integrated science, and English language, numbering 346. The census method was used to engage all the 346 teachers in the study. Questionnaire validated using IRT and CFA was used for the study. The data collected were analysed using frequencies, percentages, mean and standard deviation, Structural Equation Model (SEM), and mediation analysis with Hayes' PROCESS, with bootstrap samples. It was revealed that the teachers had high level of skills in test construction, and they were somehow committed to their work. The teachers also had positive attitude towards test construction. It was further revealed that commitment to work and attitude towards test construction do not serially mediate the relationship between test construction skills and adherence to test construction principles. However, attitude towards test construction, but not commitment to work solely mediates the relationship between test construction skills and adherence to test construction principles. It was concluded that positive attitude towards test construction is a requirement for adherence to test construction principles. The Ministry of Education, Ghana (MoE), Ghana Education Service (GES), and heads of the various SHSs, as part of their training programmes or workshops for teachers, are entreated to continue and intensify the acquisition of skills by teachers in test construction to improve their skills.

ACKNOWLEDGEMENTS

iv

DEDICATION

To my parents and siblings.

TABLE OF CONTENTS

LIST OF TABLES

x

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

The global agenda of promoting quality education, as specified in the Sustainable Development Goal (SDG) 4, seeks to ensure inclusive and equitable quality education and promote learning opportunities for all (UNESCO, 2017). In line with this agenda, member countries such as Ghana has come out with several policies to regulate and bind the teaching and learning activities, which are geared towards promoting quality education in Ghana (Ministry of Education, MoE-Ghana, 2017). The Target 1 of the SDG 4, states that "by 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes". It is noteworthy to state that teachers are key actors as far as quality education is concerned. Teachers are the main implementers of the curriculum and for that matter issues that bother on curriculum implementation activities such as teaching and assessment of students' learning must be given the needed attention.

The target of education leading to relevant and effective learning outcomes, can only be realised through effective and comprehensive assessment of students' learning. Assessment is the means through which learning can be made effective and relevant to learners. The issue of teacher assessment of

students' learning in Ghana, has over the years not been properly done, as it has manifested in many studies reporting poor testing skills/competencies among teachers (Amedahe, 1993; Anhwere, 2009; Quaigrain, 1992; Quansah, Amoako, & Ankomah, 2019). This happens to be the focus of the current study. The Nation Teachers' Standards for Ghana Guidelines of the Ministry of Education, MoE-Ghana (2017) spells out three competencies teachers should demonstrate. One of these competencies is professional practice, which has to do with managing the learning environment, teaching and learning, and assessment. Under assessment, the teacher is expected to incorporate a number of assessment techniques into teaching, in order to make teaching more effective. This requires teachers to be skilled in order to carry out this mandate. This study assesses teachers' skills in test construction, which will enable them deliver as have been specified in the Nation Teachers' Standards for Ghana Guidelines.

Background to the Study

The quality of education system and the profession of educators depend mainly on teachers who are responsible for guiding, and carrying most of the important tasks and activities in educational institutions (Swarnalatha, 2016). Teachers play a significant role in supporting and promoting learning in the class. Amedahe (2014) indicated that the importance of the teacher in the education system cannot be overemphasised; in view of that "the teacher needs to be well trained, motivated and provided with the needed resources to be able to perform his/her duties properly" (p. 56).

2

Many studies have shown that quality of teachers, their motivation and attitude influence the performance of the students and their academic achievement (Amedahe, 2014; Swarnalatha, 2016; Ebinye, 2001). This requires a lot of commitment from the teacher. Teachers should have commitment to the teaching profession and should justify public trust and confidence by providing quality education to all students. The teacher who is committed to the profession has more efficacy, job satisfaction and high competences. The committed teacher strives more for better performance and keeps abreast with latest skills which can be useful in classroom instruction for teaching learning material (Swarnalatha, 2016).

Students' assessment is a sine qua non in the teaching and learning processes. One of the purposes of assessing students, as indicated by Amedahe (2014), is to motivate and/or encourage students to learn, and also help the teacher to evaluate his/her effectiveness as an educator. This, therefore, implies that much attention has to be given to the media or means through which students are being assessed. Teachers are therefore required to have adequate knowledge and skills in order to assess students' achievement in an accurate and precise manner. Teachers must, thus, have the capability in the science and art of test constructing (D'Agostino, 2007).

At all levels of education globally, teachers construct and administer tests as a means of assessing the amount of students' learning and skills students have acquired (Quaigrain & Arhin, 2017). These tests are mostly teacher-made tests. According to Asamoah-Gyimah (2002), classroom or teacher-made tests

are frequently used as a major evaluating device of students' progress in schools in Ghana. Tests play an important role in providing feedback to teachers on their educational actions and for that matter the quality of the test is pertinent. Quaigrain and Arhin (2017) posit that strict adherence to the principles of test construction, test administration and analyses, and reporting is very essential. According to Anamuah-Mensah and Quaigrain (1998), teachers in Ghana are placed in a sensitive and central role in the testing and evaluation process, and this makes it imperative for the classroom teacher to be adequately conversant in the formalized testing techniques to ensure sound facilitation of the herculean task that impinges on their profession.

Silker (2003) noted that test construction requires utilization of skills that can enable a teacher to develop a test with precision, appropriateness of language use, objective communication, items validation and good grading scales. Silker argued that teachers need not be experts in measurement and evaluation to construct valid and reliable instrument needed. However, they need to acquire the general test construction skills to ensure that items are structured to elicit clear and bring out noticeable differences among learners only on whatever construct learners are being assessed on. The argument of Silker to some extent is reasonable in the sense that, generally in Ghana, professionally trained teachers have undertaken at least a course in measurement and evaluation or assessment in the course of their training (Anhwere, 2009), which does not make them expert in measurement and evaluation. These teachers, however, engage in test construction and assessment

4

of students (Quaigrain & Arhin, 2017). Teachers' lack of test construction skills may result in poor performance and false assessment of students' achievement.

Several studies have investigated assessment practices of teachers. For example, studies by Simon (2002) and Ebinye (2001) found that poor test construction have warranted examination malpractices, academic dishonesty in most secondary schools in Nigeria. Chan (2009) observed that classroom teachers pay little attention to the design and development of reliable assessment tools. Chan found that the inappropriate construction of a test instrument was due to lack of knowledge or low level of awareness on the part of teachers' inadequate experience in the preparation of a test blue-print. 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, Hamman-Tukur, and Hamafyelto (2015) discovered that Senior High School (SHS) teachers in Borno State, Nigeria, constructed items which focused on lower cognitive operations. Ololube (2008) also evaluated test construction skills of professional and nonprofessional teachers in Nigeria and reported that professional teachers tend to construct effective evaluative instruments more than the non-professional teachers. Ololube's study further revealed that professional teachers have the tendency to employ the various assessment techniques correctly, which is unlikely to happen in the case of non-professional teachers.

In Ghana, the situation is, however, not different as studies have equally found limited skills and competence in the knowledge of construction of tests

among teachers and subsequently non-adherence to recommended principles in test construction (Amedahe, 1993; Anhwere, 2009; Quagrain, 1992). It has also been found that teachers who have received training in construction of tests followed recommended principles in test construction better than those who did not receive training (Amedahe, 1993; Anhwere, 2009; Oduro-Kyireh, 2008; Quagrain, 1992). Oduro-Kyireh (2008), for example, found statistically significant differences between respondents who received instruction in testing and those who did not receive instruction in testing on five out of the 10 principles. The study concluded that respondents who received instruction in testing applied the principles more frequently than their counterparts who did not receive instruction in testing. The following are the principles that were duly followed:

1. Relate the instructional objectives of the subject matter to the test.

2. Select the test format suitable for testing the stated objectives.

3. Write the test items in advance of the test date to permit reviews and editing.

4. Review the test items after they have been set aside for a few days by reading over them.

5. Write clear and concise directions for the entire test and sections of the test.

The following were, however, not followed:

1. Define the purpose of the test.

2. Use a table of specifications to determine the items on the test.

3. Prepare more items than needed in the test.

4. Prepare marking scheme as soon as the test items are written.

5. Evaluate the test as a whole on the criteria of clarity, validity, practicality and fairness.

From the findings of the previous studies in Ghana, it is clear that teachers have limited knowledge in test construction, and also they do not adhere to recommended principles in test construction. It is possible that teachers do not adhere to recommended principles in test construction because of their limited skills in test construction, however, the methods employed by the previous studies were problematic. This calls for further investigation into test construction skills and adherence to test construction principles.

It is interesting to note, however, that other factors such as teacher commitment and attitude towards test construction could have influence on teachers' adherence to test construction principles. That is to say, a teacher may have the requisite skills and competencies in test construction, but when he/she is not committed to his/her teaching profession, he/she may not follow recommended principles in test construction. Amedahe (2014) stressed the importance of teacher's commitment and attitude towards work. According to Amedahe, the poor attitude and commitment to work by some teachers affect the quality of education. He further added that while some teachers work diligently, others do not show any serious commitment to the work, for the possible reason that "the teaching profession, to some extent, has become the last resort for many; they are in to get a living not to earn it" (p. 57). In relation

7

to this, understanding of the teachers' level of commitment and attitude towards test construction is crucial because they reflect the teachers' involvement in school work and the degree to which the teachers devote time to their teaching profession. This, therefore, triggers the need to further investigate teachers' test construction skills, commitment to work, and attitude towards test construction.

Statement of the Problem

Achievement tests are relevant for measuring important aspects of a subject and also in measuring appropriate level of student's knowledge in a school subject (Frey, 2007). These tests can be either classroom-based, which are produced by an individual classroom teacher (teacher-made tests), or standardized, which are developed by test experts for more global purposes (Amedahe, 2014). The most instructionally-relevant achievement tests, however, are classroom-based and if carefully constructed, provide teachers with accurate and useful information about the knowledge retained by their students in particular school subjects (Childs, 1989). According to Agu, Onyekuba and Anyichie (2013), quality classroom-based assessment means adherence to standard procedures for test construction. Every classroom teacher is expected to possess and apply requisite skills in construction of good items for student assessment. A good test must yield both valid and reliable results. Lack of test construction skills by teachers might result in false assessment of students' achievements and examination malpractice (Anhwere, 2009; Esomonu, 2002; Paulson, 2003; Ebinye, 2001; Quaigrain & Arhin, 2017).

The competency in test construction is an essential tool needed by every teacher if learning and instructional objectives are to be effectively attained. Several studies have been conducted on assessment practices in Ghana (Amedahe, 1993; Amedahe, 1989; Anhwere, 2009; Oduro-Kyireh, 2008; Quagrain, 1992; Quansah et al., 2019). The findings of these studies revealed that teachers do not adhere to recommended principles in test construction and subsequently attributing that to their limited knowledge in test construction. Interestingly, in attempts to measure teachers' test construction skills, the previous studies resorted to asking teachers what principles they follow in developing test items. This, in my view, does not clearly indicate teachers' skills. A teacher might possess the skills, but for reasons best known to him/her, he/she may feel reluctant to adhere to recommended principles. In such an instance, it would be problematic to substitute adherence to test construction with skills or competence in test construction. There is, therefore, the need to conduct this study to ascertain the skills that teachers possess, and determine whether or not teachers' skill would predict adherence to test construction principles.

The issue of teacher commitment to work and attitude towards test construction are of prime concern. Commitment and attitude on the part of teachers are key factors that can affect the quality of education (Amedahe, 2014). A teacher, for example, may have the requisite skills and competencies in test construction, but when he/she is not committed to his/her teaching profession, he/she may not follow recommended principles in test construction.

Similarly, a teacher who has a negative or poor attitude towards test construction and perceive it as burdensome is more likely not to adhere to recommended principles of test construction (Ebinye, 2001). Examining critically the literature, it appears that some studies have investigated commitment among teachers (Swarnalatha, 2016; Dumay & Galand, 2012; Nir, 2002; Park, 2005; Somech & Bogler, 2002), and attitude towards testing (Monsaas & Engelhard, 1994; Kistashvili, 2014; Quansah & Amoako, 2018). Amazingly, apart from Quansah and Amoako (2018) which was conducted in Ghana, most of the studies were conducted in the western world. The study by Quansah and Amoako found a negative attitude towards test construction among Senior High Sschool (SHS) teachers in the Cape Coast Metropolis. Even though the study found a negative attitude, the study did not examine the influence of attitude towards test construction on adherence to test construction principles.

Deliberating further on the literature, several questions come to mind: Do teachers' skills in test construction predict their adherence to test construction principles? Does the level of teachers' commitment to work predict their adherence to test construction principles? Does teachers' attitude towards test construction predict their adherence to test construction principles? Does teachers' level of commitment to work explains the relationship between teachers' test construction skills and their level of adherence to test construction principles? Does teachers' attitude towards test construction explains the relationship between teachers' test construction skills and their level of adherence to test construction principles? These are mind-boggling questions

with no readily available answers. This study, therefore, found answers to the aforementioned questions.

Conceptual Framework

Based on the literature reviewed, the study proposes relationships among teachers' test construction skills, commitment to work, attitude towards test construction and adherence to test construction principles. Figure 1 presents the model depicting relationships among the variables.
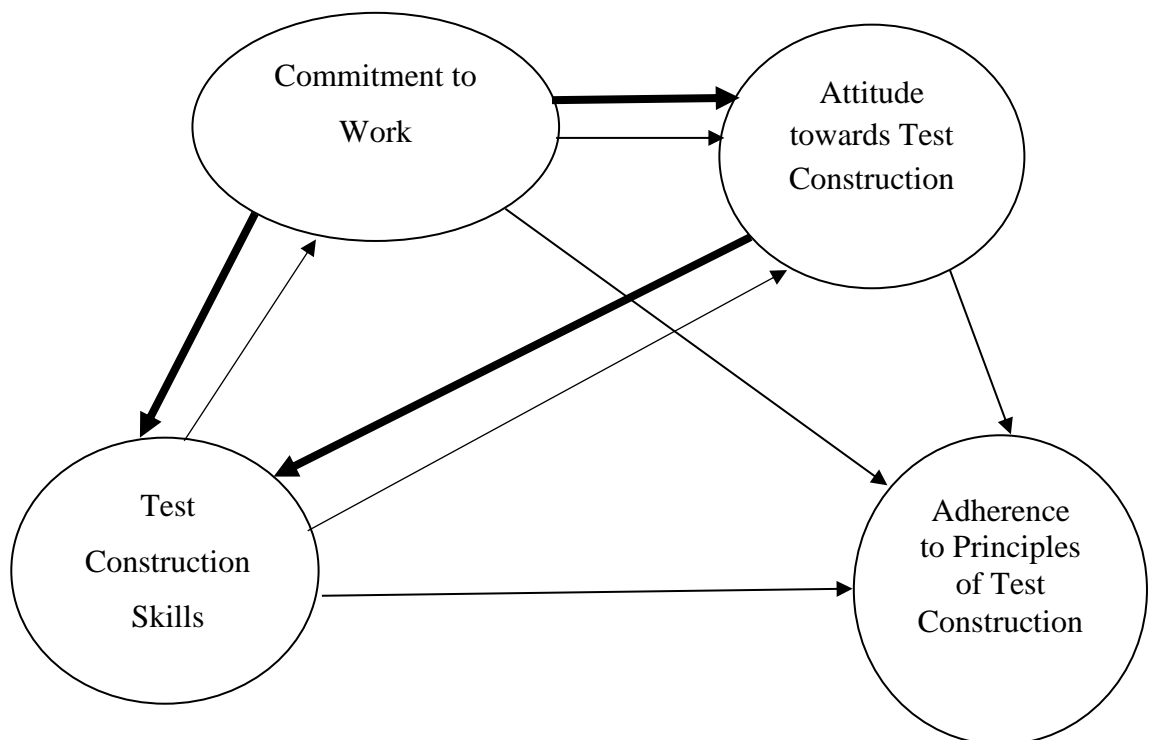
Figure 1- Relationships among test construction skills, commitment to work, attitude towards test construction, and adherence to principles of test construction
Source: Author's construct

As shown in Figure 1, the study proposes a direct relationship between teachers' test construction skills and their level of adherence to principles of test construction. That is, teachers' level of skills will predict their level of

adherence to recommended principles of test construction. Again, it is proposed that teacher commitment has a direct influence on their level of adherence to principles of test construction. Similarly, teachers' attitude towards test construction may have a direct influence on their level of adherence to principles of test construction.

Figure 1 further shows that the relationship between teachers' test construction skills and their level of adherence to principles of test construction will be mediated by both their level of commitment and attitude towards test construction. The study further proposes that commitment will mediate the relationship between test construction skills and teachers' attitude towards test construction. That is, teachers who have higher test construction skills will be committed to their work, and this will influence their attitude towards test construction.

Purpose of the Study

The study sought to examine teachers' test construction skills, commitment to work, and attitude towards test construction as predictors of adherence to test construction principles among SHS teachers in the Sekondi-Takoradi Metropolis. Specifically, I examined the influence of teachers':

1. test construction skills on their adherence to recommended principles of test construction,

2. level of commitment to work on their adherence to recommended principles of test construction,

12

3. attitude towards test construction on their adherence to recommended principles of test construction.

I also examined the role of teachers':

4. level of commitment to work and attitude towards test construction (serial) in the relationship between teachers' test construction skills and adherence to recommended principles of test construction,

5. attitude towards test construction in the relationship between teachers' commitment to work and their level of skills in test construction.

6. examine whether the data collected will fit the proposed model.

Assumptions of the Study

The following assumptions underpinned the study:

1. Teachers with higher test construction skills will follow recommended principles in constructing test items.

2. Teachers who are committed to their work will follow recommended principles in constructing test items.

3. Teachers with positive attitude towards test construction will follow recommended principles in constructing test items.

4. Teachers who are committed to their work will have a positive attitude towards the construction of test items.

5. Teachers with higher test construction skills, higher commitment, and positive attitude towards test construction will develop quality test items.

13

Research Questions/Hypotheses

Research questions

The study was guided by the following research questions;

1. What is the level of skills of teachers in test construction among SHS teachers in the Sekondi-Takoradi Metropolis?

2. What is the level of commitment to work among SHS teachers in the Sekondi-Takoradi Metropolis?

3. What is the attitude of SHS teachers in the Sekondi-Takoradi Metropolis towards test construction?

Hypotheses

The study tested the following hypotheses:

1. H_0: Teachers' level of skills in test construction will not predict their adherence to recommended principles of test construction.

   H_1: Teachers' level of skills in test construction will predict their adherence to recommended principles of test construction.

2. H_0: Teachers' level of commitment to work will not predict their adherence to recommended principles of test construction.

   H_1: Teachers' level of commitment to work will predict their adherence to recommended principles of test construction.

3. H_0: Teachers' attitude towards test construction will not predict their adherence to recommended principles of test construction.

   H_1: Teachers' attitude towards test construction will predict their adherence to recommended principles of test construction.

14

4. H_0: Teachers' level of commitment to work and attitude towards test construction will not mediate (serial) the relationship between teachers' test construction skills and adherence to recommended principles of test construction.

    H_1: Teachers' level of commitment to work and attitude towards test construction will mediate (serial) the relationship between teachers' test construction skills and adherence to recommended principles of test construction.

5. H_0: Teachers' attitude towards test construction will not mediate the relationship between teachers' commitment to work and their level of skills in test construction.

    H_1: Teachers' attitude towards test construction will mediate the relationship between teachers' commitment to work and their level of skills in test construction.

6. H_0: The data collected will not fit the proposed model.

    H_1: The data collected will fit the proposed model.

Significance of the Study

The findings of this study would be beneficial to the Ministry of Education, Ghana (MoE), Ghana Education Service (GES), and heads and teachers in the various SHSs in the Sekondi-Tarkoradi Metropolis. The findings of this study would provide evidence that would bring clarity on some factors that may influence teachers' adherence to the principles of test construction. This would help the aforementioned stakeholders of education in strengthening

15

existing policies in order to help address the issue of non-adherence to test construction principles. Knowledge on the possible causes of non-adherence to test construction principles is a step in the right direction towards providing antidote to the situation. This study has contributed to knowledge by adding to existing literature in the area of testing practices in Ghana.

Delimitation

This study focused on only public SHS teachers in Sekondi-Takoradi Metropolis. Again, the study concentrated on teachers who teach mathematics, integrated science, and English language. There might be other possible factors that may influence teachers' adherence to recommended principles of test construction, but this study, however, focused on teachers' test construction skills, commitment to work, and attitude towards test construction.

Limitations

Even though the study examined predictors of adherence to test construction principles, these predictors could not be assumed as the cause of non-adherence to test construction principles due the descriptive nature of the study. The use of self-report measures subjects the validity of the results to the accuracy of the responses provided by the respondents.

Operational Definitions

For the purpose of this study, teachers' test construction skills, commitment to work, attitude towards test construction, and adherence to test construction principles were defined as follows:

16

Teachers' test construction skills: This refers to teachers' knowledge on the principles of test construction. This was conceptualised in terms of language use, content coverage, item organisation, and test guidance.

Commitment to work: This refers to the extent to which teachers are engaged in their teaching work. Specifically, this study conceptualised commitment as teachers having stronger psychological ties to their school, students, teaching, and profession.

Attitude towards test construction: This refers to teachers' thinking or feelings about the construction of test items. This study conceptualised attitude in terms of teachers' feeling and thinking about test planning, item construction, item review, and item assembling.

Adherence to test construction principles: This refers to the principles that teachers follow when constructing test items. This was conceptualised in terms of language use, content coverage, item organisation, and test development.

Organisation of the Study

The study is organised in five chapters. Chapter One provided an introduction which comprised the background of the study, problem statement, conceptual framework, purpose of study, and significance of the study, delimitations. Chapter Two reviewed related literature on test construction skills, commitment to work, attitude towards test construction, and adherence to test construction principles. Theoretical and conceptual review were also presented. In addition, empirical findings of some studies were reviewed. Chapter Three presented the research methodologies that were used to conduct

17

the study. These comprised the research design, population, sampling procedures, data collection instruments, data collection procedures, and data processing and analysis. Chapter Four presented the results and discussion of the findings. Finally, Chapter Five provided a summary of the entire study, conclusions and recommendations.

CHAPTER TWO

LITERATURE REVIEW

Introduction

The purpose of this study was to examine teachers' test construction skills, commitment to work, and attitude towards test construction as predictors of adherence to test construction principles among SHS teachers. This chapter presents a review of related literature. The literature comprises conceptual issues, theoretical and empirical review. The chapter is organised under the following sub-headings:

1. Conceptual review

    a. The Concept of Assessment

    b. Process/Principles of Test Construction

    c. The Concept of Commitment and Teacher Commitment to Work

    d. Teacher Attitude towards Testing

2. Theoretical review

    a. Validity

    b. Classical True Score Theory – Reliability

3. Empirical review

    a. Teachers' Test Construction Skills

    b. Teachers' Attitude Towards Test Construction

19

c. Teachers' Commitment to Work

d. Test Construction Skills, Adherence to Test Construction Principles, Commitment to Work, and Teachers' Attitude towards Test Construction.

Conceptual Review

This section provides some definitional issues to some concepts or variables used in this study.

The Concept of Assessment

Assessment is defined as "the process of gathering, interpreting, recording, and using information about the pupils' response to educational tasks" (Lambert & Lines, 2000, p. 4). It is a helpful and constructive tool because the information derived from it is used to adapt and modify teaching and learning techniques (Black & William, 1998). Pelligrino, Chudowsky, and Glaser (2001) suggest that assessment has a fundamental role in providing information to help students, teachers, administrators, and policy makers arrive at decisions. Assessment has been described as an instructional tool (Bailey, 2004) and important in the teaching-learning process (Malone, 2013). The use of assessment as it relates to students, is very much emphasised because it is used in teaching and learning; some have even seen assessment and teaching and learning as two sides of the same coin. This use of assessment, however, is greatly dependent on teachers' assessment practices (Cumming, 2001).

Assessment can have a more noticeable influence on students' behaviour and performance than teaching does (Miller & Parlett, 1974).

20

Assessment is still the most influential factor in formal education and, if not approached properly, may subvert the positive aspects of both the teaching and learning process. Kellough and Kellough (1999) have characterised six purposes of assessment, to: assist student learning; identify students' strengths and weaknesses; determine and improve teaching effectiveness; ascertain the effectiveness of a particular instructional strategy; determine and improve the effectiveness of curriculum programmes; and provide information useful in decision making.

From these purposes, it is evident that the concept of assessment cannot be looked down upon as far as teaching and students' learning are concerned. Assessment practice is the totality of the steps and procedures taken by teachers throughout an assessment. It includes preparation, administering, grading, recording and reporting of the assessment information (Malone, 2013). The traditional testing procedures used in assessment include multiple choice, matching, true/false, short answer and essay questions. Alternative assessments such as observation, conferences, portfolios, peer and group assessments techniques are equally used.

Test Construction Skills/Competency

The teachers' standards set specific expectations on how teaching is supposed to be practiced in order to produce good effects especially on improving students' learning (MoE-Ghana, 2017). According to Magno (2013), teaching competencies also serve as benchmarks on how teaching should be done within a specific scale of acceptable level. The teachers' competency

21

reflects the quality and caliber of teachers of a country. Stoof, Martens, and Van Merrienboer (2000) define teaching competencies as the level of integration of knowledge, skills, and attitudes. It is interesting to note that teaching and assessment are complementary activities. That is, the process of teaching begins with assessment and also ends with assessment. Teachers, on one hand, informally assess their students at the beginning of the instructional process. This assessment helps the teacher to know the level of students, so that instruction could be structured to suit students' knowledge and ability level. On the other hand, teachers assess students at the end of an instructional period. Information on this assessment helps the teacher to ascertain the amount of students' learning after an instructional period. Such information can be used to plan future instruction and also to improve on students' strengths and weaknesses.

There is a growing concern in studying more in-depth the specific components of teaching standards such as teachers' competency to assess student learning. Teachers need to be more competent in the rigors of assessing student learning to determine how effective their instruction is in developing learning targets. The development of assessment competencies depends on the kind of experience and training that teachers receive in their preservice years (Cizek, Fitzgerald, & Rachor, 1996). Teachers' assessment and evaluation competency, most particularly, teachers' test construction competency and quality lesson preparation, are important to the overall development of the education system (Jiraro, Sujiva, & Wongwanich, 2014). These are the factors

22

that allow a sustainable education reform. Therefore, the role of teachers is highly crucial in terms of preparation of tests or assessments that match the learning objectives as well as other supportive evidence for such results.

Test construction skills include competencies needed for developing quality tests based on stipulated principles of test construction. These competencies are outlined by Agu, Onyekuba and Anyichie (2013) as: objectivity, communication, item validation skills and skills for applying appropriate strategies for ascertaining the reliability of test instruments. Skills in test construction enable teachers construct tests with precision, appropriateness of language use, objectivity and good grading scales. These skills help teachers to: (a) structure items to elicit clear and concise answers from students; (b) construct tests that will be appropriate for learners of different ages, abilities, and gender; and (c) set tests so that students finish within a given stipulated time and do not grow scared of tests (Ali, 1999).

Sanders et al. (1990) posit that the scope of a teacher's professional role and responsibilities for student assessment may be described in terms of a trinity concept. This concept includes activities occurring; (a) prior to instruction, (b) during instruction, and (c) after instruction. According to Sanders et al. (1990), prior to instruction, teachers are expected to understand the cultural backgrounds, interests, skills and abilities of students; clarify and articulate the performance outcomes expected of pupils, and plan the instruction for individuals or groups of students. During instruction, teachers are expected to monitor pupils' progress toward instructional goals; identify the gains and

difficulties pupils are experiencing; adjusting instruction; and judging the extent to which pupils have attained instructional outcomes. Finally, during the final stage, teachers are expected to describe the extent to which each pupil has attained both the short- and long-term instructional goals; providing information on strengths and weakness based assessment results to students and parents; determining effectiveness of instruction; and evaluating effectiveness of the curriculum and materials in use among others. From these activities, the role of assessment cannot be underestimated. The implication of these activities is that teachers have to possess some level of competency in order to carry out these mandates, which are primarily based on assessment.

For the purpose of this study, emphasis would be laid on activities which occur during and after instruction which are purely based on students' assessment. Teachers are expected to have adequate assessment knowledge or skill in order to perform well in their evaluation of students' achievement. According to American Federation of Teachers, National Council on Measurement in Education, and National Education Association (1990), teachers should demonstrate skills at selecting, developing, applying, using, communicating, and evaluating student assessment information and student assessment practices. In terms of developing assessment, the aforementioned bodies identified the following as skills and competencies of teachers.

1. Teachers should be skilled in choosing assessment methods appropriate for instructional decisions. According the authors, skills in choosing appropriate, useful, administratively convenient, technically adequate,

and fair assessment methods are prerequisite to good use of information to support instructional decisions. Teachers, therefore, need to be conversant with the kinds of information provided by a broad range of assessment alternatives and their strengths and weaknesses. In particular, they should be familiar with criteria for evaluating and selecting assessment methods in light of instructional plans.

2. Teachers should be skilled in developing assessment methods appropriate for instructional decisions. Teachers in the classroom are expected to develop assessment instruments that would aid in measuring the amount of students' learning regarding some instructional objectives.

Darling-Hammond (2010) refers to teacher competence when she stresses that the quality of test questions depends on the quality of the teacher. Shukla (2014), however, notes that professional competency alone does not result in effective performance. Teacher effectiveness is not automatically ensured by professional competencies and practical skills only. Well-trained and effective teachers are those who are both competent and committed to the profession. Teachers have to acquire different skills and abilities to make their teaching competent. Chau (1996) maintains that a teacher's level of competence is one of the factors that directly affect the quality of his/her test questions. This implies that the quality of assessment instrument given by a teacher determines its ability to provide information needed regarding students' performance. Well-written items allow for accurate and consistent measures of students' mastery

25

of content taught in class. Results of such tests allow teachers to measure to some degree, how effective their instruction has been. Conversely, poorly designed test items can lead to inaccurate measurements of learning and provide false information regarding student performance as well as instructional effectiveness (Education Up Close, 2005). Frey (2007) stated that any characteristics of an assessment instrument item which distracts the test-taker from the major point or focus reduces the effectiveness of that item.

Process/Principles of Test Construction

Adamu, Dawha, and Kamar (2015) assert that in order to construct good assessment instruments, teachers should be able to set out targets and write objectives, choose an assessment items and technique(es), administer assessment and analyze the data and share the results with students. Chidolue (cited in Agu et al., 2013) outlined the following competencies that classroom teachers should possess to construct good assessment instrument item:

1. determining the purpose of each assessment exercise;

2. stating specific, measurable educational objectives;

3. making good content outline;

4. preparing assessment plan which will guide item construction;

5. choosing appropriate assessment instrument item formats;

6. constructing clear, precise and unambiguous items;

7. constructing items that focus the attention of a group of students, often with widely varying background experiences, on a single idea;

8. constructing items with appropriate difficulty and discriminative indices;

9. developing marking guide suited for the assessment;

10. performing item analysis of their items;

11. developing assessment instrument that are economical in time and money;

12. giving clear directions on how the instrument should be administered and taken;

13. reviewing the instrument in order to correct any errors made during item construction.

Koksal (2004) outlined factors that are inherent in poorly designed assessment instruments which if well-handled will lead to quality classroom-based assessment instruments. These are:

1. Non specification of the target audience, what skill or area of ability the assessment instrument is intended to measure, how much time is allocated for each item, and what points the test-takers would get for each correct response.

2. Separate sections not clearly stated.

3. Assessment instrument items having more than one possible answer because they were not conceptualized.

4. Not stating time allocated for each task on the papers. Only the total time available to perform all the tasks that was given.

27

5. Non consideration of level of students in assessment instrument construction.

6. Unclear instructions.

7. Tasks students are expected to perform not being in concert with the tasks they are told to do during the classroom instructions.

8. Lack of representativeness of items concerning what the teacher intends to assess.

9. College-production of some assessment instrument items.

When teachers follow the aforementioned principles or procedures when developing test items, the quality test items are assured. Quality test would elicit information that can be relied on for useful decisions.

The Concept of Commitment and Teacher Commitment to Work

Commitment means being loyal to one's profession. Vandenberg and Scarpello (1994) define professional commitment as "a person's belief in and acceptance of the values of his or her chosen occupation or line of work, and a willingness to maintain membership in that occupation" (p. 535). Meyer, Allen, and Smith (1993) define professional or occupational commitment based on three distinct components. These components are namely: affective professional commitment (APC), continuance professional commitment (CPC), and normative professional commitment (NPC). Affective professional commitment (APC) refers to identification with, involvement in, and emotional attachment to the profession. Thus, employees with strong affective professional commitment remain members of their profession because they

28

want to do so. For example, professionals with a strong sense of affective commitment to their profession will keep up with developments in their profession, subscribe to trade journals, attend professional meetings, and participate in their professional association.

Continuance professional commitment (CPC) refers to commitment based on the employee's recognition of the costs associated with leaving their profession. Employees with strong continuance commitment remain with their profession because they realise that they have much to lose by not doing so. For example, professionals with high levels of continuance commitment might be less inclined to involve themselves in professional activities other than those required to retain membership of their profession (Meyer et al., 1993). Normative professional commitment (NPC) refers to commitment based on a sense of obligation to the profession. Employees with strong normative professional commitment remain members of their profession because they feel they ought to do so. Normative professional commitment may develop because of effective professional socialisation or the sacrifices involved in becoming a member of a particular profession (Meyer et al., 1993). All three components of professional commitment have implications for an employee's staying with (or leaving) their profession. Common to the three components is the understanding that commitment is a psychological state that (a) characterises the employee's relationship with their profession and (b) has implications for the employee's decision to continue or discontinue membership of their profession. Given these

29

differences, commitment is most meaningfully assessed using three separate measures.

It is obvious that these three types of commitment are psychologically inclined to a worker's perception and beliefs about the workplace, duties and responsibilities, missions and goals of the organisation, the organisation itself (the composition of the organisation), and the image and reputation of the organisation in the external environment as well as other competitors. Though the three-component conceptual model remains the bedrock for other related and improved models, different perspectives had also been considered to explain well why some employees may want to, need to or ought to remain with a particular employer including organisational goals, values, mission and culture (Olffen & Roe, 2008).

Olffen and Roe (2008) employed the attitude-behaviour model to employee commitment by constructing a five-component commitment model. Here, two dimensions are added, and they are habitual commitment and forced commitment. While habitual commitment stresses continual practice or doing a job that leads to commitment, forced commitment looks at the kind of commitment driven by job descriptions and role plays in the workplace. Swarnalatha (2016) posit that commitment requires acquisition of the special knowledge that characterises a particular profession and application of that knowledge to achieve certain ends. Professional commitment refers to the measure of strength of the employees' identification with the goals and values

30

of their profession, having strong faith in it and showing considerable effort to continue in the membership of the profession (Nazari & Emami, 2012).

In the context of education, and for that matter curriculum implementation, teacher commitment cannot be assumed. Without such commitment, any form of implementation will surely fail. Achieving the goals of education depends on better understanding the nature and development of a teacher's commitment. The teacher is the only one who can transform purposes into action in the classroom. It is, therefore, pertinent to be much concerned about whether the teacher is doing this effectively. Commitment is important for teachers because it reflects a personal interpretation of work experience as absorbing and meaningful. It is a significant factor in efforts to improve school outcomes, especially student academic achievement (Hussen, Tegegn, & Teshome, 2016).

Teachers' professional commitment refers to the willingness of the teachers to 'go the extra mile' to ensure that students can be successful. Professionally committed teachers assist students in their extra time, cooperate with and willing to work with parents, and utilize class time profitably (Simpson & Hood, 2000). Teachers should have commitment to the teaching profession and should justify public trust and confidence by providing quality education to all students. Teachers should be committed to students and maintain professional relations with students as well as share responsibilities and prove themselves as successful facilitators of learning for the learners placed under their guidance and supervision.

31

According to Simpson and Hood (2000), a committed teacher reflects certain behavioural characteristics. He/she shows that professional development is a top priority; reflects excitement about teaching and learning; connects with students; shows positive attitude about students; and is perceptive about student motives, strengths, need and situation. Teaching is a profession that requires personal commitment to maintain enthusiasm for being actively involved in the work (Day, 2004). Le-Compte and Dworkin (1991) view commitment as an investment of personal resources and it has long been associated with the professional characteristics of a teacher. Teacher commitment is thought to be important because it is seen as significant in achieving quality teaching, the ability of teachers to adapt to change, teacher attendance, burnout, staying in the profession, organisational health of the institutional and student attitudes and learning outcomes (Firestone, 1996).

A committed teacher is focused and more involved in her/his profession. Hussen et al. (2016) noted that even though the line of work has been chosen voluntarily by teachers, commitment tend to decrease after sometime. Lower commitment affects the effectiveness of schools and causes teachers to be less successful in their professional performance or to leave the profession in extreme cases. VanAmelsvoot, Hendriks and Scheerens (2000) disclosed that the teaching profession is demoralised and suffering from diminishing job satisfaction and decreased commitment, a situation that will affect the quality of teaching and the adequacy of teacher supply in the long term globally. This situation is, however, not different in the Ghanaian context, even though much

32

has not been documented. This study therefore attempts to assess teachers' level of commitment to the teaching profession.

Teacher commitment, for the purpose of this study can be viewed in the lenses of four dimensions, namely, commitment to: students, teaching, school, and the profession.

Commitment to students

Commitment to student learning includes teacher dedication to helping students learn regardless of their academic difficulties or social background. Conceptually, this type of commitment speaks to increasing student engagement in learning and academic achievement, particularly for students who are academically at risk (Kushman, 1992). A teacher, who is committed to students, engages students in interaction, uses active learning techniques, uses assessment to evaluate both teaching and students' learning, demonstrates flexibility and responsiveness, demonstrates knowledge of students, creates an environment of respect, and manages students' behaviour.

Commitment to teaching

Warr, Cook, and Wall (1979) viewed commitment as the extent to which a person desires to be engaged in an organisation. Lodahl and Kejner (1965) defined commitment to work as the extent to which a person identifies his or her work psychologically. Teacher commitment to teaching is conceptualized as the psychological link between teachers and their teaching. This is reflected through teachers' willingness to exert their effort in providing effective

33

teaching, to show greater enthusiasm in teaching the subject matter, and willingness to devote extra time to students (Tyree, 1996).

Commitment to school

Teacher commitment to school is conceptualized as teachers' belief and acceptance of the goals and values of the school. Teachers' efforts for actualization of these goals and values and teachers' strong desires to keep up membership in the school (Thien, Razak, & Ramayah, 2014). Teachers who are highly committed to their school are expected to engage in school activities to achieve the school goals, exert considerable effort beyond minimal expectations, and remain working within the organisation.

Commitment to the profession

Teachers are trusted professionals by the society to shoulder the responsibility of shaping the present generation for the future through the process of teaching and learning. Committed professionals are expected to adopt various innovative method of teaching taking into consideration how best to learn and bring about effective learning (Thien et al., 2014). This would be manifested when teachers themselves show commitment to the profession to improve through both guided as well as self-directed learning. Teachers who are committed to the profession hold responsible other school-related activities entrusted to their care. They maintain accurate records of students or the school, communicates with parents and families of students concerning students' learning, participate in professional bodies' activities, demonstrate

34

professionalism, demonstrate knowledge of content and pedagogy, and meaningfully manage classroom procedures.

Teacher Attitude towards Testing

Teachers' attitudes toward testing are part of their attitudes toward teaching as a whole and probably affect their behaviour in the classroom. Assessment plays a vital role by linking teaching and learning. It is the tool that allows students to provide evidence of their learning and also the avenue that gives teachers the well needed assurance that their hard work in classrooms is not in vain (Wiliam, 2014). Because of this important role, teachers must ensure that their assessment practices are aimed at providing the right kinds of evidence, and are not influenced by personal beliefs or biases. Teachers have to be willing to embrace any assessment practice that will benefit their students' learning.

Assessment tends to generate negative sentiments by teachers (Wright, 2001). The need to set aside time to construct, administer, mark students' work and provide feedback may dampen teachers' enthusiasm towards assessment and serve as a deterrent from giving too many assessment tasks. Brooks (2002) proposed a view that teachers' attitudes towards assessment are often shaped from what they have "caught" by observing their own teachers while being students. This is of great concern because it suggests that unsuccessful or unsuitable practices could easily be learnt and passed on. Biased beliefs, perspectives and practices could also be upheld by teachers who have pre-formed unhealthy assessment attitudes.

35

Teachers' attitudes toward the tests they give and toward the practice of testing can influence many facets of education: the quality of tests given, the meaning in test scores, the way in which information from tests is used, the evaluations made by students (and parents) as well as by the teacher, and the students' perceptions of themselves, the school, and the instructional process.

In summary, assessment in this context, is general is the process of gathering information which aids in determining the amount of students' learning on particular content or segment of a course. Most of these assessments are test-based. The construction of these tests require much competency and expertise on the part of the classroom teacher. Competency in this case implies familiarity with principles of test construction. In order to design quality test items which will yield reliable results, it is required that teachers duly follow the recommended principles. Adherence to the recommended principles partly depends on teachers' level of skills in test construction, commitment to work, and attitude towards test construction. The foregoing concepts provide a sound basis to conceptualise the relationship among teachers' level of skills in test construction, commitment to work, and attitude towards test construction, which is focal point of this study. The next section presents the theoretical base within which the study is grounded.

Theoretical Review

This section presents theoretical background to this study. Basically, two measurement theories are considered for the purpose of this study. The validity theory and classical true score theory – reliability. These two theories

36

form the foundation of educational measurement, and for that matter, assessment of students' learning. The following paragraphs explain these theories and how they are related to this study.

Validity

Formerly, validity was best described as when an instrument measures what it is meant to measure. However, validity theory has a far broader scope, and views on validity and the validity framework have undergone several changes during the 1900s and early 2000s. Validity was introduced in the 1920s and initially focused solely on what later became known as criterion validity. One of the earliest formal descriptions of validity was presented by Cureton (1951), who proposed that validity is linked to a particular purpose for which a test score is used, instead of being a static attribute of a test. Cureton also discussed an aspect of validity, which he called "relevance", and this concept was similar to what later became known as construct validity.

However, Cureton did not endorse relevance, but instead dismissed it due to concerns that the inherent subjective nature of the term could have a negative effect on validity, and focused instead on criterion validity. Thurstone (1955) labelled criterion validity as obsolete. He suggested that criterion validity should be replaced with internal reliability in validity studies, at least those used to assess personality tests, and thus proposed that reliability should be incorporated into validity. Thurstone's idea, however, did not receive much support, and reliability and validity thus remained separate concepts. The American Educational Research Association [AERA], American Psychological

37

Association [APA], National Council on Measurement in Education, and National Council on Measurements Used in Education (1954) further presented four types of validity – predictive, concurrent, content, and construct validity. Since that time, the concept of validity has undergone several transformation to contemporary times.

Validity refers to the evidence presented to support or refute the meaning or interpretation assigned to assessment results. All assessments require validity evidence (Messick, 1989). Validity is the sine qua non of assessment, as without evidence of validity, assessments in education have little or no intrinsic meaning. Validity is not a property of the test or assessment as such, but rather of the meaning of the test scores.

In its contemporary conceptualization, validity is a unitary concept, which looks to multiple sources of evidence (American Educational Research Association, American Psychological Association, National Council on Measurement in Education (Standards), 1999; Messick, 1989). These evidentiary sources are typically logically suggested by the desired types of interpretation or meaning associated with measures. All about validity is the construct validity in this current framework, described most eloquently by Messick (1989), and embodied in the current Standards of Educational and Psychological Measurement. The earliest unitary conceptualisation of validity as construct validity dates to 1957 in a paper by Loevinger (American Educational Research Association, American Psychological Association, National Council on Measurement in Education [NCME], 1999). In the past,

validity was defined as three separate types: content, criterion and construct, with criterion-related validity usually subdivided into concurrent and predictive depending on the timing of the collection of the criterion data.

According to the Standards (1999), "Validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests" (p. 9). The current Standards fully embrace this unitary view of validity, following closely on Messick's work, that considers all validity as construct validity, which is defined as an investigative process through which constructs are carefully defined, data and evidence are gathered and assembled to form an argument either supporting or refuting some very specific interpretation of assessment scores (Kane, 1992, 2001). Historically, the methods of validation and the types of evidence associated with construct validity have their foundations on much earlier work by Cronbach (1971, 1988), Cronbach and Meehl (1955), and Messick (1984).

The comprehensive view of validity integrates considerations of content, criteria, and consequences into a construct framework for empirically testing rational hypotheses about score meaning and utility (Messick, 1984). Fundamentally, then, score validation is empirical evaluation of the meaning and consequences of measurement. As such, validation combines scientific inquiry with rational argument to justify (or nullify) score interpretation and use.

Validity issues of score meaning, relevance, utility, and social consequences are many faceted and intertwined. They are difficult if not

39

impossible to disentangle empirically, which is why validity has come to be viewed as a unified concept (APA, 1985; Messick, 1989). For example, social consequences provide evidence contributing to score meaning, and utility is both validity evidence and a value consequence. The essence of unified validity is that the appropriateness, meaningfulness, and usefulness of score-based inferences are inseparable and that the integrative power derives from empirically grounded score interpretation.

According to Messick (1989), referring to validity as a unified concept does not imply that validity cannot be usefully differentiated conceptually into distinct aspects to underscore issues and nuances that might otherwise be downplayed or overlooked, such as the social consequences of performance assessments or the role of score meaning in applied use. The intent of these distinctions is to provide a means of addressing fictional aspects of validity that help disentangle some of the complexities inherent in appraising the appropriateness, meaningfulness, and usefulness of score inferences.

Messick (as cited in Nitko, 2001) outlines four principles of validation that help in deciding the degree to which assessment results are valid:

1. The interpretations (meanings) you give to your students' assessment results are valid only to the degree that you can point to evidence that supports their appropriateness.

2. The uses you make of your assessment results are valid only to the degree to which you can point to evidence that supports their correctness and appropriateness.

40

3. The interpretations and uses of your assessment results are valid only when the values implied by them are appropriate.

4. The interpretations and uses you make of your assessment results are valid only when the consequences of these interpretations and uses are consistent with the appropriate values.

These principles, according to Nitko (2001), must all be present before judgment on validity could be made. From these, the implication is that, judgment on validity is not based on a single evidence. The American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999), and Messick (1989) categorise validity evidences as content-related, criterion-related, and construct-related. Content-related validity evidence is related to how adequately the content of a test and the responses to the test samples the domain about which inferences are to be made. Criterion-related validity evidence is concerned with the empirical method of studying the relationship between the test scores or other measures and some independent external measures. Construct-related validity evidence is the degree to which one can infer certain constructs in a psychological theory from the test scores (Amedahe & Asamoah-Gyimah, 2015a).

Teachers, as part of their duties, construct tests, administer, score, and interpret the results. This study bothers on the test construction aspect of the teacher's duties. Generally, there are principles which teachers are to follow in constructing test items. It has empirically been proven that when teachers follow

41

these principles they develop quality test items (Agu et al., 2013). It has been established in the background and problem statement that teachers, specifically, in Ghana, do not follow recommended principles when constructing test items. The implication of this is that their test items may not be of quality and may lack validity. This study seeks to determine whether teachers' test construction skills, commitment to work, and attitude towards test construction will predict adherence to test construction principles. The argument, therefore, is that when teachers lack test construction skills, they will develop tests of low quality. In a similar way, when teachers have the skills but are not committed and do not have positive attitude towards test construction, they will as well develop tests of low quality. Tests of low quality can manifest in ways such as unclear definitions, poor sentence structure, ambiguous items, inadequate time limits, difficult test, mismatch between learning objectives and test items, a test with few items, improper arrangement of items, and identifiable patterns of answers among others (Amedahe & Asamoah-Gyimah, 2015a).

When the aforementioned factors are at play, the validity of the test results is reduced. That is to say, the interpretations and uses the test results will be put to will be flawed. As noted by Amedahe (2014), most instructionally-relevant tests are teacher-made tests since they are directly tailored with instruction. The results of these tests mirror students' abilities and instructional effectiveness. These results are largely relied on to make several educational decisions. When the assessment process is flawed, the ensuing decisions are baseless because the means through such information was gathered is

42

problematic. In all, it can be said that when teachers do not have the skills, or have the skills, but due to poor attitude and lack of commitment they do not follow recommended principles, the quality of the test they produce has serious implications for validity.

Classical True Score Theory – Reliability

History has it that Classical True Score Theory evolved from the early works of Edward Lee Thorndike in 1904 in his first textbook on test theory. Since that time, the study of test theory became part of graduate students' curriculum in psychology and education. Thorndike's works were expanded and added to the body of knowledge known as classical test theory. This theory has provided the theoretical foundation for the development of most aptitude, achievement, personality, and interest measures used in this century (Crocker & Algina, 2008). Between 1904 and 1913, Spearman argued mathematically and logically that test scores are imperfect measures of attributes of humans, and as such the correlation between the imperfect test scores and the true values is low (Spearman, 1904). Spearman (1907, 1913) repeatedly explained the terms imperfect measures and true values. These provided the basis for the classical true score model.

According to the theory, any observed test score could be envisioned as the composite of two hypothetical components: a true score and a random error. This is expressed mathematically as: $X = T + E$; where X is the observed test score, T is the true score of the individual, and E is the random error. The observed is the score that is seen on the test paper. The true score is the expected

43

value of the observed value of the observed score over repeated measures. The error score is the discrepancy between the individual's observed score and his/her true score.

Reliability is theoretically defined as the ratio of the variance of the true score to the variance of the observed score (Amedahe & Asamoah-Gyimah, 2015). This is mathematically represented as: $\rho^2{}_{xx} = \dfrac{\sigma^2 T}{\sigma^2 X}$

This implies that reliability tells the extent to which the observed score variance is due to true variance. If a test is perfectly reliable, the true score and the observed score variances are equal and the test reliability is +1. Based on this, it can be deduced that as the error score decreases, the test reliability increases. The APA, AERA, and National Council on Measurements Used in Education (1985) define reliability as "the degree to which test scores are consistent, dependable, repeatable, that is, the degree to which they are free from errors of measurement" (p. 93). From this, it is clear that reliability is inversely related with errors of measurement.

In relation to this study, it can be said that, when teachers construct tests which have the following characteristics: unclear definitions, poor sentence structure, ambiguous items, inadequate time limits, difficult test, mismatch between learning objectives and test items, test with few items, improper arrangement of items, and identifiable patterns of answers among others (Amedahe & Asamoah-Gyimah, 2015a); they introduce error scores to students' scores and thus by their actions reduce the reliability of the test scores. The presence of some of these factors, such as unclear definitions, poor sentence

structure, and ambiguous items interferes with students' ability to identify clearly what is being measured. This in turn create difficulty for the students and tend to lower their scores. When these errors are introduced to students' scores, their scores cannot be relied on for any useful decision making.

Empirical Review

Teachers' Test Construction Skills

Test construction competence and quality are essential tools required by any teacher if teaching and learning goals are to be achieved. The significance of tests in a school system is bountiful since it is the means by which any meaningful attainment of educational goals are determined (Hamafyelto et al., 2015). The studies on test construction skills have been presented based the approaches the researchers used in their investigation. Basically, three approaches were used by the researchers. One group of researchers used the documentary analysis approach, where samples of tests were evaluated. The next group of researchers solicited for factual information on test construction, but with the use of Likert-type scale. Finally, the last group of researchers also resorted to asking teachers which test construction principles they follow.

The potency of learning objectives, embedded in a school curricula remain the most cardinal sign post for educational growth, institutional excellence and individual aspirations. Teachers who are custodians of knowledge must be seen to be competent in measuring learning goals with precision and accuracy. Similarly, the tools with which these learning goals are measured must also be precise and accurate to be able to measure what the

45

teacher intends to measure and evaluate. These cannot be possible without teachers themselves being competent in the art and science of handling the tool, which are tests and examinations (D'Agostino, 2007).

Among the researchers who adopted the documentary analysis approach, Quansah, Amoako, and Ankomah (2019) explored the test construction skills of Senior High Schools (SHS) teachers in the Cape Coast Metropolis using a qualitative documentary analysis, samples of End-of-Term Examination papers between 2015 and 2018 in Integrated Science, Core Mathematics and Social Studies in three randomly (lottery method) selected SHS in the Cape Coast Metropolis. Specifically, five samples of question papers were selected based on each subject from each school. In total, 15 samples of examination papers were taken from each of the three schools. In all, 45 samples of examination papers were sampled from the three schools. The assessment tasks on the sampled instruments were critically examined by experts in the area of Educational Measurement and Evaluation. The results of the study revealed that the teachers have limited skills in the construction of end-of-term examination. This was evident as issues were found with the content representativeness and relevance of the test, low reliability, and unfairness of the assessment tasks which were evaluated.

In another study, Marmah and Impraim (2013) examined university lecturers' competency in constructing multiple choice test item questions. A sample of 63 lecturers was chosen from five departments from the College of Technology Education-Kumasi of the University of Education, Winneba,

Ghana. The main instrument used to collect data was a questionnaire. Data was also obtained through content analysis of selected end-of-semester examinations containing multiple choice items the lecturers have constructed for the end of semester examinations from 2010 to 2012. The analysis was done under (1) appropriate use of multiple choice items, (2) format of the test, (3) quality of the items, and (4) appropriateness of the options. The findings indicated that most of the time, the teachers do not follow the principles laid down by measurement experts in developing the multiple choice item. For example, the lecturers did not prepare table of test specification as part of planning the item, and they did not either underline or capitalize negative words in the items.

Koskal (2004), similarly, examined teachers' testing skills in English Language Teaching (ELT) among teachers of English in Turkey. Data was collected during the academic years 2001-2002 and 2002-2003. Fifty-six (56) different classroom test samples (progress, quiz and achievement) were collected from the state schools where teachers from pursuing ELT had teaching practice. Evaluation of the sampled tests revealed, among others, the following flaws:

1. The tests did not tell who they targeted, what skill or area of ability they intended to measure, how much time was allocated; or what points the test-takers would get for each correct response.

2. The test items had more than one possible answer because they were not contextualised. For instance, the test-takers were asked to choose between the two:

"I *decided/have decided* to move to a better job.

I *worked/ have worked* in an office in Japan."

In these two questions, both options are correct since the test items have no context.

3. The time allocated for each task was not stated on the test papers, most instructions were wrong and unclear.

4. Most items were full of mistakes in spelling and grammar. Some items were also not based on the learning objectives.

5. There were some tests claiming to test speaking skills instead of writing skills. The test takers were asked to fill in the missing parts of dialogues.

Based on the findings, the study concluded that the teachers had limited skills in test construction and therefore recommended that teachers be trained in testing and assessment.

In their study, Hamman-Tukur and Kamis (2000), examined three categories of students' examination questions in University of Maiduguri (200, 300, and 400 levels) designed for biochemistry students. The study revealed that most of the examination questions assessed simple learning outcomes of knowledge and comprehension categories of the cognitive domain at the expense of learning outcomes that call for synthesis and evaluation. Based on the finding, the study concluded that teachers have limited test skills. The study

48

recommended that teachers be sensitized on the importance of setting questions that assess these complex learning outcomes.

The findings of the previous researchers (Quansah et al., 2019; Marmah & Impraim, 2013; Koskal, 2004; Hamman-Tukur & Kamis, 2000) are unique in the sense that the various authors examined and evaluated samples of tests developed by the teachers used in their investigations. Interestingly, it emerged that the test samples were deficient in terms of content representativeness and relevance of the test, reliability, fairness of the assessment tasks, preparation of table of specification, underlining or capitalising negative words in the items, development of options for multiple-choice items, and spelling and grammar, among others. Obviously, these are clear indications of undue diligence on the part of the teachers in discharging their duties. Certainly, one is more likely to share in the conclusion of the aforementioned researchers that 'teachers have limited skills in test construction'.

In a quick rebuttal, what I find problematic was the fact that these researchers did not clearly establish that the teachers have limited skills by way of assessing their level of skills, but rather judging from samples of tests developed by these teachers. I find it difficult to agree with these results, because of the fact that the results could be explained by other factors other than limited skills in test construction. Merely evaluating a sample of a teacher's test does not give much evidence of his/her skills in test construction. This is in the sense that, a teacher may possess the skills and competency in test construction, but for reasons such as lack of commitment to the teaching work and poor

49

attitude towards test construction could explain the state of their test samples. That is to say, a teacher who is not committed to the teaching work, may not follow recommended principles in test construction. In this case, it could be possible that due to lack of commitment on the part of the teachers used in the studies, they did not see it so important to waste time to construct good tests, hence they result to using past questions or already made items in textbooks.

Again, it must be noted that test construction appears to be a quite a tedious task, if a teacher is to follow the recommended principles in doing so. Teachers who perceive test construction as burdensome will simply resort to the use of past questions or questions available from textbooks. From these arguments, it is problematic to conclude that teachers have limited test construction skills by merely evaluating samples of their tests. The current study, however, assesses teachers' test construction skills. By so doing, teachers responded to specific statements concerning test construction which are factual in nature. The assumption is that, teachers who are more knowledgeable should be in the position to indicate whether those statements are true or false, and thus, leading to accumulation of higher scores to depict possession of higher test construction skills.

Among the second group of researchers who used the Likert-type scales were Afemikhe and Imobekhai (2016) who examined whether teachers use appropriate construction procedures when developing test items. The descriptive survey design was employed to carry out the study. The population of the study was composed of teachers in primary and secondary schools in

50

Benin Metropolis in Nigeria. A sample of 500 teachers made up of 250 each from primary and secondary school levels was used for the study. A questionnaire with a Cronbach's alpha of .75 was used to assess the steps in constructing a valid achievement test. The results of the study indicated that teachers decided on the purpose of test, outlined the content that the test would cover, they specified objectives to be tested, wrote out items well ahead of time, they found out if the constructed test items were of good quality, and they edited the items for grammatical errors, among others. It was, however, found that the teachers, generally, did not: (a) generate a test specification table or test blue print, (b) use only objective test items, (c) find out how difficult test items are for the examinees, (d) seek opinion of other teachers of the subject on quality of test items, (e) use only essay questions, and (f) make sure that information in one question does not provide a clue to another question. The study concluded that even though teachers did not follow some approved steps, they utilised most of the approved procedures in constructing achievement tests, hence they have higher test construction skills.

Adamu et al. (2015) similarly assessed teachers' competencies in constructing assessment instruments. The study employed instrumentation research design. Using the census method, 96 technical teachers in six technical colleges in Gombe State, Nigeria were engaged in the study. The Scheme for Assessing Technical Teachers' Competencies for Constructing Assessment Instruments (SATTCCAI) which was made up of 26 items on a 5-point Likert-type scale with a Cronbach's alpha reliability coefficient of .82 was used to

gather information from respondents. In terms of competency in test construction, the SATTCCAI looked at language usage, content coverage, item organization, and test guidance. The study revealed that the technical teachers considered all the 26 items as being important competency statements for inclusion in the SATTCCAI. Based on this finding, the study concluded that the teachers are competent in constructing test items.

Agu et al. (2013) also conducted a study to assess teachers' competencies in constructing classroom-based tests in Nigerian secondary schools. The study developed and validated a Test Construction Skill Inventory (TCSI) for assessing test construction competencies. Using the stratified proportionate sampling method, 543 secondary school teachers in Onitsha education zone, Anambra State, Nigeria were engaged. The results of the study revealed that secondary school teachers were competent in terms of language use, content coverage, item organisation, and test guidance. The study concluded that teachers had adequate competency in test construction.

It is apparent from the foregoing studies (Adamu et al., 2015; Afemikhe & Imobekhai, 2016; Agu et al., 2013) that teachers' test construction skills was assessed. The authors asked specific questions, which were factual and directly related to test construction. However, these researchers used Likert-type scale to measure test construction skills among teachers. It must be noted that skill is an ability which depicts what a teacher can do (Miller, Linn, & Gronlund, 2009; Crocker & Algina, 2008). In the measurement of constructs of this nature, teachers are to get as high score as possible, where high scores show more of

52

what the teachers can do in terms of test construction. Likert-type scales are mostly used to measure typical performance or score, which depict what a teacher will do rather than what he/she can do. Typical Likert-type scale scores are mostly used in the measurement of constructs such as feelings, interest, attitude, perception, personality traits, and reactions to issues, among others. With these constructs, the emphasis is on obtaining representative responses rather than high scores. This current study, however, assesses test construction skills by using 18 factual items with dichotomous responses for which teachers are supposed to indicate either 'True' or 'False' to each of the items. The items will then be scored as '1' or '0', where the correct response will be awarded '1'. In all, the total possible scores will range from '0' to '18', with higher scores indicating higher skills, and vice versa. This study, however, addresses the inefficiency in the previous studies.

The third category of researchers on test construction skills included Anhwere (2009) and Oduro-Kyireh (2008). To begin with, Anhwere investigated the management of assessment practices with regard to basic principles in the construction, administering and scoring of classroom or teacher-made tests with particular reference to the scoring of essay test items in the teacher training colleges in Ghana. A descriptive sample survey was conducted in 20 selected teacher training colleges using 310 respondents comprising 230 males and 80 females. Questionnaire was used to gather data from respondents. The study showed that, teacher training college tutors did not follow the basic principles of testing in the construction of teacher-made or

53

classroom tests, and that they perceived the management of assessment practices in the colleges as extra load to their teaching activities. It was concluded from the study that, on the whole, tutors in the teacher training colleges had limited skills and competence in the knowledge of construction, administration of classroom/teacher made tests and scoring of essay-type tests.

In a similar study, Oduro-Kyireh (2008) investigated the testing practices of Senior Secondary School (SSS) teachers of English Language, Core Mathematics and Integrated Science with respect to the construction, administration and scoring of their classroom achievement tests and the interpretation of the results of these tests. The study aimed at finding out whether the procedures used by teachers in the construction, administration and scoring of classroom achievement tests and the interpretation of the results of these tests were in line with the principles and guidelines prescribed by measurement specialists. A sample of 10 districts was randomly selected from the 21 districts in the Ashanti Region. Using the cluster sampling method, 26 SSSs from a total of 56 SSSs were sampled from the 10 districts. The sample for the study comprised 265 teachers teaching the three subjects. Questionnaire was designed to gather data from respondents.

The study revealed that out of 10 principles of test construction, teachers followed seven, which include: defining the purpose of the test; relating the instructional objectives of the subject matter to the test; selecting the test format suitable for testing the stated objectives; writing the test items in advance (at least two weeks) of the test date to permit review and editing; preparing the

54

marking scheme as soon as the test items are written; reviewing the test items after they have been set aside for a few days; and writing clear and concise directions for the entire test and sections of it. The principles not used frequently by teachers included: the use of a table of test specifications to determine the items on the test; preparing more items than needed in the test; and evaluating the test as a whole on the criteria of clarity, practicality, validity, efficiency and fairness. It is clear from the findings that the teachers do not have much skills in terms of test construction.

In all, Anhwere (2009) and Oduro-Kyireh (2008) were very clear in their aims and tried to achieve that. Data were obtained by means of questionnaire. However, in the quest to measure skills in test construction, the authors asked teachers to indicate which principles they follow, and this was used to determine their level of test construction skills. The use of such an approach is quite questionable in the sense that a teacher who may be judged as having limited skills on the basis of him/her not following principles in test construction, could be as a result of other factors such as lack of commitment and poor attitude. It is possible that a teacher may have much skills in tests construction, but because he/she is not either committed to the teaching work or has a poor attitude towards test construction, he/she does not follow the recommended principles in test construction. In this case, it is problematic to substitute non-adherence to test construction principles with limited skills or knowledge in test construction. This triggers the aim of this current study to actually assess teachers' test

construction skills by asking them to respond to factual statements which bother on test construction.

It is, therefore, clear that measurement of test construction skills is problematic in the sense that approaches used by the previous researchers, in my view, led to spurious conclusions on the level of test construction skills possessed by teachers. In addition, it is worth noting that factors such as skills in test construction, commitment to work and attitude towards test construction could explain why teachers do not adhere to test construction principles. Teachers with higher test construction skills are more likely to adhere to recommended principles of test construction. Again, teachers' commitment to work and attitude towards test construction could serve as intermediary variables which could explain why teachers with higher test construction skills do not adhere to recommended principles or otherwise. These variables could serve as mediators explaining the link between test construction skills and adherence to principles of test construction.

Teachers' Attitude towards Test Construction

This section presents a review of studies on attitude of teachers towards test construction. First, Fakeye (2016) investigated teachers' and students' attitudes towards formative assessment in English Language in Ibadan Metropolis. The study adopted descriptive research design of survey type. The population of this study comprised all English language teachers and students in senior secondary schools in Ibadan North East Local Government Area of Oyo State. The participants were 420 students and six teachers. Questionnaire

56

on Teachers' Attitudes towards Formative Assessment and Corrective Feedback in English Language, with an alpha (α) = .78 was used to gather data from the teachers. The findings of the study revealed that teachers had moderately positive attitude to formative assessment and corrective feedback. Even though the study was on attitude towards formative assessment, it is related to the current study. This is in the sense that teachers construct tests to assess students formatively, hence if they have positive attitude towards formative assessment, it is possible they may equally have positive attitude towards achievement test construction. It must be noted, however, that formative assessment does not necessarily mean the use of tests. Techniques such as oral questioning, observation, and interview, among others, can be used formatively but these are not test-based, and do not require formal test construction. It is, therefore, necessary to assess teachers' attitude towards test construction. In addition, it was observed that the study used only six teachers. This sample size of six teachers is not adequate enough to yield a good result, hence the need to use a larger sample of teachers.

Quansah and Amoako (2018) assessed teachers' attitude towards test construction. The study 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 developed instrument was administered to 432 Senior High School (SHS) 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 which 32 items which were on a four point Likert scale were maintained. Results of the study indicated that SHS teachers have negative attitude towards the following: planning of classroom tests (M=2.43, SD=.74); item construction (M=1.90, SD=.89); items review (M=2.03, SD=.66); and assembling (M=2.14, SD=.63). Generally, it was concluded that teachers in SHS in the Cape Coast Metropolis have an overall negative attitude towards test construction (M=2.13, SD=.72).

The authors were clear in their investigation by indicating the various methods used in carrying out the study. The study was conducted in the Cape Coast Metropolis, and hence its findings cannot be meaningfully extended to Sekondi-Takoradi Metropolis teachers. This is because even though the teachers appear to be similar in some characteristics, workshops are organised differently based on regions and districts. So if teachers in the Sekondi-Takoradi Metropolis receive, on regular bases, in-service training and workshops on assessment, it is possible that their attitude towards test construction may vary from those in Cape Coast Metropolis. The current study, therefore, assesses teachers' attitude towards test construction in the Sekondi-Takoradi Metropolis.

In a similar study, Awofala and Babatunde (2013) examined the attitudes of 339 pre-service Science, Technology and Mathematics (STM) teachers towards continuous assessment practices in Nigeria using a descriptive survey research design in a conventional university in the South-Western part of Nigeria. Results of the study showed that a higher proportion of the pre-

58

service STM teachers seemed to display positive attitude towards most of the continuous assessment practices while their attitude towards some assessment practices tended to be either negative or neutral. The findings of this present study is more likely to vary. First, the study was conducted in Nigeria. The conditions of service in Nigeria are not the same as in Ghana. Again, pre-service teachers were used for the investigation. Pre-service teachers are not yet teachers, hence their attitude towards test construction is more likely to be positive. It is possible that the attitude of these pre-service teachers would be changed when they are fully engaged in service because of the challenges involved. In view of this, the findings of pre-service teachers cannot be extended to in-service teachers, and for that matter teachers in Ghana.

Further, Bramwell-Lalor and Rainford (2016) in their study assessed advanced level biology teachers' attitudes towards assessment and their engagement in assessment for learning in secondary schools from rural and urban areas in Jamaica using the mixed method design. Quantitative data were collected from 40 teachers using the Teachers' Attitudes towards Assessment Questionnaire, whereas qualitative data were collected through focused group discussions and interviews from six teachers who implemented assessment for learning strategies. The findings revealed that advanced level biology teachers displayed highly favourable attitude towards assessment. The majority of the teachers (97.5%) indicated that the main purpose of assessments is to monitor the teaching and learning process rather than to give grades. It was not surprising that the majority of the teachers (75%) believed that assessment

59

should be integrated into the teaching process rather than being a terminal exercise. They (67.5%) also, generally, believed that there were other ways besides classwork, homework and tests that could indicate what a student has learnt.

The findings further revealed that teachers, however, did not regularly implement the strategies in their classrooms. Positive attitude, therefore, did not necessarily lend themselves to "best" practices in assessment. Despite these favourable views of the purpose of assessment, the interview data of Bramwell-Lalor and Rainford's (2016) seemed to suggest that the teachers mainly used assessments to prepare students for external exams. A teacher, for example, stated that a regular practice of his is to "spot" (predict) questions coming on examinations and use them to prepare his students. From the findings of Bramwell-Lalor and Rainford's study, even though there were some lapses, teachers, generally, had a positive attitude towards assessment. They really acknowledged the purpose of classroom assessment of students' learning. It must be noted that the authors did not assess teachers' attitude towards the test construction aspect of assessment. Merely showing positive attitude in assessment in general does not necessarily mean that teachers will have positive attitude on all components of assessment. There is, therefore, the need to assess the attitude of teachers towards test construction. Doing this will help in determining likelihood of teachers' attitude towards test construction in predicting the extent to which they would adhere to test construction principles.

Teachers' Commitment to Work

Professional commitment, likely, has an impact on people's working behaviours such as their observable attitudes, their judgments about the work and their involvement in professional groups. Professional commitment is related to positive behaviours which are useful for organizations. People with high levels of professional commitment subsequently are less involved in activities that are harmful to the organization (Nazari & Emami, 2012).

In a study, Adu-Gyamfi, Adu-Oppong, and Boahen (2017) examined job satisfaction and commitment among the non-teaching staff of the College of Technology of the University of Education Winneba – Kumasi (COLTEK). The study adopted the quantitative approach to research. The study was descriptive and exploratory in nature. The researchers used convenience sampling technique to select 213 workers of the non-teaching staff of which 86 workers were selected from the senior staff and 127 from junior staff. A sample of 213, representing 72.9% of the population was considered representative of the population of non-teaching staff. The main instruments for collecting data were questionnaire and observation. The study revealed that respondents believed the University is a place of personal meaning for them, and also they felt a strong sense of belonging to the university. Respondents further indicated they were not leaving the organisation because of scarcity of alternatives and disrupted life. Respondents again reported that job appointments are given to employees in order to commit them to tasks and other responsibilities.

61

The findings of the researchers are, however, not surprising as their study used non-teaching university staff. Generally, in Ghana, university staff and those in the secondary schools happen to be engaged by different bodies. University staff are engaged by the universities while teachers in secondary school are engaged by the Ghana Education Service; and for that matter the conditions of service for both categories of employees are different. It was observed that staff of the university were committed to their work. This cannot be extended to those in the secondary school. Perhaps, they may have a different feeling. The current study assesses secondary school teachers' level of commitment to work.

In a similar study, Butucha (2013) investigated teachers' perceptions of professional commitment in Ethiopia. The descriptive-comparative research design was employed to conduct the study. Respondents were 381 secondary school beginning teachers in East Shoa and West Arsi Zones of Oromiya regional state, Ethiopia. The perception was assessed using the 18-item three-factor Professional Commitment Scale (PCS) developed by Allen and Meyer (1993). The study revealed that teachers were emotionally connected to the teaching profession (affective commitment). The teachers, for example, agreed to statements such as "My teaching profession is important to my self-image," and "I am proud to be in my teaching profession." Respondents further indicated that changing teaching profession now would be difficult for them to do, and too much of their life would be disrupted if they were to change their teaching profession. The respondents also reported they felt a sense of moral

62

responsibility to stay in the teaching profession and would feel guilty if they left their teaching profession because of their sense of loyalty to it.

In another study, Shoaib and Khalid (2017) conducted a study on commitment among teacher educators of Government Colleges for Elementary Teacher (GCETs) in Punjab. The study was descriptive in nature and a survey type research. Three hundred and twenty (120 female and 200 male) teacher educators from GCETS were selected as sample from population of 540 teacher educators by using proportionate random sampling technique. Professional Commitment Questionnaire (PCQ) – a five point rating scale, was used for data collection. The findings of the study indicated that teachers were committed to their profession, with a mean of ($M = 4.11$, $SD = .35$) out of a total mean of 5.0. It was further revealed that the teachers were a little high on commitment to teaching profession ($M = 4.35$, $SD = .53$) than commitment to work ($M = 4.30$, $SD = .43$), which implies that the teachers were more committed to the actual teaching work than other activities that are associated with the teaching profession. The findings of this study are similar to that of Swarnalatha (2016), however, Swarnalatha's population was made up teachers in a private school. Also, Swarnalatha used work commitment scale, a scale which is general for all organisations, but not necessarily for teachers.

Swarnalatha (2016) examined work commitment among private high school teachers of Secunderabad and Hyderabad in India. A sample of 150 teachers were selected randomly. To measure level of teachers' work commitment, a Commitment Scale (WCS) was used. The WCS scale consisted

of 15 items on the pattern of a Likert type scale having 7-point response category ranging from '1' strongly disagree to '7' strongly agree, and a split-half reliability coefficient of .80. The study revealed that 21.3% of the respondents had high level of commitment to work, 42% had average levels of work commitment, while 36.7% had low level of work commitment. In all, the study concluded that teachers were committed to their teaching work.

Hussen, Tegegn, and Teshome (2016) assessed the secondary school teaching graduates commitment towards students learning, the community and their profession in Eastern Ethiopia. The study employed mixed research design particularly concurrent embedded qualitative and quantitative methods. The population of the study included teachers of three groups, mentors, principals, and supervisors. In all, 170 respondents were involved in filling out the questionnaires. Out of these, 90 respondents were secondary school teachers (sample PGDT graduates of three batches) and the remaining 80 respondents were secondary school leaders (mentors, principals, and supervisors). The study was carried out on 14 secondary schools. Four data collection instruments namely; questionnaire, interview form, documentary analysis and focus group discussion (FGD) were used to collect the necessary data.

The study revealed that school leaders and teachers' commitment to their students learning were 2.16 and 1.97, respectively out of 5.0. This indicates that, both groups of respondents were not committed. This was confirmed by interview as one participant said, "due to low motivation in their profession as a result of low salary and less incentives, teachers do not cater to encourage

64

students to work hard in order to achieve their individual goals in the development of their physical and intellectual aspects". The study concluded that, teachers' commitment to learning, the community and their profession were low due to low salary, low respect, low attitude towards the teaching profession, less motivation and less incentives.

The studies by Butucha (2013), Shoaib and Khalid (2017), and Swarnalatha (2016) found that teachers felt a sense of moral responsibility to stay in the teaching profession and would feel guilty if they left their teaching profession because of their sense of loyalty to it. In the case of Butucha it was realised teachers were committed because the disruption it would cause when they change their teaching profession. On the contrary, Hussen et al. (2016) found that teachers' professional commitment were low due to low salary, low respect, negative attitude towards teaching profession, less motivation and less incentives. The findings of these researchers cannot be held-on to as same in the case of Ghana. Policy variations and differences in conditions of service may bring disparity in the results of the studies conducted in Ethiopia, Punjab, and India. The current study, therefore, examines teachers' level of commitment in Ghana, specifically, in the Sekondi-Takoradi Metropolis.

Test Construction Skills, Adherence to Test Construction Principles, Commitment to Work, and Teachers' Attitude towards Test Construction

This section presents a review of studies on the relationship among test construction skills, commitment to work, attitude towards test constructions, and adherence to test construction principles. The following paragraphs present details of the studies reviewed.

David (2010) examined the relationship between chemistry teachers' knowledge of test construction and quality of their tests in private secondary schools in Maiduguri Metropolis, Borno State. The study used all chemistry teachers in 10 selected private secondary schools. Teacher made tests (TMT) chemistry results for five years each were used for the study. Pearson product-moment correlation coefficient was used to establish relationship between teachers' knowledge of test construction and quality of their tests, TMT. The results revealed high content validity, high level of conformity to guidelines, but found no significant relationship between teachers' knowledge and quality of their tests TMT.

Zhang and Burry-Stock (2003) examined teachers' self-perceived assessment skills and classroom assessment practices. A sample of 297 teachers in two school districts in a south-eastern state in the US was engaged in the study. Assessment Practices Inventory (API) was used as the instrument for data collection. The findings of the study revealed that self-perceived assessment skills and assessment practices overlap. The overlap between self-perceived assessment skills and assessment practices was also reflected in a Pearson product–moment correlation coefficient of .71 that explained 50% of the shared variance between self-perceived assessment skills and assessment practices. This implies as scores on self-perceived assessment skills increase, that of assessment practices also increase. Thus, teachers who possess higher assessment skills are more likely to engage in better assessment practices. According the Zhang and Burry-Stock, the difference between assessment

practices and self-perceived assessment skills mainly stems from the fact that the former is largely behaviour-based and thus internally coherent, whereas the latter reflects teachers' perception of their ability to perform classroom assessment. The findings of this study contradicts that of the foregoing study. This contradiction could probably emanate from the fact that foregoing study examined already made tests. There is a possibility that the samples of tests used were copied from already made textbooks or past questions, hence high content validity, and high level of conformity to guidelines.

In a study on teachers' attitude and adherence, Monsaas and Engelhard (1994) investigated how teachers' attitude toward testing practices affect the way teachers prepare and administer standardized tests in Georgia. One hundred and eighty-six (186) classroom teachers from Georgia took the testing practices instrument developed to measure the three variables: behaviour, attitude, and pressure. Teachers' attitude negatively correlated with behaviour; teachers who felt that the testing practices were dishonest were less likely to engage in them. Pressure to increase standardized test scores positively correlated with behaviour; the greater the perceived pressure to increase test scores, the greater the likelihood that teachers had engaged in more test preparation activities. These findings clearly show that attitude is associated with adherences to testing principles among teachers, and thus positive attitude leads to high level of adherence. This finding contradicts that of Kitiashvili (2014).

Kitiashvili (2014) conducted a study on teachers' attitudes toward assessment of students' learning and their assessment practices in Georgia's

general educational institutions. Georgia has 2300 general educational institutions and about 559,400 students. The study included 106 schools, 928 surveyed teachers, and 292 observed classes. The results showed that, overall, teachers have positive attitude toward using assessment methods that require cognitive complexity from students as well as asking questions that require higher level reasoning to answer. Although, teachers' attitudes are related to their willingness toward using such assessment approaches in their classes to some extent, this study reveals the lack of correspondence between teachers' attitudes and their assessment practices.

On the relationship between competency and commitment, Shukla (2014) examined teaching competency, professional commitment and job satisfaction among primary school teachers in Lucknow city, India. The sample for the study consisted of 100 primary school teachers, from 10 schools (5 private and 5 government schools) of Lucknow, using the convenience sampling method. The questionnaire was made up of teacher's teaching competency scale, teacher's professional commitment scale, and Teacher's Job Satisfaction Scale (TJSS). The results of the study indicated that professional commitment and teaching competency have very low positive correlation, $r =$ .10, $p > .05$. Based on the results, it was concluded that teachers who are committed need not be equally competent in the field as it depends on their training, among others. Again, teachers who are competent need not be committed as it depends on their work environment and facilities which assure them to have commitment towards the profession, but it is not essential that

professional commitment and teaching competency have any relationship with one another or they are dependent on each other.

In sum, the findings of the studies reviewed showed assessment skills overlap with assessment practices. This overlap was positive as assessment skills and assessment practices share 50% variances. This suggests that teachers with higher skills in assessment are more likely to engage in better assessment practices, which include construction of quality test items. Could it be that due to the higher skills of the teachers, they follow recommended principles in assessment? David (2010), however, found no significant relationship between teachers' knowledge and quality of their tests. Could this also be that the quality of teachers' tests does not depend on their level of skills? It appears there is no clarity on this. Based on this, I hypothesised that teachers' level of skills will predict their level of adherence to recommended principles of test construction.

It was also found that teachers' attitude negatively correlated with their behaviour, that is, teachers who felt that the testing practices were dishonest were less likely to engage in them (Monsaas & Engelhard, 1994). Teachers in the US resorted mostly to the use of standardised testing. Could this be due to their negative attitude? Hence they want an already-made test with its manual on how to score and interpret, since this will save them the cost of constructing tests by themselves. It was also found that teachers' attitude (positive) is related to their willingness toward using such assessment approaches that require cognitive complexity from students as well as asking questions that require higher level reasoning to answer (Kitashvili, 2014). Obviously, developing such

69

test requires some kind of competency and commitment. I, therefore, hypothesised that teachers' attitude towards test construction has a direct influence on their level of adherence to principles of test construction.

It was further revealed that professional commitment and teaching competency have positive correlation (Shukla, 2014). This implies that teachers who are committed are more likely to be competent in executing their duties including test construction. This is because such teachers may strive to enrich themselves with information that would enable them execute their duties diligently. By so doing, their level of skill is enhanced and such teachers are more likely to follow recommended principles in assessment. In this case, if teachers have the skills but do not adhere to recommended principles, then probably their attitude or commitment may be responsible for that. This would imply that teachers' level of commitment to work and attitude explain the relationship between teachers' test construction skills and their level of adherence to test construction principles. I hold the view that teachers who have higher test construction skills will be committed to their work, and this will influence their attitude towards test construction, hence their compliance to recommended principles of test construction. Based on this assertion, I hypothesized that teachers' level of commitment to work and attitude towards test construction will mediate serially the relationship between teachers' test construction skills and adherence to recommended principles of test construction. This means that teachers' test construction skills will have an indirect effect on adherence to recommended principles of test construction,

70

through commitment to work, which then influences attitude towards test construction.

Chapter Summary

The chapter reviewed literature on test construction skills, commitment to work, attitude towards test construction, and adherence to principles of test construction. It emerged from the literature that teachers do not have much skills in test construction. The review shows that, generally, three approaches were used by the previous authors in determining teachers' test construction skills. The first approach took samples of tests developed by teachers, evaluated these tests and subsequently made their conclusions in terms of skills based on the quality of tests examined. The second approach resorted to asking respondents factual questions about test construction, but with the use of Likert-type scale. The last approach resorted to merely asking teachers what principles they follow in test construction, and subsequently making inference from that as their level of skills force.

Majority of the studies reviewed shows that teachers were highly committed to their work as teachers while few others showed low level of commitment. Similar findings were revealed in terms of attitude of teachers. The review shows a split in terms of the relationship between knowledge in test construction and adherence to assessment principles. Commitment to work and attitude are seen to have a significant role as far as the work of teachers are concerned. Validity and the classical test theories were reviewed to underpin the study.

71

The present study examines the predictors of adherence to test construction principles. Specifically, the study assesses the role of teachers' commitment to work and attitude towards test construction in the relationship between teachers' test construction skills and adherence to principles of test construction. In view of that, study examined the mediation effect of teachers' level of commitment to work and attitude towards test construction. Basically, the study examined whether teachers with higher skills in test construction would have to be committed to their work, which will lead to development of positive attitude towards test construction, which will then lead to higher level of adherence to test construction principles. The next chapter presents the research methods employed in carrying out the study.

CHAPTER THREE

RESEARCH METHODS

The study sought to examine teachers' test construction skills, commitment to work, and attitude towards test construction as predictors of adherence to test construction principles among SHS teachers in the Sekondi-Takoradi Metropolis. This chapter presents the research methods that were employed to carry out the study. Basically, the chapter encompasses the research design, study area, population, sampling procedure, data collection instruments, data collection procedures, and data processing and analysis.

Research Design

Research design is a researcher's overall plan for obtaining answers to the research questions or for testing the research hypotheses (Amedahe & Asamoah-Gyimah, 2015). The research design provides a framework on which the entire research activity is executed. The research design spells out the basic strategies that the researcher adopts to develop information that is accurate and interpretable. There are different types of research design, however, the choice of a particular design mainly depends on the nature of the research problem to be addressed, the research questions and/or hypotheses involved, and the study group among, others. In line with this study's aim of examining the extent to which teachers' test construction skills, commitment to work, and attitude

towards test construction predict adherence to test construction principles among SHS teachers, the descriptive research design was deemed as appropriate to carry out this investigation.

The study, therefore, adopted the descriptive survey design, specifically, the cross-sectional design. This is a type of design where researchers are often interested in describing the attitudes and behaviours of a large group of people about a particular topic or issue at a particular point in time (Fraenkel, Wallen, & Hyun, 2012). In survey designs, investigators administer a survey instrument to a sample or to the entire population of people to measure the attitudes, behaviours, or characteristics of the population. In this procedure, survey researchers collect either quantitative data using questionnaires and analyse the data to describe trends about responses to questions and to test research questions or hypotheses. Amedahe (2002) maintains that, in descriptive research, the condition or relationships to be described should exist, and that accurate description of activities, objects, processes and persons is the objective. Falling on Amedahe's assertion, this study examined relationships among teachers' test construction skills, commitment to work, attitude towards test construction, and adherence to test construction principles. These phenomena or variables exist. That is to say, SHS teachers are engaged in the construction of tests. Obviously, teachers exhibit some traits in terms of test construction skills, commitment to work, attitude towards test construction, and adherence to test construction principles. Exhibition of these traits clearly shows the existence of these phenomena, hence an endorsement of the descriptive design.

74

This descriptive design was appropriate for this study since this study sought to gather information from SHS teachers and describe them as it were, without any form of manipulation. Descriptive design allows participants to be observed in a natural and unchanged environment. Descriptive studies may be pre-cursors to future research because they can be helpful in identifying variables that can be tested. The finding of a descriptive study may point the researcher to specific variables that may warrant further study. Again, descriptive studies result in rich data, that is, collected in large amounts from large populations. They provide a meaningful picture of events and seek to explain the relationships among teachers' test construction skills, commitment to work, attitude towards test construction, and adherence to test construction principles at a point in time.

In spite of the strengths of descriptive design, I really acknowledge the challenges associated with its use. They are susceptible to distortions through the introduction of biases in the measuring instruments (Amedahe & Asamoah-Gyimah, 2015). For example, errors due to the use of questionnaires might distort a research finding. Descriptive studies do not determine cause and effect relationships. Despite descriptive survey's wide scope of coverage, the data that are produced are likely to be deficient with respect to the detail or depth on the problem investigated. Moreover, the prominence on wide coverage limits the degree to which the researcher can check on accuracy and honesty of responses. Counterbalancing the pros and cons associated with the use of descriptive

75

survey design, I realised that the pros outweighed the cons. I, therefore, used the descriptive survey design to conduct the study.

Study Area

The study area for this research was Sekondi-Takoradi Metropolis. Sekondi-Takoradi Metropolis is located at the south-eastern part of the Western Region. The Metropolis is bordered to the west by Ahanta West District and to the east by Shama District. At the south of the Metropolis is the Atlantic Ocean and at the northern part is Wassa East District. The Metropolis covers a land size of 191.7 km^2 and Sekondi-Takoradi is the regional administrative capital. Though the smallest district in terms of land size in the region, the Sekondi-Takoradi Metropolis is the most urbanised among the 22 districts in the region. The population of Sekondi-Takoradi Metropolis, according to the 2010 Population and Housing Census, was 559,548 representing 23.5% of the region's total population. Males constitute 48.9% and females represent 51.1%.

Population

The population for this study comprised all teachers in public SHSs in Sekondi-Takoradi Metropolis. Records from the Metropolitan Directorate of Education (2018) indicates 11 public SHSs with a total of 998 teachers in the Metropolis. Out of the total population, 592 are males while 406 are females. The distribution of the teachers are presented in Table 1.

Table 1- *Distribution of Public SHS Teachers in STMA*

School	Teaching Staff		
	Male	Female	Total
St. John's School	65	25	90
Ghana Secondary Technical School	60	13	73
Ahantaman Girls Senior High School	50	38	88
Archbishop Porter's Girls Secondary School	53	40	93
Sekondi College	68	39	107
Fijai Senior High School	72	42	116
Takoradi Senior High School	55	36	91
Diabene Senior High School	33	42	75
Bompeh Secondary Technical School	23	49	72
Adiembrah Senior High school	50	40	90
Methodist Senior High School	63	42	105
Total	592	406	998

Source: GES, Sekondi-Takoradi Metropolitan Directorate (2018)

The study targeted teachers who teach mathematics, integrated science, and English language. Records from the Metropolitan Directorate indicate a total of 346 teachers of the aforementioned three subjects. Teachers from these subject areas were considered because in Ghana, these subject areas where students mostly fail. The use of teachers in these areas, would, in a way, help to address the lapses in assessment.

77

Sampling Procedures

Sampling refers to the process of selecting a portion of the population to represent the entire population (Amedahe & Asamoah-Gyimah, 2015). It is the sample from which information is collected and generalised to the population. Sarantakos (as cited in Amedahe & Asamoah-Gyimah, 2015) identified some instances under which conduct of a sample survey is necessary. Among these reasons are: (1) difficulty in having a complete coverage of the population; (2) when complete coverage may not offer substantial advantage over a sample survey (sampling provides a better option since it addresses the survey population in a short period of time and produces comparable and equally valid results); and (3) sample survey is more economical, since it contains fewer people and requires less printed material, fewer general costs (travelling, and accommodation) and fewer experts.

Based on these guidance, and considering the size of the target population of 346 teachers, a complete coverage of the population was possible. This was because it would offer substantial advantage over a sample survey. The census method, therefore, was used to engage all teachers who teach mathematics, integrated science, and English language. This method was appropriate in the sense that there was no sampling error since every member of the target population was involved in the study. Table 2 shows distribution of teachers with respect to the three subjects.

From Table 2, 128 of the teachers teach mathematics, 110 teach English and 108 teach integrated science. Out of this number, 56 of them were not trained teachers, while 290 were trained teachers.

Table 2- *Distribution of Teachers in STMA by Subjects*

School	Teaching Staff			
	Mathematics	English	Science	Total
Single-sex (boys)				
St. John's School	12	8	10	30
Ghana Secondary Technical School	15	8	9	32
Single-sex (girls)				
Ahantaman Girls Senior High School	13	12	12	37
Archbishop Porter's Girls Secondary School	8	8	9	25
Mixed				
Sekondi College	13	13	6	32
Fijai Senior High School	17	19	19	55
Takoradi Senior High School	12	10	9	31
Diabene Senior High School	9	9	7	25
Bompeh Secondary Technical School	10	8	12	30
Adiembrah Senior High school	10	10	9	29
Methodist Senior High School	9	5	6	20
Total	128	110	108	346

Source: GES, Sekondi-Takoradi Metropolitan Directorate (2018)

Data Collection Instruments

Questionnaire was used to gather information for this study. The questionnaire was adapted from various scales developed by some researchers

(Agu et al. 2013; Quansah & Amoako, 2018; Thien et al., 2014). The questionnaire was made up of five sections labelled from 'A' to 'E'. Section 'A' was made up of items soliciting information on demographic characteristics of respondents. The demographic variables include gender, age, educational qualification, and years of teaching experience. Section 'B' was made up of 18 items soliciting information on teachers' test construction skills. These items were adapted from Agu et al. Section 'C' was made up of 12 items adapted from Thien et al., and this solicited information on teacher commitment to work. Section 'D' was made up of 29 items adapted from Quansah and Amoako, and this measured attitude towards test construction. Section 'E' was made up of 24 items adapted from Agu et al., which solicited information on adherence to test construction principles.

Test Construction Skills

Test Construction Skills Inventory (TCSI) was developed by Agu et al. (2013) in Nigeria. The study was an instrumentation research geared purely towards the development and validation of an instrument for assessing test construction skills of secondary school teachers in Nigeria. 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 made 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).

The items of the Test Construction Skill Inventory were developed from review of literature and survey instruments. The initial instrument was a 30-item inventory structured on a 4-point Likert-type scale with the following options: strongly agree (SA), agree (A), disagree (D), and strongly disagree (SD). After the preliminary validation of the instrument, it was administered on the research sample. The data collected for the study was subjected to factor analysis. The results revealed that out of the five factors extracted during the analysis, only four factors (language use, content coverage, item organisation, and test guidance) have items sufficiently loaded on them, that is, up to four items. Only one item loaded on Factor 3 so it was not considered as the number of items for accepting a factor is four and the minimum loading of any item should be 0.35 (Baker, 2003). Four of the items were not loaded on any factor and another item was loaded on two factors, and these were dropped. In all, four factors with 25 items associated were considered to be factorially valid. Factor 1 (language use) was made up of 5 items; factor 2 (content coverage) was made up of 10 items; factor 4 (item organisation) had 6 items; and factor 5 (test guidance) had 4 items. All the items had factor loadings and communality above

0.35. The Cronbach's coefficient alpha of the test construction skill inventory is 0.73. A mean score of 2.5 was estimated by summing responses to the five responses and dividing by the number of items. Mean score above 2.5 shows the presence of the skill, while a mean score below 2.5 shows absence of the skill.

Teacher Commitment to Work

The Teacher Commitment Scale (TCS) was developed by Lei Mee Thien, Nordin Abd Razak, and T. Ramayah in 2014 in Malaysia. The target population was Malaysian primary school teachers who are working in three types of primary school: National Primary School (NPS), National Type Chinese Primary School (NTCPS), and National Type Tamil Primary School (NTTPS) in the state of Penang. The teacher sample was selected randomly using multistage cluster stratified sampling technique to ensure data representativeness and generalizability. The multistage cluster stratified sampling technique involved two levels of sample selection. At the school level, a sample was selected from NPS (25), NTCPS (10), and NTTPS (5) using quota sampling procedure. At the teacher level, a consensus number of 15 teachers were selected from each school of NPS, NTCPS, and NTTPS due to the limited number of teachers in NTTPS. In all, a total of 600 teacher sample were selected with the first half of 300 for exploratory factor analysis (EFA) and the second half of 300 for confirmatory factor analysis (CFA).

The questionnaire was developed to measure the dimensions of teacher commitment. A total of 17 items were adapted from various sources based on

82

literature. The questionnaires were trial tested to assess the instrument's validity and reliability. Five raters were engaged to review the suitability of the items that captured the underlying dimension of teacher commitment. The raters comprised three lecturers from different disciplines in local higher education institutions and two Malaysian primary school language teachers. Several inappropriate items were found to suffer from problems related to item structure, grammar, and ability to generalise to the broader context based on the comments from raters. For instance, 'I always try to help my students' was improved to yield 'I try to do my best to help unsuccessful students'. The items deemed problematic based on the raters' comments were refined, and no single item was excluded. Finally, the initial 17 items were retained, and the teacher commitment scale was formed. Prior to designing the questionnaire, a preliminarily test was conducted on four primary school teachers to ensure the appropriateness of the question content, wording, sequence, format, layout and instruction. This sample was excluded from participation in the study. Following this, a set of questionnaires in the Malaysian national language was administered in the pilot study, with responses based on a six-point Likert-type scale: 1 (very strongly disagree), 2 (strongly disagree), 3 (disagree), 4 (agree), 5 (strongly agree) and 6 (very strongly agree). The teachers' mean total scores were estimated by summing responses to the five items and dividing by the number of items. Ethical issues such as privacy, anonymity, and confidentiality were adhered to.

Data collected for the final study were analysed in two stages: (1) EFA and (2) CFA. Principal components extraction method and varimax orthogonal rotation were used in the first stage to extract the factors and their factor loadings. Four factors were extracted from EFA and this represented four dimensions of the Teacher Commitment: Commitment to School, Commitment to Students, Commitment to Teaching, and Commitment to Profession. The loadings of the items ranged from .61 to .77, above the cut-off value of .50 as recommended by Hair, Black, Babin, and Anderson (2010). Four items out of the 17 had loadings below .50 so these items were discarded. Finally, 13 items were maintained. These comprises 3 items each on Commitment to School, Cronbach's alpha = .89; Commitment to Students, Cronbach's alpha = .82; Commitment to Teaching, Cronbach's alpha = .73; and 4 items on Commitment to Profession, Cronbach's alpha = .71. The average variance extracted (AVE) estimates for Commitment to School, Commitment to Students, Commitment to Teaching, and Commitment to Profession are above the cut-off value of .50. Overall, the results presented good reliability of the scale measured and signified the convergent validity of Teacher Commitment.

Attitude towards Test Construction

Attitude towards Test Construction scale (ATC) was developed by Frank Quansah and Isaac Amoako in 2018 in Cape Coast Metropolis, Ghana. The instrument was developed based on literature as well as personal experiences of the researchers. The instrument was developed based on the behaviours exhibited by teachers in various schools. These behaviours were

84

observed by the authors of the instrument. Literature was, further, reviewed to obtain information on the test construction behaviours of teachers. Items were then crafted based on literature and observations made by the researchers. Initially, 41-items were crafted but only 32-items remained after the instrument had gone through several review and factor analysis. The items were on a 4-point Likert-type scale of agreement (SD- strongly disagree, D-disagree, A-agree, SA- strongly agree). After the items were crafted and reviewed, a pilot testing was conducted among 100 teachers from some selected SHS in the Sekondi-Takoradi Metropolis. Some items were modified after the pilot testing of the instrument. Items like "Learners decide item format to be used" was changed to "I prefer the item format of a classroom test to be decided by the learners". In all, 4 items were reworded after the pilot testing.

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 item assembling. A confirmatory factor analysis using the Maximum Likelihood Method was then conducted to examine the factor loadings of the items. After critical evaluation, the items on the instrument one was rejected while remained 32 which were on a four point Likert scale. The Cronbach's alpha reliability estimates of the subscales are as follows: planning (13 items) = .81, item construction (10 items) = .79, item review (5 items) = .70 and item assembling (4 items) = .71. The overall reliability estimate of the instrument was .85.

The ATC scale has both positive and negative questions of which responses are measured on 4-point scale. In scoring the items on the instrument, negative items are scored on point score from 1-4. That is, strongly agree is valued 1-point, agree for 2-points, disagree is 3-points, and strongly disagree for 4-points. For positive items, strongly agree is 4-point, agree 3-point, disagree 2-point and strongly disagree 1-points. Apart from items 14, 16, 23, and 32, the rest of the items are negative questions. For the overall attitude, the responses from all the items are added and divided by the number of questions. The same computational method is applicable to the sub-scales (i.e., calculating the composite score for the responses for a particular respondent or group of respondents). In calculating for the attitude of respondents the mean of their responses is computed for and interpreted. In interpreting the attitude of respondents of a particular item (e.g., item 5), the mean score of the responses is compared with 2.5 ([1+2+3+4]/4=2.5). Mean scores less than 2.5 shows a negative attitude whereas mean scores above 2.5 shows a positive attitude to that particular item. For the interpretation of individual scores, the mean of the obtained scores is also compared with 2.5.

Adherence to Test Construction Principles

Adherence to Test Construction Principles was measured using an adapted version of Test Construction Skills Inventory (TCSI) developed by Agu et al. (2013). The response type for the TCSI is 4-point Likert type agreement scale: strongly agree (SA), agree (A), disagree (D), and strongly disagree (SD). The response type was changed to a 5-point Likert scale from Never to Always.

86

Thus, Never = 1, Not often = 2, Often = 3, Very often = 4, and Always = 5. This helped to determine the extent to which teachers adhere to test construction principles. The number of items remained the same as in the original Test Construction Skills Inventory. However, the preamble for the items "A teacher should take the following steps in constructing tests for his/her class" was changed to "I take the following steps in constructing tests for my class".

Pilot testing

Prior to the pilot test, the questionnaire was vetted by my supervisors, who are expects in the area of Measurement and Evaluation. The questionnaire was pilot-tested using a sample of 111 SHS teachers in the Cape Coast Metropolis. The pilot testing was done to determine whether the items on the questionnaire portray the intended meanings by the original researchers in Ghanaian context. Results and feedback on the pilot test were used to refine the instrument before the final data collection. The instrument was finally validated. The validation was done using the Item Response Theory (IRT), and the covariance-based SEM, which was introduced by Karl Joreskog in 1973 (Hair, Hult, Ringle, & Sarstedt, 2014). This type of SEM is very efficacious in performing a confirmatory factor analysis. In the confirmatory factor analysis, the theoretically predetermined factor structure is confirmed by the current data. In other words, in the confirmatory factor analysis, which factors will be loaded on an observed variable is predetermined. Thus, for the purpose of this study, previously discovered scales were re-confirmed with the collected data by means of confirmatory factor analysis and item response theory.

In terms of the factor analysis, convergent and divergent validity were determined. Items, which had low factor loadings (thus, below .30) were discarded before the final data collection (Pallant, 2010). In addition, convergent validity was assessed using AVE of .50 or more (Fornell & Larcker, 1981). The discriminant validity was assessed using Fornell and Larcker's (1981) criterion, which states that the square roots of AVEs should be greater than the correlation among the dimensions.

Results on pilot testing

This section comprises the results of the pilot test. A total of 111 sample was used for the pilot testing. The various sub-sections of the questionnaire (Test Construction Skills, Commitment to Work, Attitude towards Test Construction, and Adherence to Test Construction Principles) were validated before the final data collection. For Teachers' Test Construction Skills, Item Response Theory (IRT) using the 3-parameter logistic model, which comprises: item difficulty, discrimination, and probability of guessing as parameters was used to validate the items. This analysis was performed using IRTPRO software version 4.2. Morizot, Ainsworth, and Reise (2007) indicate that in order to get a reasonably unbiased estimates of parameters in IRT, a sample of 100 is appropriate for dichotomously scored items. In this case, the sample of 111 used is appropriate. In terms of Commitment to Work, Attitude towards Test Construction, and Adherence to Test Construction Principles, they were validated using Confirmatory Factor Analysis (CFA) through the use of AMOS

software using 1000 bootstrap samples. Figures 2 to 4 show the measurement model for the various scales.

Validation of Test Construction Skills Scale

The Test Construction Skills section of the questionnaire was validated using the IRT. IRT is a test theory, which evolved as a result of the inefficiencies in the Class Test Theory (CTT). CTT is a test theory based on the fallible measures of human traits as against true objective values (Spearman, 1904). The CTT operates on the assumption that an individual's observed score on a construct is a composite of true score and error score. Some of the deficiencies in the CTT, among others, are; it is based on a composite score, it is based on a group, and it does not take into how testees with different ability levels on the trait have performed on the item (Crocker & Algina, 2008). Following these inefficiencies, the IRT evolved. IRT, primarily, emphasises how individuals with different abilities on the trait respond to an item. IRT has been used for validating achievement and non-achievement tests (Morizot et al., 2007; Baker, 2001). IRT assumes that the ability scale is an interval scale with a zero midpoint, and scores ranging from positive infinity to negative infinity, but practically ranging from -3 to +3 (Baker, 2001). The negative part represents low ability, while positive represents high ability. The basic assumption of the IRT is that every testee (respondent) has some amount of the latent trait in question, which is referred to as ability (θ). Each testee has a score which places him/her on the ability scale. Any point on the ability scale has a corresponding probability that a testee with such ability will choose the correct response to the

89

item. The probability is low for testees with low ability and high for testees with high ability.

IRT has a number of models which can be used to analysed data on polytomously and/or dichotomously scored items. The models, among others, include: one-parameter logistic model (1PLM), two-parameter logistic model (2PLM), and three-parameter logistic model (3PLM), rating scale model (RSM), nominal model (NM), partial credit model (PCM), graded response model (GRM), and generalised partial credit model (GPCM). For the purpose of this study, the 3PLM was used to validate the items on test construction skills. The 3PLM consists of three parameters, namely; item difficulty, discrimination, and probability of guessing. These parameters (difficulty, discrimination, and probability of guessing) in the CTT are not the same as in the IRT.

First, difficulty in the CTT is the proportion of testees getting an item correct. This ranges from 0 to 1. However, in IRT, it is the location on the ability scale where an item functions. Difficulty is a point on the ability scale where the probability of correct response is .5 for one- and two-parameter models and $(1 + c)/2$ for a three-parameter model. The difficulty denoted by 'b', ranges from positive infinity to negative infinity. A negative difficulty index means the item functions among the low ability group; and this means the item is an easy item. Positive difficulty index shows that the item functions among the high ability group; and thus, the item is difficult. Another parameter in 3PLM is discrimination. In CTT, discrimination is the difference between the proportion of testees in the upper and lower scoring groups who had an item correct.

90

Discrimination indices range from -1 to +1. In the IRT, however, discrimination refers to the slope or steepness of the item characteristic curve, which is denoted by *'a'*. It tells how an item discriminates between low and high abilities. The discrimination indices range from 0 to positive infinity. Discrimination indices are interpreted based on the following criteria: none = 0; very low = .01 - .34; low = .35 - .64; moderate = .65 - 1.34; high = 1.35 - 1.69; very high > 1.70 (Baker, 2001). Guessing parameter is the lower bound for the item characteristic curve. It is denoted by 'c', and it is the probability of getting an item correct at all ability levels. It is the probability that connotes the tip of the item characteristic curve from below.

The 3PLM was used for this study based on the fact that items on the test construction skill scale were scored dichotomously as either '1' or '0', and there was the possibility that respondents could simply guess to get an item correct. Crocker and Algina (2008) posit that 3PLM is advantageous than 1 and 2PLMs, since the model is able to accommodate guessing. This, therefore, explains why difficulty, discrimination, and guessing were estimated for this validation. In addition, Response Factor Analysis (RFA) was performed. This forms part of the IRT output. The results of the 3-parameter model are presented in Table 3.

Table 3- *3PL Model Item Parameter Estimates for Group 1, logit: aθ + c or a(θ – b)*

Item	Discrimi.	Diffi.	Guess.	Factor Loadings	S-X² Item Statistics			New λ_2
	a	*b*	*c*	λ_1	X^2	*d.f*	*Sig.*	
2	.65	-3.89	.20	.36	5.73	2	.057	0.42
3	.91	-3.32	.20	.47	1.43	2	.490	0.53
4	.92	-3.10	.20	.48	4.71	2	.095	0.53
5	.95	-2.02	.20	.49	10.65	6	.100	0.49
6	1.35	-1.87	.20	.62	12.81	5	.025*	0.61
7	.76	-2.37	.20	.41	7.36	6	.290	0.39
8	.99	-2.36	.20	.50	7.98	4	.092	0.49
9	.71	-5.10	.20	.39	4.17	3	.245	0.50
10	1.03	-2.70	.20	.52	2.11	1	.147	0.57
11	.82	-4.14	.20	.44	5.57	4	.235	0.39
12	.78	-2.55	.20	.42	6.30	4	.178	0.40
13	1.43	-.01	.19	.64	2.96	5	.707	0.68
14	.59	-3.57	.20	.33	3.73	1	.054	0.34
15	1.02	-3.05	.20	.52	9.82	6	.132	0.54
16	.76	-1.94	.20	.41	7.11	3	.068	0.42
17	3.29	-1.39	.19	.89	3.50	5	.624	0.89
18	1.88	-1.27	.19	.74	5.02	5	.415	0.74
19	.70	-2.53	.20	.38	2.26	1	.134	0.39
20	.50	-5.91	.20	.28**	2.18	1	.141	
21	.61	-4.58	.20	.34	9.46	5	.092	0.88
22	.37	-3.98	.20	.22**	7.77	4	.100	
23	.37	-6.03	.20	.21**	4.63	6	.593	
24	.42	-3.04	.20	.24**	3.29	6	.772	
25	.36	-3.72	.20	.21**	5.73	2	.057	

Item 2 – 25 = questionnaire items (see Appendix A); **items to be discarded; M_2 (228) = 348.70, $p < .001$, RMSEA = .07.

From Table 3, the S-X^2 item fit statistics indicates that all the items, apart from item 6 ($p = .025$) fit the model. The overall goodness of fit model, M_2 (228) = 348.70, $p < .001$, RMSEA = .07. This suggests that though the test was significant, the RMSEA value was closer to 0, hence, the model can be deemed fit. This implies that the hypothesised model fit the observed data.

It was further revealed that item 1 had a variance of 0, hence, it was not included in the model. As presented in Table 3, the discrimination indices for all the items range from .65 to 3.29. Fourteen (14) of the items (items; 2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 14, 15, 16, and 19) had moderate discrimination indices, thus, they ranged from .65 – 1.34. Also, 6 of the items (20, 21, 22, 23, 24, and 25) had low discrimination indices, thus, .35 - .64; 2 items (6 and 13) had high discrimination indices, thus, 1.35 – 1.69. Finally, 2 items (17 and 18) had very high discrimination indices, thus, 1.70 and above (Baker, 2001). In terms of difficulty, the indices for all the items ranged from -.01 to -6.03, and this indicates that all the items functioned among the low ability levels, hence they are considered as easy items. The difficulty of an item describes where the item functions along the ability scale. An easy item functions among the low ability levels, whereas a difficult item functions among the high ability examinees. The probability of guessing ranged from .19 to .20. This means that at all ability levels, the probability of getting the item correct by guessing alone is approximately .20. These guessing parameters (.20 and .19) were less than the natural random guessing of .5. This implies that for all the items, there is a less probability of a respondent with low ability to guess.

93

From the 3PLM, it is evident that more than half (14) of the items moderately discriminated between high and low ability groups. However, only 6 items had low discrimination. Again, all the items were relatively easy items, but the probability for guessing was small. From these results, it can be concluded all the items, apart from item 1, are good items.

In addition to the aforementioned 3 parameters, the Response Factor Analysis (RFA) indicated that five items (20, 22, 23, 24, and 25) had factor loadings below .30, hence, they are to be deleted as recommended by Pallant (2010) and Field (2009), since they are contributing less than 9% of the variances to the construct. Based on the results, in all, 6 items have to be deleted (1, 20, 22, 23, 24, and 25). However, when the aforementioned 6 items were deleted, the new factor loadings ranged from .34 to .89. Finally, the teacher test construction skills scale was made up of 19 items, which were ideal based on the analysis. Upon careful consideration item 22 (Consider the age of learners during item writing) was retained, however, 'age' was changed to 'class level', since age was not appropriate. Again, item 7 (Subject test items to item analysis) was deleted, since item analysis are mostly used by testing agencies. The items deleted included; Item 1: 'Give clear instructions to guide the test takers'; item 20: 'Consult standard text books in the subject for guide'; item 21: 'Keep a resource bank of questions that can be used to when setting tests'; item 23: 'Submit items for vetting to the Head of Department or the principal'; item 24: 'Submit tests meant for promotional examinations for expert editing on time'; and item 25: 'Review draft of the test at least twice in two days before

administering'. Finally, 18 items were used for the data collection (see Appendix A). This was then used for the final data collection.

Commitment

This section presents the results on the validation of Teacher Commitment Scale. The details of the confirmatory factor analysis are presented in Table 4.

Table 4- *Item loadings, AVE, and Reliability (Commitment)*

Dimensions	Items	Loading	Lower	Upper	p	AVE	Alpha
School	C1	.364	.213	.498	.000	.54	.66
	C2	.816	.741	.874	.001		
	C3	.905	.864	.942	.000		
Students	C4	.877	.831	.914	.000	.78	.91
	C5	.915	.864	.952	.001		
	C6	.849	.798	.888	.000		
Teaching	C7	.173*	.010	.387	.035	.41	.46
	C8	.563	.445	.667	.000		
	C9	.937	.844	1.076	.000		
Profession	C10	-.693	-.753	-.623	.001	.50	.46
	C11	-.743	-.841	-.630	.001		
	C12	-.896	-.968	-.832	.000		
	C13	.416	.268	.555	.000		

C1 – C13 = questionnaire items (see Appendix A); *item to be discarded;
Overall Alpha = .83

As shown in Table 4, item C7 had a factor loading below .30, therefore, this item was discarded. This was the only item discarded since it had a factor loading below the recommended loading of .30 (Pallant, 2010). From Table 4, all the AVEs for the various dimensions apart from Commitment to Teaching were greater than .50, hence convergent validity established. This suggests that Commitment to Teaching dimension lacks convergent validity. In addition, convergent validity was assessed using AVE (Fornell & Larcker, 1981).

Table 5- *Discriminant Validity (Commitment)*

Commitment	School	Students	Teaching	Profession
1. School	(.73)*			
2. Students	.80	(.88)*		
3. Teaching	.66	.75	(.64)*	
4. Profession	-.15	-.05	.06	(.71)*

*Values in parenthesis are square roots of AVEs

The discriminant validity was assessed using Fornell and Larcker's (1981) criterion. The result showed that apart from the correlation between Commitment to School and Commitment to Students, all the square roots of AVEs were greater than the correlation among the dimensions (see Table 5). Based on this it was concluded that discriminant validity was established. This, therefore, implies that all the items on the commitment scale have to be grouped under their respective dimensions. Figure 2 shows the model for commitment scale.
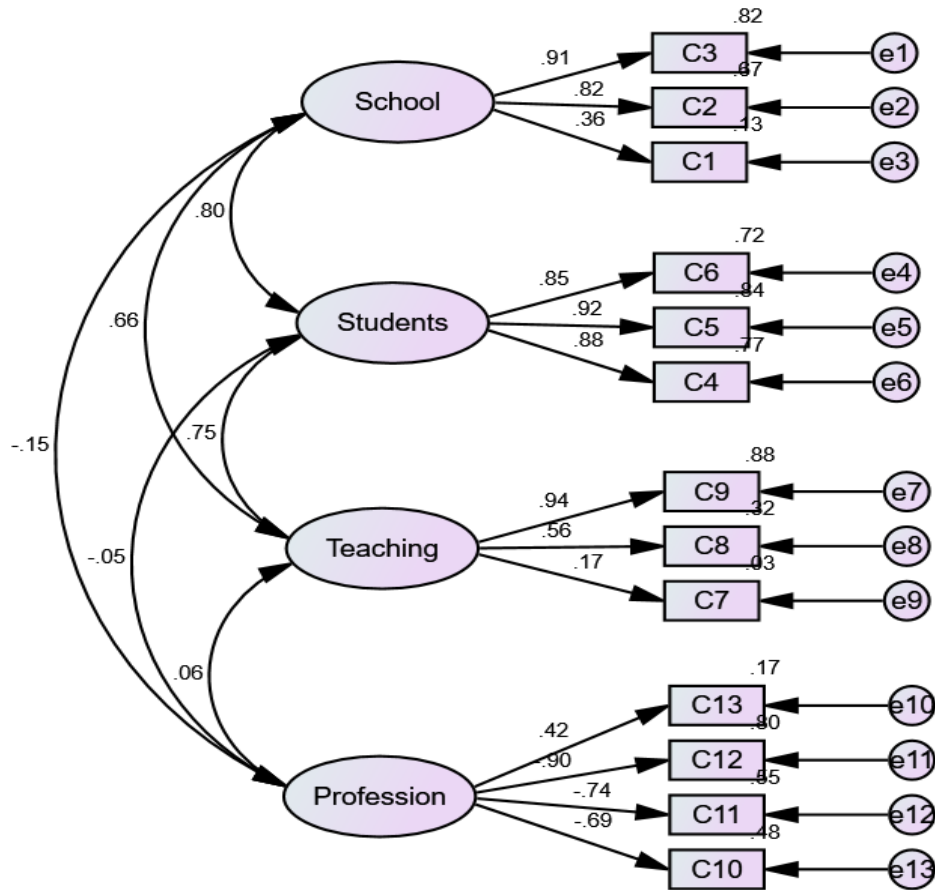
Figure 2- Commitment scale

Attitude towards Test Construction (ATTC)

This section presents the results on the Attitude towards Test Construction, using the confirmatory factor analysis. Table 6 presents details of the results.

The result in Table 6 show that item 9 (I prefer the item format of a classroom test to be decided by the learners) had a factor loading below .30, therefore, this item is not contributing much to the construct. However, all the items had loadings above .30. Item 9 was deleted. In addition, all the AVEs for the various dimensions Attitude towards Test Construction were below .50,

97

hence convergent validity was a problem. This suggests that the Attitude towards Test Construction scale lacks convergent validity.

Table 6- *Item loadings, AVE, and Reliability (Attitude towards Test Construction)*

Constructs	Items	Estimate	Lower	Upper	P	AVE	Alpha
Planning	D1	.665	.466	.809	.002	.39	.88
	D2	.655	.487	.769	.004		
	D3	.732	.576	.825	.004		
	D4	.767	.634	.862	.003		
	D5	.716	.570	.838	.003		
	D6	.698	.517	.814	.002		
	D7	.560	.309	.747	.004		
	D8	.766	.608	.872	.002		
	D9	.292*	.010	.518	.041		
	D10	.528	.316	.685	.003		
	D11	.420	.097	.683	.009		
	D12	.329	.042	.542	.025		
	D13	.717	.572	.804	.005		
Item construction	D14	.723	.343	.849	.005	.37	.81
	D15	.462	.057	.672	.027		
	D16	.506	.141	.751	.003		
	D17	.672	.428	.835	.003		
	D18	.392	-.018	.644	.076		
	D19	.751	.410	.875	.009		
	D20	.688	.329	.886	.004		
	D21	.574	.340	.734	.002		
Item review	D22	.442	.074	.710	.017	.36	.73
	D23	.390	-.026	.830	.066		
	D24	.716	.113	.936	.020		
	D25	.598	.337	.843	.001		
	D26	.771	.177	1.019	.008		
Item assembling	D27	.544	.246	.742	.006	.29	.61
	D28	.577	.164	.815	.018		
	D29	.567	.186	.892	.004		
	D30	.446	.123	.748	.009		

*item to be discarded (see Appendix A); Overall Alpha = .89

The discriminant validity was assessed using Fornell and Larcker's (1981) criterion. The results are presented in Table 7.

Table 7- *Discriminant Validity (Attitude towards Test Construction)*

ATTC	Planning	Item construction	Item review	Assembling
Planning	(.63)*			
Item construction	.80	(.61)*		
Item review	.02	.02	(.60)*	
Assembling	.41	.32	.64	(.54)*

*Values in parentheses are square roots of AVEs

The result showed that the correlation between Test Planning and Item Construction (.80); and that of Item Review and Item Assembling (.64) were greater than the square roots of AVEs the respective dimensions as presented in Table 7. Based on this, it was concluded that discriminant validity was partially established. In all, 29 items were retained for the final data collection. Figure 3 shows the hypothesised model for attitude towards test construction with the factor loadings of each item.
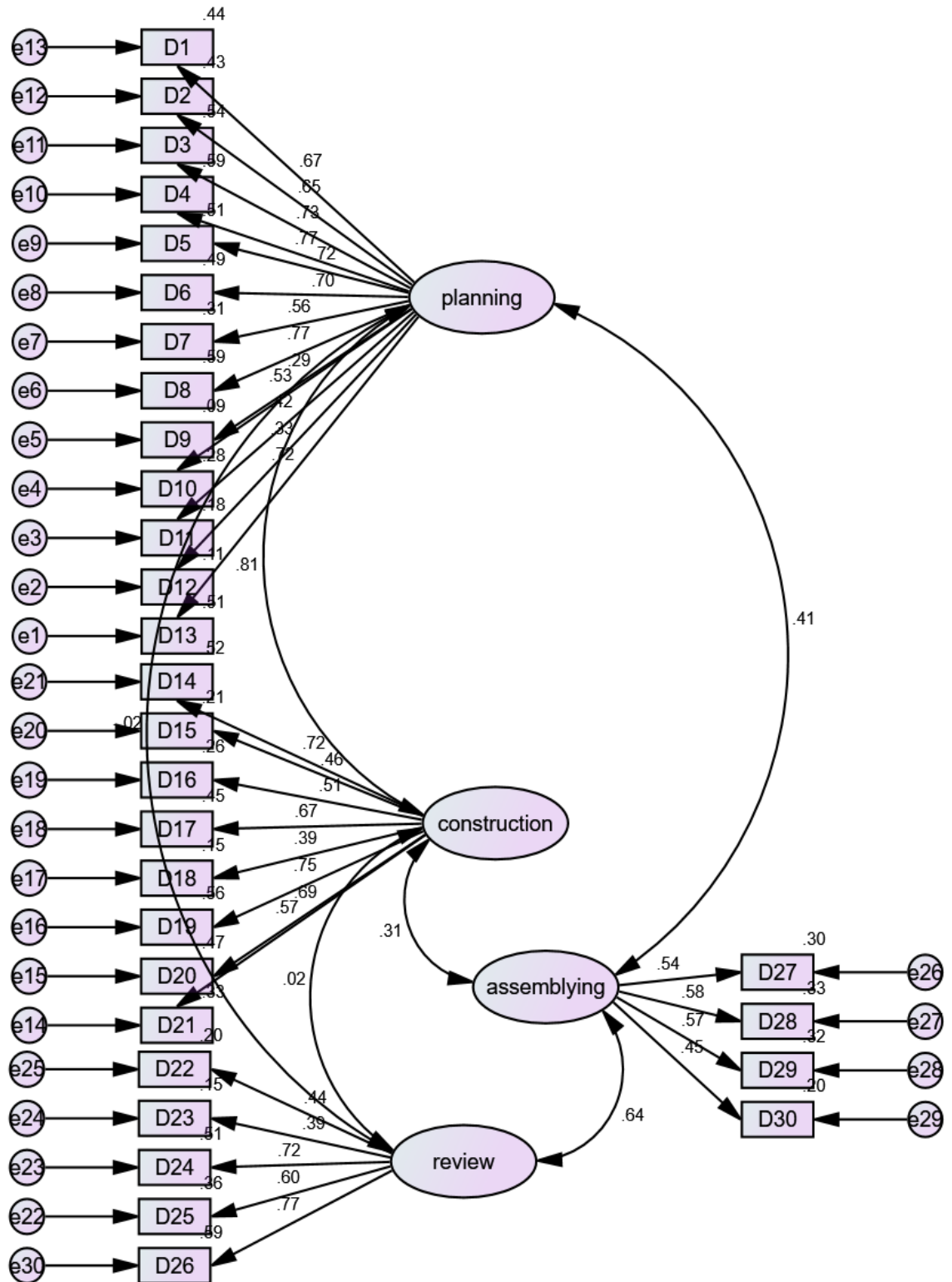
Figure 3- Attitude towards test construction

100

Adherence to Test Construction Principles (ATTC)

This section presents the results on the Adherence to Test Construction Principles using confirmatory factor analysis. Details of the results are presented in Table 8 and Figure 4.

Table 8- *Item loadings, AVE, and Reliability (ATTC)*

Dimensions	Items	Estimate	Lower	Upper	P	AVE	Alpha
Language use	EE1	.794	.646	.888	.003	.56	.86
	EE2	.682	.412	.908	.002		
	EE3	.754	.553	.878	.002		
	EE4	.688	.503	.817	.002		
	EE5	.820	.720	.910	.001		
Content coverage	EE6	.739	.569	.844	.003	.60	.92
	EE7	.705	.562	.816	.003		
	EE8	.842	.722	.919	.002		
	EE9	.859	.776	.913	.004		
	EE10	.846	.736	.919	.004		
	EE11	.773	.645	.865	.002		
	EE12	.815	.667	.908	.003		
	EE13	.590	.429	.708	.002		
Item organisation	EE14	.680	.472	.851	.003	.62	.91
	EE15	.827	.715	.908	.003		
	EE16	.734	.520	.905	.002		
	EE17	.947	.858	.986	.002		
	EE18	.924	.816	.974	.002		
Test development	EE19	.725	.556	.834	.006	.52	.86
	EE20	.713	.536	.823	.006		
	EE21	.747	.592	.856	.002		
	EE22	.602	.347	.787	.003		
	EE23	.670	.445	.832	.001		
	EE24	.642	.427	.812	.001		
	EE25	.684	.486	.822	.001		

EE1 – EE25 = questionnaire items (see Appendix A); Overall Alpha = .96

As presented in Table 8, all the items had factor loadings above .30. In addition, all the AVEs for the various dimensions were greater than .50, hence

convergent validity established. The discriminant validity was assessed using Fornell and Larcker's (1981) criterion. The results are shown in Table 9.

Table 9- *Discriminant Validity (ATTC)*

ATTC	Language use	Content coverage	Item organisation	Test development
Language use	(.75)*			
Content coverage	.87	(.77)*		
Item organisation	.86	.84	(.79)*	
Test development	.79	.85	.82	(.72)*

Values in parenthesis are square roots of AVEs

From the results, the correlations among all the various dimensions were greater than the square roots of AVEs for the respective dimensions (see Table 9). Based on this it can be concluded that discriminant validity was not established, therefore, the scale should be treated as unidimensional rather than multidimensional. Based on the findings, all the items were retained, however, the scale was treated as a unidimensional scale. Figure 4 shows the model for the adherence scale with the factor loadings of each item.

The reliability (internal consistency) estimates for the final version of the scales on the questionnaire (see Appendix A) were estimated using Kuder-Richardson 20 and Cronbach's alpha coefficients. The coefficients are presented in Table 10.
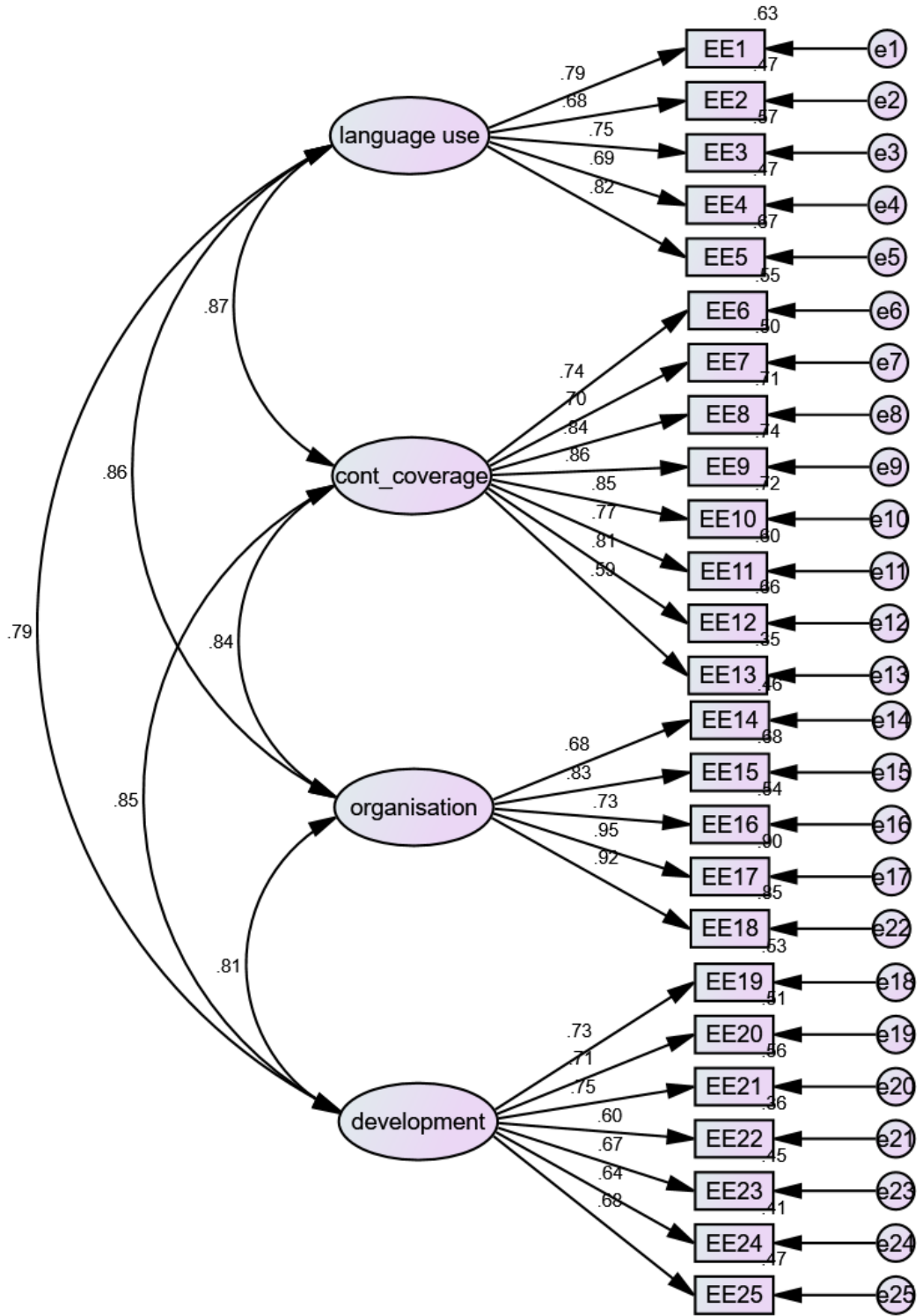
Figure 4- Adherence to Test Construction Principles

Table 10- *Reliability Estimates*

Scale/Sub-dimension	No. of items	Coefficient
Test Construction Skills	18	.65
Commitment to Work		
Commitment to school	3	.48
Commitment to students	3	.58
Commitment to teaching	2	.38
Commitment to profession	4	.71
Overall (commitment)	12	.65
Attitude towards Test Construction		
Test planning	12	.81
Item construction	8	.74
Item review	5	.62
Item assembling	4	.38
Overall (attitude)	29	.86
Adherence to Test Construction Principles	24	.89

Source: Field survey (2019)

From Table 10, the internal consistency for test construction skills was, KR20 = .65. KR20 was used because the scale was scored dichotomously. Commitment to Work, α = .65; Attitude towards Test Construction was, α = .86; and Adherence to test construction Principles, α = .89. Cronbach's alpha was used because the scales were scored polytomously. These coefficients are appropriate, and show minimal errors.

Data Collection Procedures

An introductory letter was taken from the Department of Education and Psychology to introduce the researcher as student (see Appendix C). Necessary ethical clearance and permissions were sought from the appropriate quarters throughout the study, whenever necessary (see Appendix D). The introductory letter was sent to the Metropolitan Directorate of Education, Sekondi-Takoradi Metropolis to seek permission for the conduct of the study. Upon being given approval, arrangements were made with the heads of the various schools on the specific dates and time for the data collection. On the prearranged dates, myself and three other research assistants visited the various schools for the data collection. The questionnaire were distributed personally by hand to the teachers. Afterwards, those that were responded to within the same day were retrieved. However, the remaining were left into the care of the assistant headmaster, and were to be collected in two weeks' time. Unfortunately, on arrival on second week, three out of the eleven schools had completely misplaced the questionnaire. New questionnaires were printed for those schools and retrieved in two weeks' time. On the whole, the data collection took place within four weeks. At the end of the data collection, a return rate of 335 representing 96.8% was achieved.

Ethical Consideration

The conduct of the study adhered strictly to ethical issues regulating the conduct of any research. Confidentiality, anonymity, privacy were adhered to. Consent of respondents were sought before data collection commenced. First,

the purpose of the study was thoroughly explained to the respondents, after which their consent were sought by completing the consent form. Respondents were made aware that engagement in the study is by volition and not compulsion. The data that were collected kept confidentially, names of respondents or their schools were not be disclosed in any part of the work, pseudonyms were used. The data collected were analysed as a group, and, for that matter, it was not possible to trace responses to respondents. The data collected were well-managed and kept secretly to avoid the accessibility of other people.

Data Processing and Analysis

Data collected for the study were coded and entered with the help of SPSS software version 25.0. Structural Equation Modeling (SEM) was performed to validate the items on the instrument with the help of AMOS software. SEM is a second generational multivariate technique that is well suited for testing various hypothesised or proposed relationships between variables. SEM enables an explicit modeling of measurement error in order to obtain unbiased estimates of the relationships between variables. This allows researchers to remove the measurement errors from the correlation/regression estimates. This is conceptually the same as correcting for measurement error (or correcting for attenuation), where measurement error is taken into account (Khine, 2013).

Data collected using the three research questions were analysed using descriptive statistics – frequencies, percentages, mean and standard deviation.

Specifically, for Research Question 1, each of the items were scored dichotomously as '1' or '0', where '1' were awarded for correct response and '0' for an incorrect response. In all, the scores of each respondent ranged from 0 to 18. Scores from 0 – 8.9 were described as low level of skills (thus, below 50%), scores from 9 – 14.3 (thus, 50% to 79.9%) were described as moderate level of skills, while scores 14.4 – 18 (thus, 80% or more of the total scores) were described as high level of skills in test construction. The mean score for the respondents was then estimated. In addition, the scores of the respondents were categorised, and described in terms of frequency counts and percentages.

For Research Question 2, there were 12 items on 6-point Likert-type scale which were scored from 1 to 6 (thus, from very strongly disagree = 1, strongly disagree = 2, disagree = 3, agree = 4, strongly agree = 5, and very strongly agree = 6). A mean score of 3.5 was used. This is the mean of the responses (mean of means). The mean scores of all the items were summed and divided by the number of responses to get the mean of means. A mean score above 3.5 was considered as high level of commitment, while mean score below 3.5 was considered as low level of commitment. The standard deviation showed the variations of each of the scores away from their mean score.

In terms of Research Question 3, there were 29 items on a 4-point Likert-type scale, which were scored as: strongly disagree = 1, disagree = 2, agree = 3, and strongly agree = 4. In determining the attitude of respondents, the mean of their responses were computed and interpreted. With the exception of Item 28 'It is essential to identify behaviours to represent a construct when crafting test

107

items', all the other items were reverse-coded, such that: strongly disagree = 4, disagree = 3, agree = 2, and strongly agree = 1. In interpreting the attitude score of respondents of a particular item (e.g., Item 5), the mean score of the responses is compared with 2.5 (thus, [1+2+3+4]/4 = 2.5). Mean scores less than 2.5 show a negative attitude, whereas mean scores above 2.5 show a positive attitude to that particular item. For the interpretation of individual scores, the mean of the obtained scores is also compared with 2.5. The standard deviation showed the variations of each of the scores away from their mean score.

Hypotheses 1 to 3 were tested individually using SEM path analysis, with 5000 bootstrap samples with bias corrected accelerated confidence intervals. The bootstrap results were interpreted in terms of confidence intervals. For a particular result to be significant, the bootstrap upper and lower confidence interval should not contain '0', thus, both the upper and lower confidence intervals should be of the same sign ('+ +' or '- -'). This means the confidence interval does not include 0, which means the regression coefficient cannot be 0.

Hypotheses 4 was tested using the double mediation model (serial) of PROCESS by Hayes (2018) using 10000 bootstrap samples with 95% bootstrap confidence intervals. PROCESS is a regression-based model that is used in estimating model parameters using an ordinary least squares (OLS) regression. Linear regression model using the OLS criterion generates many possible pairs. The pair of values for the regression constant and the coefficient that an OLS regression produces minimizes the residual sum of squares. Prior to the use of

PROCESS and other second generational statistical procedures, mediation analyses were based on Baron and Kenny's approach (Baron & Kenny, 1986). This approach to mediation analysis was built on three causal relationships, using a regression model to estimate for each relationship. First, the predictor variable should be a significant predictor of the criterion variable. Having satisfied the first condition, then the predictor variable should be a significant predictor of the mediator variable, and finally, the mediator variable should be a significant predictor of the criterion variable. Based on the relationships, all the conditions have to be met before mediation could be assumed as taken place. The flaw with procedure is that it is not able to estimate the indirect effect and its significance. Again, mediation is possible without the predictor variable being a significant predictor of the criterion variable. Following these flaws, the current form of mediation analysis, PROCESS, was introduced (Hayes, 2018).

PROCESS is one of the current ways of estimating direct, indirect, and conditional interactions, with 92 models. This procedure utilises the bootstrapping method, which is a method of resampling with replacement which estimates better the standard error of the mean and thereby providing a better estimate of the confidence intervals and the critical ratio. This analysis was done using the double mediation model (serial). This procedure is based on the assumption of more than one mediators. This is such that one mediator influences the other, which in turn influences the criterion variable. This model was used because I hypothesized in Hypothesis 4 that teachers' test construction skills will have an indirect effect on adherence to principles of test construction,

109

through commitment to work, which then influences attitude towards test construction. Hayes (2018) recommended the use of 10000 bootstrap samples since the mediator variables are more than one. This, he indicated provides a better estimate of the model. Specifically, model 6 of PROCESS was used for the double mediation analysis.

Hypothesis 5 was tested using the simple mediation model of PROCESS with 5000 bootstrap samples. This hypothesis was tested using model 4 of PROCESS which deals with simple mediation. In this model, attitude towards test construction was used as a mediator in the relationship between teachers' level of commitment to work and their level of skills in test construction.

Hypothesis 6 was tested using Structural Equation Modeling (SEM). Structural equation modelling (SEM) is among the most useful advanced statistical analysis techniques that have emerged in the social sciences in recent decades. SEM is a class of multivariate techniques that combine aspects of factor analysis and regression, enabling the researcher to simultaneously examine relationships among measured variables and latent variables as well as between latent variables (Hair et al., 2014). There are two approaches to the use of SEM: the covariance-based and the variance-based. The covariance-based are primarily for confirmatory analyses, while the variance-based are mostly used for exploratory analyses (Hair et al., 2014). The confirmatory approach is used when testing the hypotheses of existing theories and concepts, while the exploratory is used when there is no or only little prior knowledge on how the variables are related. The estimation procedure for PLS-SEM is an ordinary

least squares (OLS) regression-based method rather than the maximum likelihood (ML) estimation procedure for CB-SEM.

Hypothesis 6 was specifically tested using the covariance-based approach, specifically, with the use of AMOS. This is because the model, is to determine whether the relationships that I have hypothesised (model) would fit the actual data collected. In other words, the hypothesis sought to determine whether there would be a discrepancy between the model proposed (conceptual framework) and the model predicted (results obtained). Results on Hypothesis 6 will serve as a foundation for future studies in this area.

Chapter Summary

The study examined teachers' test construction skills, commitment to work, and attitude towards test construction as predictors of adherence to test construction principles among SHS teachers in the Sekondi-Takoradi Metropolis. The study was carried out employing the descriptive survey design, specifically, the cross-sectional design. The population for this study comprised all teachers in public SHSs in Sekondi-Takoradi Metropolis – with a total of 998 teachers. The study, however, targeted teachers who teach mathematics, integrated science, and English language, who 346 in number. Using the census method, all the 346 teachers were engaged in the study. Questionnaire was used for the data collection. The scales on the questionnaire were validated using IRT and CFA with the help of IRTPRO and AMOS software. The data collected were analysed using descriptive statistics – frequencies, percentages, mean and standard deviation; and inferential statistics Structural Equation Model (SEM)

and mediation analysis with Hayes' PROCESS. The bootstrap approach was used for all the inferential analyses.

CHAPTER FOUR

RESULTS AND DISCUSSION

The study examined teachers' skills in test construction, commitment to work, and attitude towards test construction as predictors of adherence to principles of test construction among SHS teachers in the Sekondi-Takoradi Metropolis. This investigation was carried out using the descriptive cross-sectional design, with the census survey. Questionnaire was used to gather data from the respondents. Three hundred and forty-six (346) questionnaires were administered. However, 335 representing 96.8% were deemed valid for the purpose of data analysis. This percentage is above Amedahe and Asamoah-Gyimah's (2015) assertion that 5% to 20% of the population size is enough for generalisation purposes. This chapter presents the results and discussion of the results. In terms of the results, the demographic characteristics of the respondents were presented first, and this was followed by the results of the research questions, and, subsequently, the hypotheses.

Demographic Characteristics of Respondents

This section presents results on the respondents based on demographic distribution. The demographic information include gender, age-category, professional qualification, and years of teaching experience. Table 11 presents details of the demographic distribution.

Table 11- *Distribution of Respondents by Demographic Characteristics (N = 335)*

Variable	Frequency	Percentage (%)
Gender		
Male	240	71.6
Female	95	28.4
Age-category		
20 years & below	50	14.9
21 – 30 years	22	6.6
31 – 40 years	158	47.2
41 – 50 years	89	26.6
51 – 60 years	16	4.7
Professional qualification		
Degree with education	286	85.4
Degree without education	49	14.6

Source: Field survey (2019)

From Table 11, there were more males 240 (71.6%) than females 95 (28.45). Most 158 (47.2%) of the respondents were from the ages of 31 to 40 years, while few 16 (4.7%) were from 51 to 60 years. The respondents have taught from 1 to 29 years ($M = 11.35$, $SD = 7.44$). Generally, from the demographic distribution, the disproportionate nature of gender makes the result susceptible to gender bias. In addition, it appears only few of the respondents were approaching their mandatory retirement age (60 years), however, majority were in their active ages. This suggests that respondents are in a better position

114

to execute effectively their teaching mandate by doing everything possible, such as engaging in proper assessment of students' learning to make their teaching effective. Again, majority of the respondents had professional education background, which presupposes that they have been formally trained on the rudiments of students' assessment, and for that matter, they stand tall as far as assessment of students' learning is concerned. Teaching for about 11 years is quite reasonable to infer that respondents are experienced teachers.

Research Question 1

What is the level of skills of teachers in test construction?

The research question aimed at determining the level of skills of teachers in test construction. Respondents were asked to respond to 18 items, which were factual statements about test construction. Their responses were scored as '1' and '0' for correct and incorrect response respectively. The total scores for all the items, therefore, ranged from 0 to 18. A score of 80% or more (thus, 14.4 or more) out of 18 were classified as possessing high level of skills, scores from 9 – 14.3 out of 18 (thus, 50% to 79.9%) were classified as possessing moderate level of skills, and scores from 0 – 8.9 were described as possessing low level of skills (thus, below 50%). Table 12 presents details of the distribution of respondents based on their level of skills.

The scores of respondents in test construction ranged from 8 to 18 with a mean ($M = 15.12$, $SD = 2.37$) out of 18. This mean score falls within the high of skills category (see Table 12), indicating that respondents, on the average, have high level of skills in test construction. Similarly, this result was confirmed

115

by more than half (62.1%) of the respondents belonging to the high level of skills category.

Table 12- *Distribution of Respondents by Level of Skills in Test Construction*

Level of skills	Score range (out of 18)	Frequency	Percentage (%)
Low	0 – 8.9	2	.6
Moderate	9 – 14.3	125	37.3
High	14.4 – 18	208	62.1

Source: Field survey (2019); *Maximum = 18, Minimum = 8*

It can, therefore, be said that respondents, generally, have higher skills in test construction. This result could emanate from the fact that majority of the respondents were professional teachers and had taken at least a course in assessment or measurement and evaluation.

Research Question 2

What is the level of commitment of teachers towards work?

This research sought to determine teachers' level of commitment towards their job. A 6-point Likert-type scale with 12 items was used to assess commitment to work among teachers. A mean score of 3.5 was used. This is the mean of the responses (mean of means). The mean scores of all the items were summed and divided by the number of responses to get the mean of means. A mean score above 3.5 was considered as high level of commitment, while mean score below 3.5 was considered as low level of commitment. Details of the level of commitment are presented in Table 13.

116

Table 13- *Respondents' Level of Commitment to Work*

Sub-dimension	No. of items	Mean	SD
Commitment to school	3	3.78	.98
Commitment to students	3	4.68	.83
Commitment to teaching	2	4.34	.99
Commitment to profession	4	3.77	1.21
Mean of means		4.14	1.00

Source: Field survey (2019)

The overall mean of respondents on commitment to work was, $M = 4.14$, $SD = 1.00$, out of 6.0. This implies respondents were committed. Even, though respondents' were committed, their level of commitment to work was not high. This implies that teacher were not highly committed. Additionally, when the mean scores of the various dimensions were computed, it was evident that respondents were highly committed to their students ($M = 4.68$, $SD = .83$). This implies that the respondents were highly committed to ensuring that their students succeed academically, fostering cordial relationships among them, as well as mediating differences among the students. This was followed by commitment to teaching ($M = 4.34$, $SD = .99$), which was characterised by respondents showing more interest in teaching and always of planning future lessons. Further, respondents were commitment to school ($M = 3.78$, $SD = .98$), and this was manifested as respondents shared in the values and vision of their schools, and discussed their schools as great to others. Lastly, respondents were least committed to the profession ($M = 3.77$, $SD = 1.21$), which implies that they did not like the teaching profession, even though they had the interest of the

students at heart. Respondents wished they got a job different from being a teacher, even if they would take the same pay.

Research Question 3

What is the attitude of teachers towards test construction?

The aim of this research question was to determine the attitude of teachers towards test construction. Respondents were asked to respond to 29 items on a 4-point Likert-type scale. In interpreting the attitude score of respondents of a particular item (e.g., Item 5), the mean score of the responses is compared with 2.5 (thus, [1+2+3+4]/4 = 2.5). Mean scores less than 2.5 show a negative attitude, whereas mean scores above 2.5 show a positive attitude to that particular item. For the interpretation of individual scores, the mean of the obtained scores is also compared with 2.5. The responses of the respondents are presented in Table 14.

Table 14- *Respondents' Attitude towards Test Construction*

Sub-dimension	No. of items	Mean	SD
Planning	12	2.97	.49
Item construction	8	2.86	.57
Item review	5	2.66	.43
Item Assembling	4	2..61	.45
Mean of means		2.83	.49

Source: Field survey (2019)

As shown in Table 14, respondents had a more positive attitude towards test planning (*M* = 2.97, *SD* = .49), followed by item construction (*M* = 2.86, *SD* = .57), item review (*M* = 2.66, *SD* = .43), and item assembling (*M* = 2.62,

118

SD = .45). Generally, it is clear that respondents had a positive attitude towards test construction (*M* = 2.83; *SD* = .49) out of 116. Even though respondents had a positive attitude towards test construction, it was not strong.

Hypotheses Testing

The study tested six hypotheses. Prior to these hypotheses testing, the normality assumption, which is the fundamental of all parametric assumptions was tested. This was tested using mean, median, 5% trimmed mean, z-skewness, and the normal Q-Q plot. Details of the results are presented in Table 15.

Table 15- *Test for Normality*

Parameters	Skills	Attitude	Commitment	Adherence
Mean	15.12	2.83	4.14	3.89
Standard deviation	2.37	.49	1.00	.63
5% Trimmed mean	15.38	2.84	4.09	3.90
Median	16.00	2.84	4.08	3.92
Skewness	-.599	-.412	.119	-.090
Std. Error	.133	.133	.133	.133
$Z_{skewness}$	-4.50	-3.10	.89	-.68

As presented in Table 15, the mean, median, and 5% trimmed mean of teachers' test construction skills, attitude towards test construction, commitment to work, and adherence to principles of test construction were approximately equal. This implies that the distribution of scores of the aforementioned variables was normally distributed. Further examination of the $Z_{skewness}$ shows that the coefficients of attitude (Z_{skew} = -3.10), commitment (Z_{skew} = .89), and adherence (Z_{skew} = -.68) were within the range of +3.29 and -3.29 (Tabachnick

& Fidell, 2007), hence they are normally distributed. However, the coefficient of skills (Z_{skew} = -4.50), is an indication of negative skewness.

Additionally, the normal Q-Q plots for all the variables were examined. The normal Q-Q plot for skills showed that the distribution was normal (see Appendix B). From Appendix, the normal Q-Q plots for all the variables showed that the distribution of all the scores were closer to the straight line (see Appendix B). To further enhance the precision of the tests, bootstrap was performed for all the tests. This procedure would take care of the anomaly of the distribution of scores on skills.

Hypothesis 1

H_0: Teachers' level of skills in test construction will not predict their adherence to recommended principles of test construction.

H_1: Teachers' level of skills in test construction will predict their adherence to recommended principles of test construction.

This hypothesis sought to determine the influence level of test construction skills of teachers on their level of adherence to test construction principles. This hypothesis was tested using SEM path analysis, with 5000 bootstrap samples with bias corrected accelerated confidence intervals. The bootstrap results were interpreted in terms of confidence intervals. For a particular result to be significant, the bootstrap upper and lower confidence interval should not contain '0', thus, both the upper and lower confidence intervals should be of the same sign ('+ +' or '- -'). This means the confidence interval does not include 0, which means the regression coefficient cannot be 0.

120

The endogenous (criterion) variable was the adherence to test construction principles. This was measured on continuous basis using a scale. The exogenous (predictor) variable was test construction skills, which was measured on continuous basis. Table 16 and Figure 5 present the regression coefficients.
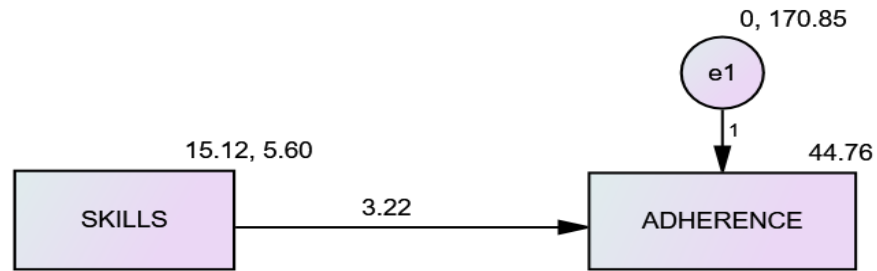
Figure 5- Path model for test construction skills and adherence

From Figure 5, test construction skills predicts adherence to test construction by 3.22, with error variances of 5.60 for skills and 170.85 for adherence. Table 16 presents results of test of significance prediction.

From Table 16, teachers' test construction skills explained 25.3% of the variance in adherence to test construction principles. Test construction skills of teachers is a significant predictor of adherence to test construction, B = 3.22, *Boot 95%CI* (2.55, 3.87). This result (B = 3.22) implies that test construction skills positively predict adherence to test construction. The result indicates that an additional increase in teachers' test construction skills would lead to 3.22 increase in adherence to principles of test construction by teachers.

121

Table 16- *Regression Model for Test Construction Skills and Adherence*

Model	B	Std. Error	CR	95% Confidence Interval	
				Lower	Upper
(Constant)	44.76	4.63	9.67	34.75	55.00
Skills	3.22*	.30	10.65	2.55	3.87

*Significant, $p < .05$; R = .503; $R^2 = .253$

Generally, based on the results, it can be said that higher test construction skills would lead to higher level of adherence to test construction principles. In this case, teachers with higher test construction skills would adhere to recommended principles in test construction. Higher test construction skills of teachers implies that teachers are knowledgeable in the rudiments of assessment, and specifically in test construction. Based on the results of this study, the null hypothesis that "Teachers' level of skills in test construction will not predict their adherence to recommended principles of test construction" is rejected in favour of the alternative hypothesis.

Hypothesis 2

H_0: *Teachers' level of commitment to work will not predict their adherence to recommended principles of test construction.*

H_1: *Teachers' level of commitment to work will predict their adherence to recommended principles of test construction.*

This hypothesis sought to determine whether teachers' level of commitment will predict their level of adherence to test construction principles. The SEM path analysis with 5000 bootstrap samples, with bias corrected accelerated confidence intervals. The exogenous (predictor) variables were the

122

sub-dimensions of teachers' commitment to work (commitment to school [ComSch], commitment to students [ComStu], commitment to teaching [ComTeach], and commitment to the profession [ComPrefess]) which were measured on continuous basis. The criterion variable was adherence to test construction principles, which was also measured on a scale. Details of the results are presented in Figure 6 and Table 17.
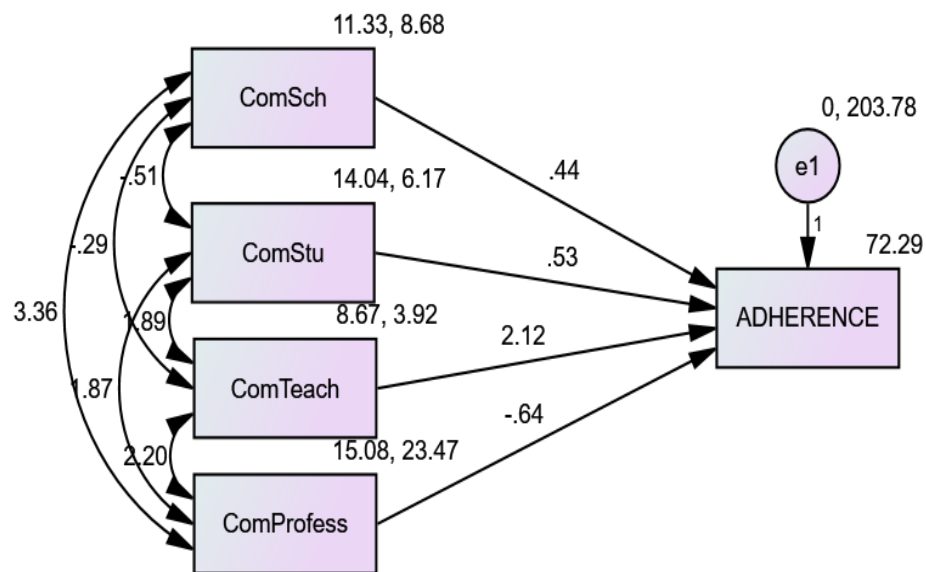
Figure 6- Path model for commitment and adherence

As presented in the path analysis (Figure 6), commitment to school predicts adherence to test construction by .44, commitment to students predicts adherence to test construction by .53, commitment to teaching predicts adherence to test construction by 2.12, and also commitment to profession predicts adherence to test construction by -.64; all with error variance of 203.78 for adherence. Table 17 presents the significance of the path model.

123

Table 17- *Regression Model for Commitment to Work and Adherence*

Model	B	Std. Error	CR	95% Confidence Interval	
				Lower	Upper
(Constant)	72.29	5.71	12.66	61.13	83.24
School	.44	.27	1.63	-.10	.97
Student	.53	.33	1.61	-.07	1.20
Teaching	2.12*	.40	5.30	1.30	2.88
Profession	-.64*	.15	4.27	-.95	-.35

*Significant, $p < .05$; R = .33; R^2 = .11

Note: School = Commitment to school; Student = Commitment to students; Teaching = Commitment to teaching; and Profession = commitment to profession

The results in Table 17 show that commitment to school, students, teaching, and profession together explained 11% of the variances in adherence to test construction principles. The results further revealed that only commitment to teaching, B = 2.12, *Boot 95%CI* (1.30, 2.88); and commitment to profession, B = -.64, *Boot 95%CI* (-.95, -.35) were significant predictors of adherence to test construction. However, commitment to school, B = .44, *Boot 95%CI* (-.10, .97); and commitment to students B = .53, *Boot 95%CI* (-.07, 1.20) were not significant predictors of adherence to test construction principles. The implication of these results is that teachers who are committed to their schools and students do not necessarily adhere to test construction principles. This means that these teachers need more than commitment to school and students in order to adhere to test construction principles.

While commitment to teaching was positive predictor of adherence to test construction principles, commitment to profession was a negative predictor. The results imply that an additional increase in teachers' commitment to teaching would lead to 2.12 increase in adherence to test construction principles, while controlling for influence of the other dimensions. Teachers who are highly committed to their teaching are more likely to adhere to recommended principles of test construction. Teachers being committed to teaching means they take seriously activities which are directly involved in their classroom teaching. On the contrary, an increase in teachers' commitment to the profession, while controlling for influence of the other dimensions, would lead to -.64 decrease in adherence to test construction principles. This implies that as teachers become more committed to activities which generally have to do which their profession, they are not likely to adhere to test construction principles. The results of the study has provided evidence to reject the null hypothesis that "Teachers' level of commitment to work will not predict their adherence to recommended principles of test construction" in favour of the alternative hypothesis "Teachers' level of commitment to work will predict their adherence to recommended principles of test construction".

Hypothesis 3

H_0: Teachers' attitude towards test construction will not predict their adherence to recommended principles of test construction.

H_1: Teachers' attitude towards test construction will predict their adherence to recommended principles of test construction.

This hypothesis sought to determine whether or not teachers' attitude towards test construction would predict their level of adherence to test construction principles. The hypothesis was tested using SEM path analysis with 5000 bootstrap samples with bias corrected accelerated confidence intervals. The exogenous (predictor) variables (attitude towards test planning, item construction, item review, and item assembling) were measured on scale. The endogenous (criterion) variable was adherence to test construction principles was measured on continuous basis. Figure 7 and Table 18 presents details of the results.

From the path model (Figure 7), attitude towards test planning predicts adherence to test construction principles by 1.03, attitude towards item construction predicts adherence to test construction principles by .37, attitude towards item review predicts -1.18 of adherence to test construction principles, and attitude towards item assembling predicts adherence to principles of test construction by -.64. The significance of these path coefficients are presented in Table 18.
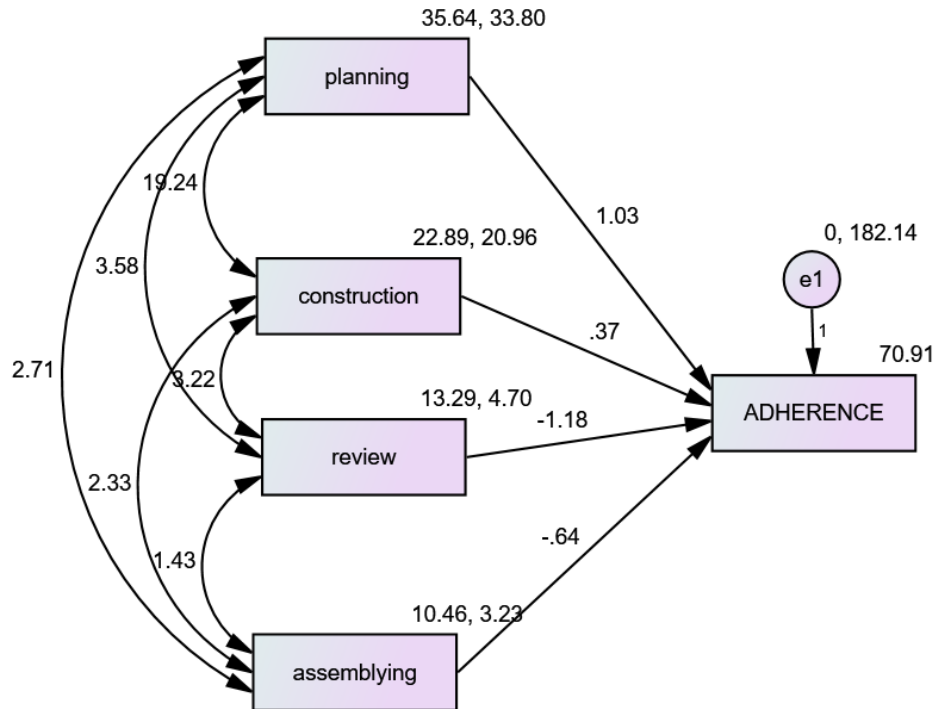
Figure 7- Path model for attitude towards test construction and adherence

Table 18- *Regression Model for Attitude towards Test Construction and Adherence*

Model	B	Std. Error	CR	95% Confidence Interval	
				Lower	Upper
(Constant)	70.91	6.96	10.19	57.27	84.03
Test planning	1.03*	.23	4.48	.59	1.47
Item construction	.37	..35	1.06	-.33	1.07
Item review	-1.18*	.37	3.19	-1.92	-.45
Item assembling	-.64	.47	1.36	-1.63	.26

*Significant, $p < .05$; R = .45; $R^2 = .20$

The results in Table 18 indicate that 20% of the variance in adherence

to principles of test construction was explained by all the dimensions of attitude

127

towards test construction. Among all the dimensions of attitude towards test construction, attitude towards test planning, B = 1.03, *Boot 95%CI* (.59, 1.47); and attitude towards item review, B = -1.18, *Boot 95%CI* (-1.92, -.45); were the only significant predictors of adherence to principles of test construction. Attitude towards test planning and attitude towards item assembling were predictors adherence to test construction principles.

Based on the results of this study, the null hypothesis which states that: "Teachers' attitude towards test construction will not predict their adherence to recommended principles of test construction" was rejected in favour of the alternative hypothesis (Teachers' attitude towards test construction will predict their adherence to recommended principles of test construction).

Hypothesis 4

H_0: Teachers' level of commitment to work and attitude towards test construction will not mediate (serial) the relationship between teachers' test construction skills and adherence to recommended principles of test construction.

H_1: Teachers' level of commitment to work and attitude towards test construction will mediate (serial) the relationship between teachers' test construction skills and adherence to recommended principles of test construction.

This hypothesis aimed at determining whether teachers' level of commitment to work and attitude towards test construction will serially mediate the relationship between teachers' test construction skills and adherence to

128

recommended principles of test construction. The implication of this is that, teachers' test construction skills will indirectly influence teachers' level of adherence to test construction principles through commitment to work, which will then influence their attitude towards test construction. This hypothesis was tested using the serial mediation analysis, specifically, model 6 of Hayes' PROCESS. This model uses 10000 bootstrap samples because the mediator variables are two, and are causal (one predicts the other). The predictor variable, teachers' test construction skills was measured on continuous basis. The mediator variables were commitment to work and attitude towards test construction, which were all measured on continuous basis. The criterion variable was adherence to test construction principles. This hypothesis basically tested the direct effect of test construction skills (X) on adherence to principles of test construction (Y). In addition, three indirect effects were tested. The indirect effect is the influence or effect of test construction skills on adherence to test construction principles through either commitment to work or attitude towards test construction or both (thus, mediators). Table 19 presents details of the results.

From Table 19, the model for test construction skills and commitment to work was not statistically significant, $F(1, 333) = .16$, $p = .692$, $R^2 = .001$. Test construction skills was, therefore, not a significant predictor of commitment to work, B = .07, *Boot 95%CI* (-.27, .39).

Table 19- Regression Coefficients for Test Construction Skills, Commitment to Work, and Attitude towards Test Construction

Model	Variable	B	BSE	CR	$BLLCI$	$BULCI$	Model Summary				
							R^2	F	df1	df2	p
One	Constant	48.06	2.55	18.85	43.18	53.16	.001	.16	1	333	.692
	Test construction skills	.07	.17	.41	-.27	.39					
Two	Constant	42.69	4.86	8.78	32.84	52.00	.15	30.41	2	332	<.001
	Test construction skills	1.42*	.24	5.92	.94	1.91					
	Commitment to work	.37*	.08	4.63	.22	.52					
Three	Constant	29.46	8.63	3.41	12.10	45.95	.29	44.48	3	331	<.001
	Test construction skills	2.84*	.35	8.11	2.16	3.51					
	Commitment to work	-.01	.09	-.11	-.18	.16					
	Attitude towards test construction	.26*	.08	3.25	.10	.43					

Criterion: Model 1- Commitment to work; Model 2- Attitude towards test construction; Mode 3- Adherence to test construction principles
*Significant, $p < .05$ level

In model two, both test construction skills, B = 1.42, *Boot 95%CI* (.94, 1.91); and commitment to work, B = .37, *Boot 95%CI* (.22, .52) were significant predictors of attitude towards test construction (see Table 19). In model three, test construction skills, B = 2.84, *Boot 95%CI* (2.16, 3.51) and attitude towards test construction, B = .26, *Boot 95%CI* (.10, .43) were significant predictors of adherence to test construction principles. In contrast, commitment to work, B = -.01, *Boot 95%CI* (-.18, .16) was not a significant predictor of adherence to test construction principles. Details of the indirect effects (mediation analysis) are presented in Table 20.

As shown in Table 20, the direct effect of test construction skills on adherence to test construction principles was statistically significant, B = 2.84, *Boot 95%CI* (2.23, 3.46). This implies that without any of the mediator variables (commitment to work or attitude towards test construction), teachers' test construction skills solely predicted adherence to test construction principles by 2.84 for an increase in test construction skills by a score of 1. This suggests that as teachers' skills in test construction is enhanced, their level of adherence to test construction principles will increase.

Further, when the two mediators (commitment to work and attitude towards test construction) were introduced into the relationship, as shown in indirect effect 3, the relationship was not statistically significant, B = .01, *Boot 95%CI* (-.03, .04). This implies that commitment to work and attitude towards test construction do not serially mediate the relationship between test construction skills and adherence to test construction principles.

Table 20- *Indirect Effect, Direct Effect, and Total Effect of Test Construction Skills*

	Effect	BSE	CR	p	Confidence Interval Lower Limit	Upper Limit
Total effect of X on Y	3.22*	.30	10.63	<.001	2.62	3.81
Direct effect of X on Y	2.84*	.31	9.11	<.001	2.23	3.46
Indirect effect of X on Y	Effect	*BSE*	*BootLLCI*		*BootULCI*	
Total	.38*	.12	.15		.63	
Indirect effect 1	-.001	.02	-.03		.04	
Indirect effect 2	.37*	.12	.15		.62	
Indirect effect 3	.01	.02	-.03		.04	
Completely standardised indirect effect (c'_{cs})	c'_{cs}	*BSE*	*BootLLCI*		*BootULCI*	
Total	.06	.02	.02		.10	
Indirect effect 1	-.001	.003	-.01		.01	
Indirect effect 2	.06	.02	.02		.10	
Indirect effect 3	.001	.002	-.004		.01	

Completely standardised effect (c'_{cs}): Total effect = .50; Direct effect = .44
X-Test construction skills; Y-Adherence to test construction principles
*Significant, *p* < .05 level
Indirect effect key:
Indirect effect 1 – Test construction skills ➡ Commitment ➡ Adherence
Indirect effect 2 – Test construction skills ➡ Attitude ➡ Adherence
Indirect effect 3 – Skills ➡ Commitment ➡ Attitude ➡ Adherence

Based on the result, I failed to reject the null hypothesis that "Teachers' level of commitment to work and attitude towards test construction will not mediate (serial) the relationship between teachers' test construction skills and adherence to recommended principles of test construction".

Even though the result of this study (Hypothesis 4) led to a decision of failing to reject the null hypothesis, it provides other interesting results. These

results manifested in Indirect Effect 1 and Indirect Effect 2. In the case of Indirect Effect 1, there was no significant effect of commitment to work on the relationship between test construction skills and adherence to principles of test construction, B = -.001, *Boot 95%CI* (-.03, .04). This implies that commitment to work does not play any role as far as test construction skills and adherence to principles of test construction are concerned. That is, commitment does not explain the relationship between the two variables. In effect, teachers who have higher skills in test construction do not necessarily have to be committed to their work before they adhere to recommended principles in test construction.

It was further revealed (Indirect Effect 2) that attitude towards test construction significantly mediated the relationship between test construction skills and adherence to test construction principles, B = .37, *Boot 95%CI* (.15, .62). The implication of this result is that, attitude towards test construction explains the relationship between test construction skills and adherence to test construction principles. When all the effects (regression coefficients) were standardised, comparatively, the completely standardised effect for the direct effect (c'_{cs} = .44) was higher than completely standardised effect for indirect effect, c'_{cs} = .06, *Boot 95%CI* (.15, .62). This implies attitude of teachers towards test construction reduced the effect of test construction skills on adherence to test construction principles by .38. This result, therefore, suggests that irrespective of teachers' level of skills in test construction, their level of adherence to test construction principles depends on their attitude towards test construction.

Hypothesis 5

H0: Teachers' attitude towards test construction will not mediate the relationship between teachers' commitment to work and their level of skills in test construction.

H1: Teachers' attitude towards test construction will mediate the relationship between teachers' commitment to work and their level of skills in test construction.

The purpose of this hypothesis was to examine whether teachers' attitude towards test construction would explain the relationship between test construction skills and adherence to test construction principles. The simple mediation analysis with Hayes' PROCESS, with 5000 bootstrap samples was performed to test the hypothesis. The predictor variable was teachers' commitment to work, and this was measured on continuous basis. The criterion variable was test construction skills, which was also measured on continuous basis. Details of the results are presented in Tables 21 and 22.

From Table 21, the model for commitment to work and attitude towards test construction was statistically significant, $F(1, 333) = 23.76$, $p < .001$. Commitment to work was a significant predictor of attitude towards test construction, B = .38, *Boot 95%CI* (.23, .53). In model two, commitment to work did not significantly predict test construction skills, B = -.02, *Boot 95%CI* (-.05, .01), attitude towards test construction was, however, a significant predictor of test construction skills, B = .07, *Boot 95%CI* (.04, .09).

Table 21- Regression Coefficients for Attitude towards Test Construction, Test Construction Skills, and Commitment to Work

Model	Variable	B	BSE	CR	$BLLCI$	$BULCI$	Model Summary				
							R^2	F	df1	df2	p
One	Constant	63.68	4.03	15.80	55.73	71.42	.07	23.76	1	333	<.001
	Commitment to work	.38*	.08	4.75	.23	.53					
Two	Constant	10.55	1.19	8.87	8.13	12.83	.10	17.41	2	332	<.001
	Commitment to work	-.02	.02	-1.00	-.05	.01					
	Attitude towards test construction	.07*	.01	7.00	.05	.09	.001	.16	1	333	.692
Three	Constant	14.79	.84	17.61	13.15	16.44					
	Commitment to work	.01	.02	.50	-.03	.04					

Criterion: Model 1- Attitude towards test Construction; Model 2- Test Construction Skills; Mode 3- Test Construction Skills
*Significant, $p < .05$ level

135

In model three, commitment to work was not a significant predictor of test construction skills, B = .01, *Boot 95%CI* (-.03, .04, see Table 21). Table 22 presents the indirect effect (mediation).

Table 22- *Indirect Effect, Direct Effect, and Total Effect of Commitment to Work*

	Effect	BSE	CR	*p*	Lower Limit	Upper Limit
					Confidence Interval	
Total effect of X on Y	.01	.02	.50	.692	-.03	.04
Direct effect of X on Y	-.02	.02	-1.12	.265	-.05	.01
Indirect effect of X on Y	Effect	*BSE*	*BootLLCI*		*BootULCI*	
Attitude towards Test Construction	.03*	.01	.01		.04	
Completely standardised indirect effect (c′$_{cs}$)	c′$_{cs}$	*BSE*	*BootLLCI*		*BootULCI*	
Attitude towards Test Construction	.08	.02	.05		.12	

Completely standardised effect (c′$_{cs}$): Total effect = .02; Direct effect = -.06
X-Commitment to Work: Y- Test construction skills
*Significant, *p* < .05 level

As presented in Table 22, the direct effect of commitment to work on test construction skills was not statistically significant, B = -.02, *Boot95%CI* (-.05, .01). This implies that teachers who are committed to their work would not necessarily have higher or lower skills in test construction. On the contrary, the effect of commitment to work on test construction skills was statistically significant, B = .03, *Boot95%CI* (.01, .04) when attitude towards test construction was mediated in the relationship. Comparatively, the completely standardised direct effect (c′$_{cs}$ = -.06), was negative and not significant.

136

However, upon the introduction of a third variable, that is, attitude towards test construction, the mediator variable, the completely standardised indirect effect became positive (c'_{cs} = .08). This, in effect, implies that irrespective of teachers' level of commitment to work, their attitude towards test construction determines their level of skills in test construction. This result suggests that commitment to work alone does not determine teachers' level of skills in test construction. However, attitude towards test construction explains why teachers, irrespective of their level of commitment to work, would have higher skills in test construction. Based on the result, the null hypothesis that "Teachers' attitude towards test construction will not mediate the relationship between teachers' commitment to work and their level of skills in test construction" is rejected in favour of its alternative hypothesis.

The implication of this result is that, on one hand, teachers who are highly committed to their work would not necessarily have higher skills in test construction. These teachers are not likely to learn more in terms of the rudiments of test construction. On the other hand, teachers who have positive attitude towards test construction are more likely to have higher skills in test construction. This result emphasises the essence of teachers' attitude as far as commitment to work and test construction skills are concerned. Generally, teachers who have positive attitude towards test construction are in a better position to learn more about test construction and its modalities, which in turn lead to improvement in test construction skills.

Hypothesis 6

H$_0$: The data collected will not fit the proposed model.

H$_1$: The data collected will fit the proposed model.

The aim of this hypothesis was to determine whether the model proposed (serial model) would fit the data collected. Thus, the hypothesis sought to determine the discrepancy between the hypothesised model (conceptual framework) and the actual model predicted (results obtained). This would determine whether the data gathered on the variables would reproduce the hypothesised model. This hypothesis was tested using SEM with Analysis of Moment Structures (AMOS), a covariance-based approach. Model fit indices such as NFI, TLI, CFI, PNFI, PCFI, among others were estimated. Details of the results are presented in Figures 8 and 9, and Table 23.

Table 23- *Model Fit Indices*

Fit Measure	Range	Index	Recommended cut-off
Absolute fit indices			
Chi-square (χ^2)		0	
Incremental fit indices			
Normed Fit Index (NFI)	0 – 1	1.0	.90 or greater
Incremental Fit Index (IFI)	0 – 1	1.0	.90 or greater
Tucker-Lewis Index (TLI)	0 – 1	1.0	.90 or greater
Comparative Fit Index (CFI)	0 – 1	1.0	.90 or greater
Parsimony-Adjusted Measures			
Parsimony Normed Fit Index (PNFI)	0 - 1	0	
F minimum (FMIN)	0 - 1	0	

Model 1: NPAR = 14; Non-redundant observations = 14; *df* = 0
Model 2: NPAR = 6, Non-redundant observations = 6; *df* = 0

138

From Table 23, the chi-square $(\chi^2) = 0$, implies that there is no difference between the expected data and the actual data. In terms of chi-square, large values suggest poor fit. In addition, the NFI, TLI, CFI, and IFI indices for both models were equal to 1.0, respectively, that indicates perfect fit of data and the model. For NFI, TLI, CFI, and IFI, smaller values suggest good fit. Further, the degrees of freedoms (*df*) for Model 1 and Model 2 are equal to 0. The implication of this is that both models are saturated model. Saturated models perfectly fit or reproduce the data. In sum, the null hypothesis that "The data collected will not fit the proposed model" is rejected in favour of the alternative hypothesis, therefore the model fit the data. Figures 8 and 9 present the models.

Figure 8- Model 1

139
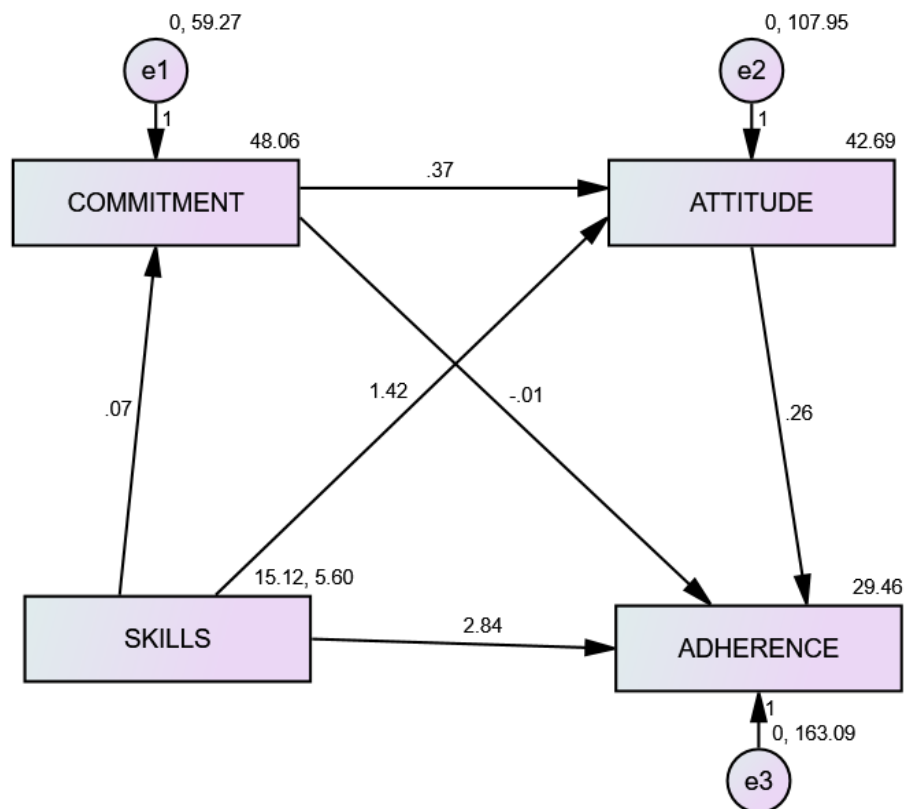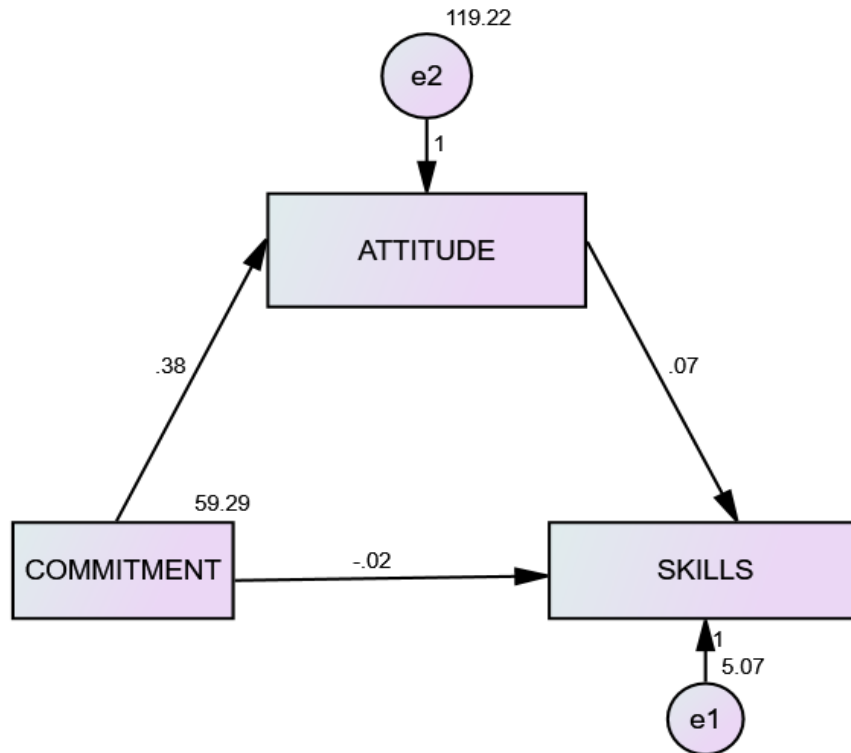
Figure 9- Model 2

The results for Model 1 and Model 2 suggest a perfect fit model which precisely reproduces the data. By implication, these models can be tested with other samples in order to validate their usefulness.

Discussion

This section discusses the results of the study. The discussion was organised under the following topical issues:

a.   teachers' test construction skills;

b.   teachers' commitment to work;

c.   teachers' attitude towards test construction; and

140

> d. test construction skills, commitment to work, and teachers' attitude towards test construction as predictors of adherence to test construction principles.

Teachers' Test Construction Skills

The study revealed that teachers had high skills in test construction. More than half of the respondents had a score of at least 14.4 out of 18 in test construction. This result has implications on teachers' assessment of students' learning, and the effectiveness of teaching and learning, at large. High skills in test construction puts teachers in a better position to deliver diligently up to task. With this, teachers are capable of engaging in proper assessment of students' learning. Teaching/learning and assessment are complementary ways of describing the same activity. That is to say, teaching without assessment is not effective, likewise, assessment without teaching is baseless, thus, teachers assess what they have taught. Clearly, high skills of teachers in test construction presupposes that teaching and learning are effective, even though that was not the position of this current study, and also not within its mandates. As indicated earlier, for classroom teachers, assessment provides them with much information tailored to the entire teaching and learning activities. Through assessment, teachers are able to determine whether a particular lesson has to be retaught, probably, with same or different method. Information on assessment provides feedback to students on which segments/topics they have mastered, and where they are deficient.

Teachers having high skills in test construction also implies that they are knowledgeable in test construction. By so doing, teachers know that in developing test items, they should make sure that both upper and lower performing groups would derive the same meaning from the test items. In addition, the test items should only discriminate abilities in terms of those who are knowledgeable and those who are not knowledgeable on the item. This means that test items should be devoid of defects and also giving clues to testees. Teachers with high skills in test construction would have in mind that test items they develop should not overlap in the alternatives for multiple choice items, and several other principles, among others.

The findings of this study, that teachers have high skills in test construction is in harmony with a couple of studies (Afemikhe & Imobekhai, 2016; Adamu et al., 2015; Agu et al., 2013). The findings of the aforementioned studies equally found high level of test construction skills among teachers in Nigeria, and even though their approach was quite different from the approach of the current study. The previous authors used Likert-type scales in measuring teachers' skills in test construction. In strict technical sense, a higher score on such a Likert-type scale depicts what teachers will do rather than what they can do. That is to say, in the case of Adamu et al., for example, the teachers strongly agreed to all the 26 items on test construction. In principle, agreement to such statements, in any way, does not tell what teachers know as far as test construction is concerned. Perhaps, agreement to such statements could be teachers' perception or opinion, which are not necessarily their skills in test

construction. The finding of the current study is unique as it clearly articulates what teachers can do in terms of test construction.

The finding of the current study, however, disconfirms some earlier studies (Quansah et al., 2019; Marmah & Impraim, 2013; Koskal, 2004; Hamman-Tukur & Kamis, 2000). The authors found that teachers do not have adequate skills in test construction. Interestingly, the authors, in their quest to investigate test construction skills, evaluated samples of test developed by teachers. Indeed, one obviously agrees to such an approach as it directly deals with work samples of the teachers. The issues, however, are: first, whether those tests were truly developed by the teachers solely and not in collaboration with other teachers; and two, were the questions collated from past questions which have duly been crafted? In my candid opinion, such approaches, though appear promising, there are other factors which could explain results away from such approaches. In contrast, with the finding of the current study, teachers' skills were assessed and described appropriately. I do not, in any way, dispute the findings of the previous studies, I, however, evaluated their findings within the context of their study. This study, in a way, addresses the inefficiencies in assessing test construction skills.

Teachers' Commitment to Work

One of the sub-purposes of this study was to describe the level of commitment among teachers. The result of the study revealed that teachers were committed to their work. It was, noted, however, that teachers' commitment was not so high. The fact that teachers are committed to their work could imply that

143

they would be up to the task as teachers. Generally, committed teachers are more likely find their values to be similar to the values of their schools. Paramount among the sub-dimensions of commitment to work was commitment to students. Teachers who are committed to their students have the responsibility to ensure good relations among their students. They also tend to have the students' interest at heart, and they enjoy teaching.

The finding of this study agrees with a couple of studies (Adu-Gyamfi et al., 2017; Butucha, 2013; Shoaib & Khalid, 2017; Swarnalatha, 2016). The findings of Adu-Gyamfi et al. are, however, not surprising as their study used non-teaching university staff. Generally, in Ghana, university staff and those in the secondary schools happen to be engaged by different bodies. University staff are engaged by the universities while teachers in secondary school are engaged by the Ghana Education Service; and for that matter, the conditions of service for both categories of employees are different. It was observed that staff of the universities were committed to their work. The current study also found secondary school teachers to be committed to their work. What could probably 'cause' teachers in Sekondi/Takoradi Metropolis to be committed to their work? It could probably be that the teachers are enjoying some incentives, or probably they are satisfied with their work as teachers. The finding of this study was coherent with Shoaib and Khalid (2017), who found teachers in their study to be generally committed in Punjab.

In contrast with the findings of this study, Hussen et al. (2016) found low commitment among teachers in Ethiopia. The possible causes of the low

144

commitment, among others, included low salary, lack of respect, poor attitude towards the teaching profession, and lack of incentives and motivation.

Indeed, when teachers are committed, they would do whatever it takes to make their job effective. From this, it can be said that teachers in Sekondi/Takoradi Metropolis would be more efficacious and provide more service to their students, by way of engaging them in effective teaching and learning activities. Based on the finding of this study, I would say teachers are committed because most of them had taught for quite a long time (about 11 years or more). It could also be that the teachers had the interest of their students at heart.

Teachers' Attitude towards Test Construction

The current study investigated the attitude of teachers towards test construction, and it was evident in its finding that teachers have a positive attitude towards test construction. Teachers' attitude towards test planning was paramount compared to the other dimensions of attitude towards test construction. Teachers having a positive attitude towards test construction implies that they do not see test construction as burdensome, and in this case they are more likely to adhere to recommended principles of test construction. It was revealed that majority of the teachers had taken a course in assessment suggesting that this could probably explain why teachers have positive attitude towards test construction.

This finding is consistent with Fakeye (2016) whose study revealed that teachers have a positive attitude towards formative assessment. The current

study, therefore found a positive attitude among teachers. Generally, classroom teachers are primarily engaged in developing tests for both formative and summative purposes. This, therefore, suggests that if teachers have positive attitude towards formative assessment, then there is the possibility that they may equally have positive attitude towards achievement test construction. In a similar study, Awofala and Babatunde (2013) found out that majority of respondents in their study had positive attitude towards continuous assessment. However, pre-service teachers were used for the study. Pre-service teachers are not fully engaged teachers and for that matter they having a positive attitude is not surprising. In addition, Bramwell-Lalor and Rainford (2016) found teachers have a positive attitude towards assessment in general. It must be noted that the authors did not assess teachers' attitude towards the test construction aspect of assessment. Merely showing positive attitude in assessment in general does not necessarily mean that teachers will have positive attitude on all components of assessment. There is the possibility that when teachers were assessed on other aspects of assessment, their attitude may change.

In contrast, Quansah and Amoako (2018) found a negative attitude towards test construction among SHS teachers in the Cape Coast Metropolis. Teachers, in theirs study, had negative attitude on all the dimensions of testing, namely, test planning, item construction, items review, and item assembling. The current study, however, revealed positive attitude on the aforementioned dimensions. Variations in the results of these two studies could emanate from the fact that each study was conducted in different regions in Ghana. This is

146

because even though the teachers appear to be similar in some characteristics, workshops are organised differently based on regions and districts. There is the possibility that teachers in Sekondi-Takoradi Metropolis receive, on regular bases, in-service training and workshops on assessment, hence their positive attitude towards test construction. It must be noted that the foregoing studies, apart from Quansah and Amoako, were conducted outside Ghana, therefore, contextual differences could also account for some variations in the studies.

Test Construction Skills, Commitment to Work, and Teachers' Attitude towards Test Construction as Predictors of Adherence to Test Construction Principles

This part discusses the results on the relationships among test construction skills, commitment to work, attitude towards test construction, and adherence to test construction principles. First, the result revealed that test construction skills is a significant predictor of adherence to test construction principles. Test construction skills, which positively predicted adherence to test construction principles, explained about a quarter of the variation in adherence to test construction principles. The result lead to the rejection of the null hypothesis that "Teachers' level of skills in test construction will not predict their adherence to recommended principles of test construction". The implication of the result is that, as teachers' skills in test construction improves, their level of adherence to test construction principles would rise. Higher test construction skill means teachers are more knowledgeable in the rudiments and dynamics of test construction. Generally, it is expected that teachers who are knowledgeable in test construction would follow the recommended principles,

since they know what steps they have to go through. Heads and other stakeholders such as the Sekondi/Takoradi Metropolitan Directorate of Education, are entreated to intensify their workshop activities, and stress much on issues on test construction, since doing that would act as refresher course for teachers. In addition, it would serve as knowledge impartation for non-professional teachers. The result generally underscores the relevance of test construction skills in order for teachers to adhere to recommended test construction principles.

Zhang and Burry-Stock (2003) found a significant positive relationship between assessment skills and assessment practices among teachers. Assessment skills explained 50% of the variances in assessment practices. The finding of the current corroborates the finding of Zhang and Burry-Stock. The current study equally found a positive influence of test construction skills on adherence to principles of test construction. In the current study, 25% of the variations in adherence to test construction principles was explained by test construction skills. Even though results from these studies are similar, the former investigated assessment in general, whereas the latter investigated specifically the test construction aspect of assessment. Again, both studies were conducted in different contexts. While the current study was conducted in Ghana, Zhang and Burry-Stock conducted their study in the US.

In contrast with the finding of this study, David (2010) found no significant relationship between teachers' knowledge and quality of their teacher-made test. This contradiction could probably emanate from the fact that

148

the foregoing study examined already made tests. There is a possibility that the samples of tests used were copied from already made textbooks or past questions, hence the high content-related validity and high level of conformity to guidelines.

This study also revealed that commitment to teaching and commitment to profession were the only significant predictors of adherence to test construction. Commitment to teaching was positive predictor of adherence to test construction principles, while commitment to the teaching profession was a negative predictor. The results imply that an increase in teachers' commitment to teaching would lead to an increase in their level of adherence to test construction principles, while controlling for influence of the other dimensions. Teachers who are highly committed to their teaching are more likely to adhere to recommended principles of test construction. Teachers who are committed to teaching means they take seriously activities which are directly involved in their classroom teaching. On the contrary, an increase in teachers' commitment to profession, while controlling for influence of the other dimensions, would lead to a decrease in adherence to test construction principles. This implies that as teachers become more committed to activities which, generally, have to do with their profession, they are not likely to adhere to test construction principles. This result sounds so weird, but it could possibly mean that when teachers are more committed to professional activities in teaching such as high engagement in GNAT or NAGRAT activities, among others, they become preoccupied with their activities at the expense of the core mandate of teaching in the classroom.

In this instance, engagement in professional association activities could hinder effective deliverance of core mandates. The results of the study has provided evidence to reject the null hypothesis that "Teachers' level of commitment to work will not predict their adherence to recommended principles of test construction" in favour of the alternative hypothesis "Teachers' level of commitment to work will predict their adherence to recommended principles of test construction".

The finding of this study that commitment to work predicts adherence to test construction, supports Shukla's (2014) finding that teachers who are committed are more likely to be competent in executing their duties. This is because such teachers may strive to enrich themselves with information that would enable them execute their duties diligently. Consequently, these teachers would follow recommended principles in test construction, and this, in the end would help improve the quality of their test items.

It was further revealed that one-fifth of the variance in adherence to principles of test construction was explained by all the dimensions of attitude towards test construction. Among all the dimensions of attitude towards test construction, attitude towards test planning and attitude towards item review were the only significant positive and negative predictors of adherence to principles of test construction respectively. For, an additional increase in teachers' attitude towards test planning would lead to 1.03 increase in adherence to test construction principles, while controlling for influence of the other dimensions. This implies that as teachers develop more positive attitude towards

150

test construction, their level of adherence would increase – thus, positive attitude towards test planning would lead to high adherence. On the contrary, a unit increase in attitude towards item review, would lead to 1.18 decrease in adherence to test construction principles. The result implies that teachers with positive attitude towards item review would not adhere to recommended test construction principles. This result could be explained from the fact that respondents attitude towards item review was not so strong, even though it was positive. Based on the results of this study, the null hypothesis which states that: "Teachers' attitude towards test construction will not predict their adherence to recommended principles of test construction" was rejected in favour of the alternative hypothesis.

Kitiashvili (2014) found that teachers' attitudes are not related to their willingness toward using such assessment approaches in their classes to some extent which indicates lack of correspondence between teachers' attitudes and their assessment practices. This result is in contravention with the finding of the current study. The difference in these results could be due to difference in the specific focus of these studies. While Kitiashvili focused on assessment in general, the current study was on test construction. It is possible that teachers may have a positive attitude towards item construction, but not test scoring or administration.

The study also examined whether teachers' test construction skills will indirectly influence teachers' level of commitment to work, which will influence their attitude towards test construction, and finally, on their level of

adherence. It was revealed that commitment to work and attitude towards test construction do not serially mediate the relationship between test construction skills and adherence to test construction principles. However, there was a direct relationship between test construction skills and adherence to test construction principles. I would, therefore say that, teachers with higher skills in test construction would adhere to test construction principles. Teachers, however, do not necessarily have to be committed to their work, which would in turn influence their attitude towards test construction, then finally to adherence to test construction principles. Simply put, commitment and attitude concurrently are not requirements for teachers to adhere to test construction principles. This means that commitment and attitude concurrently do not explain the relationship between teachers' test construction skills and adherence to test construction skills. Based on the result, I failed to reject the null hypothesis that "Teachers' level of commitment to work and attitude towards test construction will not mediate (serial) the relationship between teachers' test construction skills and adherence to recommended principles of test construction".

The finding of this study agrees with Shukla (2014), who found no relationship between commitment and competency in assessment. Based on the result, Shukla concluded that teachers who are competent need not be equally committed before they engage in better practices in assessment. This implies that commitment is not a requirement for better assessment practices.

Aside the non-rejection of the null hypothesis, interestingly, the result revealed no significant effect of commitment to work on the relationship

between test construction skills and adherence to principles of test construction. This implies that teachers who have higher skills in test construction do not necessarily have to be committed to their work before they adhere to recommended principles in test construction. This could explain why serial mediation was not significant. On the contrary, attitude towards test construction significantly mediated the relationship between test construction skills and adherence to test construction principles. This means that attitude towards test construction explains the relationship between test construction skills and adherence to test construction principles; that is it reduced the effect of test construction skills on adherence to test construction principles by .38 (see Table 20). In effect, attitude towards test construction plays a significant role as far as adherence to test construction principle is concerned. The result of this study underscores the need for teachers to have positive attitude towards test construction. This result has implications for teachers and other educational stakeholders such as the Ministry of Education, Ghana (MoE), Ghana Education Service (GES), and heads of the various SHSs involved in this study.

The result of this study provided evidence in support of the proposed model. It was revealed that the data collected perfectly reflected the proposed model. There was no discrepancy between the actual model and the proposed model. Based on this, I can confidently say, from the data collected and analysed, the relationship between teachers' test construction skills and adherence to test construction principles can be explained by their attitude towards test construction. I, therefore, recommend that the issue of teachers'

153

attitude towards test construction should not be looked down upon, for better and effective assessment of students' learning. Effective teaching thrives on effective assessment of students' learning. There is the need to give critical attention to the attitude of teachers towards test construction.

Chapter Summary

The study examined predictors of adherence to test construction principles. The results revealed that teachers had high level of skills in test construction. Teachers were commitment to their work, even, though respondents' level of commitment to work was not very high. It was further revealed that teachers, generally, had a positive attitude towards test construction. Teachers' test construction skills was a significant positive predictor of adherence to test construction principles. Commitment to work, specifically, commitment to teaching and commitment to the profession were significant predictors of adherence to test construction principles. In addition, attitude towards test construction, specifically, attitude towards test planning, and item review were significant predictors of adherence to test construction principles. Teachers' commitment to work and attitude towards test construction do not serially mediate the relationship between test construction skills and adherence to test construction principles. Attitude towards test construction was found to be a significant mediator in the relationship between commitment to work and test construction skills. Finally, the proposed model perfectly fits the data collected. The model was saturated, and showed no discrepancy between the hypothesised model and the predicted model.

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

The purpose of this study was to examine teachers' test construction skills, commitment to work, and attitude towards test construction as predictors of adherence to test construction principles among SHS teachers in the Sekondi-Takoradi Metropolis. This chapter presents a summary of the study, conclusions made and recommendations made, therefrom.

Summary

Overview of the study

The study examined teachers' test construction skills, commitment to work, and attitude towards test construction as predictors of adherence to test construction principles among SHS teachers in the Sekondi-Takoradi Metropolis. The study was driven by six purposes which were translated into three research questions and six hypotheses. The study was carried out employing the descriptive survey design, specifically, the cross-sectional design. The population for this study comprised all teachers in public SHSs in Sekondi-Takoradi Metropolis, with a total of 998 teachers. The study, however, targeted teachers who teach mathematics, integrated science, and English language, who were 346 in number. Using the census method, all the 346 teachers were engaged in the study.

155

Questionnaire which comprised scales from other researchers were adapted and used for the study. These scales are namely, Test Construction Skills Scale, Commitment to Work Scale, Attitude towards Test Construction Scale, and Adherence to Test Construction Principles. The scales were validated using IRT and CFA with the help of IRTPRO and AMOS software, respectively. The final version of the questionnaire had internal consistency as follows: Test Construction Skills, KR20 = .65; Commitment to Work, α = .65; Attitude towards Test Construction was, α = .86; and Adherence to test construction Principles, α = .89.

The data collected were analysed using descriptive statistics – frequencies, percentages, mean and standard deviation; and inferential statistics Structural Equation Model (SEM) and mediation analysis with Hayes' PROCESS. The bootstrap approach was used for all the inferential analyses.

Key findings

The following findings emerged from the study:

1. Teachers used in the study had high level of skills in test construction. The teachers, on the average, had a score of 15 and above out of 18 items used in assessing their skills.

2. Teachers were committed to their work. However, teachers' level of commitment to work was not high.

3. Teachers, generally, had a positive attitude towards test construction. The extent of their attitude, though positive, was not so high.

4. Teachers' test construction skills was a significant positive predictor of adherence to test construction principles. Teachers' test construction skills explain about a quarter (25.3%) of the variances in adherence to test construction principles.

5. Commitment to work, specifically, commitment to teaching and commitment to the profession were significant predictors of adherence to test construction principles. The former positively predicted adherence to test construction principles, while the latter was a negative predictor. Commitment to student and commitment to school were not significant predictors of adherence to test construction principles. Jointly, all the four sub-dimensions of commitment explained a little about a tenth (11%) of the variances in adherence to test construction principles.

6. Attitude towards test construction, specifically, attitude towards test planning, and item review were significant predictors of adherence to test construction principles. While attitude towards test planning was a significant predictor, attitude towards item review was a negative predictor. In all, the four sub-dimensions of attitude towards test construction, altogether, contributed 20% to the variations in adherence to test construction principles.

7. Teachers' commitment to work and attitude towards test construction do not serially mediate the relationship between test construction skills and adherence to test construction principles. However, attitude towards

157

test construction solely mediate the relationship between test construction skills and adherence to test construction principles. Commitment to work, on the other hand does not solely mediate the relationship between test construction skills and adherence to test construction principles.

8. Attitude towards test construction was a significant mediator in the relationship between commitment to work and test construction skills. Attitude towards test construction had a positive effect on the relationship between the two variables.

The proposed model perfectly fits the data collected. The model was saturated, and showed no discrepancy between the hypothesised model and the predicted model. Based on this finding, I present my final model in Figure 10.
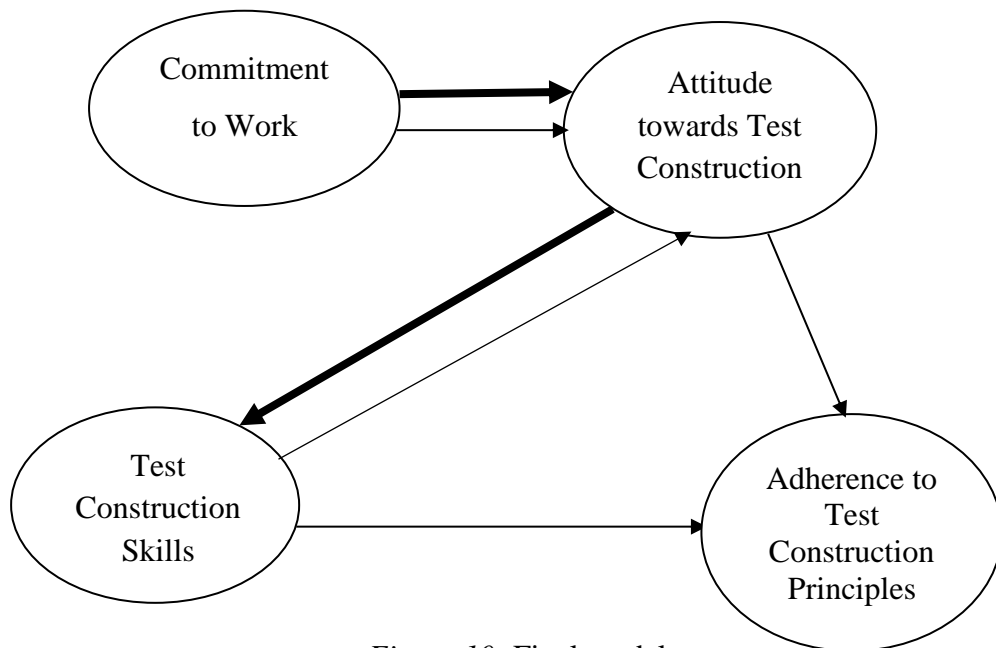

Figure 10- Final model

158

Conclusions

From the findings, it can be concluded that teachers had higher level of skills in test construction. This could possibly be due the educational background of teachers. It could also be that because teachers had some kind of positive attitude towards test construction, they had the zeal to learn more about the rudiments of test construction, and thereby, having high skills in test construction.

It can also be concluded that teachers do not necessarily have to be committed to their work before they have positive attitude towards test construction, which will in turn influence their adherence to test construction. However, a positive attitude towards test construction would solely influence teachers' adherence to principles of test construction. That is to say, commitment to work is not a requirement for teachers to adhere to principles of test construction. I would conclude that commitment to work could not explain why teachers with a particular level of skills in test construction would not adhere to recommended principles of test construction. Thus, even teachers who are not committed, but have a positive attitude towards test construction, would be motivated to learn more about test construction, which in the end lead to high adherence to recommended principles.

I further concluded that high teacher commitment to work, even though influences attitude towards test construction, it does not influence either test construction skills or adherence to test construction principles.

Recommendations

Following the findings of the study, the following are recommended:

1. The Ministry of Education, Ghana (MoE), Ghana Education Service (GES), and heads of the various SHSs in the Sekondi-Takoradi Metropolis as part of their training programmes or workshops for teachers should continue and intensify the acquisition of skills by teachers in test construction to improve their skills.

2. The Ministry of Education, Ghana (MoE), Ghana Education Service (GES), and heads of the various SHSs in the Sekondi-Takoradi Metropolis are encouraged to organise special workshops and symposia, which would sensitize teachers on the essence of developing quality test items, and this would at the end help teachers develop positive attitude towards test construction.

3. Teachers in the various SHSs in the Sekondi-Takoradi Metropolis are entreated to develop a strong positive attitude towards test construction; and also embrace it as an essential component of their core duties as teachers.

4. The heads of the various SHSs in the Sekondi-Takoradi Metropolis are encouraged to periodically assess teachers' attitude towards test construction, since this would enable the heads identify, and also offer the necessary assistance and guidance to the teachers in question.

5. As part of the process of recruiting teachers, the Ministry of Education, Ghana (MoE), Ghana Education Service (GES), and heads of the various

SHSs in the Sekondi-Takoradi Metropolis, are encouraged to assess teachers' level of commitment to the profession, since doing so can help identify potential teachers who are committed and would be more likely to give in their best.

6. The Ministry of Education, Ghana (MoE), Ghana Education Service (GES), and heads of the various SHSs in the Sekondi-Takoradi Metropolis are encouraged to enrol activities that would entice teachers to be highly committed to their work than before.

Suggestions for Further Research

1. Future study in this regard is recommended to assess teachers' commitment towards test construction, and determine its role as far the relationship between test construction skills and adherence are concerned.

2. It is recommended that this study be replicated in other regions in Ghana, since variations in some internal policies and workshops at the various districts and regions could bring about variations in results.

3. Further studies in this area is recommended to assess teachers skills in test construction by using the test construction skills scale, and in addition, observe these teachers while they develop test items, would establish criterion-related validity evidence.

161

REFERENCES

Adamu, G. G., Dawha, J. M., & Kamar, T. S. (2015). A scheme for assessing technical teachers' competencies in constructing assessment instruments in technical colleges in Gombe State. *ATBU, Journal of Science, Technology & Education (JOSTE), 3*(2), 1-8.

Adu-Gyamfi, M., Adu-Oppong, A. A., & Boahen, S. (2017). Job satisfaction and commitment among non-teaching staff of universities in Ghana. *International Journal of Innovative Research & Development, 6*(5), 54-64.

Afemikhe, O. A., & Imobekhai, S. Y. (2016). *Nigerian teachers' utilization of test construction procedures for validity improvement of achievement tests*. Unpublished paper, Institute of Education, University of Benin, Benin City, Nigeria.

Agu N. N., Onyekuba, C. & Anyichie, A. C. (2013). Measuring teachers' competencies in constructing classroom-based tests in Nigerian secondary schools: Need for a test construction skill. *Academic Journals, 8*(8), 431-439.

Ali, A. A. (1999). *Basic research skills in education*. Enugu: Orient Printing and Publishing.

Allen, N. J., & Meyer, J. P. (1993). The measurement and antecedents of affective, continuance and normative commitment to the organization. *Journal of Occupational Psychology, 63*, 1-18.

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

Amedahe, F. K. (1993). Test construction practices in secondary schools in Central Region of Ghana. *The Oguaa Educator, 2*, 52-63.

Amedahe, F. K. (2002). *Fundamentals of educational research methods.* Cape Coast: Mimeograph.

Amedahe, F. K. (2014). *The issue of falling educational standards in Ghana: A perception or reality?* Cape Coast: University of Cape Coast Press.

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

Amedahe, F. K., & Asamoah-Gyimah, K. (2015). *Introduction to educational research.* Cape Coast: UCC Printing Press.

Amedahe, F. K., & Asamoah-Gyimah, K. (2015a). *Introduction to measurement and evaluation.* Cape Coast: UCC Printing Press.

American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCMB). (1985). *Standards for educational and psychological testing.* Washington, D.C.: American Psychological Association.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999).

Standards for educational and psychological testing. Washington, DC: American Educational Research Association.

American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). *Standards for teacher competence in educational assessment of students*. Washington, DC: Author.

Anamuah-Mensah, J., & Quaigrain, K. A. (1998). Teacher competence in the use of essay test. *The Oguaa Educator University of Cape Coast, 12*, 31–42.

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

APA, AERA, & National Council on Measurements Used in Education. (1954). *Technical recommendations for psychological tests and diagnostic techniques*. Washington, D.C.: APA.

Asamoah-Gyimah, A. (2002). *An evaluation of the practice of continuous assessment in the senior secondary schools in the Ashanti Region of Ghana.* Unpublished Master's thesis, University of Cape Coast, Cape Coast, Ghana.

Awofala, A. O. A., & Babatunde, V. F. T. (2013). Examining attitude towards continuous assessment practices among Nigerian pre-service STM Teachers. *Journal of Education and Practice, 4*(13), 37-49.

Bailey, K. M. (2004). *Learning about language assessment: Dilemmas, decisions and directions*. Boston, MA: Heinle & Heinle Publishers.

Baker, F. B. (2001). *The basics of item response theory*. College Park, MD: ERIC Clearinghouse on Assessment and Evaluation. Original work published in 1985. Retrievedfrom http://echo.edres.org:8080/irt/baker/

Baker, J. O. (2003). *Testing in modern classrooms*. London: George Allen and Urwin Ltd.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.

Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7-74.

Bramwell-Lalor, S., & Rainford, M. (2016). Advanced level biology teachers' attitudes towards assessment and their engagement in assessment for learning. *European Journal of Science and Mathematics Education, 4*(3), 380-396.

Brooks, V. (2002). *Assessment in secondary schools: The new teacher's guide to monitoring, assessment, recording, reporting and accountability*. Buckingham: Open University Press.

Butucha, K. G. (2013). Teachers' perceived commitment as measured by age, gender and school type. *Greener Journal of Educational Research, 3*(8), 363-372.

Chan, K. K. (2009). *Using test blueprint in classroom assessments: why and how*. Paper presented at the 35th International Association for Educational Assessment (IAEA) Annual Conference, Brisbane, Australia.

Chau, T. N. (1996). *The quality of primary schools in different development contexts*. Paris: UNESCO Publishing International Institute for Educational Planning.

Childs, R. A. (1989). *Constructing classroom achievement tests*. ERIC Digest. ED 315426.

Cizek, G. J., Fitzgerald, S. M., & Rachor, R. E. (1996). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational Assessment, 3*, 159-179.

Crocker, L., & Algina, J. (2008). *Introduction to classical and modern test theory*. New York, NY: Cengage Learning.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed), *Educational measurement* (2nd ed.) (pp. 443-507). Washington, DC: American Council on Education.

Cronbach, L. J. (1988). Five perspectives on validity argument. In H. Wainer, H. Braun (Eds), *Test validity* (pp. 3-17). Hillsdale, NJ: Lawrence Erlbaum.

Cronbach, L. J., Meehl, P. E. (1955). Construct validity in psychological tests. *Psychol Bull, 52*, 281-302.

Cumming, A. (2001). ESL/EFL instructors' practices for writing assessment: Specific purposes or general purposes? *Language Testing, 18*, 207-224.

Cureton, E. E. (1951). Validity. In E. F. Lindquist (Ed.), *Educational measurement* (pp. 621–694). Washington, D.C.: American Council on Education.

D'Agostino, J. V. (2007). *Quantitative research, evaluation, and management section*. Ohio State University. 29 Woodruff Avenue, Columbus, OH 42310-1177. USA.

Darling-Hammond, L. (2010). *Performance counts: Assessment systems that support high-quality learning*. Council of Chief State School Officers, Washington DC.

David, R. (2010). Chemistry teachers' knowledge of test construction and quality of their tests in private secondary schools in Maiduguri, Borno State, Nigeria.

Day, F. (2004). *A passion for teaching.* Routledge Falmer.

Dumay, X., & Galand, B. (2012). The multilevel impact of transformational leadership on teacher commitment: Cognitive and motivational pathways. *British Educational Research Journal*, 38, 703-729.

Ebinye, P. O. (2001). Problems of testing under the continuous assessment Programme. *Journal of Quality Education, 1*(1), 12-19.

Education Up Close (2005). *Writing effective tests: A guide for teachers*. Retrieved on November 17, 2018 from http://www.glencoe.com/sec/ teachingtoday/phtml/print/40.

Esomonu, N. P. (2002). Assessment of non-cognitive behaviours in secondary schools for national development. In: R. C. Ebenebe, & L. R. Akudolu (Eds.). *Education for National Development and Integration*. Awka: John Best Publishing.

Fakeye, D. O. (2016). Secondary school teachers' and students' attitudes towards formative assessment and corrective feedback in English Language in Ibadan Metropolis. *Journal of Educational and Social Research, 6*(2), 141-148.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: SAGE Publications Ltd.

Firestone, W. A. (1996). Images of teaching and proposals for reform: A comparison of ideas from cognitive and organisational research. *Educational Administration Quarterly, 32*(2), 209-235.

Fornell, C. G., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed.). McGrawHiil.

Frey, B. B. (2007*). An introduction to quality test construction. Retrieved on October 19, 2007* from http://www.specislconnection.ku.edu/cgi bin/cgiwrap/speccom/mainph.

Frey, B. B. (2007). *An introduction to quality test construction*. Retrieved on November 19, 2018 from http://www.specislconnection.ku.edu/iwrap/

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Hair, J. F., Hult, G. T., Ringle, C. M., & Sarstedt, M. (2014). *A primer on partial least squares structural equation modeling (PLS-SEM)*. London: SAGE Publications, Inc.

Hamafyelto, R. S., Hamman-Tukur, A., & Hamafyelto, S. S. (2015). Assessing teacher competence in test construction and content validity of teacher made examination questions in commerce in Borno State, Nigeria. *Journal of Education, 5*(5), 123-128.

Hamman-Tukur, A., & Kamis, A. B. (2000). Content analysis of B.sc Bio-chemesrty examination questions implication for testing, teaching and development. *African Journal Research in Education, 1*, 59-62.

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). New York, NY: The Guildford Press.

Hussen, A. A., Tegegn, S. A. W., & Teshome, T. Z. (2016). Teachers' professional commitment towards students learning, their profession and the community in Eastern Ethiopian Secondary Schools. *Journal of Teacher Education and Educators, 5*(3), 289-314.

Jiraro, S., Sujiva, S., & Wongwanich, S. (2014). An application of action research for teacher empowerment to develop teachers' test construction competency development models. *Procedia - Social and Behavioral Sciences, 116*, 1263-1267.

Kane, M. T. (1992). An argument-based approach to validation. *Psychol Bull, 112*, 527-535.

Kane, M. T. (2001). Current concerns in validity theory. *J Educ Measure, 38*, 319-342.

Kellough, R. D., & Kellough, N. G. (1999). *Secondary school teaching: A guide to methods and resources; planning for competence*. Upper Saddle River, New Jersey: Merrill.

Khine, M. S. (2013). *Contemporary approaches to research in learning innovations: Application of structural equation modelling in educational research and practice*. Boston: Sense Publishers.

Kitiashvili, A. (2014). Teachers' attitudes toward assessment of student learning and teacher assessment practices in general educational institutions: The case of Georgia. *Improving Schools 17*(2), 163-175.

Koksal, D. (2004). Assessing teachers' testing skills in ELT and enhancing their professional development through distance learning on the net. *Turkish Journal of Distance Education. 5*(1), 1-11.

Kushman, J. W. (1992). The organizational dynamics of teacher workplace commitment: A study of urban elementary and middle schools. *Educational Administration Quarterly, 28*(1), 5-42.

Lambert, D., & Lines, D. (2000). *Understanding assessment: Purposes, perceptions, practice*. London, New York: Routledge Falmer.

Le-Compte, M. D., & Dworkin, A. G. (1991). *Giving up on school, student dropouts and teacher burnouts.* Newbury Park, CA: Corwin Press.

170

Lodahl, T. M., & Kejner, M. (1965). The definition and measurement of job involvement. *Journal of Applied Psychology, 49*, 24-33.

Magno, C. (2013). Standards of teacher competence on student assessment in the Philippines. *The Assessment Handbook, 10*, 42-53.

Malone, M. E. (2013). The essentials of assessment literacy: Contrast between testers and users. *Language Testing, 30*(3), 329-344.

Marmah, A. A., Impraim, A. K. (2013). University lecturer's competence in the construction of multiple choice test items: A case study of Coltek-Kumasi. *The International Journal of Humanities & Social Studies, 1*(4), 1-9.

Messick, S. (1984). The psychology of educational measurement. *J Educ Measure, 21*, 215-237.

Messick, S. (1989). Validity. In R. L. Linn (Ed), *Educational measurement* (3rd ed.) (pp. 3-104). New York, NY: American Council on Education and Macmillan.

Meyer, J. P., Allen, N. J., & Smith, C. A. (1993). Commitment to organizations and occupations: extension and test of a three-component conceptualization. *Journal of Applied Psychology, 78*(4), 538-551.

Miller, C. M. I., & Parlett, M. (1974). *Up to the mark: A study of the examination game*. Guildford: Society for Research into Higher Education.

Miller, M. D., Linn, R. L., & Gronlund, N. F. (2009). *Measurement and assessment in teaching* (10th ed.). Ohio: Pearson Education, inc.

Ministry of Education, Ghana (2017). *Nation Teachers' Standards for Ghana Guidelines*. Accra: Author.

Monsaas, J. A., & Engelhard, G. J. (1994). Teachers' attitudes toward testing practices. *Journal of Psychology, 128*(4), 469-477.

Morizot, J., Ainsworth, A. T., & Reise, S. P. (2007). Toward modern psychometrics: Application of item response theory models in personality research. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of Research Methods in Personality Psychology* (pp. 407-423). New York: Guilford.

Nazari, K., & Emami, M. (2012). Analysis of relation between organizational commitment and professional commitment. *Elixir Human Res. Mgmt., 51*, 10661-10664.

Nir, A. E. (2002). School-based management and its effect on teacher commitment. *International Journal of Leadership in Education, 5*, 323-341.

Nitko, J. A. (2001). *Educational assessment of students*. New Jersey: Prentice Hall.

Oduro-Kyireh, G. (2008). *Testing practices of senior secondary school teachers in the Ashanti Region of Ghana*. Unpublished master's thesis, University of Cape Coast, Cape Coast, Ghana.

Olffen, W., & Roe, R. A. (2008). Beyond the three-component model of organizational commitment. *Journal of Applied Psychology, 93*(1), 70–83.

Ololube, N. P. (2008). Evaluation competencies of professional and non-professional teachers in Nigeria. *Studies in Educational Evaluation (SEE), 34*(1), 44-51.

Onyechere, I. (2000). New face of examination malpractice among Nigerian youths. *The Guardian Newspaper July 16*.

Pallant, J. (2010). *A step by step guide to data analysis using the SPSS program: SPSS survival manual* (4th ed.). Crows Nest: Allen & Unwin.

Park, I. (2005). Teacher commitment and its effects on student achievement in American high schools. *Educational Research and Evaluation, 11*(5), 461-485.

Paulson, T. U. (2003). *The study of teaching*. New York: Holt, Rinehart and Winston Inc.

Pelligrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington DC: National Academic Press.

Quaigrain, A. K. (1992). *Teacher competence in the use of essay tests: A study of secondary schools in the Western Region of Ghana*. Unpublished Master's Thesis, University of Cape Coast, Ghana.

Quaigrain, A. K., & Arhin, A. K. (2017). Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation. *Cogent Education, 4,* 1-11.

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), 25-30.

Quansah, F., Amoako, I., & Ankomah, F. (2019). Teachers' test construction skills in Senior High Schools in Ghana: Document analysis. *International Journal of Assessment Tools in Education, 6*(1), 1– 8.

Sanders, R. J., Hills, J. R., Nitko, A. J., Merwin, J. C., Trice, C., Dianda, M., & Schneider, J. (1990). *Standards for teachers' competence in educational assessment of students*. American Federation of Teachers, National Council on Measurement in Education, and National Education Association.

Sekondi-Takoradi Metropolitan Directorate (2018). *Annual report on teachers*. Author: Sekondi Metropolitan Press.

Shoaib, H., & Khalid, M. I. (2017). Professional commitment of teacher educators: Future of nation builders. *Pakistan Vision, 18*(2), 163-177.

Shukla, S. (2014). Teaching competency, professional commitment and job satisfaction-a study of primary school teachers. *IOSR Journal of Research & Method in Education, 4*(3), 44-64.

Silker, R. T. (2003). *Teachers and tests*. London: Basil Blackwell.

Simon, G. M. (2002). *Testing to destruction: A problem in a small state*. Unpublished Seminar Paper.

Simpson, D., & Hood, F. (2000). Changing perception and practice through the professional development model. In M., Belok, & T. E. Deering (Eds.) *Teacher education*, Meerut: Anne Books.

Somech, A., & Bogler, R. (2002). Antecedents and consequences of teacher organizational and professional commitment. *Educational Administration Quarterly, 38*, 555-577.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology, 15*, 72-101.

Spearman, C. (1907). Demonstration of formulae for true measurement of correlation. *American Journal of Psychology, 18*, 161-169.

Spearman, C. (1913). Correlations of sums and differences. *British Journal of Psychology, 5*, 417-426.

Stoof, A., Martens, R. L., & Van Merrienboer, J. J. G. (2000). *What is competence? A constructivist approach as a way out of confusion*. Paper Presented at the Conference of the Dutch Educational Research, Leiden.

Swarnalatha, S. S. (2016), Work commitment of secondary school teachers. *International Journal of Indian Psychology, 3*(4), 84-89.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson Education.

Thien, L. M., Razak, N. A., Ramayah, T. (2014). Validating teacher commitment scale using a Malaysian sample. *Sage Open*, 1-9.

Thurstone, L. L. (1955). The criterion problem in personality research. *Educational and Psychological Measurement, 15,* 353–361.

Tyree, A. K. (1996). Conceptualising and measuring commitment to high school teaching. *Journal of Educational Research, 89*, 295-304.

United Nations Educational, Scientific and Cultural Organization [UNESCO] (Oct 2017 Edition) Unpacking Sustainable Development Goal (SDG4) Education 2030 Guide. Retrieved from http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E

VanAmelsvoot, G., Hendiks, M., & Scheerens, J. (2000). Selection and development of international indicators on Staffing. *Educational Economics, 8*(1), 17-36.

Vandenberg, R. J., & Scarpello, V. (1994). A longitudinal assessment of the determinant relationship between employee commitment to the occupation and the organization. *Journal of Organizational Behavior, 15*, 535-547.

Warr, P., Cook, J., & Wall, T. (1979). Scales for the measurement of some work attitudes and aspects of psychological wellbeing. *Journal of Occupational Psychology, 52*, 129-148.

Wiliam, D. (2014, April). *Formative assessment and contingency in the regulation of learning processes*. Paper presented in a Symposium entitled Toward a Theory of Classroom Assessment as the Regulation of Learning at the annual meeting of the American Educational Research Association, Philadelphia, PA.

Wright, A. (2001). The ABCs of assessment. *The Science Teacher, 68*(7), 60-64.

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

APENDICES

APPENDIX A

FINAL QUESTIONNAIRE

Questionnaire for Teachers

This questionnaire seeks to solicit information on your experiences and practices as a teacher. The information you provide in this questionnaire is strictly for academic purposes. The information provided will be treated as a group, and for that matter, you will not be associated with any of the responses. Participation in this study is not compulsory. Any information you provide will be kept anonymous and confidential. Please provide responses as frankly as possible.

Instruction: Please check (√) or write where necessary.

SECTION A – Demographic Information

1. Gender
   a. Male   [   ]
   b. Female   [   ]

2. Age
   a. 20 years & below [   ]
   b. 21 – 30 years   [   ]
   c. 31 – 40   years   [   ]
   d. 41 – 50 years   [   ]
   e. 51 – 60 years   [   ]
   f. Above 60 years   [   ]

3. Professional qualification
   a. Degree with education   [   ]
   b. Degree without education [   ]

4. Years of teaching experience…………………………………

SECTION B

Please read carefully the following statements and indicate to the best of your knowledge whether the statements are true or false.

No.	Items	True	False
	A teacher should take the following steps in constructing tests for his/her class		
1.	Write test items so that both high and low achievers can understand.		
2.	Avoid gender stereotypes in the test items.		
3.	Ascribe scores for each test item.		
4.	Avoid too long questions or phrases in item writing.		
5.	Outline the content covered for the term before setting test from them.		
6.	Add enough test items to cover all the requisite levels of cognitive domain.		
7.	Ensure that the items are measuring the determined objectives.		
8.	Set essay items that elicit creative and imaginative answers from the students.		
9.	Prepare a marking guide while constructing the test.		
10.	Add sufficient items to cover the appropriate instructional units.		
11.	Limit essay tests to high level objectives.		
12.	Organize test items in a logical manner.		
13.	Set tests with due regard to the time available for testing.		
14.	Avoid the use of clues in multiple choice questions.		
15.	Avoid the use of overlapping items.		
16.	Avoid overlapping alternatives in writing objective tests.		

180

17.	Prepare a test blueprint as a guide in the test construction.		
18.	Consider the class level of learners during item writing.		

SECTION C

Please read carefully the following statements and indicate the extent to which you agree or disagree to each of the following statements as they apply to you, using the following scale: **VSD - Very strongly disagree; SD - Strongly disagree; D – Disagree; A –Agree; SA - Strongly agree; VSA - Very strongly agree.**

No.	Items	VSD	SD	D	A	SA	VSA
---	**Commitment to school**						
1.	Often, I find it difficult to agree with school policies on important matters relating to teachers						
2.	I talk up about this school to my friends as a great school to work for						
3.	I find that my values and the values of the school are very similar						
---	**Commitment to students**						
4.	All students can succeed, and it is my mission to ensure their success						
5.	It is my responsibility to ensure good social relations among my students						

6.	I feel obliged to mediate among rival groups of students						
---	**Commitment to teaching**						
7.	Occasionally, I lie awake at night thinking about the next day's lesson						
8.	I enjoy teaching						
---	**Commitment to profession**						
9.	If I could get a job different from being a teacher and paying the same amount, I would take it						
10.	If I could do it all over again, I would not choose to work in the teaching profession						
11.	I am disappointed that I ever entered the teaching profession						
12.	Working as a teacher is the best decision I ever made						

SECTION D

Please read carefully the following statements and indicate the extent to which you agree or disagree to each of the following statements as they apply to you, using the following scale: **SD - Strongly disagree; D – Disagree; A –Agree; SA - Strongly agree**

No.	Items	SD	D	A	SA
---	**Planning**				
1.	To be honest, it is a waste of time trying to outline the purpose of a test when planning the test.				

2.	I just need my textbook to start writing test items.				
3.	I mostly do not prefer using test specification table in crafting questions.				
4.	I prefer to finish crafting the test before considering the thinking skills those items measure.				
5.	Since I am the classroom teacher, I do not need to specify the content area I want to test.				
6.	Planning a test is needless as a teacher.				
7.	I prefer writing items based on what learners are expected to know whether taught or not.				
8.	As a teacher, there is nothing wrong with crafting items without considering the learning objectives.				
9.	It is not possible to always craft new questions for learners.				
10.	Crafted items do not necessarily have to match learning objectives.				
11.	I like to write tricky questions to test my students understanding.				
12.	Arranging of the options to multiple-choice items logically is not compulsory				
---	**Item construction**				
13.	There is the need to take items verbatim from textbooks used in teaching.				
14.	I usually construct test items few days for the paper to be written.				
15.	It is optional to review constructed items before it is administered.				

183

16.	Checking for the item difficulty and discrimination after the test has been constructed is not necessary.				
17.	It is essential to present more difficult items before less difficult items in assembling crafted items				
18.	It is optional to number all the items on a test.				
19.	It is optional to provide clear directions for examinees on the test instrument.				
20.	It is right to arrange options of test items horizontally.				
---	**Item review**				
21.	It is better to rely on past questions when constructing a test.				
22.	I like to prepare marking scheme after the test have been administered.				
23.	It is necessary to check for the clarity of crafted items				
24.	I prefer preparing marking scheme two or more days after constructing the test.				
25.	I always like to arrange questions into sections based on their nature or type.				
---	**Item assembling**				
26.	I select questions from topics I think students have understood.				
27.	I think test specification table should be prepared by test experts and not the classroom teacher.				
28.	It is essential to identify behaviours to represent a construct when crafting test items.				

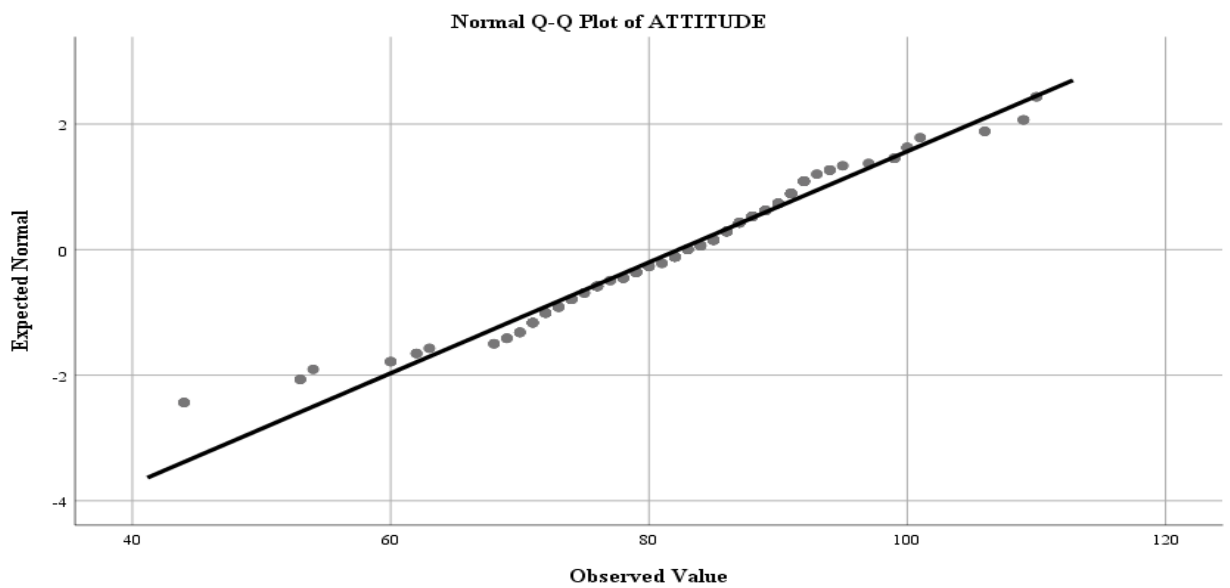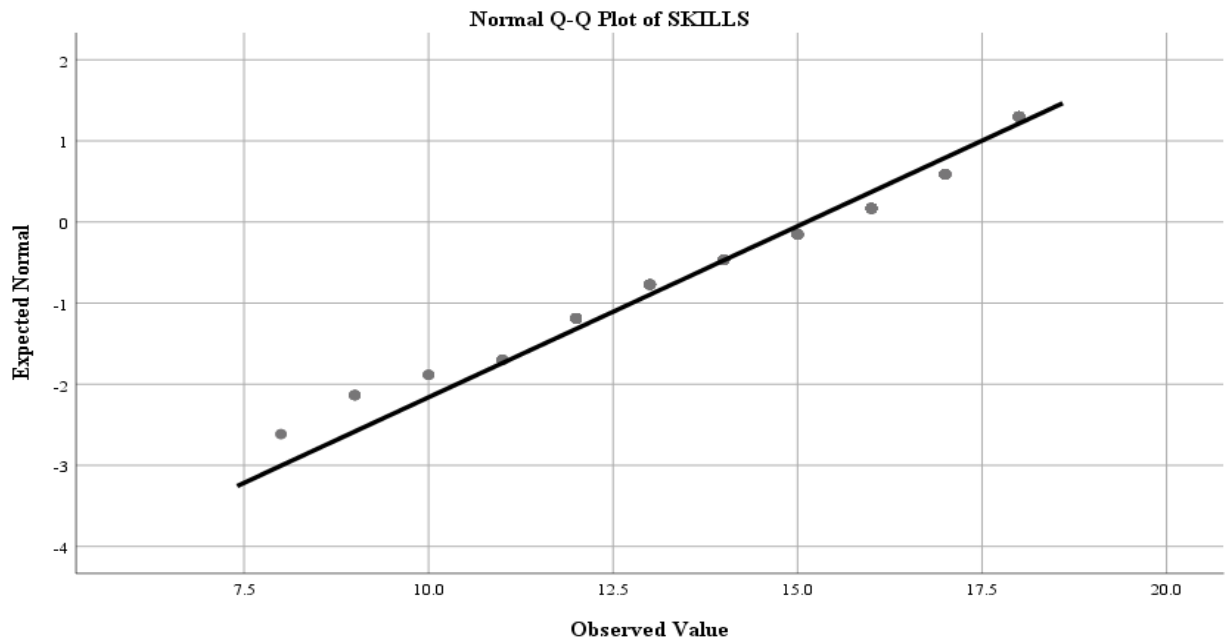| 29.	I do not think it's necessary to craft more items than actually needed.				

SECTION E

Please read carefully the following statements and indicate the extent to which each of the following statements applies to you, using the following scale: **N – Never; NO – Not often; O – Often; VO – Very often; A - Always**
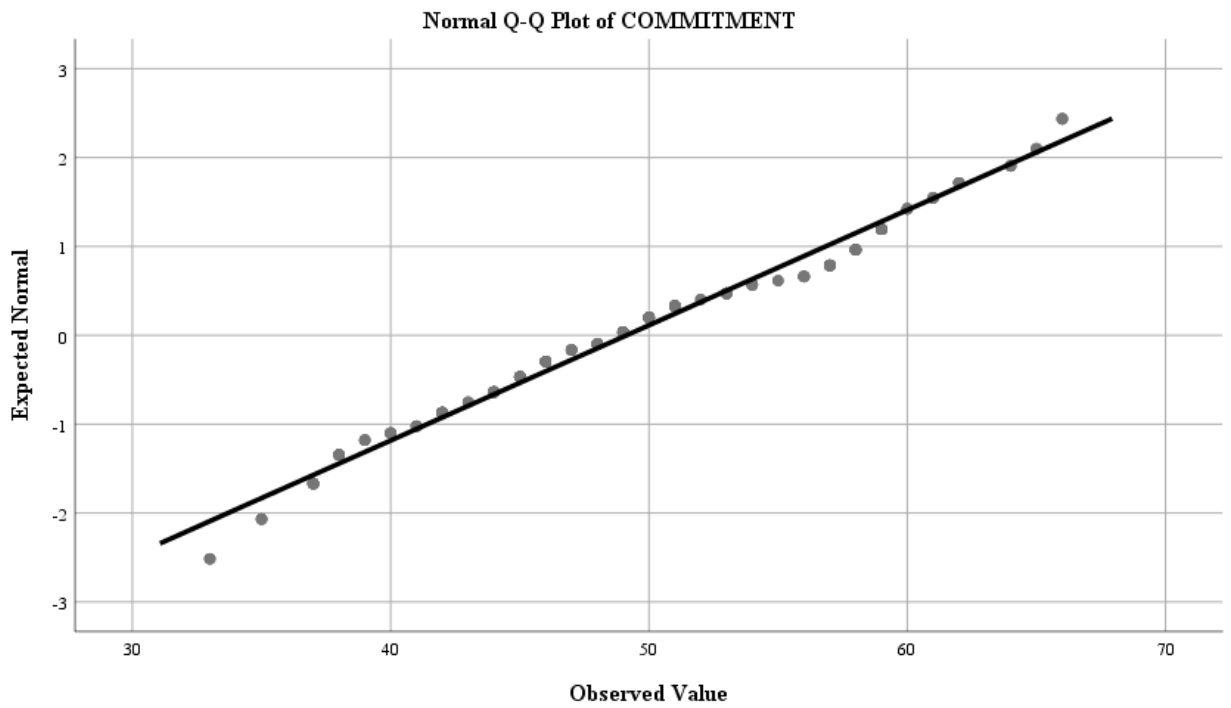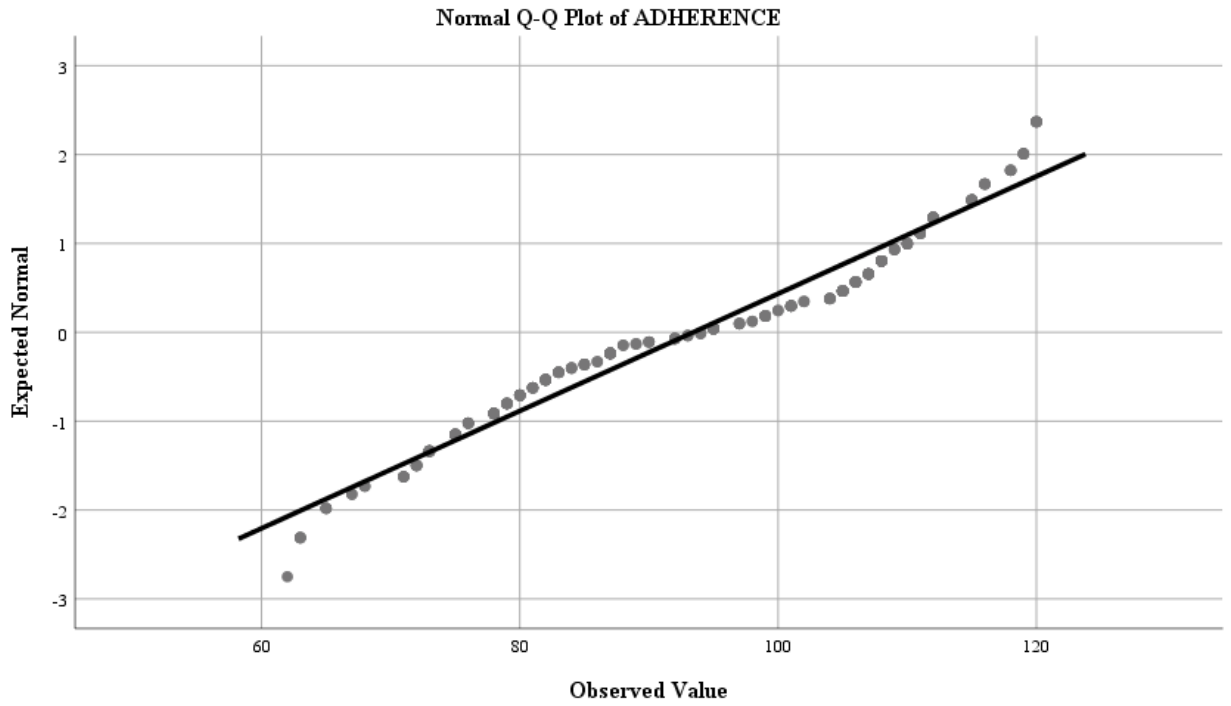
No.	Items	N	NO	O	VO	A
	I take the following steps in constructing tests for my class					
1.	Give clear instructions to guide the test takers.					
2.	Write test so that both high and low achievers can understand.					
3.	Avoid gender stereotypes in the test items.					
4.	Ascribe scores for each test item.					
5.	Avoid too long questions or phrases in item writing.					
6.	Outline the content covered for the term before setting test from them.					
7.	Add enough test items to cover all the requisite levels of cognitive domain.					
8.	Ensure that the items are measuring the determined objectives.					
9.	Set essay items that elicit creative and imaginative answers from the students.					
10.	Prepare a marking guide while constructing the test.					

185

11.	Add sufficient items to cover the appropriate instructional units.					
12.	Limit essay tests to high level objectives.					
13.	Organize test items in a logical manner.					
14.	Set tests with due regard to the time available for testing.					
15.	Avoid the use of clues in multiple choice questions.					
16.	Avoid the use of overlapping items.					
17.	Avoid overlapping alternatives in writing objective tests.					
18.	Prepare a test blueprint as a guide in the test construction.					
19.	Consult standard text books in the subject for guide.					
20.	Keep a resource bank of questions that can be used to when setting tests.					
21.	Consider the age of learners during item writing.					
22.	Submit items for vetting to the Head of Department or the principal.					
23.	Submit tests meant for promotional examinations for expert editing on time.					
24.	Review draft of the test at least twice in two days before administering.					

APPENDIX B

NORMALITY TEST

Normal Q-Q Plot of SKILLS

Normal Q-Q Plot of ATTITUDE

187

Normal Q-Q Plot of ADHERENCE

Normal Q-Q Plot of COMMITMENT

188

APPENDIX C

INTRODUCTORY 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

4th March, 2019

Our Ref:

Your Ref:

TO WHOM IT MAY CONCERN

Dear Sir/Madam,

THESIS WORK
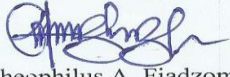**LETTER OF INTRODUCTION**
MR. FRANCIS ANKOMAH

We introduce to you Mr. Ankomah, 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 he is currently at the thesis stage.

Mr. Ankomah is researching on the topic:

"**Predictors of Adherence to Test Construction Principles among Senior High School Teacher.**"

He has opted to collect or gather data at your institution/establishment for his Thesis work. We would be most grateful if you could provide him the opportunity and assistance for the study. Any information provided would be treated strictly as confidential.

We sincerely appreciate your co-operation and assistance in this direction.

Thank you.

Yours faithfully,

Theophilus A. Fiadzomor
Senior Administrative Assistant
For: **HEAD**

189

APPENDIX D

ETHICAL CLEARANCE

UNIVERSITY OF CAPE COAST
COLLEGE OF EDUCATION STUDIES
ETHICAL REVIEW BOARD

UNIVERSITY POST OFFICE
CAPE COAST, GHANA

Our Ref: CES-ERB/ucc-edu/V3/19/02

Date: March 4, 2019

Your Ref: ....

Chairman, CES-ERB
Prof. J. A. Omotosho
jomotosho@ucc.edu.gh
0243784739

Vice-Chairman, CES-ERB
Prof. K. Edjah
kedjah@ucc.edu.gh
0244742357

Secretary, CES-ERB
Prof. Linda Dzama Forde
lforde@ucc.edu.gh
0244786680

Dear Sir/Madam,

ETHICAL REQUIREMENTS CLEARANCE FOR RESEARCH STUDY

The bearer, Francis Ankomah, Reg. No. EF/MEP/17/0006 is an M.Phil. / Ph.D. student in the Department of Education and Psychology .... in the College of Education Studies, University of Cape Coast, Cape Coast, Ghana. He / She wishes to undertake a research study on the topic:

Predictors of adherence to test construction principles: The case of Senior High School teachers in the Sekondi-Takoradi Metropolis.

The Ethical Review Board (ERB) of the College of Education Studies (CES) has assessed his/her proposal and confirm that the proposal satisfies the College's ethical requirements for the conduct of the study.

In view of the above, the researcher has been cleared and given approval to commence his/her study. The ERB would be grateful if you would give him/her the necessary assistance to facilitate the conduct of the said research.

Thank you.
Yours faithfully,

Prof. Linda Dzama Forde
(Secretary, CES-ERB)