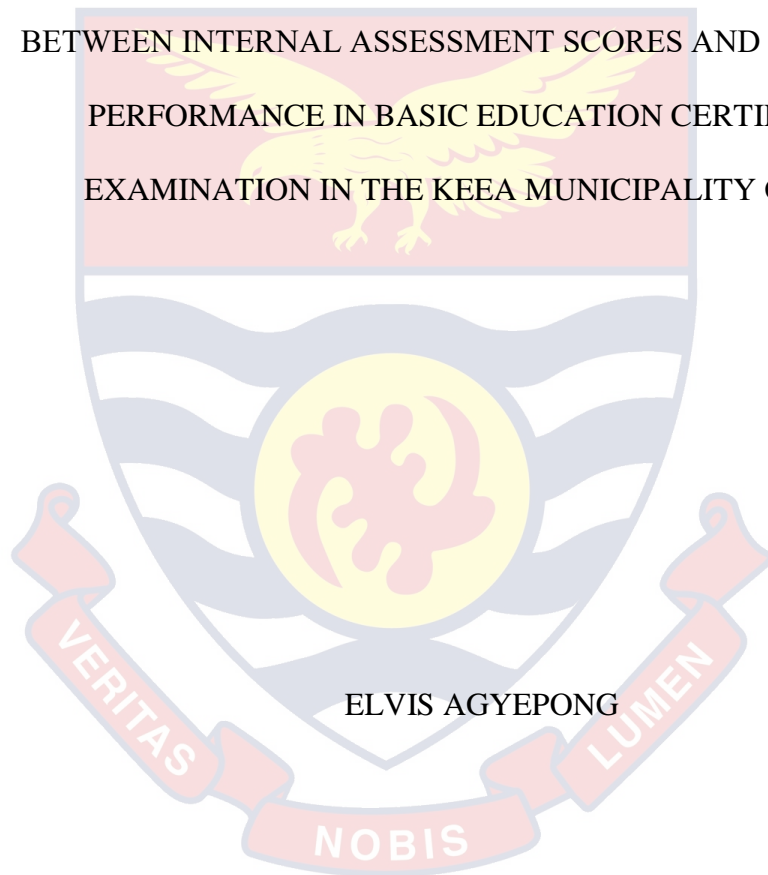


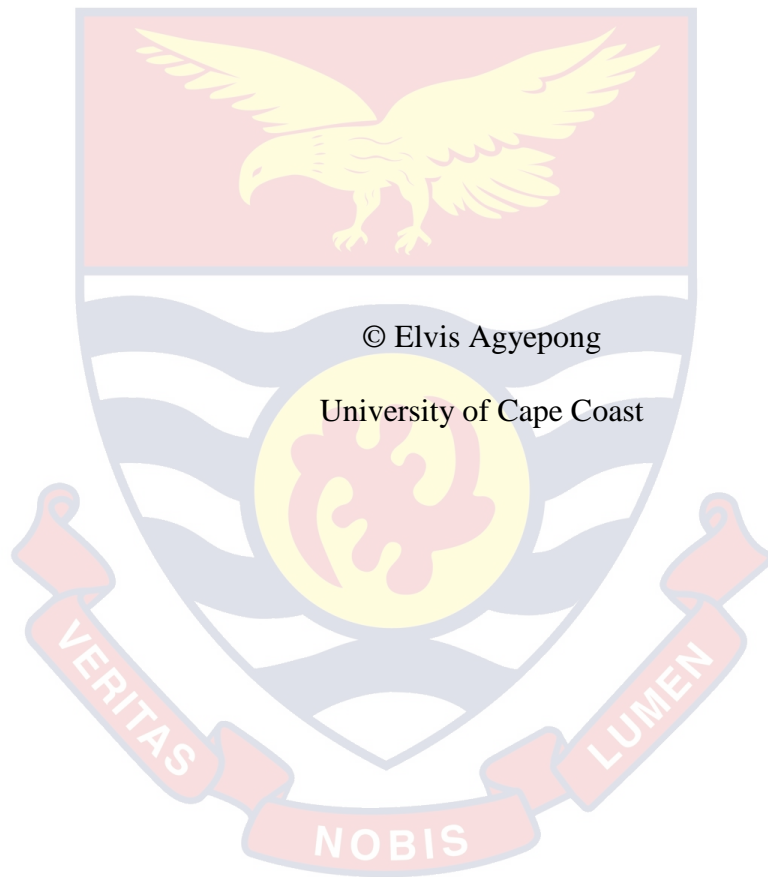
UNIVERSITY OF CAPE COAST

QUALITY (QUANTITATIVE) AND PREDICTIVE RELATIONSHIP  
BETWEEN INTERNAL ASSESSMENT SCORES AND STUDENTS'  
PERFORMANCE IN BASIC EDUCATION CERTIFICATE  
EXAMINATION IN THE KEEA MUNICIPALITY GHANA



ELVIS AGYEPONG

2021



© Elvis Agyepong

University of Cape Coast

UNIVERSITY OF CAPE COAST

QUALITY (QUANTITATIVE) AND PREDICTIVE RELATIONSHIP  
BETWEEN INTERNAL ASSESSMENT SCORES AND STUDENTS'  
PERFORMANCE IN BASIC EDUCATION CERTIFICATE  
EXAMINATION IN THE KEEA MUNICIPALITY GHANA



BY  
ELVIS AGYEPONG

This 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

JUNE 2021

DECLARATION

**Candidate's Declaration**

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

Candidate's Signature ..... Date .....

Name .....

**Supervisors' Declaration**

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

Principal Supervisor's Signature ..... Date .....

Name .....

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

Name .....

## ABSTRACT

It is expected that students' performance in internal assessment at a particular grade level will not be significantly different from their performance in external assessment of the same grade level. Grounded in the expectancy theory of motivation by Vroom (1964), this study was carried out to examine the quality (quantitative) of the District Based Mock Examination and the predictive relationship between internal assessment scores and students' performance in the Basic Education Certificate Examination (external) for the 2019 candidates in Komenda Edina Eguafo Abrem Municipality in the Central Region of Ghana. The study adopted the Ex-post facto design and used the multistage sampling technique to select 602 out of 2,545 of the 2019 Basic Education Certificate Examination candidates. Ordinal regression was used to test the hypotheses. The findings of the study revealed that the item difficulty and discrimination indices of the District Based Mock Examination were within the recommended coefficients. However, the reliability indices of the items were deficient of the recommended coefficients ( $\geq .80$ ). Further, with the exception of the Integrated Science, which had a significant negative predictive relationship with the BECE, the School Based Assessment for the other three core subjects did not show a significant predictive relationship with the Basic Education Certificate Examination. Interestingly, all the District Based Mock Examination scores showed a significant negative predictive relationship with the Basic Education Certificate Examination. It was, therefore, suggested that the practice of the District Based Mock Examination should be sustained and regulated by Ghana Education Service while the Ghana Education Service device mechanisms to check adherence of the School Based Assessment guidelines by teachers.

## KEY WORDS

Assessment

External

Internal

Performance

Quality



## ACKNOWLEDGEMENTS

I would like to express my outmost gratitude to my principal supervisor Dr. Kenneth Asamoah-Gyimah, for his patience, professional guidance and goodwill throughout my M. Phil programme. I am really grateful.

Also, I am greatly indebted to Professor Eric Anane, my co-supervisor, for his expert advice and suggestions in the course of this study. I have learnt to be independent through my interaction with him.

Again, I wish to express my sincere appreciation to professor Godwin Awabil of the Counselling services Centre, and to all the lecturers and staff of the Department of Education and Psychology, especially Dr. Regina Mawusi Nugba and Dr. Ruth Annan-Brew. My invaluable further extends to the Agyepong and Appartaim families for their support and encouragement in diverse ways. Finally, I gratefully acknowledge Mr. Ebenezer Buerthey and Ebenezer Ansah and the entire staff of KEEA Municipal Directorate of Education for providing the needed data for this study.

DEDICATION

To my family: Anita B., Akua Nhyiraba, Kwabena Ayeyi and Akosua

Anuonyam Agyepong





TABLE OF CONTENTS

	Page
DECLARATION	ii
ABSTRACT	iii
KEY WORDS	iv
ACKNOWLEDGEMENTS	v
DEDICATION	vi
TABLE OF CONTENTS	vii
LIST OF TABLES	xi
LIST OF ABBREVIATIONS	xiii
CHAPTER ONE: INTRODUCTION	
Background to the Study	2
Statement of the Problem	5
Purpose of the Study	6
Objectives of the Study	7
Research questions and Hypotheses	7
Research Questions	7
Hypotheses	8
Significance of the Study	8
Delimitations	9
Limitations	10
Definition of Terms	11
Organization of the Study	11
CHAPTER TWO: LITERATURE REVIEW	
Introduction	13

Theoretical Framework	13
Vroom Expectancy Value Theory (1964)	13
Concept of Predicting Academic Performance	16
Concept of educational assessment	17
Formative Assessment	18
Summative assessment	19
External teacher made assessment	20
School Based Assessment	22
The purpose of the SBA in the BECE	23
Challenges of implementing the SBA	24
Qualities of a Good Measuring Instrument	26
Validity of Teacher-Made Test	27
Forms of validity	30
Reliability of test scores	34
Forms of Reliability	36
Factors that can simultaneously affect validity and reliability of a test	41
Test Evaluation Parameters (quantitative)	42
Item Difficulty	44
Item Discrimination Index	45
Students' BECE performance evaluation	47
Review of Empirical Studies	48
Test quality (quantitative) evaluation	48
Predictive power of mock examination	50
Predictive power of the SBA	55

### CHAPTER THREE: RESEARCH METHODS

Introduction	58
Research Design	58
Study Area	59
Population	60
Sample and Sampling Procedure	61
Data Collection Instrument	62
Ethical Consideration	62
Data Collection Procedure	63
Data Processing and Analysis	64
The Ordinal regression model	66
Inferences about the Proportional Odds Model (POM)	67
The test of parallel lines	67
Goodness of fit test	68
Parameter estimates	68
<b>CHAPTER FOUR: RESULTS AND DISCUSSION</b>	
Introduction	70
Analysis of results	70
Research Questions One	71
Research Questions Two	73
Research Questions Three	75
Research Questions Four	77
Discussion	95
Qualities (quantitative) of the DBME Test (core) Items	95
Predictive relationship between the DBME and the BECE	98

Predictive relationship between the SBA and the BECE	101
Students' Performance in the 2019 BECE at KEEA	104
CHAPTER FIVE: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	
Overview	106
Summary of Findings	107
Conclusions	108
Recommendations	108
Suggestions for Future Research	109
REFERENCES	110
APPENDICES	
A Inventory Form	125
B Ethical Clearance	126
C Introductory Letter	127
D Data Collection Permission	128
E Second Mock Examination (KEEA, 2019)	129
F Second Mock Examination (KEEA, 2019)	135
G Second Mock Examination (KEEA, 2019)	142
H Second Mock Examination (KEEA, 2019)	149

LIST OF TABLES

Table	Page
1 Reliability Coefficients (Cronbach’s alpha)	41
2 Difficulty Index and its Corresponding Interpretations	45
3 Discrimination Index and its Respective Interpretations	47
4 Number of Schools by Ownership	60
5 Summary of English Language DBME test Parameters	72
6 Item Difficulty and Discrimination Values of the English Language Test	73
7 Item Difficulty and Discrimination Values of the DBME Mathematics Test	74
8 Summary of the Mathematics DBME test Parameters	75
9 Item Difficulty and Discrimination Indices of the DBME Int. Science Test	76
10 Summary of the Integrated Science DBME Test Parameters	76
11 Item Difficulty and Discrimination Indices of the DBME Social Studies Test	77
12 Summary of Social Studies DBME Test Parameters	78
13 Frequency Distribution for the 2019 BECE Performance in English Language in KEEA	79
14 Model Fitting Information of the 2019 BECE (English Language)	80
15 Goodness-of-Fit for the 2019 BECE (English Language)	80
16 Test of Parallel Linesa for the 2019 BECE (English Language)	81
17 Parameter Estimates of the Predictive Powers of the DBME and the SBA on BECE (English Language)	82

18	Performance of Students in Mathematics in the 2019 BECE at KEEA	83
19	Model Fitting Information for the 2019 BECE (Mathematics)	84
20	Goodness-of-Fit for 2019 BECE (Mathematics)	84
21	Test of Parallel Linesa for the 2019 BECE (Mathematics)	84
22	Parameter Estimates of the Predictive Powers of the DBME and the SBA on BECE (Mathematics)	86
23	Frequency Distribution in 2019 BECE Performance in Integrated Science in KEEA	87
24	Model Fitting Information for the 2019 BECE (Integrated Science)	88
25	Goodness-of-Fit for the 2019 BECE (Integrated)	88
26	Test of Parallel Linesa for the 2019 BECE (Integrated Science)	89
27	Parameter Estimates of the Predictive Powers of the DBME and the SBA on BECE (Integrated Science)	90
28	Performance of students in Social Studies in the 2019 BECE at KEEA	91
29	Model Fitting Information for the 2019 BECE (Social Studies)	92
30	Goodness-of-Fit for the 2019 BECE (Social Studies)	92
31	Test of Parallel Linesa for the 2019 BECE (Social Studies)	93
32	Parameter Estimates of the Predictive Relationship of the DBME and the SBA on BECE (Social Studies)	94

## LIST OF ABBREVIATIONS

AEAA:	Association for Educational Assessment in Africa
CRDD:	Curriculum Research and Development Division
CSEE:	Certificate of Secondary Education Examination
DBME:	District based mock examination
FTSEE:	Form Two Secondary Education Examination
HKPTU:	Hong Kong Professional Teachers' Union
JHS:	Junior High School
JSC:	Junior Secondary Certificate
KCSE:	Kenya Certificate of Secondary Education Examinations
KEEA:	Komenda Edina Eguafu Abirem
M&E:	Monitoring and Evaluation
MMDED:	Metropolitan, Municipal and District Directorate of Education
MoE:	Ministry of Education
SBA:	School Based Assessment
SHS:	Senior High School
SSCE:	Senior Secondary Certificate Examinations
STMA:	Sekondi/Takoradi Metropolitan Assembly
WAEC:	West African Examination Council
WASSCE:	West African Senior School Certificate Examination

## CHAPTER ONE

### INTRODUCTION

All over the world the process of obtaining information about students' achievement and progress in academic institutions has not remained the same. The Ghana Education Service (GES), for example, in 2012 officially adapted and introduced two new assessment practices dubbed the School Based Assessment (SBA) which replaced the Continuous Assessment system, and District Based Mock Examination (DBME) which final year students at the pre-tertiary level take prior to taking the Basic Education Certificate Examination (BECE) as well as the West African Senior School Examination (WASSCE). The SBA guidelines stipulate that students' performance in the BECE is based on their performance at school level. Also, the Monitoring and Evaluation Division of the Ministry of Education report (2011), indicated that, mock examinations contributed significantly to BECE performance of schools that practiced it, hence its introduction in 2012. Interestingly, while some districts, including the Ellembelle district, are attributing the improvement in BECE performance to the adaptation of these assessment practices, same cannot be said about the other districts such as the Komenda Edina Eguafo Abrim (KEEA) Municipal Assembly where students' performance in the BECE continue to decline. This raises concern about quality and reliability issues of the SBA and the DBME respectively. This study, therefore, seeks to fill the lacuna by investigating these issues.



## Background to the Study

In Ghana, at the pre-tertiary level, a policy of merging internal and external assessment scores for grading students was adopted in the 1980s when the “new structure and content of education” was introduced at the pre-tertiary levels of education. “The assessment-policy framework at the beginning required that the final grading of pre-tertiary students should be made up of 40% continuous -assessment scores and 60% external-examination scores” (Amedahe, 2001). Amedahe further indicated that the weighting of the two assessment scores was later reviewed to 30% and 70%, respectively. This weighting raises the issue of dependability of the continuous assessment scores.

According to Amedahe (2001), although it is educationally meaningful to combine internal assessment scores with external assessment scores for final determination of students’ performance, the practice is fraught with reliability issues. He raised an issue concerning predictive relationship between teacher-made assessment scores presented to the WAEC at the end of school programme for first and second cycles education programme and students’ performance in the programme. He again, raised a concern about the quality of teacher made tests with regard to the number of items on tests, the difficulty level of the items as well as the quality of grading essays. This is because teachers craft their own test instruments without any validation by testing experts (Amedahe, 2001).

Consequently, another assessment system referred to as the School Based Assessment was officially introduced and adopted by the Ghana Education Service in 2008 and 2012 respectively that sought to deal with the

challenges posed by the Continuous Assessment System. The SBA policy is a all-inclusive, systematic, continuous, diagnostic and integrative teacher-supervised process which also reduces the workload of the student and the teacher by 64% and 53% respectively (Ukwuije & Opara 2013; CRDD, 2012). The Curriculum Research and Development Division (CRDD) indicated that the SBA should be a continuous process and integrated within the three domains of learning, thus, cognitive, affective and psychomotor.

Also, Opara, Onyekuru and Njoku (2015), posited that SBA scores in any subject for a particular student in a real sense, should mirror his or her attainment in external examination of similar content and difficulty level. They further indicated that scholars such as Orubu (2013) and Ogunkola (2007) examined SBA scores as predictors of student's final grades in Delta and Ogun States and establish that, the SBA scores predicted students' performance in Junior Secondary Certificate Examination (JSCE) mathematics and integrated science respectively. It is, therefore, expected that the SBA policy will help achieve similar results as captured in the SBA guidelines. Unfortunately, despite the formal introduction of the SBA in the Ghanaian education system more than a decade ago, there is still growing concern among stakeholders on the predictive relationship between the SBA scores and candidates' performance in the BECE.

Interestingly, the Ministry of Education (MoE), instead of focusing on the implementation bedeviling the implementation of the SBA, launched another form of assessment in 2012 at the district level dubbed District Based Mock Examination (DBME) project in Accra. The DBME, according to the Monitoring and Evaluation Unit of the MoE, would help improve students'

performance in the BECE (GNA, 2012). As a result, the DBME has been adapted by almost all district, municipal and metropolitan directorates of education including some district chief executives and non-governmental organisations in the country. Consequently, the DBME is currently receiving a lot of financial and political attention from the District Directorates of Education and some non-governmental organisations.

In the Western Region, for example, the Sekondi/Takoradi Metropolitan Assembly (STMA) allocated GH¢74,000 to assist the metro education directorate to conduct two different mock examinations for the 2017/2018 Basic Education Certificate Examination (BECE) candidates in the public schools within the metropolis. The Metropolitan Chief Executive, at a press briefing, indicated the initiative had been taken in wake of abysmal performance of candidates in the BECE examinations. The move embarked on to have an improved performance by the Metropolitan Chief Executive may have been inspired by the Ellembelle District Directorate of Education which attributed their improved performance (76%) in the 2017 BECE to the efforts of the Member of Parliament (MP) for the area who liaised with the Education Office to organize two separate district based mock examinations prior to the commencement of the BECE that year. Also, similar observation was made by the Assin South Directorate of Education in August 2017 on the impact of the DBME on students' performance in the BECE. Interestingly, same cannot be said about the KEEA. The situation raised concerns about the quality and usefulness of the SBA and the DBME in improving students' performance in the BECE in the municipality. To ensure quality and fairness of a test, Kargar, Tarmizi and Bayat (2010), suggested that effective analysis of test items

should be undertaken. Thus, item analysis helps in identifying potential errors in scoring, ambiguous items, and malfunctioning distractors that could be reviewed to improve the quality of test for future use.

### **Statement of the Problem**

Academic performance of junior high school (JHS) students in Ghana, especially the public schools, over the last decade has not been encouraging (Etsey, 2005). Similarly, Okyerefo, Fiaveh and Lamprey (2011); Gyan, Mabefam and Baffoe (2014) indicated that that has been the trend in the (BECE) for quite some years now in the public schools, particularly those in rural Ghana. Interestingly, the Department of Monitoring and Evaluation (M&E) of the Ministry of Education report (2011) corroborated the findings of Etsey and Gyan *et. al* (2014) and indicated that some of the schools that performed well in the BECE implemented a rigorous schedule of mock examination for the final year students, but in many underperforming schools this was not witnessed.

Subsequently, the MoE in 2012 officially launched a local assessment scheme dubbed DBME upon establishing that schools that performed well in the BECE implemented rigorous schedule of mock examinations. This was in an effort to improve BECE performance nationwide (MoE, 2011). Then in 2008, the MoE officially introduced the school based assessment which was purposed to reduce the work load of the student and the teacher by 64% and 53% respectively, and to ensure that the scores students obtained at the school level related positively to their performance on the BECE. Curriculum Research and Development Division (CRDD) (2012). Orubu (2013) and Ogunkola (2007), asserted that SBA scores for a candidate in any subject

should reflect his or her performance in external examination of same or similar content.

Interestingly, whereas the timely intervention of the DBME augmented students' performance in some other districts, it is not same when it comes to students' performance in the BECE at the KEEA Municipality despite the formal introduction of the SBA and DBME in the municipality. In fact, statistics available at the district education office show evidence of a remarkable drop in the academic performance of students in the municipality over the last eight years even with the adoption of the DBME in the Municipality. In 2012 and 2013, for example, the Municipality recorded 40.2% and 41.0% passes respectively. Unfortunately, in 2014, BECE performance in the KEEA Municipality further dropped to 32.0% out of a total of 2,043 candidates presented and, in 2018 the municipality was ranked 20<sup>th</sup> out of 20 districts in the Central Region.

It is clear that, despite the introduction of the SBA and the DBME, students' performance in the BECE at the KEEA Municipality continues to decline. Hence, the launch of this study to assess the quality of the DBME as well as the predictive powers of the SBA and the DBME scores on students' performance in the 2019 BECE in the KEEA Municipality of the Central Region.

### **Purpose of the Study**

The purpose of this study is to evaluate the quality (quantitative) of the DBME and the predictive relationships between the SBA, DBME and students' performance in the 2019 BECE at the KEEA Municipality of the Central Region of Ghana.

## Objectives of the Study

Guided by the following, the study sought to:

1. evaluate the quality (quantitative) of the 2019 DBME (English language, Mathematics, Integrated Science, and Social Studies) test items by subjecting them to item analysis.
2. establish the predictive relationship between the SBA score of the core subjects and students' performance in the 2019 BECE at the (KEEA) Municipality.
3. establish the predictive relationship between the DBME score and students' performance in the 2019 BECE at the (KEEA) Municipality.
4. establish the impact of the SBA and DBME on students' performance in students' performance in the (KEEA) Municipality.

## Research questions and Hypotheses

The following research questions and hypotheses were posed and tested to guide the study;

### Research Questions

1. What are the qualities (quantitative) of the 2019 DBME English Language test held at KEEA?
2. What are the qualities (quantitative) of the 2019 DBME Mathematics test held at KEEA?
3. What are the qualities (quantitative) of the 2019 DBME Integrated Science test held at KEEA?
4. What are the qualities (quantitative) of the 2019 DBME Social Studies test held at KEEA?

## Hypotheses

The study was guided by the following hypotheses.

- i.  $H_0$ : Students' scores in the DBME and SBA do not significantly predict their performances in the 2019 BECE in English Language at KEEA.  
 $H_1$ : Students' scores in the DBME and SBA significantly predict their performances in the 2019 BECE in English Language at KEEA.
- ii.  $H_0$ : Students' scores in the DBME and SBA do not significantly predict their performances in the 2019 BECE in Mathematics at KEEA.  
 $H_1$ : Students' scores in the DBME and SBA significantly predict their performances in the 2019 BECE in Mathematics at KEEA.
- iii.  $H_0$ : Students' scores in the DBME and SBA do not significantly predict their performances in the 2019 BECE in Integrated Science at KEEA.  
 $H_1$ : Students' scores in the DBME and SBA significantly predicts their performances in the 2019 BECE in Integrated Science at KEEA.
- iv.  $H_0$ : Students' scores in the DBME and SBA do not significantly predict their performances in the 2019 BECE in Social Studies at KEEA.  
 $H_1$ : Students' scores in the DBME and SBA significantly predicts their performances in the 2019 BECE in Social Studies at KEEA.

## Significance of the Study

This study was designed to establish the predictive relationship between school and or district level assessment scores and students' performance in national examinations. A number of studies assert that measures of predictability of results obtained in an examination at a particular level can be used to make decisions about which subjects to study or which courses to take at a higher level. Thus, if a predictive relationship is

established between internal and the external assessment results, the internal assessment results could be used as an evidence to validate the use of the internal assessment results in times of need or doubt.

Consequently, if the relationship between the SBA or DBME and the BECE is well established, it could be used as evidence for or against placement decisions of BECE graduates into senior high schools and programmes. Also, it could help reduce situations where candidates are placed in programmes above or below their ability levels. Next, the findings could help identify students at risk of attrition in order to provide necessary support early enough for them with the goal of increasing retention. Again, the established relationship could be of noteworthy to the KEEA Municipal Directorate of Education and other relevant stakeholders in the sense that they would be in a better position to base their decisions and arguments on the findings of this study.

Further, the findings of this study would also be of immense help to head teachers, teachers as well as the pupils on how predictive their DBME scores could be and how to derive maximum benefit out of them. Finally, since enough attention has not been given to this topic, this study may lay the foundation for further research by other researchers and academicians who might wish to launch further enquiry into this area.

### **Delimitations**

The scope of the study was delimited to students who sat for the 2019, DBME and the BECE in the KEEA Municipality of the Central Region. The first reason for selecting the KEEA is that the district education office was ready to provide the full complement of the SBA and DBME scores for the



years under consideration. Also, BECE performance of the municipality had been declining over the last eight years.

Secondly, the study was delimited to English, Mathematics, Integrated Science and Social Studies since they are all core subjects of which every candidate is required to pass before being placed in any of the Government assisted senior high schools.

Thirdly, the study was delimited to only the public schools in the municipality because the private schools in the assembly performed relatively better than the public schools. Finally, the quality of assessment of the test was delimited to multiple choice aspect of the test since it is less subjective to evaluate.

### **Limitations**

The data for the study was an existing (secondary data) DBME and SBA scores obtained from the KEEA municipality. Unfortunately, the data available at the municipal education office at the time of data collection were incomplete which left the researcher with the option of waiting for the 2019 data on the SBA, DBME and the BECE results. In addition, the test administration conditions for the DBME and the BECE were not the same because whereas candidates were invigilated by their own teachers during DBME, it was not the same during the BECE, where candidates had to be conveyed to examination centres, supervised by unknown invigilators which could have caused some level of nervousness and uneasiness among the candidates and may have influenced the BECE results.

## Definition of Terms

This section gives a definition of the key terms as used in the framework of this study;

**External Examination:** These are examinations which are designed and conducted by recognised examination bodies such as the WAEC.

**Internal Assessment:** A form of formative assessment involving appraisals and feedbacks to students by the respective schools based on their curriculum.

**Mock Examination:** an examination set as a prototype or made to look exactly in content and cognitive requirement as the Basic Education Certificate Examination (BECE).

**Performance:** this is the actual accomplishment showing the learner's ability as measured by a score or grade in the BECE.

**Predictive Validity:** It is concerned with the relation of internal test scores to predict students' performance in the BECE.

**Quantitative Qualities:** the use of item difficulty and discrimination indices to evaluate test items.

**School Based Assessment (SBA):** a system for collecting comprehensive performance data on students in the course of their work at school.

## Organization of the Study

The entire study was ordered into five chapters. Chapter one presented the introduction and background to the problem, statement of the problem, purpose of the study, research objectives, research questions and hypotheses, significance of the study, delimitation of the study, limitation of the study, definition of terms and the organization of the study. Chapter two presented the literature review of the study, while chapter three presented the detail

description of the methodology adopted to carry out the study. Chapter four presented the analysis of data and discussion of results, and finally chapter five dealt with the summary, conclusion and recommendations for policy makers and other relevant stakeholders.



## CHAPTER TWO

### LITERATURE REVIEW

#### Introduction

One of the vital features of a scientific research report is the literature review as it lays a concrete foundation for advancing knowledge, abridges concepts and theory development, closes areas where great volume of research exists, and reveals areas where new or further investigation is required (Webster & Watson, 2002). As a result, this chapter is focused on review of relevant theories and concepts as well as other empirical findings on the subject of reliability and predictability of test scores. Specifically, this chapter shall review related literature under the following sub-headings:

1. Theoretical Framework
2. Concept of Predicting Academic Performance
3. Concept of Educational Assessment
4. Qualities of a Good Measuring Instrument
5. Test Evaluation Parameters (Quantitative)
6. Students' Performance Evaluation in BECE
7. Review of Empirical studies

#### Theoretical Framework

##### Vroom Expectancy Value Theory (1964)

Vroom's 1964 expectancy theory of motivation asserts that motivation is a function of the professed probability that effort is directly proportional to performance or achievement in any endeavor. Subsequently, the expectancy

theory, according to Vroom, is a subjective probability of an effort or action (e) yielding an outcome or performance (p) expressed as  $e \rightarrow p$ . In practice, expectancy has also been measured as the professed relationship between an action and an outcome. Moreover, expectancy has been interpreted as the subjective probability that effort leads to the outcome of performance or second-level outcome (o) expressed as  $e \rightarrow o$ . Thus, performance will yield certain outcomes, and that these outcomes are valued (Issac, Wilfred, & Douglas; Porter & Lawler as cited in Friedman & Mandel, 2009).

Thus, an individual's perceived probability that effort is related to performance is a function of past experiences in similar situations. Previous occurrences where effort led to a desired outcome may increase an individual's probability with similar tasks in the future. Also, past experiences where effort led to poor performance may lower expectancy. The performance to outcome expectancy is the perceived probability that good performance will lead to a desired outcome (Ramos, Nunez, Unda & Mejia, 2016)

Although the expectancy theory has been predominantly used to study employee workplace motivation, the theory has not been tested extensively in looking at student motivation, performance, and retention (Isaac, Wilfred & Douglas, 2001). According to Friedman and Mandel (2010), prior to college, students often go through a wide range of experiences that influence their motivation to academically and socially succeed in school. These experiences are characterized by variations and similarities among students with respect to school systems, teachers, administrators, financial requirement per students, educational models such as varied reading models as well as quality and level of parental and community support. Since school experiences are a unique

blend of these and other influences, students' effort and performance, and valence levels will vary. For instance, students whose teachers are competent and attend to their unique learning needs may increase students' belief that they are more likely to achieve good grades if they exert effort. Similarly, students who are not rewarded for good performances from teachers or parents may experience low performance expectancies, and therefore may have lower motivation than students whose past performances were duly rewarded (Gibson, Ivancevich, Donnelly, & Konopaske, as cited in Friedman & Mandel, 2010). Likewise, students who excel academically in internal examinations are expected to replicate such performances in similar examinations administered by external examiners such as the West African Examination Council.

Applying Vroom's expectancy theory to the use of internal assessment to predict students' performance in external examinations, there is the need for internal assessment tools to be of good quality in order to yield reliable scores. Thus, for the DBME and the SBA to predict students' performance in the BECE, the DBME and the SBA assessment tools must be of good quality.

Kimberlin and Winterstein, (2008) opined that the quality of a measurement instrument is described by three quality domains. These are validity, reliability, and responsiveness. Thus, in a test, any score obtained by a good measuring instrument is the sum of the 'true score', which is unknown, and 'error' which is obtained during the measurement process. If the error margins are insignificant and the scores are of high standards, the test instrument and the test scores are considered reliable and could be the bases for a valid decision about the use of the assessment scores. Consequently, the dependability of internal assessment score to predict future performance in

external examination depend so much on the quality of the instrument used and the reliability of the test scores obtained from the internal assessment.

### **Concept of Predicting Academic Performance**

Accurate prediction, valid selection and placement decisions are of high concern for parents, teachers, counsellors and other stakeholders. However, despite different frameworks and models used by various practitioners and researchers across many institutions for predicting performance, rate of success that has been recorded in terms of accuracy, efficiency and validity of selection and placement decisions is negligible (Adejo & Connolly 2017).

Nonetheless, prediction of academic performance is continually being used to professed students who will succeed in further academic endeavours (Omirin & Ale, 2008). Omirin and Ale indicated that predictions of examination results are meant to reveal how successful teachers' instructions have been mastered. Thus, teachers use examinations such as mock examinations to ascertain how students are progressing and identify where students are facing challenges. The outcome of the construct assessed could be used to make instructional intervention plans such as re-teaching, trying alternative instructional approaches in order to improve performance.

Ochieng (2012) asserts that a student's performance is critical in ensuring the supply chain is continued. Educational institutions as a result, use predictive powers of examinations to make decisions regarding the capabilities of students in a related future situation. Mock examinations for example are often used as tools for predicting academic performances and ultimately for making admission decisions into academic programmes in some jurisdictions.

His assertion is in line with the notion that the main products of educational institutions are the students. Upon graduation from one grade level to another, the students may either continue into the higher level of the education ladder or join the manpower of industry, government and private sector based on their performance in an examination (predictor) (Zaidah & Daliela 2007).

Prediction of a future performance in an examination could be made with reasonable success on the basis of the results of previous examinations (Daniel & Schouten 1970). They indicated that, 'Predictive examinations are designed to determine each student's likelihood of meeting some criterion score on the end of year tests'. A good example of such predictive examinations in Ghana is the DBME, which is conducted by District Directorate of Education in Ghana to help boost performance in the BECE. Also, it helps to predetermine performance of candidates in the BECE.

### **Concept of educational assessment**

Assessment is a huge topic that encompasses everything from nationwide standardized tests to district wide teacher made tests otherwise referred to as District Based Mock Examination (DBME) in this study or everyday classroom tests. In order to grapple with what seems to be an abuse of testing, educators should frame testing as assessment and that assessment is basically information gathering. Thus, the more information we gather about a student or system, the clearer the picture we have about their achievement and where intervention may be needed.

According to Nitko (as cited in Opara, Onyekuru & Njoku, 2015), assessment can be defined as a process of gathering information about pupils, curricula, programmes and educational policy used for making decisions to



improve learning outcomes. Similarly, Allen (as cited in Brempong, 2019) avers that, assessment entails the gathering of information about the student, curricula, programmes and other relevant information for decision making. Juxtaposing the various assertions made by these researchers, it could be concluded that assessment is a means to an end and not an end in itself.

To this effect, Amedahe and Asamoah-Gyimah (as cited in Brempong, 2019) indicated that any assessment process adapted to collect information about a system or an individual should be guided by the purpose of conducting the assessment. Likewise, one of the three main purposes of the SBA is to ensure that the grades students obtain at the BECE are based on their performance on the internal tests administered in school and on the external examination administered by WAEC (CRDD, 2011). Assessments of students, however, is often in two forms, namely formative and summative assessment.

### **Formative Assessment**

Formative Assessment is integral part of an instructional process. It is often integrated into classroom practice to provide the information needed to make teaching and learning effective. In effect, it enlightens teachers on students understanding of a particular instruction given and the effectiveness of the teachers' instructional methods. According to Clark (as cited in Akyina & Oduro-Okyireh, 2019) formative assessment is best seen as a unifying theory of instruction, which directs instructional practice and enhances classroom learning process. This is achieved by developing Self-Regulated Learning (SRL) strategies among learners (Akyina & Oduro-Okyireh, 2019). Akyina and Oduro-Okyireh, indicated that the ultimate goal of formative assessment is to improve classroom instruction. Their assertion reiterated

Clark's position, which posited that formative assessment if used strategically could help develop SRL in learners through metacognition, strategic action, and motivation to learn which is a precursor to academic performance. Interestingly, many associate formative assessments not only to summative assessments but also with standardized tests such as national examinations. These tests form an integral part of district and classroom programs (Garrison & Ehringhaus, 2010). Similarly, the SBA forms an integral part of final assessment of BECE candidates in determining final grades obtained by a candidate in the BECE.

### **Summative assessment**

Generally, teachers rely on two forms of summative assessment; teacher constructed, also referred to as teacher made test and standardized test. Teacher made assessment is the most common form of assessment found in classrooms. The purpose is to provide data for appraising student performance at the end of term or year examinations, however; it is prone to bias. Standardized assessment, on the other hand, is designed to overcome many of the biases that can taint teacher made assessment tools. Juxtaposing the purpose of teacher made and standardized forms of summative, one could conclude that both types of summative assessment have a place in an effective education system, but for real positive impact they should be employed to meet the specific needs for which they were employed.

Final examinations are external standardized examinations open to the general public and conducted by accredited examining bodies using tests that have appropriate psychometric properties (Obioma & Salau, 2007). Further, Adeyegbe (2004) indicated that these tests used by the various examination

boards are often better than those of the teacher-made tests since they are standardized. By implication, standardised tests are forms of examination that satisfy the required psychometric properties of a test and holding all other things constant, they ensure that the curriculum and learning outcomes are assessed in a vigorous and fair manner and such tests lead to reliable and valid exit certificates. Standardized tests, however, undesirably put a lot of pressure on the classroom teacher because emphasis is often shifted from teaching for attainment of knowledge to teaching to pass examination.

The BECE conducted by the WAEC is an example of standardized final examination in Ghana. The results of the BECE are used to place students in senior high schools based on merit (performance). A student who is placed in a particular school with a given programme is assumed to possess the requisite abilities and skills to cope with the academic challenges of the programme and school he or she is placed in (Othun,1994).

#### **External teacher made assessment**

External teacher-made tests are tests that are designed and conducted by ‘expert’ examination organisation (Adeyemi, 2008). The DBME, for example, is one of such tests, in the sense that, the examining organisations of the test did not organize instructional sessions nor prepare students for the examinations. The examinations are designed and administered under specified conditions based on norms that are regarded as standard. They are designed to assess school candidates prior to the final or national examination (Ochieng, 2012). Ochieng’s assertion suggests that all external examinations assume the nature of summative examination regardless of the source, and are used for placement of students in higher levels of academic institutions for

further education as well as for monitoring and predicting students' performance. An example of such examination is a mock examination.

Fundamentally, mock examinations are forms of examination set to follow a pattern or made to look closely in content and cognitive requirement as the main examination or standardized examination (Madu & Ebere, 2016). According to Adesoji, and Kemmi (2013), it is an examination designed as a look-alike to the final examination day experience. In doing so, it is structured with essentially the same topic areas in terms of content as well as the difficulty level of the main examination. Thus, a mock examination is administered as closely as possible to the main examination in terms of curriculum and syllabus. It is, therefore, an attempt to permit candidates planning to write a major examination to have similar experiences prior to the main examination (Amadioha & Uko, 2019).

Additionally, Ramatlala and Nenty (2012) asserted that the performance of students in a mock examination is an essential indicator that could possibly relate to the performance of a student in a national or external examination. This supports Omirin and Ale's (2008) assertion that mock examination as trial examination is selective, predictive and diagnostic in nature; revealing how teachers' instructions have been grasped by the candidates, and revealing in advance those candidates who are likely to pass or fail in the final examination. Thus, the stronger the correlation between the mock examination and the final examination, the higher the degree of predictive validity of the mock examination and vice versa. Consequently, the debate on prediction of academic success using mock examinations has attracted the attention of many researchers and this study seeks to contribute to

the debate by investigating the quality and predictive power of the District Based Mock Examination (DBME) in the KEEA Municipality

### **School Based Assessment**

According to Ifeanyieze and Aneke (2013), the school assessment is a type of assessment which includes various categories of stakeholders (teachers, parents and peers) in the process of determining the results of the learning to encourage children to become interested and to make constant progress. This method of evaluation permits learners to obtain support from individuals and help them to engage in the evaluation process, thereby increasing individual preparation both at home and in the classroom. SBA is an integral, standardized, continuous, diagnostic and integrative evaluation method for the instructor, according to Opara, Onyekuru and Njoku (quoted in Ukwuije & Opara, 2013). It begins with school circumstances that promote active participatory participation and student engagement, with an emphasis not on scores and grades but on learning.

Okpala, Afemikhe and Anyanwu (2012) also noted that SBA is a dynamic mechanism requiring the use of various modes of evaluation processes to obtain and/or provide information on education-related issues for decision-making. In line with Okpala et. al claim, the SBA guidelines (2012), indicated that the SBA is a method for gathering performance data on students during their school work. A spectrum of practical work, as well as theoretical work defines education in the school system. Therefore, the SBA offers schools the opportunity to ensure that students' success on both practical skills and theoretical knowledge is expressed in the marks earned by students in the SBA and in their end-of-term exams. The guideline further suggested that

education is a broad concept and students are required to learn from a variety of sources; from their teachers, from their peers and elders, from the library and, in general, from their community, indicating that learning does not take place only in the classroom. All these sources contribute to a young person's education. Thus, in the three areas of learning, SBA evaluate of students continuously at intervals; cognitive, affective and psychomotor using various instruments such as research, assignment, observation, interview, questionnaire and project

### **The purpose of the SBA in the BECE**

It has been generally established that students tend to perform well on tests administered in a familiar environment because they are more relaxed when taking tests in such environment. The familiarity with their classroom and school surroundings for example help them to perform better in tests administered in school. However, the same testing conditions is not guaranteed during external formative and summative examinations. This is because during such examinations, students' work is supervised by people who are not known to the students which tend to create anxiety and tension for some students. This anxiety may tend to reduce the level of performance on the part of some students. It is, therefore, generally accepted that the performance of students on tests administered by their teachers in their classrooms better reflects the true performance of the student.

It is for this reason that, the Ministry of Education decided that a percentage of the marks for the BECE should come from performance of students' SBAs. Consequently, 30 percent of the final examination mark comes from schools while the remaining 70 percent mark is an external mark

based on students' performance at the BECE. (CRDD 2011). The purpose of the SBA, therefore, is to ensure that the grades students obtain at the BECE are reflections of their performances on internal tests administered in school and the external examination administered by other stakeholders (CRDD 2011). A second purpose of SBA is to give schools the chance to ensure that all aspects of the education and training of the young person are taken into account in the assessment of the performance of the young person while he/she is in school. Finally, the SBA provides students the opportunity to show the quality of their learning they have received in and out of school. The SBA, in effect, gives students the chance to receive a broad education.

### **Challenges of implementing the SBA**

The adoption of the SBA ideology implies a shift in the philosophy of assessment 'from a "testing" model to a broader model of "educational assessment"' (Yung, as cited in Lee & Cheng 2010). The SBA is believed to encourage student-centred learning which ensures overall development of the student and enhances validity of assessment (Cheng & Lee, 2010). The SBA is a forward-looking approach that motivates students to engage in individual learning paths. It is, moreover, an avenue to meet the needs of students with diverse learning styles (Keightley & Coleman, 2003). On the other hand, despite the usefulness of the SBA in improving teaching and learning, it is bedeviled with some implementation challenges.

Yip and Cheung (2005) in their study reported that teachers might not necessarily have the skills needed to achieve the objectives of SBA. They indicated that the differences in students' ability and learning attitude may also influence the effective implementation of SBA. In addition, they enumerated a

number of challenges that teachers face in the process of implementing SBA.

These challenges include;

1. making assessment appropriate to students;
2. diagnosing the learning needs of students accurately, and responding accordingly;
3. creating a supportive and non-competitive assessment environment and
4. helping students to develop a personal learning path

Ironically, a study conducted by Ahenkora (2019) involving 200 basic school teachers in the KEEA Municipality, reported that teachers showed positive attitudes towards the application of SBA guidelines in schools. This suggests that teachers are well abreast of the SBA implementation guidelines and as a result should be able to contain some of the challenges listed above. Unfortunately, that may not be the situation in all cases as other studies suggest that teachers are not well versed with the SBA guidelines. For example, a study conducted by Veloo, Krishnasmany and Ali (2015) in a study involving 260 teachers from Nigeria to assess teachers' knowledge and readiness towards the implementation of SBA in Nigeria revealed that, teachers have relatively more knowledge but are lackadaisical in implementing the SBA.

According to Nitko (as cited in Ahenkora, 2019), one of the major challenges encountered by basic school teachers in an attempt to implement the SBA at their various schools is the large class sizes, especially in the rural areas. This was corroborated by a report from the Hong Kong Professional Teachers' Union (HKPTU) in 2013, which indicated that the workload brought by SBA is overwhelming for both teachers and students. It indicated,



that for students, the demand for SBA is high in terms of quality and quantity. On the contrary, according to the 2012 SBA guidelines, one of the benefits of the SBA is that it reduces the workload of the teacher in terms of the number of exercises and activities to be assessed within a term.

### **Qualities of a Good Measuring Instrument**

In recent times, the quality of education has focused an immense attention on student performance in national or final examination. One of the ways by which such performances are measured is the extent to which students' performance in teacher-made tests is able to predict their future performance in a high stake tests such as the BECE (Notar, Zuelke, Wilson, & Yunker 2004). Logically, it is expected that there would be a strong positive correlation between a student's performance in a teacher-made test and the student's performance in final or national tests. Consequently, the role of testing as a part of teaching and learning is inevitable in gathering reliable information about the students for teachers' use. It provides the bases for informed judgments about the learner progress in relation to a task criterion as well as to provide the feedback for teachers about how they teach and for the students about how they learn. 'To be effective, assessment should be designed and developed based on the principles of effective assessment practices'. Thus, developing effective assessment instrument as stipulated by Nitko, Brookhart and Susan, (2007) requires some established principles, such as identifying the learning goals and targets, planning how and what to assess, writing good quality items, using marking guides, and selecting assessment forms. They further indicated that whether a test is a standardized or teacher-

made, it should apply the qualities of a good test which are; validity, reliability and usability.

### **Validity of Teacher-Made Test**

The concept of test validity as a quality indicator in testing is attributed to the manner in which the test scores are interpreted and used. The interpretations tend to be given little value in a norm referenced test such as teacher-made tests where a criterion for scoring each item is pre-set. On the other hand, in high stake testing, the criterion for what is the correct answer to an item is determined by the available scientific evidence about a phenomenon. Thus, there is only one interpretation given to a test score, that is, the interpretation is influenced by attributes possessed by the test takers (Borsboom, 2005).

The interpretation-oriented proponents consider validity as making sense in a norm-referenced test because the scores are interpreted as per the observed norm. That is, a test will remain as it is and it is only the interpretations made about the test within a given norm that make it either valid or invalid (Hathcoat, 2013). This means that both instrument-based and interpretation-based approaches to validity are appropriate. However, the interpretation-based approach tends to have much wider application because it can be used in diverse contexts of testing, but the instrument based testing is applicable in contexts where specific attributes are being measured (Hathcoat, 2013). However, in the context of SBA the instrument-based approach to validity is what is best applicable to testing at schools (Hathcoat, 2013).

Consequently, to meet the demands of quality education, classroom assessment cannot be independent of national or final examinations since it is

used to gather information to predict the future performance of a student. Unfortunately, various studies have reported a mismatch between teacher-made tests and national examinations. Fives and DiDonato-Barnes, (2013) for example, reported that there is a disparity between the content learned in class and the material assessed at the end of unit test. Undoubtedly, such a test is likely to be invalid due to wrongful usage of assessment tools to gather information about the students.

However, for validity to hold, Afolabi and Faleye (as cited in Ochieng, 2012) assert that a valid test must produce consistent results, and be resistant to external factors that influence test scores. They also assert that validity is the most important quality measure of a test and that a valid test has to measure what it is intended to measure. Their assertion, ironically is outdated, because currently validity is not limited to just the intended measure of a test, but the degree to which theory and reliable evidence support the interpretation and intended use of the test result.

According to Giacomini and White (2006), it is assumed that teachers have undergone the required training to become professionals in the classroom and, therefore, have the capability of teaching as well as designing valid and reliable tools for assessment and to test students' academic achievement. Consequently, one would expect that the question of validity of a teacher-made test as a measure against the psychometric properties of standardised test is answerable on scientific grounds by referring to the course content of teacher trainees (Messick, as cited in Ochieng 2012). However, since the training teachers acquire at college does not always reflect in their output, it is

uncommon for a teacher made-test to lack the required psychometric properties of a test, thus validity and reliability.

Validity of a test is the extent to which the results are truthful (Pallant, as cited in Mohajan 2017). Pallant explained that validity encompasses the entire measurement concept, and establishes whether the scores of the test meet the criteria of psychological testing. Thus, adding to the assertion that validity of a test is the extent at which psychological testing method has been followed during the process of generating test items (Oliver, 2010). Therefore, validity is not a property of the instrument, but of the scores obtained using the instrument, interpretations and the extent to which evidence validates the interpretation and usage of the test results. Thus, it is best viewed as the process by which evidence is gathered to justify proposed inferences of a test result (Messick, 1989).

Validity of a test has two vital parts known as internal and external validity. Internal validity indicates whether the test scores are credible due to how the test items were sampled, recorded and analysed. Thus it gives an indication of the 'replicability' of the test scores (Willis, 2007). External validity, however, indicates the transferability of the test score to alternate forms of the test (Last, 2001). Deductively, to ensure transferability of a test score, there should be adequate representativeness of test items on the construct being tested as well as the heterogeneity of the group, the use of non-reactive measures and precise instruction to allow replication of the test scores (Mohajan, 2017). It is sometimes mind boggling to determine whether a test measures the right construct or not (Shekharan & Bougie as cited in Mohajan, 2017). Consequently, validity demand that any instrument used to

collect data is reliable. However, a reliable instrument does not guarantee validity of a test (Kimberlin & Winterstein, 2008).

### **Forms of validity**

Validity of a test is generally classified into four categories as indicated by (Kinyua and Okunya 2014). They are; content validity, face validity, construct validity and finally criterion-related validity.

Content Validity refers to the degree to which all potential items that may have been tested reflect the item on the instrument and the scores obtained from a test (Creswell, 2005). It ensures that the notion is tapped by representation of the set of items on the test. Therefore, the more the test reflect the domain of the being measured, the greater the validity of the test (Sekaran & Bougie, 2010). DeVellis, (2006) and Messick (1998) also aver that it is related to a type of validity in which various components, abilities and behaviours are sufficiently and effectively evaluated. It is important to note that there is no statistical test to assess whether a measure adequately covers in a particular content area, leaving judgment of content validity decision to the discretion of experts in the field of psychological testing.

Relatedly, Notar, Zuelke, Wilson, and Yunker as cited in Osabede, 2015 suggested that teachers should make good use of a Table of Specification (TOS) in their test construction process as it helps to identify not only the content areas to be covered but also the success expectation at each stage of the Bloom's Taxonomy's cognitive domain. Thus, whenever a TOS is prepared for a test, it projects that the assessment instrument is likely to be valid based on the proof of the quality of the test and dependable decision-making response (Fives & DiDonato-Barnes, 2013).

Moreover, Osabede (2015) called for structured test preparation to identify the instructional and behavioural targets to be measured; identifying the test content areas the test format decision and a TOS. For example, in a study to help develop a valid Economics test for secondary school students, Osabede indicated that knowledge, understanding, implementation, analysis, synthesis and assessment as postulated by Bloom 1956. He also indicated that to ensure the content validity of an instrument, a TOS must be drawn vertically and horizontally to determine the content areas and behavioural objectives for item writing in order to guarantee content validity of the instrument. From the examined literature, it is suggestive that for a test to be valid, the content is as vital as the construction of table of specification.

Similarly, Fives and DiDonato-Barnes (2013) and Notar et. al. (2004) indicated that in test development, one of the key quality indicators is the preparation of a TOS which is also referred to as the test blueprint since it helps test constructors to align objectives, instruction, activity and assessment. They argued that the TOS can be used in conjunction with lesson and unit planning to help teachers to easily to easily connect preparation, teaching, activity and evaluation in conjunction with lesson and unit planning. Deductively, the purpose of a TOS in test construction is to ensure content validity of a test achieved by a teacher or Psychometricians. On the other hand, Crocker and Algina, (1986) have suggested four steps /processes that could be adapted to achieve content validity of a test; define and outline the area of interest, gather resident domain experts, establish clear matching methodology, and finally evaluate the matching task results (Allen & Yen, 1979).

Face Validity: as the name suggests is considered as the basic requirement of content validity however, it is defined after the test is constructed (Allen & Yen, 1979). The ideas of content and face validity bear a seeming similarity, but they are actually very dissimilar. Face validity relates to the extent to which a test seems to measure what it claims to measure (Leedy & Ormrod, 2004). It is a general response to a quick evaluation of what is being evaluated. Thus, it is the easiest and least reliable determination approach that relies on assessor's experience and knowledge of the subject matter (Nwana, 2007). That is, it is usually used to describe the perception of validity without empirical testing and as a result, it is considered as the weakest form of validity (Cook & Beckman, 2006).

Criterion-related validity which is strongly related to the purpose of this study, is the next form of validity. It is used for predicting future or present performance. This correlates test findings with another interest criterion (Burns, Morris, Periard, LaHuis, Flannery, Carretta & Roebke, 2017). It also deals with the relation between test scores and some unique observable parameters such as the DBME and BECE scores in this case. Again, it examines how individuals are differentiated by the test on criterion that it is supposed to predict (Pallant, 2011). That is, when we predict a future output based on the scores obtained by the measure at present by correlating, for example, the scores obtained in the DBME with the performance in the BECE (Messick, 1989). Again there are two forms of the criterion validity; the concurrent validity, and the predictive validity.

Construct validity according to Martella, Nelson, March and Martella cited in Shillingburg (2016) "refers to the skills, attitudes, or characteristics of

individuals that are not directly observable but are inferred on the basis of their observable effects on behavior”. Similarly, Kinyua and Okunya (2014), explained construct validity as measure that seeks to ensure that a test is actually measuring the intended construct and not any other extraneous attributes or constructs. For example, if a chemistry test is crafted using terminology beyond the level of the learner, such a test will be described as having low construct validity because it is measuring other constructs besides the intended chemistry construct.

Concurrent Validity, generally, is the degree to which the scores on a test are compared to the scores on another. It is intended to calculate the same construct of a test administered simultaneously or to some other relevant criterion available at the same time. It is appropriate when a skills assessment test is designed in order to replace a less efficient one in use (Denga, as cited in Mahajan, 2017). It is created by correlating one item with another that has previously been validated an expect (Okoro, 2002). Thus Concurrent Validity examines the validity of a test on highly theoretical bases (Messick, 1989).

Predictive validity is the ability of an assessment tool to predict future performances in a similar or same activity or in another assessment of the same construct (Brown & Coughlin, 2007). DeVellis, (2006) defined predictive validity as the degree at which one can predict another measure of the same or similar construct in the future. Similarly, Rothstein (2004) asserted that predictive validity refers to the degree of correlation between the measure of a construct and some future measure of the same construct. It requires the assessment instrument to have same evidence of the reasonableness of the proposed interpretation. However, Rothstein indicated



that because of the passage of time, the correlation coefficients between the two are likely to be fairly lower for predictive validation studies. That notwithstanding, being a form of criterion validity, it is used to gather evidence in relation to a known criterion. Thus, for a test to have predictive validity, there must be a statistically significant relationship between test scores and the criterion being used to measure (Rothstein, 2004). As a result, the higher the association between the criterion and the predictor variables, the higher the predictive power of the predictor variable.

### **Reliability of test scores**

Reliability is a term widely used by test experts to refer to the consistency of test scores, implying that, all other things being equal, it is the scores that have an error in them, not the test. The concept of reliability of test scores has long been the primary condition to be met in order to produce a sound measurement. The classical test of reliability theory by Parkes (2000), rely on reliability coefficients and standard errors of measurement. The reliability coefficient is the piece of evidence that operationalizes the values of accuracy, dependability, stability, consistency and precision in test scores.

More importantly, the quality of a measurement instrument is described by three quality domains including reliability. In a standardised test, any score obtained by a measuring instrument is the sum of the ‘true score’ which is unknown, and ‘error’ in the measurement process. If the error margins are insignificant and reporting of results of a score is of high standards, the test instrument and the test scores are considered dependable (Kimberlin & Winterstein, 2008). Therefore, the dependability of the test

score to make a valid decision relies so much on the reliability and validity of the scores and the instrument used.

Test scores are one of the ways teachers obtain information about their students when making instructional decisions and for evaluation of student progress. Some of the decisions are so important that teachers and test users cannot afford to assume that the test scores obtained from students are dependable enough. As a result, the extent to which measurement errors compromise the quality of test scores needs to be estimated to decide on whether to rely on the test scores or not (Crocker & Algina, 1986). Validity is often assessed along with reliability. Reliability is a measure of stability, dependability, trustworthiness and consistency of a test measuring the same construct at each given time (Kinyua & Okunya, 2014). Hence, reliability of a test and the criterion place prominence on the possible correlation between the two. Instinctively, this reflects the fact that reliability denotes freedom from random error which does not correlate with one another (Ochieng, 2012).

Similarly, Chakrabarty (2013) posited that reliability is a measure of consistency, precision, repeatability, and trustworthiness of test scores. Thus, it indicates the degree to which it is error free and hence assures consistent measurement across time and the various items on the test. However, some qualitative researchers use the term 'dependability' instead of reliability. Further, with Twycross and Shields (2004), reliability refers to the consistency, stability and repeatability of results, that is, the result of a test taker is considered reliable if similar scores are consistently obtained in identical situations but different circumstances. All the explanations given on the concept of reliability suggest that reliability issues are closely linked to

repeatability of a score of a test administered at different times to the same individuals over a period of time. This by implication suggests that, the more accurate a test score, the higher the probability of making valid decisions about the test taker. However, according to Downing (2004) reliability is a prerequisite but not a sufficient condition for the validity of a test result.

### **Forms of Reliability**

Reliability is categorised into two main types: stability, and internal consistency (reliability). Stability of a test score refers to the ability of a test score to remain similar over a period of time despite uncontrolled testing conditions or the test takers (Mohajan, 2017). Further, it refers to how much a person's score can be expected to vary from one administration of a test to another (Allen & Yen, 1979). Mohajan asserted that a perfectly stable measure will produce exactly the same scores time after time and moreover indicated that the two major methods of measuring stability are test-retest and parallel-forms of a test. Nonetheless, the role of inter-rater reliability cannot be overlooked especially in this study because it always precedes the actual rating of the BECE candidates.

Test-retest reliability is obtained by re-administering the same test for a second time, to verify the similarity of scores obtained by the test taker (Graziano & Raulin, 2006). If the coefficients obtained after the repetition is above .7, it is considered fairly reliable, and coefficients above .8, are considered very good (Sim & Wright, 2005). The method assesses the external consistency of a test. It is determined by the variations in coefficients which also suggest the degree of measurement errors of the test.

Parallel-forms reliability is a measure of reliability attained by administering different forms of an assessment tool to the same group of examinees. The scores obtained from the different forms can then be correlated in order to assess the consistency of the scores across alternate forms. If the scores are highly correlated, then they are referred to as parallel-form reliability (DeVellis, 2006).

Inter-Rater Reliability, according to Keyton, King, Mabachi, Manning, Leonard, and Schill (2004) refers to the extent and consistency to which information about a student is gathered by multiple raters. It establishes the uniformity of scores obtained with an instrument when used by different observers. The Inter-Rater reliability is useful in psychological testing because human observers will not essentially interpret answers the same way; consequently, raters may disagree as to how well certain responses or material demonstrates knowledge of the construct or skill being assessed. For example, before rating the BECE and WASSCE scripts by the raters engaged by WAEC in Ghana, the raters are made to rate a ‘dummy’ script to rate and to ensure consistency among the raters before the actual rating is commissioned.

Internal Consistency Reliability, on the other hand, is a measure of reliability used to assess the degree to which different test items probe the same construct to produce similar results. That is, it examines whether or not the items within a particular construct are homogeneous (DeVellis, 2006). It is often established in one testing situation, thus it eliminates the problems associated with repeated testing found in other reliability estimates (Allen & Yen, 1979).

The most common internal consistency measure is the Cronbach's alpha ( $\alpha$ ), which is typically interpreted as the mean of all possible coefficients. The coefficient obtained is a function of the average inter-correlations of items, and the number of items in the scale (Mohajan, 2017). The reliability coefficient classically varies between 0 and 1, where 0 shows no relationship among the items on a given scale, and 1 indicates absolute internal consistency (Tavakol & Dennick, 2011). Alpha values above 0.7 are generally considered acceptable and satisfactory, above 0.8 are usually considered quite well, and above 0.9 are considered to reflect excellent internal consistency and often observed in professionally prepared standardised tests.

Split-half reliability is another measure of internal consistency which measures the degree of consistency by examining one half of the results of a set of scores against the other half (Ganesh, 2009). It involves only one administration, and especially appropriate when the test is very long. It is done by relating the results of one half of a test with the results from the other half. A test can be split into halves in several ways, for example, first half and second half, or by odd and even numbered items. If the two halves of the test provide analogous results, this would suggest that the test has internal reliability. It is a quick and easy way to establish reliability (Mohajan, 2017). Interestingly, it provides a simple remedy to the problem that the parallel form faces. However, the split-half method is not appropriate for tests which measure different constructs at the same time (Chakrabartty, 2013).

Cronbach's alpha ' $\alpha$ ' reliability is another form of internal reliability estimate which is generally used as a measure of consistency of a test score for

a sample of examinees. It was first named as alpha by Lee Cronbach in 1951, as he had intended to continue with further coefficients. The measure can be viewed as an extension of the Kuder-Richardson formula (KR20), which is an comparable measure of dichotomous items. Cronbach had given a formula to find out the reliability by using standard deviations of the scores not only of odd and even items but also the total number of items as follows.

$$\text{Cronbach's } \alpha = 2[1 - (\sigma_{\text{odd}}^2 + \sigma_{\text{even}}^2)] \sigma_{\text{total}}^2$$

However, according to the Technical White Paper Series (2007), there is an apparent confusion among some researchers concerning the calculation of internal consistency reliability for dichotomous response items whose cumulative sum forms the scale score for a test. It is indicated that some researchers suggest that the Kuder – Richardson (KR– 20) coefficients must be used instead of a Cronbach alpha. This is entirely incorrect because the KR–20 is mathematically equivalent to the formula for coefficient alpha. What appears to be perplexing these individuals is that the formula for the variance of a binary item variable looks quite dissimilar to that of a continuous valued variable. (Technical White Paper Series, 2007). The formulae for coefficient and KR – 20 are presented below.

The KR-20 formula is:

$$KR_{20} = \frac{k}{k-1} \left( 1 - \frac{\sum_{i=1}^k p_i q_i}{\sigma_{test}^2} \right) \quad \text{where:}$$

$k$  = number of items in the test

$p_i$  = the proportion of respondents answering an item  $i$  in the keyed direction [1]

$q_i$  = the proportion of respondents answering an item  $i$  in the non-keyed direction [0]

$\sigma_{test}^2$  = the test-score variance

The Cronbach alpha formula is:

$$\alpha = \frac{k}{k-1} \left( 1 - \frac{\sum_{i=1}^k \sigma_i^2}{\sigma_{test}^2} \right) \quad \text{where:}$$

$k$  = number of items in the test

$\sigma_i^2$  = the variance of item  $i$

$\sigma_{test}^2$  = the test-score variance

Lastly, the Rulon Method is an alternate method for finding Split- Half reliability and it was developed by Philip J. Rulon in 1939. It requires only the variance of differences between each student's scores on the two half- test ( $SDd^2$ ) and variance of total scores ( $SDx^2$ ). These two values are substituted in the following

$$\text{formula: } r_{tt} = 1 - \frac{SDd^2}{SDx^2}$$

Where,  $r_{tt}$  = Reliability of the test

$SDd$  = SD of difference of the scores

$SDx$  = SD of the scores of the whole test

Table 1-Reliability Coefficients (Cronbach's alpha)

Reliability	Interpretation
.90 and above	Excellent reliability; at the level of the best standardized tests
.80 - .90	Very good for a classroom test
.70 - .80	Good for a classroom test; in the range of most. There are probably a few items which could be improved.
.60 - .70	Fairly low. This test needs to be supplemented by other measures (e.g., more tests) to determine grades. There are probably some items which could be improved.
.50 - .60	Suggests need for revision of test, unless it is quite short (ten or fewer items). The test definitely needs to be supplemented by other measures (e.g., more tests) for grading.
.50 or below	Problematic reliability. This test should not contribute heavily to the course grade, and it needs revision

Source:

[http://www.washington.edu/oea/services/scanning\\_scoring/scoring/item\\_analysis.html](http://www.washington.edu/oea/services/scanning_scoring/scoring/item_analysis.html).

### Factors that can simultaneously affect validity and reliability of a test

According to Kinyua and Okunya (2014), there are three notable factors that could influence the validity and reliability of a formative assessment: the test, the environment and the test taker. They indicated that one of the measures that is used to determine the characteristics of the test-taker is the use of a statistical measurement called a confluence score. This is where scores of an individual test taker is paired with other scores of the



individual of the same or similar tests to determine the consistency and reliability of the test taker which can affect the validity and reliability of the tests. For example, a teacher may design a test so that a percentage of the questions would be asked seeking the same information in a different way, and it is expected that the student's responses to those questions should be consistent with the earlier questions asked. A reliable test taker under such situation should consistently score the item(s) right or wrong (Kinyua & Okunya, 2014).

Another factor that affects the validity and reliability of a test is the testing environment. According to Griswold (1990), if the testing environment is distracting, noisy or has a poor lighting system, the test-taker will find it difficult to remain consistent throughout the testing process. These factors, according to Griworld, are largely aspects of test administrative procedures that are external to the test itself. This is because even in the absence of these factors, individual differences in performance will still exist. This means that, the single most important factor that affects the validity and reliability of formative assessment is the quality of tests themselves. As a result, the test length or number of items as well as the use of table of specification and carefully written tests should be taken seriously since they usually provide a representative sample and less distorted scores.

### **Test Evaluation Parameters (quantitative)**

Commercialisation of psychological tests is gradually gaining grounds in Ghana especially at the education level. This could be attributed to the inception of the DBME. The DBME is a form of teacher-made test and whether the strict adherence to the principles of test construction, test

administration, analyses and reporting are observed by developers of the DBME is a source of worry to stakeholders in the education industry because the reliability and validity of these tests are not fully known. As a result, it is imperative for teachers to be well versed in testing techniques to enable them evaluate the validity and reliability of tests as well as student's progress (Quaigrain & Arhin 2017).

Consequently, there have been several attempts to identify a model for test evaluation performance prediction, but many studies advocate a detailed analysis that covers other independent variables and approaches (Zahn Saguil, Artino, Dong, Servey & Durning, 2012). Generally, a student's test performance is evaluated in a variety of ways in the classroom. Very often instructors conduct a test and rate it for instructional purposes (Kargar, Tarmizi, & Bayat, 2010). In such circumstances, item analysis can help in identifying potential mistakes in scoring, ambiguous items, and defective distractors that did not function well on the test. Mehrens and Lehmann (as cited in Amedahe & Asamoah-Gyimah, 2016), defined an item analysis as the process of examining students' responses to each item on a test to evaluate the quality of the items. Thus, it is a statistical technique that helps instructors identify the effectiveness of their test items (Shakil, 2008).

McCowan and McCowan (2012) in their book highlighted the procedure and the importance of conducting item analysis. They advocate that an item analysis assists a test developer in the process of test quality enhancement which characteristically includes endorsement, improvement and elimination of a particular item. Problematic items due to ambiguous wording and for being wrongly keyed are reviewed based on the calculated difficulty

index and discrimination coefficient values to improve the overall quality of the test (Shakil, 2008).

### **Item Difficulty**

Statistically, P-values are used to determine difficulty levels of test items by measuring the proportion of students who answer the item correctly. P-values range from 0.00-1.00; items with p-values closer to 0.00 are considered difficult items while high p-values indicate the item is easy. In some cases, items will have a high p-value and a low point-biserial value. Such information reveals that the test item is problematic and is not a good fit, indicating something was wrong with the wording which made the student to answer the item wrongly, even though the same student answered more difficult items correctly (Schwarz, 2011).

According to Ding and Beichner (2009) item difficulty is a measure of easiness of an item. The item difficulty index is one of the most useful and most frequently reported item analysis indices. It is a measure of the fraction of the test takers who answer the item correctly; for this reason, it is frequently called the p-value. Thus, for the proportion of test takers that got the item right, the p-value might more properly be called the item easiness index, rather than the item difficulty. It ranges between 0.0 and 1.0 with a higher value indicating that a greater proportion of examinees responded to the item correctly and vice versa.

For ideal validity and reliability, the optimal item difficulty level is 0.50. However, this does not mean every item should have a difficulty level of 0.50, simply that the average of all items should be around .50. Optimally, an item will encourage a widespread distribution of scores if its difficulty

index is approximately 0.5 (i.e. 50% of the students got it right). Table 2 provides the acceptable ranges of difficulty index and their corresponding interpretations.

Table 2-*Difficulty Index and its Corresponding Interpretations*

Percentage Range	Difficulty index	Interpretation
75% - 100%	0.75 – 1.0	Easy
26% - 74%	0.25 – 0.75	Average
0% - 25%	0.25 or below	Difficult

Source: Ding and Beichner (2009)

### **Item Discrimination Index**

The Discrimination index is also another measure which assesses the quality of test items included in an assessment. According to Amedahe and Asamoah-Gyimah (as cited in Asamoah & Ocansey, 2019), the discrimination power of a test item is the capacity of the item to differentiate between those who know (the upper group) and those who do not know (the lower group). They further indicated that the purpose of the discrimination index is to tell the evaluator if an item really is showing differences between knowledgeable students and less knowledgeable ones. To estimate the discrimination index, the upper and lower 27% rule is commonly used based on Kelley's (1939) derivation. The difference between the correct responses as a percentage of the upper 27% and lower 27% of the total group is used to determine whether an item effectively discriminated the high scorers from the low scorers of a particular test.

The discrimination indices obtained after the estimation could be positive, negative, or non-discriminatory. If test items discriminate positively,

it means the high scorers more often answered the items correctly than the low scoring students did. On the other hand, a negative discrimination is realised when low scoring students answer the question correctly more often than those students who score highly on the assessment as a whole. Non-discriminating items are those items that failed to discriminate between the high and low scorers of the test. Similarly, Quaigrain and Arhin (2017) asserted that item discrimination indices must always be interpreted in the context of the type of test which is being analyzed. They further indicated that items with low discrimination indices are often ambiguously worded and should be examined to determine why a negative value was obtained. In such situations, test item adjustments may be required in order to make test items better. Additionally, items with low item discrimination scores should be reviewed and transformed to make them more functional (Hong, Purzer, & Cardella, 2011). The discrimination index is expressed as;

$$D = \text{Discrimination index } T (\%) - B (\%) / (T+B) \times .5$$

T (27) = proportion of examinees getting the item correct from the top group (e.g., top 27%)

B (27) = proportion of examinees getting the item correct in the bottom groups (e.g., bottom 27%).

Test items with discrimination indices of 0.40 and above are excellent test items; between 0.30-0.39 are good, but could possibly use adjustments; between 0.11- 0.29 are fair and may need adjustments; and 0.11 and below are poor items that must be rejected or altered (Popham, 2008). Table 3 provides the recommended discrimination index and their respective interpretations.

Table 3-*Discrimination Index and its Respective Interpretations*

Discrimination Index	Interpretation
0.4 and above	Excellent
0.30 – 0.39	Good
0.11 – 0.29	Fair
0.00 – 0.10	Poor
Negative values	Flawed or miskeyed

Source: Kelley (1939)

### **Students' BECE performance evaluation**

In the absence of any pre-existing criteria for determining student performance in the BECE, the Monitoring and Evaluation Unit of the Ministry of Education under the guidance of international measures of performance and underperformance categorized students' performance in the BECE into 'above average', 'average' 'and below average' and considered it a suitable approach for assessing students' performance. Subsequently, a minimum level of achievement is pegged and any student falling below or above this level is deemed as having underperformed or performed in the examination.

Students' performance in the BECE is currently measured by categorizing the outcome of the test result as above average, average and below average. A candidate is considered to have passed the BECE if he or she obtained an aggregate score of 5-25 in the four core subjects and one best performing subject. In the BECE scoring methodology, a score of 1 is deemed excellent and 5 is average. The national pass rate for BECE in 2010 is 59%, which forms a minimum benchmark. Thus any pupil scoring higher than 5 in any subject is deemed to be a below average performance. According to this

measure, any school in which less than 3 in every five pupils pass BECE is considered to be underperforming (MoE, 2011).

Again, according to the MoE (2011) report, there is another measure of success for the BECE candidates which is a little more stringent qualification criterion which is used to determine eligibility (though not placement) in SHS. The criteria require an aggregate score of 5 to 25 in the four core subjects and a fifth best performing subject but not worse than a score of 6 in any core subject. The benchmark value for qualification rate is 49% and thus any school in which less than one in every two pupils qualifies for SHS is deemed to be underperforming. As a result, this study adopted the criteria in assessing the predictive relationship between the DBME and the SBA on students' performance in the Basic Education Certificate Examination (BECE).

## **Review of Empirical Studies**

### **Test quality (quantitative) evaluation**

Conducting item analysis is essential in test development, evaluation and improving upon items which are to be kept in item bank for future use. It can also be used to eliminate misleading items in a test. In multiple test evaluation, item statistics (difficulty, discrimination index) for each item is assessed quantitatively to determine the quality of the test items (Quaigrain, & Arhin, 2017).

In conjunction with some teachers of a senior high school in the northern region of Ghana on the subject of test evaluation using item analysis, Asamoah and Ocansey (2019), launched a study which aimed at providing a thorough analysis on item discrimination and distractor functionality of a thirty multiple choice mathematics (core) achievement test. The test was

conducted at an unnamed senior high school in the Northern Region of Ghana. The test was purposed to measure senior high school students' achievement in mathematics (core) after the following areas: linear equations, linear inequalities, simultaneous equations, percentages, vectors and indices have been taught successfully by their respective classroom teachers. The items were crafted by the teachers who taught the areas mentioned above and for each item, four options were provided for the students to select the best option as an answer. Analysis of the items after the test revealed that 27 out of the 30 items analyzed discriminated positively, 2 discriminated negatively and 1 failed to discriminate between the high and low scoring groups.

Relatedly, Quaigrain and Arhin (2017), in their study which focused on using reliability and item analysis to evaluate the quality of teacher-made test explored the relationship between item difficulty index (*p-value*) and discrimination index (DI) as well as the distractor efficiency (DE), administered fifty multiple-choice questions as an end of semester examination in Educational Measurement course to 247 first-year students pursuing Diploma in Education at the Cape Coast Technical University. According to Quaigrain and Arhin, the internal consistency reliability of the test yielded 0.77 per the Kuder–Richardson (KR-20) coefficient. The mean score was 29.23 with a standard deviation of 6.36. Mean difficulty index (*p*) value and DI were 58.46% (SD 21.23%) and 0.22 (SD 0.17), respectively. From their report, it was observed that, 30 (60%) of the test items were reasonably good or fell within acceptable value ranges of 0.2–0.49.

Similarly, a study conducted by Bichi (2015), on the use of item analysis to evaluate the quality of multiple choice chemistry test items for 530



senior secondary school two (SSII) students was subjected to the classical analysis using a sample of 530 students who sat for Kano State Qualifying Examination in July 2014. The study revealed that, out of the 40 items on the test, 12 (30%) items performed below the set standard for item difficulty and discrimination index while 28 (70%) of the items met the set standards. Also, the study reported a significant positive correlation between item difficulty and item discrimination indices. Juxtaposing the findings of Asamoah and Ocansey (2019), Quaigrain and Arhin (2019) and Bichi (2015) on the report of their studies, it is suggestive that, highly experienced teachers as well as teachers who are trained in test construction and analysis can indeed craft good quality test items as against the general perception that teacher-made tests are unreliable (Magno, 2003).

#### **Predictive power of mock examination**

The predictability of examinations especially high stake examinations has been of great concern to scholars of educational measurement (Ochieng, 2012). Afolabi and Faleye, (as cited in Ochieng, 2012) revealed increasing concern among stakeholders about the predictive validity of the state version of the JSCE for the Senior Secondary School Certificate Examination (SSCE) which necessitated a study on the predictive validity of Osun State Junior Secondary Certificate Examination (JSCE). The study was to explore the relationship between the overall performance of students in the JSCE and their performance in the Senior Secondary School Certificate Examinations (SSSCE). The sampled schools were drawn from Osun State, Nigeria. Promotion examination scores of the students in Senior Secondary School (SSS) one and SSS two as well as their SSSCE in six major subjects were

correlated with the corresponding JSCE scores using correlation analysis procedures. The results showed that, majorly, Osun State JSCE is a poor prediction of students' performance in the SSCE. However, the JSCE English language and Mathematics were found to have a greater capacity to predict performance in SSCE English language and other subjects ( $r=0.32$ ,  $p<0.05$  and  $r=0.22$ ,  $p<0.005$  respectively). They concluded that, the overall performance in JSCE across the six subjects investigated is a poor predictor of SSCE performance (except English and Mathematics). The trend according to the Nigerian Ministry of Education (NMoE) could be attributed to regulatory challenges blaming departments in-charge of examinations in the state for not ensuring that the psychometric properties of the predictor tests were standardized. Further, the test results could not correlate because the testing conditions for the predictor and the main examination were not homogeneous. Likewise, Leonard (2004) reported that conditions under which mock exams are administered in some schools are stricter than the final examinations and may influence the predictability of the internal examination scores on external examination scores.

Relating their study to other research on predictive validity of internal examination, Omirin and Ale (2008) investigated the predictive validity of English and Mathematics Mock Examination results of senior secondary school students' performance in WASSCE in Ado-Ekiti State, Nigeria. Three hundred and sixty students were selected by a simple random sampling technique from twelve public secondary schools in six local government areas of Ado-Ekiti, Ekiti State, Nigeria. The study made use of secondary data of the results of WASSCE results of students and unprocessed raw scores of mock

examinations from selected schools. It was observed that the English and Mathematics Mock results significantly predicted the performance of students in WASSCE. Nonetheless, the English language result was a better predictor of the WASSCE as compared with the Mathematics scores. They consequently recommended that Mock examination should be made compulsory for WASSCE candidates because of its predictability in the WASSCE. The study corroborated a research by Adeyemi, (2008) which was purposed to examine the predictive strength of the Junior Secondary Certificate (JSC) Examinations on students' performance in the Senior Secondary School Certificate Examinations (SSSCE) in Ondo State and it concluded that JSC Examination was a good predictor of students' performance in the SSSCE.

Similarly, Bosson-Amedenu (2017) conducted a study to investigate the predictive validity of Mathematics Mock examination scores of students in WASSCE and BECE in Ghana. The study made use of One hundred and sixty-four (164) students who were conveniently sampled from 15 public secondary schools in Ghana for the study. He adapted the Expo-facto design as it made use of an already existing data. The data used for the study were the results of WASSCE, BECE and unprocessed raw scores of mock examination from the selected schools. The study revealed that there was no significant difference between the WASSCE and Mock grades in Core Mathematics. Thus, the Mock for Core Mathematics was found to have a high predictive power of 92% in the WASSCE. Also, the study indicated that there was no significant difference between BECE and WASSCE grades in mathematics. However, similar relationship could not be established between the BECE and

Mock for Mathematics grades because a significant difference between the two examinations was also established. Nonetheless, he concluded that there is a relationship between Mock examinations and national examinations. Therefore, mock examinations should be given the necessary attention and regulated to harness its predictive power in national or final examinations. Regrettably, his conclusion is not generalizable because he did not justify why he adapted the convenient sampling technique instead of a randomised sampling technique which is a recommended technique for such studies.

Again, Komba, Kafanabo, Tryphone and Kira (2013) examined the predictive validity of Form Two Secondary Education Examination (FTSEE), a form of mock examination before the Certificate of Secondary Education Examination (CSEE) in Biology. They made use of a sample size of 120 students from some selected secondary schools in Morogoro Municipality in Tanzania. They correlated the data collected using the Pearson's Product-Moment Correlation ( $r$ ) to determine the strength, direction and significance of the relationships between the predictor and the criterion variables for the study. The findings revealed that there was a strong relationship between the students' performance in the FTSEE and CSEE (i.e. from  $r=0.442$ ,  $p<0.01$  to  $r=0.726$ ,  $p<0.01$ ). They consequently affirmed the claim that mock examinations should be made compulsory for school candidates since they strongly correlated with the CSEE. Interestingly, they could not explain why Form Two students were used instead for final year students. Final year students are those often made to write mock examination because it is expected that they would have completed their syllabus and are well prepared for final examination.

Additionally, Andala, Digolo and Kamande (2014), conducted a study with the aim of determining whether the mock examinations are reliable enough for predicting the results of the Kenya Certificate of Secondary Education examinations (KCSE). The study, however, analysed data using Pearson's R instead of ordinal regression which is the recommended statistical for such data. The study made use of questionnaire and was administered to a sample of 65 secondary schools that epitomized all the categories of schools and the results analysed both quantitatively and qualitatively. The study revealed a high positive correlation (0.949) between the mock and KCSE examinations results. Thus, indicating that the quality assurance indicators that guide examinations in most schools in the study area were observed during the preparation and administration of mock examinations. The study also suggested that mock examinations are reliable in predicting KCSE examinations and advocated that there is the need to harmonize the structures for setting, moderation and administration of mock examinations to make it more prototype of the KCSE or final examinations.

Contrary to the findings suggesting a predictive relationship between mock and state examinations, a papers presented by Merwe (1999) at the Association for Educational Assessment in Africa (AEAA) 17<sup>th</sup> annual conference at Lusaka, Zambia, expressed dissatisfaction about the low correlation between teacher-made tests and the national examination scores. Lewin (1997) attributed the high correlation observed between mock examination and final examination to poor scoring standards of the mock examination. However, this study disagrees with Lewin because the

administration and scoring of the mock examination is sometimes stricter than that of the final examination as indicated by Leonard (2004).

### **Predictive power of the SBA**

In a study to investigate the predictive power of school based assessment scores on students' achievement in Junior Secondary Certificate Examination (JSCE) in English and Mathematics by Okpala in Nigeria as well as that of Afemikhe and Anyanwu (2012), revealed that the combination of the SBA scores with the final examination score significantly predicted students' performance in English and Mathematics achievement in JSCE. The study adopted the ex-post facto research design and tested two hypotheses at 0.05 level of significance. Also, a sample of 250 students were randomly drawn from ten (10) schools out of twenty (20) junior secondary schools in Obio-Akpor Local Government Area of Rivers State. The researchers adopted the descriptive survey design as the data already existed. Also, the researchers could have considered the application of ordinal regression to analyze the data since the independent variables and the dependent variables were continuous and ordinal in nature.

SBA scores in any subject for a particular student should reflect their achievement in the external examination (Orubu 2013; Ogunkola 2007). Their assertion was attributed to the studies they conducted to investigate the SBA scores as predictors of students' final grades in Delta and Ogun States respectively. Interestingly, they found that SBA scores predicted students' performance in JSCE Mathematics and Integrated Science respectively.

On the contrary, Omole (as cited in Opara, Onyekuru & Njoku, 2015) conducted a comparative study of students' performance in SBA and

certificate examination in the Federal Capital Territory (FCT) and reported that students' scores in SBA were significantly higher than their scores in certificate examination in English and Mathematics at the upper basic education. The discrepancies between Omole's and Okpala, Afemikhe and Anyanwu studies, is what Amedahe (as cited in Amedahe, 2001) pointed out in an article indicating that, the determination of total term score for a student is not standardized at the school level.

This section of the study reviewed related literature based on the concept of the study, which assumed that Students' performance in external examinations should relate with their performance in internal examinations covering the same content areas. The review on quality of teacher made test showed that the quality indicators, thus, reliability, difficulty and discrimination indices revealed a low reliability co-efficient although the difficulty and discrimination indices were quite good. This could be as a result of inadequate test construction skills by the teachers as indicated by Amedahe, (2001).

Generally, the literature on mock examinations as posited by Komba, Kanabo, Tryphone and Kira (2013); Andala, Didolo and Kamande (2014) and Bosson-Amedenu (2017); suggested a positive relationship between mock and final examinations despite concerns raised by Merwe (1999) and claimed that the positive relationship between mock and final examinations could be attributed to poor scoring in mock examinations. This study disagreed with Merwe because prior to marking of a well-organized mock examination such as the DBME, the raters are trained and dummy scripts subjected to inter-rater reliability before the actual marking/scoring commences. Also, the findings of

Afemikhe and Anyanwu (2012) on the relationship between SBA and final examination scores was corroborated by the 2012 SBA guidelines which indicated that grades students obtained in final examinations must be based on their performances in school assessment. It is therefore expected that the findings of this study will relate with the findings of (Afemikhe and Anyanwu 2012; Komba et. al (2013); Andala et. al (2014); and Bosson-Amedenu 2017).





## CHAPTER THREE

### RESEARCH METHODS

#### Introduction

The rationale for this study was to assess the predictive relationships between internal assessment scores (SBA & DBME) and students' performance in the BECE in the KEEA Municipality. This section, describes the methodology employed by the researcher for the collection and analysis of data to answer the research questions and test the hypotheses respectively in order to explore the predictive relationships between the variables; SBA, DBME and the BECE. The chapter has been categorized into the following subtopics: research design, study area, population, sampling technique, data collection instrument and data collection procedure and data analysis.

#### Research Design

The Expost- facto design was employed for this research because according to Maduabum, (2004), the ex-post facto correlation utilizes secondary data and that fits the purpose of this study. In this study, the design helped in assessing the predictive relationships between the internal assessment scores (SBA and DBME) and students' performance in the BECE because the data already exist and fit the purpose of the study. Also, an important use of the ex-post facto correlation design is prediction, thus, if a relationship of sufficient magnitude exists between two variables, it becomes possible to predict a score of a variable if the score of the other variable is known (Fraenkel & Wallen, 2009). The basic design of ex-post facto

correlation research is simple and involves collecting two or more already existing data on a group of subjects and computing correlation coefficients. Relationships observed from the co-efficient may simply indicate what goes with what in a given context, and may provide a basis on which to make predictions about the variables being studied (Iddrisu, 2009).

Also, a major strength of correlational research is it helps to explain our understanding of important phenomena in human behaviour by identifying relationships among variables. Many relationships have been established by researchers in education and psychology by analyzing relationships among variables using correlation. Further, studies in other jurisdictions have revealed that school level assessment scores are highly related to final examination grades. Hence, school level assessment scores can be used to predict candidates' grades in high stake/national test. Deductively, SBA and DBME scores could equally be used as predictors of the BECE. The predictor variables in this study, therefore, shall be the SBA and DBME scores while the BECE performance is the criterion variable.

The limitation of this design, however, is that correlational research only identifies what goes with what, which implies that it identifies only an association that exists between two variables and does not necessarily establish cause-effect relationships. It is also less rigorous because it exercises less control over the independent and the confounding variables (Cohen, Manion & Morrison, 2007).

### **Study Area**

The Komenda-Edina-Eguafo-Abrem (KEEA) Municipality was carved out of the Cape Coast Metropolis in 1988 and was elevated to a Municipality

status in 2008 in pursuance to LI 1857 with Elmina as Municipal Capital (Ghana Statistical Service, 2014). Statistics from the Municipal Education Directorate indicate that, there is a total of 378 schools in the Municipality from the pre-school up to the tertiary level under both public and private ownership. The second cycle schools are made up of five (5) senior high schools. At the tertiary level, the Assembly has one College of Education, in Komenda, and a polytechnic (Archbishop Potter Polytechnic) in Elmina.

Table 4-*Number of Schools by Ownership*

Schools	Pre-Sch.		Prim. Sch.		JHS		SHS		Voc./Tech		Tertiary	
	Pub	Priv.	Pub	Priv.	Pub	Priv.	Pub	Priv.	Pub	Priv.	Pub	Priv.
No. of sch.	74	62	74	59	71	40	3	2	0	2	1	1
Total	136		133		111		5		2		2	

Source: GES Directorate –KEEA-MA, Beginning Term Data as of 31st September, 2015

### Population

The target population for this study comprised all BECE candidates from the 8 circuits in the KEEA Municipality who sat for the 2019 DBME and BECE conducted by the KEEA and WAEC respectively. A total number of three thousand one hundred and thirty-three (3,133) candidates from a total of one hundred and eleven (111) schools sat for the 2019 BECE in the Municipality. The population are a group of students who had received eleven (11) years of basic education from GES approved basic schools. The average age of the candidates was sixteen (16) years and the vast majority of them come from rural and peri-urban areas of the municipality. Out of the target population, an accessible population of seventy-one (71) public schools made

up of two thousand five hundred and forty-five (2,545) candidates was used for this study.

### Sample and Sampling Procedure

The multistage stage sampling technique comprising the cluster, proportionate and simple random sampling techniques were adopted by this study to sample the participating schools and students. The cluster sampling technique was firstly used to purposely select the participating public schools which were already grouped into circuits. This was followed by proportionate sampling technique to equitably select the participating schools from each circuit. This was achieved by determining a sampling fraction for the selection of the participating schools in all the eight (8) circuits within the KEEA Municipality as follows:

$$S.F = \frac{n}{N}$$

Where SF = sampling fraction

n= number of circuits in the municipality

N = total number of public schools in the municipality.

Therefore, a sampling fraction of  $\frac{8}{71} = 0.11$  was used to multiply the number of schools in each circuit to obtain the number of school(s) needed per circuit. This was followed by numbering each of the schools in a circuit, folding and placing of numbered copies in a box, then thoroughly mixed and placed in box to randomly sample the participating schools. The procedure yielded a random sample of fourteen (14) schools out of the 71 public schools in the KEEA municipality. All the fourteen (14) schools, consisting of the six hundred and two (602) candidates out of two thousand five hundred and forty-five (2,545) were used in the final sample. The sample represented 24% of the population

of candidates for the study. The sample size obtained according to Nwanna (2007) as well as Krejcie & Morgan (1970) was adequate because it fell within the recommended percentage of (10%) for a population of “few thousand” and 330 for a population of 2600 respectively. Also, the 2019 DBME tests on English language, Mathematics, Integrated Science and Social Studies were re-administered to forty (40) students from the Elmina M.A school. The school was conveniently sampled because it is located in a peri-urban area and has mixture of the characteristics of the target population.

### **Data Collection Instrument**

The instrument used to collect data for this study was an inventory adapted from (Adeyemi, 2008) and the 2019 DBME tests (English language, Mathematics, Integrated Science and Social Studies). The inventory specifically requested secondary data on the 2019 SBA/ DBME and BECE results on the following core subjects; English language, Social Studies, Mathematics and Integrated Science. Also the tests in these four subjects were conveniently re-administered to 40 candidates from the Elmina MA school for the quality analysis of the test. The choice of the subjects was based on the fact that it is a requirement for candidates to write and pass these subjects before they are placed in any of SHS selected by the candidate, in addition, the national performance assessment of the BECE is based on these four subjects and any one best performed subject by a candidate (MoE, 2011).

### **Ethical Consideration**

Since consent, respect and moral issues are essential in Social Science research the necessary steps were taken to address ethical issues related to the study. Participants consent and respect for both participants and or moral

issues are essential in social research. Thus, ethical concerns hinged on informed consent, anonymity and confidentiality. Informed consent presents to the prospective participants the opportunity to participate or decline to participate in the study. To achieve this, a clearance letter was secured from the Institutional Review Board (IRB) and subsequently an introductory letter to the Department of Education and Psychology to seek the consent of the KEEA Municipal Director of Education. (see Appendices B, C and D). Seidman (as cited in Ahenkora, 2019), indicated that informed consent affords the participants the need to be educated on the aims, objectives and potential harm of the study. Consequently, the purpose of this study was carefully reviewed with the KEEA Municipal Directorate of Education for their consent before they released the required data for the study. More importantly, the anonymity of participants was also taken into consideration because in social research it is unequivocally essential since it gives the participants the opportunity to give the required information when they are assured of concealed identity (Oliver, 2010). To ensure anonymity, information obtained from the participating schools and students were coded using Arabic numerals. Further, other relevant information including citations and referencing related to this study was duly acknowledged both citation and referencing as required in academic writing.

### **Data Collection Procedure**

The researcher personally visited the KEEA Municipal Directorate of Education to collect data on students 2019 BECE results, as well as the SBA and DBME scores. The Municipal Examination Officer assisted in collecting the data from the selected schools using the results obtained from multiplying

the sampling fraction by the number of schools in the respective circuits. The Examination Officer was later interviewed to ascertain the source of the DBME test instrument and the challenges they encountered with the planning and administration of the test. Additionally, in order to assess the quality of the DBME, the multiple choice aspect of the tests was conveniently re-administered to the BECE candidates in Elmina M.A Basic School within the Elmina Circuit of the Komenda Edina Eguafo Abrem Municipality.

The official endorsement of the study by the then Acting Director of the Education Directorate in KEEA with the able support of the Examination Officer which helped facilitate the selection of schools and collection of data from all the eight (8) circuits yielded 602 participants for the study. The sample represented 24% of the population and this was considered adequate as indicated earlier by Nwanna (2007) and Krejcie & Morgan (1970) respectively.

#### **Data Processing and Analysis**

The data obtained were analysed with reference to the research questions and hypotheses posed tested respectively for the study. The variables for the study were the SBA, DBME scores and students' BECE grades respectively. While item analysis was used to evaluate the data obtained from research questions one to four, ordinal regression was used to test the hypotheses of the study.

Since some candidates were absent during the DBME or the BECE, the data obtained for the DBME and BECE were cleaned in order to avoid an incomplete data. After the cleaning, coded data were entered into the *Itman* and *SPSS* software for analysis.

For research questions one through to four, the test items for each test together with the keys for the various items were entered into the itman software to run the item analysis. This was done after adopting the default settings of the itman programme which defined item discrimination index,  $D_i$ , as the difference,  $p_H - p_L$ , where  $p_H$  is the item difficulty in the highest scoring (top 27%) group and  $p_L$  is the item difficulty in the lowest scoring (bottom 27%) group as recommended by Crocker and Algina (1986). Thus, the quality of the of the DBME was evaluated to answer research questions one to four using item discrimination and difficulty indices.

Further, the scores obtained from the DBME, SBA and BECE were used to determine the predictive relationship between the independent and dependent variables. The dependent variable was students' grades in the 2019 BECE and was also ordinal in nature with 9 levels (1= highest, 2=highest, 3= high, 4=high average, 5= average, 6=low average, 7= low, 8= lower, 9 lowest). These nine performance domains were hierarchically structured. Those who scored the average and above grades (1 – 5) were coded 1, and those who obtained grade six (6) and above were assigned 0 as code. After the coding, the variables were fitted into the ordinal regression model and the results from the POM were examined to determine predictive powers of the independent variables. Since the outcome variables were ordinal in nature, they were categorized and coded as;

90 -100	Highest	1
80 -89	Highest	2
70 – 79	High	3
60 – 69	High Average	4



55 – 59	Average	5
50 – 54	Low Average	6
40 – 49	Low	7
35 – 39	Lower	8
0 – 34	Lowest	9

before inputting them into the ordinal regression model to test the hypotheses for the study at 0.05 level of significance.

### **The Ordinal regression model**

The ordinal regression model is a member of the family of regression analyses. The model allows its users to build and generate predictions as well as evaluate the influence of various parameters in the model after meeting the key assumption of the model, thus, the assumption of proportional odds which assumes that the effect of the explanatory variables is the same irrespective of the thresholds (Quansah, 2013). The assumption was tested using the SPSS. As a predictive analysis tool, the ordinal regression describes data and explains the relationship between one dependent variable and two or more independent variables. Although there are other forms of the ordinal regression model, a study conducted by Adejumo, and Adetunji (2013) in a study of Students' Performance from the Faculty of Science, University of Ilorin for 2011/2012 academic session asserted that, the Proportional Odds Model (POM) is often the preferred model when conducting a predictive studies involving two or more continuous predictor variables and ordinal outcome variable.

Proportional Odds Model (POM): The POM for ordinal logistic regression provides a useful extension of the binary logistic model to

situations where the response variable takes on values in a set of ordered categories. The model may be represented by a series of logistic regressions for dependent binary variables, with common regression parameters reflecting the proportional odds assumption. Key to valid application of the model is the assessment of the proportionality assumption (Brant, 1990). Again, the proportional odds model is based on the assumption that the effect of the covariates  $x_1, x_2, \dots, x_p$  are the same for all categories, on the logarithmic scale.

The model is fundamentally expressed as;

$$\text{link}[\theta_k(X_i)] = \theta_k + [\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_q X_q]$$

Where

Link is the link function

K is the number of cut points

$\theta_k(X_1)$  is the cumulative logit up to and including category k

$\theta_k$  is the threshold up to and including category

i is the number of regression parameters

$X_1, X_2, \dots, X_q$  are the predictor variables

### **Inferences about the Proportional Odds Model (POM)**

In the POM, inferences are made about the regression parameter estimates with reference to the test of parallel lines, Wald test, Goodness of fit and model fit to test for the appropriateness of the model.

#### **The test of parallel lines**

Ordinal regression is equal at each threshold because the ordinal regression model has constrained it to be so per the PO assumption. As a result, the appropriateness of the proportional odds can be evaluated using the test of parallel lines. The test compares the ordinal model which has one set of

coefficients for all thresholds to a model with a separate set of coefficients for each threshold. If the general model gives a significantly better fit to the data than the ordinal (proportional odds) model (i.e. if  $p < .05$ ), then we are led to reject the assumption of proportional odds which implies the effect of the independent variables is same on the dependent variable (Quansah, 2013).

### **Goodness of fit test**

The goodness of fit test contains Pearson's chi-square statistic for the model as well as another chi-square statistic based on the deviance. These statistics are intended to test whether the observed data are consistent with the fitted model. They evaluate the null hypothesis to ensure the model fit is good. If we do not reject this hypothesis (i.e. if the p value is large), then we can conclude that the data and the model predictions are similar and that we have a good model. On the other hand, if we reject the assumption of a good fit, conventionally if  $p < .05$ , then the model does not fit the data well.

### **Parameter estimates**

The parameter estimates summarize the effect of each regressor. The positive or negative signs of the coefficients for the predictors may indicate the contribution of the predictors in the model. According to Quansah (2013), for covariate, positive (negative) coefficient indicates positive (inverse) relationship between predictors and outcome. Thus, an increasing value of a covariate with a positive coefficient corresponds to an increasing probability of being in the highest category. He further suggested that odd ratio estimates which can be calculated directly by the formula  $p / (1-p)$  gives a relatively simple meaning to the parameter estimates.

In summary, this chapter described the research methodology employed for this study. Specifically, the Expost- facto design was adopted for this study because the data already existed and fit the purpose of the study which was to evaluate the predictive relationships between the SBA, DBME and students' performance in the BECE. Consequently, the study adopted the multistage sampling to select 602 BECE candidates for the study. The next chapter presents the results and discussion of the study.



## CHAPTER FOUR

### RESULTS AND DISCUSSION

#### Introduction

This chapter presents the results, and discussed the findings of the study with reference to the objectives, research questions and hypotheses of the study. The main objective of the study which was aimed at evaluating the quality and predictive relationship between internal assessment (SBA and DBME) scores and students' performance in the 2019 BECE at the KEEA Municipality was sub-divided into five. The first was to use item analysis to evaluate the quality of the DBME test items and the second through to the fifth was to assess the predictive relationships between the internal assessment scores and students' performance in the BECE in social studies, integrated science, mathematics and English language.

To achieve the objectives of the study, Ex-post facto design was adopted for the study. As indicated in chapter 3, the ex-post facto correlation design was employed because it best fits the purpose and answers the research questions of the study since all the required data already exist. A period of six months was used to gather the full set of data required for the study.

#### Analysis of results

In this section, the results of the item analysis (item difficulty, discrimination index and distractor analysis) of the DBME test items in English, Mathematics, Integrated Science and Social Studies were presented and evaluated first, followed by the predictive relationship between the

DBME, SBA and the BECE in English language, Mathematics, Integrated Science and Social Studies respectively.

After scoring the test, the scored items were subjected to item analysis, an important phase in test development process. The analysis was carried out using the *iteman* software. This was done to evaluate the difficulty of the test items, discrimination index, as well as the reliability of the test and also to identify any item(s) that did not function as expected, thus, whether the item was too easy or too difficult to differentiate between the upper and the lower groups respectively. The two most common statistics reported in an item analysis are the item difficulty, which is a measure of the proportion of examinees who responded to an item correctly, and the item discrimination which is a measure of how well the item discriminates between examinees who are knowledgeable in the content area and those who are not. An additional analysis that may be reported is the distractor analysis. The distractor analysis provides a measure of how well each of the incorrect options contributes to the overall quality of the test instrument.

### **Research Questions One**

**What are the qualities (quantitative) of the 2019 DBME English Language test held at KEEA?**

To start with, the English language paper of the 2019 DBME of the KEEA was administered to a total number of forty (40) candidates. The mean score out of thirty (30) items was ( $M12.475$ ) representing 51.583% and a standard deviation of ( $SD 2.784$ ) as indicated in Table 5. The mean score for the English language suggested that the performance of the candidates in the test was “Low Average” per the grading system of the Ghana Education

Service (KEEA, 2019). Additionally, the item difficulty and discrimination indices suggested that the test was averagely difficult and fairly discriminated between the high scorers and low scorers. However, the reliability co-efficient of the DBME test did not satisfy the reliability co-efficient of teacher-made test which ranged between 0.7 and 0.8 (DIIA, 2003).

Table 5-Summary of English Language DBME test Parameters

Item (alpha)	No.	M	SD	KR20
English lang.	30	12.475	2.784	0.258
Difficulty		0.416		
Discrimination Index		0.194		

Source: Field survey (2020)

Specifically, the English language test results as indicated in Table 6 showed that 19 out of the 30 items performed with a difficulty index ranging between 0.25 and 0.75. However, items number; 2 and 8 were found to be above the recommended difficulty index level of 0.75 and as such were found to be very easy items for the candidates. Furthermore, as many as eight (8) of the items (3,6,7,17,20,23,15 and 30) were found to be very difficult with item seven being the most difficult item with a difficulty index of 0.07.

Also, to find out how well the items discriminated between the high scorers and the low scorers, the discrimination indices as captured in Table 3 were applied. According to the results, seven (7) of the items (2,4,5,10,18,24 and 25) representing (23.3%) fairly discriminated the high scorers from the low scorers as shown in Table 6. In addition, two (2) of the items (16 and 17) representing (6.7%) were found to have poorly discriminated between the high and low scorers while eight (8) of the items (6,7,8,13,15,20,23 and 29)

representing (26.7%) of the items were found to be flawed by negatively discriminating between the high and low scorers.

Table 6-Item Difficulty and Discrimination Values of the English Language Test

Item No.	Item Diff.	Item Discr	Item No.	Item Diff	Item Discr
01	0.68	0.44	016	0.57	0.02
02	0.78	0.29	017	0.13	0.06
03	0.13	0.33	018	0.35	0.11
04	0.33	0.19	019	0.33	0.39
05	0.63	0.17	020	0.20	-0.01
06	0.15	-0.08	021	0.35	0.66
07	0.07	-0.14	022	0.55	0.52
08	0.85	-0.07	023	0.10	-0.08
09	0.50	0.46	024	0.35	0.11
010	0.33	0.11	025	0.35	0.26
011	0.47	0.39	026	0.72	0.72
012	0.75	0.30	027	0.53	0.53
013	0.53	-0.04	028	0.60	0.51
014	0.55	0.45	029	0.35	-0.10
015	0.15	-0.08	030	0.13	0.33

### Research Questions Two

**What are the qualities (quantitative) of the 2019 DBME Mathematics test held at KEEA?**

Table 7 shows the distribution of difficulty and discrimination indices of the DBME mathematics items. Thirty-three (33) of the items which had a representation of (82.5%) were found to be of suitable level of difficulty with *p-value* ranging from 2.5–7.5, while seven (7) of the items (1,2,6,18,19,35 and 37) representing (17.5 %) were found to be difficult, thus, falling below a *p-value* of 2.5. Likewise, majority (80%) of the items discriminated positively between the high and low scoring group. However, items 2,4,7,19,34 and 35 recorded negative discrimination indices which signify that the low performing students had the item right rather than the ‘knowledgeable ones’.



Nevertheless, test items with negative discrimination indices are undesirable and should be reviewed or removed from the test.

Table 7-Item Difficulty and Discrimination Values of the DBME Mathematics Test

Item No.	Item Diff.	item Disc.	Item No.	Item Diff.	Item Disc.
1.	0.14	0.50	21	0.57	0.32
2.	0.06	-0.08	22	0.66	0.28
3.	0.51	0.77	23	0.80	0.21
4.	0.89	-0.02	24	0.71	0.18
5.	0.46	0.49	25	0.29	0.35
6.	0.23	0.42	26	0.54	0.69
7.	0.29	-0.03	27	0.51	0.49
8.	0.34	0.02	28	0.37	0.45
9.	0.54	0.04	29	0.46	0.47
10.	0.29	0.17	30	0.60	0.06
11.	0.60	0.52	31	0.66	0.24
12.	0.29	0.55	32	0.66	0.62
13.	0.63	0.26	33	0.57	0.14
14.	0.54	0.32	34	0.26	-0.03
15.	0.51	0.04	35	0.14	-0.05
16.	0.37	0.27	36	0.74	0.18
17.	0.69	0.28	37	0.23	0.15
18.	0.20	0.32	38	0.66	0.44
19.	0.17	-0.15	39	0.54	0.24
20.	0.69	0.52	40	0.69	0.54

Source: Field survey (2020)

Further, the mean score of the forty (40) mathematics test items was ( $M19.086$ ) and a standard deviation of ( $SD 4.889$ ) as indicated in Table 8 below. The mean score for the test suggests that the overall performance of the candidates in the DBME (mathematics) represents (47.72%) indicating a ‘low grade’ performance as shown in the grading system of the Ghana Education Service (KEEA, 2019).

Table 8-Summary of the Mathematics DBME test Parameters

Items	No.	M	SD	KR20
Analyzed	40	19.086	4.889	0.670
Difficulty Index		0.477		
Discrimination Index		0.279		

Source: Field survey (2020)

### Research Questions Three

#### What are the qualities (quantitative) of the 2019 DBME Integrated Science test held at KEEA?

The item analysis for the DBME science test in Table 9 also shows that twenty-two 22 (55%) of the items fell within the recommended difficulty level with a range of 2.5 – 7.5. Moreover, while items 15 and 30 were found to be difficult, the rest excluding the twenty-two (22) items as indicated above appeared to be easy items. Also, with the exception items 22, 23, 35 and 39 which recorded negative discrimination indices, the vast majority (87.5%) of the items discriminated positively which means most of the students in the upper group answered the items correctly. On the other hand, items 34 and 40 recorded zero discrimination index which implies either the items were too difficult or ambiguous to the high performing students. Consequently, such items should be reviewed for subsequent usage or deleted from the test.

Table 9-Item Difficulty and Discrimination Indices of the DBME Int. Science Test

Item No.	Item Diff.	Item Disc.	Item No.	Item Diff.	Item Disc.
1.	0.67	0.33	21	0.85	0.40
2.	0.82	0.40	22	0.27	-0.03
3.	0.76	0.32	23	0.64	-0.28
4.	0.33	0.67	24	0.88	0.02
5.	0.94	0.20	25	0.67	0.07
6.	0.85	0.22	26	0.70	0.33
7.	0.30	0.48	27	0.85	0.30
8.	0.61	0.33	28	0.58	0.73
9.	0.48	0.18	29	0.82	0.03
10.	0.64	0.60	30	0.21	0.15
11.	0.76	0.60	31	0.67	0.23
12.	0.91	0.20	32	0.33	0.48
13.	0.52	0.47	33	0.45	0.92
14.	0.79	0.40	34	1.00	0.00
15.	0.12	-0.12	35	0.52	-0.18
16.	0.88	0.30	36	0.42	0.65
17.	0.52	-0.18	37	0.97	0.10
18.	0.58	0.17	38	0.52	0.17
19.	0.82	0.40	39	0.48	-0.37
20.	0.73	0.42	40	1.00	0.00

The mean score out of forty (40) for the Integrated Science test items was ( $M$  25.818) with a standard deviation of ( $SD$  4.475) as indicated in Table 10. The mean score for the science test representing (64.545%) suggested that the performance of the candidates in the test was “High Average” as indicated in the grading system of the Ghana Education Service (KEEA, 2019)

Table 10-Summary of the Integrated Science DBME Test Parameters

Item (Alpha)	No.	M	SD	KR20
Analyzed	40	25.818	4.475	0.660
Difficulty Index		0.645		
Discrimination Index		0.253		

Source: Field survey (2020)

### Research Questions Four

#### What are the qualities (quantitative) of the 2019 DBME Social Studies test held at KEEA?

Table 11 presents the results of the Social Studies test. The results indicate that, ten (10) of the items were difficult while 9 were easy. Thus, 19 (47.5%) of the items were problematic with regards to the difficulty level. Also, as many as 23 (57.5) of the items poorly discriminated between the low and high performing groups. In addition, out of the poor discriminating items, items 6, 8,10, 13, 14, 23, 24, 25, and 27 showed negative discrimination. Nonetheless, thirty-one (31) of the items representing (77.5%) showed positive discrimination which indicates that candidates in the high performing group did better on the test than the low performing candidates.

Table 11-Item Difficulty and Discrimination Indices of the DBME Social Studies Test

Item No.	Item Diff.	Item Disc.	Item No.	Item Diff.	Item Disc.
1.	0.14	0.15	21	0.17	0.15
2.	0.23	0.05	22	0.86	0.40
3.	0.40	0.08	23	0.31	-0.11
4.	0.66	0.31	24	0.97	-0.06
5.	0.80	0.50	25	0.54	-0.04
6.	0.23	-0.05	26	0.17	0.25
7.	0.43	0.46	27	0.20	-0.11
8.	0.60	-0.02	28	0.89	0.14
9.	0.09	0.13	29	0.63	0.29
10.	0.17	-0.01	30	0.37	0.30
11.	0.80	0.34	31	0.20	0.38
12.	0.69	0.47	32	0.14	0.19
13.	0.54	-0.04	33	0.46	0.33
14.	0.97	-0.06	34	0.77	0.11
15.	0.57	0.45	35	0.20	0.05
16.	0.69	0.15	36	0.83	-0.03
17.	0.29	0.30	37	0.29	0.15
18.	1.00	0.00	38	0.46	0.42
19.	0.80	0.18	39	0.26	0.40
20.	0.77	0.11	40	0.49	0.10

Table 12 indicates the Social Studies items had a mean score of ( $M20.057$ ) representing 50.1% and a standard deviation of ( $SD 3.430$ ) as shown in Table 12. The mean score for the Social Studies test suggest that the performance of the candidates in the DBME was “Low Average” as indicated in the grading system of the Ghana Education Service (KEEA, 2019). Also, the item difficulty and discrimination indices suggested that the test was averagely difficult and fairly discriminated between the high scorers and low scorers as shown in Table 12.

Table 12-Summary of Social Studies DBME Test Parameters

Item (Alpha)	No.	M	SD	KR20
Analyzed	40	25.818	4.475	0.417
Difficulty Index		0.501		
Discrimination Index		0.165		

Source: Field survey (2020)

### Hypothesis One

$H_0$ : The students’ scores in the DBME and SBA do not significantly predict their performances in the 2019 BECE in English Language at KEEA.

$H_1$ : The students’ scores in the DBME and SBA do not significantly predict their performances in the 2019 BECE in English Language at KEEA.

### Descriptive statistics

From Table 13, it could be seen that (1.3%) of the students scored grade 2, (2.8%) scored grade 3, (8.5%) scored grade 4 and (17.6%) scored grade 5, which is considered as average or pass. Cumulatively, (30.2%) of the

602 candidates passed the BECE in English language while 69.8% failed the test. From the results obtained in the 2019 BECE, the performance of the candidates is not significantly different from what situation had been over the past years. In fact, the result is very similar to what occurred in 2014 in the district. Consequently, the result suggests that the introduction of the SBA and the DBME have not had any positive impact on students' performance in the BECE in the KEEA Municipality.

Table 13-*Frequency Distribution for the 2019 BECE Performance in English Language in KEEA*

BECE ENGLISH	GRADE	N	Marginal Percentage
	HIGHEST	8	1.3%
	HIGH	17	2.8%
	HIGH AVERAGE	51	8.5%
	AVERAGE	106	17.6%
	LOW AVERAGE	145	24.1%
	LOW	125	20.8%
	LOWER	117	19.4%
	LOWEST	33	5.5%
Total		602	100.0%

Source; Field survey (2020)

Table 14 shows the Model Fitting Information. Before examining the impact of each explanatory variable in the model, a model fitting test was conducted to determine whether the model advances the ability to predict the outcome. This was done by comparing the model without any explanatory variables (the baseline) against the model with all the explanatory variables (Final) to see whether it significantly fits into the data. The model fitting information in Table 14 gives the -2 log-likelihood values for the baseline and

the final model. The chi-square test performed to determine the difference between the -2LL for the two models yielded  $\chi^2(2, N=602)=28.37, p<.05$ . The significant chi-square statistic ( $p<.0005$ ) indicates that the final model gives a significant improvement and gives better predictions of the outcome variable.

Table 14-Model Fitting Information of the 2019 BECE (English Language)

Model	-2 Log Likelihood	Chi-Square	Df	Sig.
Intercept Only	182.014			
Final	153.644	28.370	2	.000

Source: Field survey (2020)

Also, a Chi-square test of goodness – of – fit was performed to determine whether the explanatory variables fits the model well by equally predicting the outcome variable. The test result in Table 15 shows that the test fits the model,  $\chi^2(33, N=602) = 36.9, p >.05$ . meaning the test fit the model adequately.

Table 15-Goodness-of-Fit for the 2019 BECE (English Language)

	Chi-Square	df	Sig.
Pearson	36.892	33	.294
Deviance	40.571	33	.171

Table 16 presents the test for the assumption of proportional odds otherwise referred to as the test of parallel lines in the ordinal regression model, the most important assumption in the model. It also indicated that, the model is good ( $p > .05$ ) so the study failed to reject the null hypothesis. Thus, the proportional odds assumption gives a significantly better fit than the general model.

Table 16-Test of Parallel Linesa for the 2019 BECE (English Language)

Model	-2 Log Likelihood	Chi-Square	Df	Sig.
Null Hypothesis	153.644			
General	141.165	12.478	12	.408

Source: field data

Also, Table 17 provides the parameter estimates which is the score of all the output in the model. It indicates specifically the relationship between the independent and the dependent variables. The output provides coefficients, their standard errors, the Wald test and associated p-values, the 95% confidence interval and odds ratios which together indicate the strength of relationship and probability of predicting students' performance in the BECE. Table 17, indicates that there is a significant negative relationship between the DBME and BECE. Specifically, the DBME recorded an ordered *log-odds* = (-1.37), *SE* = .26, *Wald* = 27, *p* < .001. Thus, the estimated odd ratio indicates a significant negative relationship of approximately  $Exp^{-1.370} = (.25)$  fold, 95% *CI* (.152, .426) compared to the SBA which did not record a significant relationship with the BECE. Thus, the odds of entering the high performing category (average and above) decrease by 0.76 for each unit increase in the DBME score.



Table 17-Parameter Estimates of the Predictive Powers of the DBME and the SBA on BECE (English Language)

Parameter	B	Std. Error	95% Wald		Hypothesis Test			Exp(B)	95% Wald Confidence	
			Confidence Interval		Wald Chi-Square	df	Sig.		Interval for Exp(B)	
			Lower	Upper					Lower	Upper
[BECEENGACT=2]	-6.136	.5825	-7.278	-4.995	110.983	1	.000	.002	.001	.007
[BECEENGACT=3]	-4.957	.5016	-5.940	-3.974	97.646	1	.000	.007	.003	.019
[BECEENGACT=4]	-3.719	.4666	-4.634	-2.805	63.537	1	.000	.024	.010	.061
Threshold [BECEENGACT=5]	-2.572	.4481	-3.450	-1.694	32.942	1	.000	.076	.032	.184
[BECEENGACT=6]	-1.524	.4399	-2.387	-.662	12.007	1	.001	.218	.092	.516
[BECEENGACT=7]	-.576	.4390	-1.436	.285	1.720	1	.190	.562	.238	1.329
[BECEENGACT=8]	1.179	.4623	.273	2.085	6.500	1	.011	3.250	1.313	8.042
DBMESCOREFORENG	-1.370	.2633	-1.886	-.854	27.072	1	.000	.254	.152	.426
SBAENGSCORE	-.068	.1051	-.274	.139	.412	1	.521	.935	.761	1.149

**Hypothesis Two**

H<sub>0</sub>: Students’ scores in the DBME and SBA do not significantly predict their performances in the 2019 BECE in Mathematics at KEEA.

H<sub>1</sub>: Students’ scores in the DBME and SBA significantly predict their performances in the 2019 BECE in Mathematics at KEEA.

**Descriptive statistics for students’ performance in the 2019 BECE**

Table 18 shows the percentage of students’ performance in mathematics in the 2019 BECE at KEEA Municipal Assembly. The results indicate that, cumulatively, only 72 (12%) out of 602 candidates who sat for the 2019 BECE scored above the average grade while 74 (12.3%) scored average. On the other hand, the majority of the students (456) representing (75.7%) performed below the average (pass) grade.

Table 18-*Performance of Students in Mathematics in the 2019 BECE at KEEA*

BECE MATHEMATICS GRADE	N	Marginal Percentage
HIGHEST	7	1.2%
HIGH	16	2.7%
HIGH AVERAGE	49	8.1%
AVERAGE	74	12.3%
LOW AVERAGE	135	22.4%
LOW	156	25.9%
LOWER	108	17.9%
LOWEST	57	9.5%
Total	602	100.0%

From Table 19, it can be seen that the difference between the two log-likelihoods and the chi square has an observed significance level of less than 0.0005. This means that there are 34 (35.4%) cells (i.e., dependent variable levels by combinations of predictor variable values) with zero frequencies.

Therefore, the result justifies the rejection of the null hypothesis that the model without predictors is as good as the model with the predictors.

Table 19-Model Fitting Information for the 2019 BECE (Mathematics)

Model	-2 Log Likelihood	Chi-Square	Df	Sig.
Intercept Only	281.596			
Final	235.900	45.695	2	.000

Source: Field survey (2020)

The *goodness – of - fit* result in Table 20 reveals that, the test fits the model very well,  $\chi^2(75, N=602) = 77.1, p = .41$ . This means that the impact of the independent variables is equally distributed to predict the outcome of the test.

Table 20-Goodness-of-Fit for 2019 BECE (Mathematics)

	Chi-Square	Df	Sig.
Pearson	74.606	75	.491
Deviance	77.135	75	.410

Source: Field survey (2020)

Also, Table 21 examined the assumption of the test of parallel lines. The test indicates that the proportional odds assumption gives significantly better fit to the data than the general model. Thus,  $p = .20$  as a result, the study failed to reject the assumption of proportional odds.

Table 21-Test of Parallel Linesa for the 2019 BECE (Mathematics)

Model	-2 Log Likelihood	Chi-Square	Df	Sig.
Null Hypothesis	235.900			
General	220.037b	15.863c	12	.198

Table 22 indicates that there is a systematic predictive relationship between the independent and dependent variables. The DBME (Mathematics) was found to contribute to the model very well. The ordered *log-odds* =  $(-.70)$ ,  $SE = .11$ ,  $Wald = 38.0$ ,  $p = < .001$ . The estimated odds ratio indicates a negative relationship between the DBME and the BECE of approximately  $Exp\ -.70 = (.50)$  fold,  $95\% CI (.033, .156)$  compared to the SBA. Thus, the odds of being entered into the high performing category is decreased by .50 for each unit increase in the DBME score. Conversely, the odds of entering into the high performing category (average and above) increases by 2 for each unit increase in the DBME score.

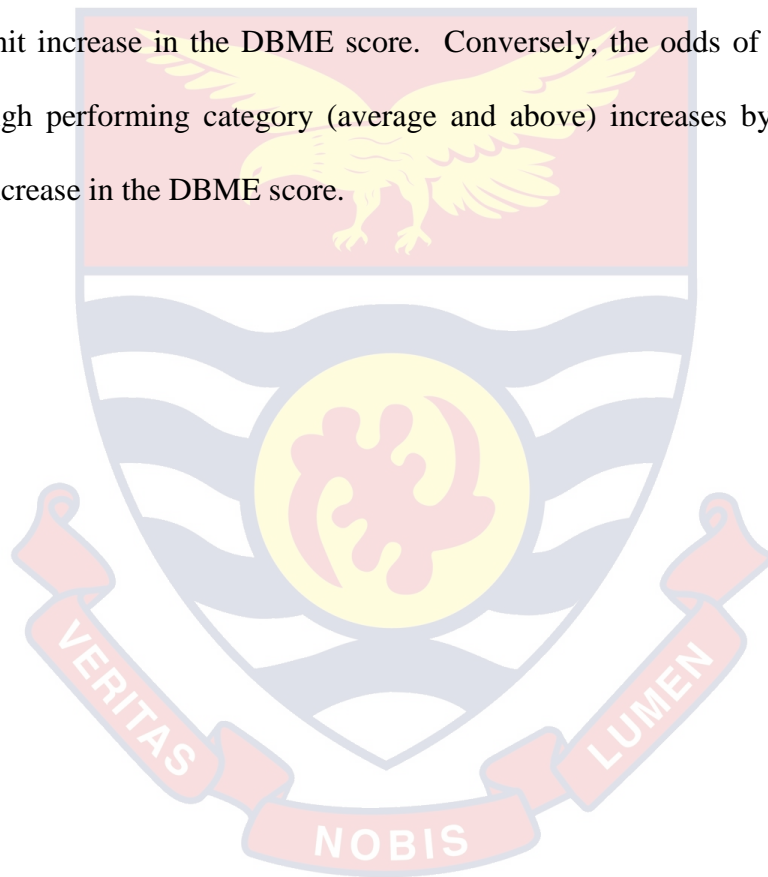


Table 22-Parameter Estimates of the Predictive Powers of the DBME and the SBA on BECE (Mathematics)

Parameter	B	Std. Error	95% Confidence Interval		Wald Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	Df	Sig.		Lower	Upper
[BECEMATACT=2]	-6.046	.5530	-7.130	-4.962	119.522	1	.000	.002	.001	.007
[BECEMATACT=3]	-4.810	.4521	-5.697	-3.924	113.216	1	.000	.008	.003	.020
[BECEMATACT=4]	-3.542	.4134	-4.352	-2.731	73.402	1	.000	.029	.013	.065
Threshold [BECEMATACT=5]	-2.639	.3999	-3.423	-1.856	43.564	1	.000	.071	.033	.156
[BECEMATACT=6]	-1.579	.3904	-2.344	-.814	16.352	1	.000	.206	.096	.443
[BECEMATACT=7]	-.422	.3860	-1.178	.335	1.194	1	.275	.656	.308	1.398
[BECEMATACT=8]	.891	.3957	.115	1.666	5.067	1	.024	2.437	1.122	5.292
DBMEMATSCORE	-.699	.1131	-.920	-.477	38.143	1	.000	.497	.398	.621
SBAMATSCORE	-.108	.1113	-.326	.111	.934	1	.334	.898	.722	1.117
(Scale)	1 <sup>a</sup>									

Source: Field survey (2020)

### Hypothesis Three

H<sub>0</sub>: Students' scores in the DBME and SBA do not significantly predict their performances in the 2019 BECE in Integrated Science at KEEA.

H<sub>1</sub>: Students' scores in the DBME and SBA significantly predicts their performances in the 2019 BECE in Integrated Science at KEEA.

### Descriptive Statistics

The performance of the candidates in the 2019 BECE in Integrated Science as shown in Table 23 indicates that a vast majority 457 (75.9%) of the candidates performed below average while only 74 (12.3%) and 71 (11.8%) candidates obtained average and above average grades respectively. Indeed, only 1 (0.2%) candidate out of the 602 scored grade 1 in the subject.

Table 23-Frequency Distribution in 2019 BECE Performance in Integrated Science in KEEA

BECE INT. SCIENCE GRADE	N	Marginal Percentage
HIGHEST	1	0.2%
HIGHEST	6	1.0%
HIGH	12	2.0%
HIGH AVERAGE	52	8.6%
AVERAGE	74	12.3%
LOW AVERAGE	113	18.8%
LOW	153	25.4%
LOWER	133	22.1%
LOWEST	58	9.6%
Total	602	100.0%

Source: Field survey (2020)

Table 24 shows the outcome of the model fitting test. From the table, it could be observed that the final two log-likelihoods yielded a  $p < .0005$ . This justifies the rejection of the null hypothesis that the model without predictors is as good as the model with the predictors.

Table 24-Model Fitting Information for the 2019 BECE (Integrated Science)

Model	-2 Log Likelihood	Chi-Square	Df	Sig.
Intercept Only	248.012			
Final	215.344	32.668	2	.000

Source: Field survey (2020)

The Chi-Square *Goodness-of-Fit* test for the Integrated Science data indicates that the test fits the model adequately well,  $\chi^2(62, N=602) = 68.1, p = .30$ . This means that the impact of the independent variables is equally distributed to predict the outcome of the test.

Table 25-Goodness-of-Fit for the 2019 BECE (Integrated)

	Chi-Square	Df	Sig.
Pearson	116.700	62	.000
Deviance	68.088	62	.278

Source: Field survey (2020)

The result of test for the assumption of proportional odds does not provide justification for the rejection of the assumption. The test indicates that the assumption gives significantly ( $p > .05$ ) better fit to the data as indicated in Table 26.

Table 26-Test of Parallel Linesa for the 2019 BECE (Integrated Science)

Model	-2 Log Likelihood	Chi-Square	Df	Sig.
Null Hypothesis	215.344			
General	192.342 <sup>b</sup>	23.002 <sup>c</sup>	14	.060

The parameter estimate result in Table 27 indicates that there is a systematic relationship between the independent and dependent variables. Both the DBME and the SBA (Integrated Science) were significant negative predictors of the BECE. For the DBME, an ordered *log-odds of* = (-.66), *SE* = .14, *Wald* = 21.20, *p* < .001 and for the SBA, the *log-odds* was = (-.31), *SE* = .11, *Wald* = 8.40, *p* < .001. However, while the estimated odds ratio indicates a negative relationship between both the DBME and the SBA, the exponent of the DBME *Exp* -.66 = (.52) fold, 95% *CI* (.39, .69) and SBA *Exp* -.31 = (.73) fold, 95% *CI* (.59, .90) indicating that the probability of entering the high performing category (average and above) decreases by .52 and .73 for each unit increase in the DBME and the SBA score respectively.



Table 27-Parameter Estimates of the Predictive Powers of the DBME and the SBA on BECE (Integrated Science)

Parameter	B	Std. Error	95% Confidence Interval		Wald Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
[BECESCIACT=1]	-8.472	1.0885	-10.606	-6.339	60.579	1	.000	.000	2.477E-005	.002
[BECESCIACT=2]	-6.501	.5701	-7.619	-5.384	130.035	1	.000	.002	.000	.005
[BECESCIACT=3]	-5.461	.4807	-6.403	-4.519	129.075	1	.000	.004	.002	.011
[BECESCIACT=4]	-4.010	.4314	-4.855	-3.164	86.410	1	.000	.018	.008	.042
[BECESCIACT=5]	-3.115	.4175	-3.933	-2.297	55.670	1	.000	.044	.020	.101
[BECESCIACT=6]	-2.230	.4096	-3.032	-1.427	29.625	1	.000	.108	.048	.240
[BECESCIACT=7]	-1.146	.4042	-1.938	-.354	8.045	1	.005	.318	.144	.702
[BECESCIACT=8]	.353	.4112	-.453	1.159	.739	1	.390	1.424	.636	3.188
DBMESCISCORE	-.655	.1422	-.933	-.376	21.182	1	.000	.520	.393	.687
SBASCISCORE	-.314	.1083	-.526	-.102	8.414	1	.004	.730	.591	.903
(Scale)	1 <sup>a</sup>									

Source: Field survey (2020)

**Hypothesis four**

H<sub>0</sub>: Students’ scores in the DBME and SBA do not significantly predict their performances in the 2019 BECE in Social Studies at KEEA.

H<sub>1</sub>: Students’ scores in the DBME and SBA significantly predicts their performances in the 2019 BECE in Social Studies at KEEA.

**Descriptive Statistics**

The outcome of the 2019 BECE in Social Studies as shown in Table 28 indicates that cumulatively, only 52 (8.7%) performed above average while 70 (11.6%) fell into the average category. Moreover, the results show that, out of the 602 candidates who sat for the 2019 BECE in Social Studies, 480 (79.7%) of them obtained grades below the average grade.

Table 28-*Performance of students in Social Studies in the 2019 BECE at KEEA*

BECE SOCIAL STUDIES GRADE	N	Marginal Percentage
HIGHEST	1	0.2%
HIGHEST	6	1.0%
HIGH	17	2.8%
HIGH AVERAGE	28	4.7%
AVERAGE	70	11.6%
LOW AVERAGE	132	21.9%
LOW	106	17.6%
LOWER	130	21.6%
LOWEST	112	18.6%
TOTAL	602	100.0%

From the results presented in Table 29, the model fits the data well and the final log-likelihoods yielded a  $p < .0005$ . This provides enough evidence to support the rejection of the null hypothesis that the model without predictors is as good as the model with the predictors.

Table 29-Model Fitting Information for the 2019 BECE (Social Studies)

Model	-2 Log Likelihood	Chi-Square	Df	Sig.
Intercept Only	320.108			
Final	265.205	54.903	2	.000

Link function: Logit.

The *goodness-of-Fit* test of the model as revealed in Table 30 indicates that, the Pearson Test marginally fits the model,  $\chi^2(78, N=602) = 98.6, p = .06$ . compared to the deviance  $\chi^2(78, N=602) = 106.5, p = .02$ . This means that the results are fairly mixed and does happen sometimes.

Table 30-Goodness-of-Fit for the 2019 BECE (Social Studies)

	Chi-Square	Df	Sig.
Pearson	98.752	78	.056
Deviance	106.501	78	.018

Source: Field survey (2020)

From the results of the test of parallel lines as indicated in Table 31, the test indicates a non-significance  $p = .24$  which means that the assumption is satisfied.

Table 31-Test of Parallel Linesa for the 2019 BECE (Social Studies)

Mode	-2 Log Likelihood	Chi-Square	Df	Sig.
Null Hypothesis	265.205			
General	247.891	17.314	14	.240

Source: Field survey (2020)

Table 32 shows the regression coefficients and significance test for the DBME and the SBA for Social Studies in the model. From the parameter estimates Table, the DBME was a significant negative predictor of interest into the high performing category. The estimates yielded an ordered *log-odds of* = (-.84), *SE* = .12, *Wald* = 51.40, *p* = < .0001. Furthermore, the estimated odds ratio of the significant predictor (DBME) *Exp* -.84 = (.43) fold, 95% *CI* (.34, .54) indicating that for every one unit increase of the DBME score, there is a predicted decrease of .43 in the log odds of being in the category of high performing students in the 2019 BECE.

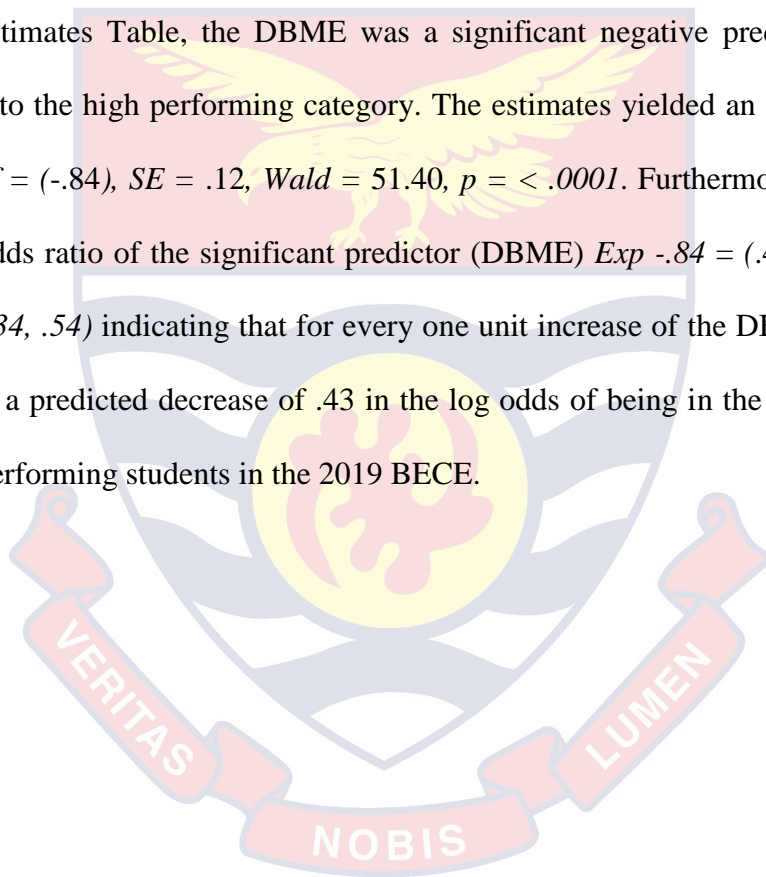


Table 32-Parameter Estimates of the Predictive Relationship of the DBME and the SBA on BECE (Social Studies)

Parameter	B	Std. Error	95% Confidence Interval		Wald Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)		
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper	
Threshold	[BECESOACT=1]	-8.023	1.0868	-10.153	-5.893	54.499	1	.000	.000	3.897E-005	.003
	[BECESOACT=2]	-6.056	.5678	-7.168	-4.943	113.732	1	.000	.002	.001	.007
	[BECESOACT=3]	-4.769	.4672	-5.684	-3.853	104.161	1	.000	.008	.003	.021
	[BECESOACT=4]	-3.910	.4382	-4.769	-3.051	79.631	1	.000	.020	.008	.047
	[BECESOACT=5]	-2.862	.4191	-3.684	-2.041	46.658	1	.000	.057	.025	.130
	[BECESOACT=6]	-1.743	.4093	-2.546	-.941	18.146	1	.000	.175	.078	.390
	[BECESOACT=7]	-.997	.4069	-1.795	-.200	6.004	1	.014	.369	.166	.819
	[BECESOACT=8]	.126	.4075	-.672	.925	.096	1	.757	1.135	.510	2.522
DBMESCOSCORE		-.840	.1173	-1.070	-.611	51.359	1	.000	.432	.343	.543
SBASOCScore		-.058	.1066	-.267	.151	.300	1	.584	.943	.765	1.163
(Scale)		1 <sup>a</sup>									

Source: Field survey (2020)

## Discussion

The findings of this study which sought to evaluate the quality (quantitative) of the District Based Mock Examination (DBME) and compared the predictive relationship between the DBME and the School Based Assessment (SBA) on the Basic Education Certificate Examination were discussed under the following sub-headings;

1. Qualities (quantitative) of the DBME test items in the four core subjects (English Language, Mathematics, Integrated Science and Social Studies)
2. Predictive relationship between the DBME and the BECE in the selected core subjects; and
3. Predictive relationship between the SBA and the BECE in the selected core subjects
4. Students' Performance in the 20019 BECE at KEEA

### Qualities (quantitative) of the DBME Test (core) Items

Juxtaposing the item difficulty indices of Ding and Beichner (2009) and Kelley (1939) with the indices obtained from the analysis of the DBME, the findings of the study revealed that, the item difficulty level of all the four different tests considered for this study had indices ranging from 0.25 – 0.75. This suggests that the tests were adequately difficult. However, according to Kelley, to maximize validity and reliability of a test, the optimal item difficulty level should be 0.50. This does not necessarily mean that every item should have a difficulty level of 0.50 but simply that the average of all items should be 0.50.

Also, the analysis shows that, the item discrimination indices for the DBME tests were within the acceptable range as recommended by Kelley

(1939) and corroborated by Popham, (2008). They asserted that test items with discrimination index of 0.40 and above are good; between 0.30-0.39 are reasonable, but could possibly use adjustments; between 0.20-0.29 are marginal and may need adjustments; and 0.19 and below are poor items that must be rejected or altered. The seemingly adequate indices recorded by the test could be attributed to the fact that, the test was constructed by a commercial producer and had gone through the recommended procedure required for constructing a valid and reliable test.

Ironically, the reliability (KR20) of the test items, which measured the internal consistency of the test to determine the extent to which the items provided consistent information about the students' level of knowledge of the test, provided contrasting information compared to the indices of the item difficulty and discrimination indices. From the results of the four tests (English Language, Mathematics, Integrated Science and Social Studies), it was found that none of the tests met the internal consistency criteria of .70 - .90 generally considered acceptable for a formative assessment as suggested by DIIA (2003). Specifically, the English language test showed a KR 20 reliability measure of (0.26), while Mathematics, Science and Social Studies showed reliability measures of (0.67), (0.66) and (0.42) respectively. The English Language test had comparably the poorest internal consistency measure while the Mathematics test had a relatively better measure than the rest of the tests. The result is an indication that the tests especially the English language and the Social Studies tests did not relate well with the content domain(s) measured. The poor levels of the reliability co-efficient could be attributed to the short length of the tests. For example, the analysis of the

Mathematics test revealed that to obtain a KR-20 Reliability of .80, the test must be 1.97 times longer to yield a total of 79 items of similar quality to those in the test now.

The suggestion regarding the influence of test length on reliability of the test is in line with the claims made by (Griswold, 1990). He asserted that, despite other factors such as testing environment and the individual test taker, the single most important factor that affects the validity and reliability of formative assessment is the quality of tests themselves. As a result, test length or number of items as well as the use of table of specification should be carefully considered when constructing tests items since they usually provide a representative sample and less distorted scores Justin and John (as cited in Kragu,2012).

To sum up, the test items were found to be adequately difficult and discriminated as indicated in the result section. However, this did not reflect candidates' performance on the test, since other factors as asserted by Agyeman (as cited in Etsey, 2005) could be responsible for the poor performance on the test. Agyeman cited teacher factor and stated that, a teacher who does not have both academic and professional qualification would undoubtedly have a negative influence on the teaching and learning of their subject as well as their students' academic performance. Relatedly, Etsey, Amedahe and Edjah as cited in Etsey (2005) reported that, there is a positive relationship between academic performance and factors such as teacher qualification, supervision, class size and motivation. These factors together with the characteristics of the DBME identified by this study should be



considered by the KEEA Municipal Education Directorate in order to help improve students' academic performance in the municipality.

In addition, to assess the strength of the predictive relationships between the DBME and SBA and the BECE, an ordinal logistic regression analysis was carried out to test the hypotheses of the study. Prior to the analysis, the predictor variables were examined a priori to ensure there is no violations of the assumptions of the statistical tool (ordinal regression) used especially the no multicollinearity and the proportional odds assumption. Also, descriptive statistics for each of the tests were examined to show the proportion of BECE candidates entered for each of the subjects considered in the study as well as the summaries of performances of the candidates in the BECE. What the study sought to achieve with descriptive statistics in the analysis was to assess whether there is a systematic relationship between the explanatory variables (DBME and SBA scores) and the dependent variable (BECE) in predicting performance of the 2019 BECE candidates in their final examination.

### **Predictive relationship between the DBME and the BECE**

Findings of the predictive relationship between the DBME and the BECE suggest that there is a significant but inverse relationship between the DBME and the BECE in all the four subjects considered in this study. Also, with regard to the predictive power of the DBME, the result indicated that for every unit increase in the DBME score there was approximately 0.7 decrease in students' performance in the BECE. The findings of the predictive relationship between the DBME and the BECE supported the findings of Omirin and Ale (2008) who investigated the predictive validity of English and

Mathematics Mock Examination results of Senior Secondary School students' performance in WASSCE in Ado-Ekiti state. They reported that the English and Mathematics Mock results significantly predicted the performance of students in WASSCE. Nonetheless, the English language result was a better predictor in the WASSCE as compared to the Mathematics scores. They consequently recommended that Mock examination should be made compulsory for WASSCE candidates because of its predictability in the WASSCE.

Similarly, Bosson-Amedenu (2017) investigated the predictive validity of Mathematics Mock Examination scores of WASSCE and BECE candidates in Ghana concluded that, the Mock examination for Core Mathematics had high predictive power of 92% in the WASSCE. However, similar relationship could not be established between the BECE and Mock for Mathematics grades because a significant difference between the two examinations was also established. Nonetheless, he concluded that there is a relationship between Mock examinations and national examinations and suggested that mock examinations should be given the necessary attention and regulation to harness its predictive power in national or final examinations.

The findings of this study also corroborated the findings of a study by Komba, Kafanabo, Tryphone and Kira (2013) which centred on Form Two Secondary Education Examination (FTSEE), a form of mock examination before the Certificate of Secondary Education Examination (CSEE) in Biology. Their findings revealed a strong relationship between the students' performance in the FTSEE and CSEE (i.e. from  $r=0.442$ ,  $p<0.01$  to  $r=0.726$ ,  $p<0.01$ ). They consequently supported the call for mock examinations to be

made compulsory for school candidates since it strongly correlated with the CSEE.

Andala, Digolo and Kamande (2014) in their study aimed at determining whether the mock examinations are reliable enough for predicting the results of the Kenya Certificate of Secondary Education examinations (KCSE), reported that there is high positive correlation (0.949) between the mock and KCSE examinations results. They, however, stated that the quality assurance indicators that guide examinations in schools in the study area were observed during the preparation and administration of mock examinations. Their study also concluded that mock examinations are reliable in predicting KCSE examinations and advocated that there was the need to harmonize the structures for setting, moderation and administration of mock examinations to make it more prototype of the KCSE or final examinations.

Contrary to findings suggesting a positive correlation between mock and state examinations, a paper on relationships between internal and external examination scores presented by Merwe (1999) at the 17th Annual Conference of Association for Educational Assessment in Africa (AEAA) held on September 26th–October 2nd, at Lusaka, Zambia. expressed dissatisfaction about the predictive relationship between mock and state examinations in one of their annual conferences where they indicated that there is low correlation between teacher-made tests and the national examination scores. Interestingly, the high correlation observed between mock examination and final examination according to Lewin (1997) is as a result of poor scoring standards of the mock examination. Unfortunately, this study disagrees with Lewin

because administration and scoring of the mock examination is sometimes stricter than that of the final examination as corroborated by Leonard (2004).

Although the findings of this study revealed a negative predictive relationship between the DBME and the BECE, juxtaposing the findings of this study with other related studies on the topic suggest that there is predictive relationship between mock examination and high state examinations. Consequently, the practice of setting local examination standards as being practised in Kenya should be adopted by municipal and district directorates of education in Ghana to help in the standardisation, and administration of mock examinations to make them more prototype of final examinations such as the BECE.

#### **Predictive relationship between the SBA and the BECE**

The findings of this study in relation to the predictive relationship between the School Based Assessment (SBA) and the Basic Education Certificate Examination (BECE) indicated that there is there no significant predictive relationship between the SBA and the BECE with the exception of the Integrated Science, which showed a significant negative relationship with the BECE. Again, the findings of this study corroborated the suspicion that some teachers “manufacture scores” for their students as reported by Amedahe (2001). Therefore, the findings of this study contradicts that of Brempong (2019) and Ahenkora (2019) which suggested that teachers are well versed in the practice of school level assessment. Indeed, if that were the case, the analysis would have shown some level of predictive relationship between the SBA and students’ performance in the BECE. On the contrary, the findings

rather indicate that the reliability of SBA scores submitted to WAEC by the various basic schools in the country seem questionable.

Relatedly, Omole (2007) also conducted a comparative study of students' performance in SBA and certificate examination in the Federal Capital Territory (FCT) of Nigeria and reported that students' scores in SBA were significantly higher than their scores in Certificate Examination in English and Mathematics at the upper basic education.

On the contrary, Okpala, Afemikhe and Anyanwu (2012), in a study to investigate the predictive power of school based assessment scores on students' achievement in Junior Secondary Certificate Examination (JSCE) in English and Mathematics in Nigeria revealed that the combination of the SBA scores with the final examination score significantly predicted students' performance in English and Mathematics achievement in JSCE. Also, Orubu (2013) and Ogunkola (2007), asserted that School Based Assessment scores in any subject for a particular student should reflect his or her achievement in the external examination. Their assertion was based on their findings after investigating the SBA scores as predictors of students final grades in Delta and Ogun States. Interestingly, they found that SBA scores predicted students' performance in JSCE Mathematics and Integrated Science respectively.

More importantly, the discrepancies observed between the findings of the study as well as that of Omole's and Okpala, Afemikhe and Anyanwu should be a cause to worry about. Interestingly, it is what Amedahe as cited in Amedahe (2000) pointed out that the determination of the total term score for a student is not standardized at the school level resulting in unreliable scores and hence a poor predictive relationship between the SBA and final

examinations. He therefore advocated for unambiguous and thorough guidelines for school administrators to restructure the practice of assessing students at the school level to ensure uniformity and reliability of school assessment. The concept of the SBA is believed to have originated from concerns raised by Amedahe and other researchers in the field of Educational Measurement.

Unfortunately, it appears the implementation of SBA is facing some challenges which need to be addressed to achieve the full benefits of the policy. Ironically, Ahenkora (2019) in a study to assess the implementation of SBA for the KEEA District of the Central Region reported that, to a great extent, teachers indicated that, they have requisite knowledge about the SBA as well as a positive attitude towards the application of the guidelines as stipulated in the SBA policy document in their respective schools. On the other hand, Yip and Cheung (2005) indicated that teachers might not necessarily have the skills needed to achieve the objectives of the SBA. They stated that, differences in students' ability and learning attitude may also influence the effective implementation of the SBA. In addition, they enumerated a number of challenges that teachers face in the process of implementing the SBA.

Additionally, a major challenge encountered by basic school teachers in the process of implementing the SBA is large class sizes especially in the rural areas (Nitko 2001 as cited in Ahenkora 2019). Similarly, the Hong Kong Professional Teachers' Union (PTU) in 2013, indicated that effective implementation of SBA policy increases the workload on both teachers and students. It indicated that for students, the demand for SBA is high in terms of

quality and quantity. The PTU report contradicts the 2012 SBA guidelines which indicated that one of the benefits of the SBA is that it reduces the work load of the teacher in terms of number of exercises and activities and that of the student as much as 53% and 64% respectively within a term.

The controversies surrounding the SBA suggest that the implementation guidelines has not gone down well with practitioners resulting in contradictory relationship reports by researchers between the SBA and final examinations. There is, therefore, a need to carry out more studies in this area, to gather enough empirical evidence on the subject matter in order to achieve the purpose for which it was made to replace the continuous Assessment of students at the school level as well as ensuring its predictive relationship on students' performance in the Basic Education Certificate Examination.

#### **Students' Performance in the 2019 BECE at KEEA**

Finally, from the findings of this study, it appears that students' performance in the 2019 BECE is not significantly different from what it has been over the last eight years. This confirms Etsey, Amedahe and Edjah's claim (cited in Etsey, 2005) that academic performance of junior high school (JHS) students in Ghana, especially the public schools, has not been encouraging over the last decade. Okyerefo, Fiaveh and Lamptey (2011); Gyan, Mabefam and Baffoe (2014) in support of the of the findings of Etsey indicated that the trend has been the same over the years in the BECE especially in the public schools and particularly for those in the rural areas.

Juxtaposing the findings of this study with that of the Ghana MoE's (2011) report which suggested that students' performance at the school level requires an output measure of education that is comparable nationally. The

report indicated that, to be considered as having performed in the BECE, the candidate must obtain grades ranging from 1 (best) to 5, with 5 being defined as average. Also, the BECE results is used to determine eligibility for placement of JHS graduates into senior high school (SHS). However, to be placed in any of the 920 public SHS in the country, a candidate is expected to have an aggregate score of not more than 25 for the four core subjects and a fifth best performing subject. Thus, the scoring methodology used in the BECE tests equates a score of 5 as representing an average performance across all students who sit for the examination.

Again, majority of the candidates who sat for the 2019 BECE in the KEEA performed below the pass mark of grade 5 in all the required core subjects as indicated in the result section of this study. Thus, had it not been the Free SHS Policy, majority of the candidates would likely have found it difficult to secure a place in any of the government assisted senior high schools. This study, therefore, strongly recommends further investigation of the predictive relationships between the SBA, DBME and students' performance on the BECE to ascertain the predictive relationships of these internal assessment tools used at the basic school level in order to have a holistic picture of their usefulness.



## CHAPTER FIVE

### SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

#### Overview

The closing remarks of this study are presented in this chapter, where the summary, conclusions and recommendations of the research endeavor are provided. The study, though, sought to be a contribution to this subject area, analysed its concomitant limitations. Recommendations are made with regard to aspects that could be explored in future research, especially within the context of predictability of internal assessment scores on students' performance in external examinations.

The study examined the predictive relationship between the school based assessments as well as the district based mock examination scores and students' performance in the 2019 BECE in the KEEA Municipality in the Central Region of Ghana. The study primarily focused on the 2019 SBA and DBME scores of BECE candidates and how they predicted students' performances in the BECE.

The study adopted the multistage sampling technique to select 602 BECE candidates from all the eight (8) circuits within the KEEA municipality. The data collected was analysed using both descriptive statistics (frequency, percentage tables, means and standard deviations) and inferential statistics (ordinal regression) after all the recommended assumptions had been tested to permit the use of ordinal regression.

## Summary of Findings

The following are the main findings from the data analysis.

1. Item Difficulty and Discrimination indices of the District Based Mock Examination were found to be within the recommended range. However, the KR 20 measure of the test items examined were found to be below the recommended coefficients for English language and Social Studies while that of Mathematics and Integrated Science were found to be relatively good.
2. There was no significant predictive relationship between the School Based Assessment scores (SBA) and the Basic Education Certificate Examination (BECE) with the exception of the Integrated Science, which showed a significant negative predictive relationship between the SBA scores and students' performance in the BECE.
3. There was a significant negative relationship between the District Based Mock Examinations scores and students' performance in the BECE in all the four core subjects investigated. To add to this, the odds ratio estimates indicated a decrease of approximately 0.7 in students' performance in the BECE for each unit increase in the DBME scores.
4. Students' performance in the 2019 BECE was not significantly different from what had been for the past eight (8) years. Thus, the majority of candidates who sat for the 2019 BECE had grades greater than five (5) which is the pass or average grade for the BECE.

## Conclusions

It is important to note that the 2019 District Based Mock Examination was of average quality. This could explain why it had a predictive relationship with students' performance in the Basic Education Certificate Examination compared to the School Based Assessment scores which had no predictive relationship with students' performance in the BECE. What was found to be problematic with the DBME was the negative relationship it had with students' performance in the BECE, which could be attributed to the poor reliability indices of the majority of the DBME tests.

Moreover, the perception that some teachers "manufacture" SBA scores for their students is supported by that fact that the SBA scores failed to show any predictive relationship with students' performance in the 2019 BECE. This is disadvantageous to students and the school selection and placement programme of the Ghana Education Service because it might give misleading information about a student's ability which may result in the wrong placement of students at the senior high schools.

## Recommendations

From the summary of the major findings of this study, it is recommended that:

1. the practice of district mock examinations be sustained by all metropolitan, municipal and district directorates of education (MMDE). However, the practice should be regulated by test experts or an oversight committee to improve its predictive relationship with the BECE.
2. head teachers and circuit supervisors of the various schools and circuits ensure teachers are abreast of the SBA guidelines and implement them

to the letter and desist from the practice of “manufacturing scores” for pupils.

3. to improve the predictive relations between mock examinations and students’ performance in BECE at the KEEA, a thorough investigation be conducted by the KEEA Municipal Directorate of Education to reveal practices/factors which might be influencing the negative relationship between the mock examinations and students’ performance in BECE.

### **Suggestions for Future Research**

1. Further research on this topic is recommended to assess teacher competence in the implementation of the School Based Assessment Policy because differences in teacher competences in the implementation of the SBA policy could bring about variations in findings.
2. Replication of this study is recommended in other metropolitan, municipal and district assemblies in Ghana to provide additional knowledge on the relationship between students’ performance in internal and external examinations.

## REFERENCES

- Adejo O., & Connolly T. (2017). *Holistic approach to predicting students' performance in higher educational institutions – a conceptual framework*. University of the West of Scotland, Paisley, United Kingdom
- Adejumo, A. O., & Adetunji A. A. (2013), Application of ordinal logistic regression in the study of students' performance. *Journal of Mathematical Theory and Modeling*, 3(11), 9 - 12
- Adesoji, F.A. & Kenni, A. M. (2013). Continuous assessment, *mock* results and gender as predictors of academic performance of chemistry students in WASSCE and NECO examinations in Ekiti State. *Journal of International Educational Studies*, 6(7), 1 –8
- Adeyegbe, S.O. (2004). Research into STM curriculum and school examination in Nigeria. The state of the part in *Science Teachers Association of Nigeria. Proceedings of the 45<sup>th</sup> annual conference* (pp 70 – 79)
- Adeyemi, T. O. (2008). Predicting students' performance in senior secondary certificate examinations from performance in junior secondary certificate examinations in Ondo State, Nigeria. *Humanity and Social Science Journal*, 3(1), 26-36.
- Ahenkora, A. T. (2019). *The Implementation of school-based assessment in KEEA District in Central Region*. (Unpublished master's thesis) Department of Education and Psychology, University of Cape Coast. Cape Coast, Ghana.

- Akyina, O. A., & Oduro-Okyire, G. (2019). Formative assessment practices of senior high school teachers in the Asante Mampong Municipality of Ghana. *British Journal of Education, 17* (1) 27 – 38
- Allen, M.J., & Yen, W. M. (1979), *Introduction to measurement theory*. Long Groye IL: Waveland Press.
- Amadioha, A., & Uko B.E. (2019). Mock examination as predictor of academic performance in Agricultural Science in Ikot Ekpene Senatorial District of Akwa Ibom State. *International Journal of innovative Social Science Education Research* 7 (2), 49 – 53.
- Amedahe, F. K. (2001). Combining teacher-assessment scores with external examination scores for certification. The Ghanaian experience. *Educational Measurement: Issues and Practice* 24 (4), 29-30.
- Amedahe, F. K., & Asamoah-Gyimah, K. (2016). *Introduction to measurement and evaluation* (7thed.). Hampton Press: Cape Coast.
- Andala H.O., Digolo .O. & Kamande. M. (2014) Reliability of mock examination for prediction of the Kenya Certificate of Secondary Examination (KCSE) results. *IOSR Journal of Research and Methods in Education, 5*(27) 50 -58.
- Asamoah, D., & Ocansey, M.K.K. (2019). Item discrimination and distractor Analysis: A technical report on thirty multiple choice core mathematics achievement test items. *International Journal of Research and Scientific Innovation, 6*(9), 2321–2705.
- Bichi, A.A., (2015). Item analysis using a derived science achievement test data. *International Journal of Science and Research (IJSR), 4*(5), 1655 – 1660.

- Bloom, B. S.; Hastings, J.T., & Madaus, G. F. (Eds) (1971) *Handbook on the formative and summative evaluation of student learning*. New York, McGraw-Hill.
- Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics*. New York: Cambridge University Press.
- Bosson-Amedenu S. (2017). Predictive validity of mathematics mock examination results of senior and junior high school students' performance in WASSCE and BECE in Ghana. *Asian Research Journal of Mathematics*, 3(4), 1-8,
- Brant, R. (1990), Assessing proportionality in the proportional odds model for ordinal logistic regression, *Biometrics*, 46(4), 1171–1178
- Brempong, D. A. (2019). Comparing traditional assessment procedures, and performance and portfolio assessment procedures. *Researchgate*. <https://doi.org/10.13140>
- Brown, R.S. Coughlin E, Ed (2007). The predictive validity of selected benchmark assessment used in the Mid-Atlantic Region (Issues and Answers report, REL 2007 – No 017). Cambridge: University of Cambridge School of Education. Washington, DC: US Department of Education. *Institute of Education Sciences. National Centre for Education Evaluation and Regional Assistance*. Retrieved from <http://ies.ed.gov/ncee/edlabs> 13/01/2020
- Burns, G. N., Morris, M. B., Periard, D. A., LaHuis, D., Flannery, N. M., Carretta, T. R., & Roebke, M. (2017). Criterion-related validity of a big five general factor of personality from the TIPI to the IPIP. *International Journal of Selection and Assessment*, 25, 213–222.

- Chakrabartty, S. N. (2013). Best split-half and maximum reliability. *IOSR Journal of Research & Method in Education*, 3(1), 1-8.
- Cheng, M. M. H., & Lee, T. T. H. (2010). Addressing the challenges in school-based assessment: Practices in the assessment of inquiry learning. *Curriculum and Teaching*, 25(1), 25–41.
- Cohen, L. Manion. L., & Morrison, K. (2007). *Research methods in education*, 5. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Cook, D. A., & Beckman, T. J. (2006). Current concepts in validity and reliability for Psychometric Instruments: Theory and application. *The American journal of medicine*, 119, 166.e7-166.e1.
- Creswell, J.W (Eds.). (2005). *Educational research: planning, conducting and evaluating quantitative and Qualitative research*. Pearson Merrill Prentice Hall.
- Crocker, L, & Algina, J. (1986). *Introduction to classical and modern test theory*. Fort Worth: Harcourt Brace Jovanovich College Publishers.
- Curriculum Research and Development Division, (2012). *Teachers' handbook on school based assessment for junior high school*. Accra, Ghana: Ministry of Education.
- Daniel. M & J. Schouten (1970), *Education in Europe; The screening of students, problems of assessment and prediction of academic performance*, Council for cultural cooperation of the council of Europe; George Harrap Co. Ltd
- Devillis, R. E. (2006). *Scale Development: Theory and application. Applied social science research method series*. Newbury Park: SAGE Publishers Inc.



Ding, L., & Beichner, R. (2009). Approaches to data analysis of multiple-choice questions. *Physical Review Special Topics-Physics Education Research*, 5(2), 20-33.

Division of Instructional Innovation and Assessment (2003), The University of Texas at Austin. Retrieved from; <https://www.utexas.edu/academic/diia>.  
13/11/2019.

Downing, S. M. (2004). Reliability: of the Reproducibility of assessment data. *Med Education*, 38, 1006-1012.

Etsey, Y. K. A. (2005). *Causes of low academic performance of primary school pupils in the Shama Sub-Metro of Shama Ahanta East Metropolitan Assembly (SAEMA) in Ghana*. A paper presented at the regional conference on education in West Africa, Dakar, 1<sup>st</sup> – 2<sup>nd</sup> November, 2005.

Fives, H., & DiDonato-Barnes, N. (2013). Classroom test construction: The power of a table of specification. *Practical Assessment, Research and Evaluation*, 18(3),1-7.

Fraenkel, J. R., & Wallen, N. E. (2008). *How to design and evaluate research in education* (7th ed.). New York, NY: McGraw-Hill.

Friedman A.B., & Mandel R.G (2009). The prediction of college student academic performance and retention: Application of expectancy and goal setting theories. Baywood Publishing Co., Inc. *J. college student retention*, 11(2), 227-246.

Garrison, C., & Ehrighau, M., (2010). Formative and summative assessment in the classroom. National Middle School Association; measured progress. 29 July, 2011. Retrieved from <http://www.amle.org/>

*publications/webexclusive/assessment/tabid/1120/default.aspx.*

*13/13/2020.*

Ghana News Agency, (2012). Mock BECE to be organized for junior high school finalists countrywide. Retrieved from [https:// www.modernghana.com/news/386388/mock-bece-to-be-organised-for-jhs-finalists-countr.html](https://www.modernghana.com/news/386388/mock-bece-to-be-organised-for-jhs-finalists-countr.html). 20/11/2019.

Ghana. Statistical Service. (2014). *2010 Population & housing census report: Districts in Ghana*. Accra, Ghana. Ghana Statistical Service

Giacomini R., & White H., (2006) Tests of conditional predictive Ability, *Econometrica*, 74(6), 1545-1578.

Graziano, A. M., & Raulin, M. L. (2006). *Research methods: a process of inquiry* (6th Ed.). Boston, MA: Allyn & Bacon.

Griswold, P. A. (1990) Assessing relevance and reliability to improve the quality of teacher-made tests. *NASSP Bulletin*, 18-24.

Gyan, C., Mabefam, M. G., & Baffoe, M. (2014). Push out or drop out? Taking a critical look at the poor performance and drop-out of students of the JSS/JHS Programme in Ghana. *Academic Journal of Interdisciplinary Studies*, 3(1), 405-409.

Hathcoat, J. D. (2013). Validity semantics in educational and psychological assessment. *Practical Assessment, Research and Evaluation*, 18, (9):1-14.

Hong, T., Purzer, Ş., & Cardella, M. E. (2011). A psychometric re-evaluation of the Design, Engineering and Technology (DET) survey. *Journal of Engineering Education*, 100(4), 800-818.

- Iddrisu, S. A. (2009). *Predictive validity of senior high school (SHS) aggregate of students' grade-point average (GPA) at the University of Cape Coast, Ghana*. (Unpublished master's thesis), Department of Education and Psychology, University of Cape Coast. Cape Coast, Ghana.
- Ifeanyieze, F.O. and Aneke, K. (2013) Competencies required by teachers in school Based assessment of student in secondary school. *Nigerian Journal of Educational Research and Evaluation*, 12, 29-37.
- Isaac, R. Z., Wilfred, J. P., & Douglas, C. (2001). Leadership and motivation: The effective application of expectancy theory. *Journal of Managerial Issues*, 13(2), 1045-3695.
- Kargar, M., Tamizi, R.A., & Bayat, S. (2010). Relationship between mathematical thinking, mathematics anxiety and mathematics attitude among university students. *Procedia-Social and Behavioral Science*, 8, 537 – 542.
- Keightley, J. V.; & Coleman, M. J. (2002). *School based assessment in a high stakes environment*. Adelaide: Senior Secondary School Board of South Australia.
- Kelley, T. L. (1939). The selection of upper and lower groups for the validation of test items. *Journal of Educational Psychology*, 30, 17-24.
- Keyton, J., King, T., Mabachi, N. M., Manning, J.; Leonard L.L.; & Schill, D. (2004). *Content analysis procedure book*. Lawrence, KS: University of Kansas.

- Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacists*, 65(1), 2276- 2284.
- Kinyua, K., & Okunya, L.O. (2014). Validity and reliability of teacher-made test: Case study of year 11 physics in Nyahururu District of Kenya. *African Educational Research Journal* 2(2), 61 – 71.
- Kiragu, K.D. (2012). *Validity and reliability of teacher made tests in Kenya: a case study of Physics form three in Nyahururu District*. (Unpublished master's thesis). School of education, Department of Psychology. University of Nairobi, Kenya.
- Komba, S. C., Kafanabo, E. J., Tryphone, D., & Kira, E. S. (2013). The Predictive Validity of Form Two Secondary Education Examination (FTSEE) on Students' Performance in the Certificate of Secondary Education Examination (CSEE) in Biology Subject: A. *Journal of Education and Practice*, 4(4), 238-242
- Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, 30, (3), 607-610.
- Last, J. (Ed.) (2001). *International Epidemiological Association, A Dictionary of Epidemiology* (4th ed.). New York: Oxford University Press.
- Leedy, P. D., & Ormrod, J. E. (2004). *Practical research*, (8th Ed.). Upper Saddle River, N.J: Prentice Hall.
- Leonard, R. (2004) Mock examinations. *Assessment in Education*, 55-67

- Lewin, K. (1997). Criterion-referenced assessment panacea or palliative? Bunde & K. Lewin (Ed.), *Improving test design*, pp 12-34. German Foundation for Development.
- Madu, A., & Ebere, C. (2016). Predictive validity of mock senior school certificate examination on NECO senior secondary school certificate examination scores in agricultural science in Abia state. *International Journal of Current Research and Academic Review*, 4(2), 237-257.
- Maduabum, M. A. (2004). *Fundamentals of educational research*. Owerri: Versatile Publishers.
- Magno, C. (2003). The profile of teacher-made test construction of the professors of University of Perpetual Help Laguna. *UPHL Institutional Journal*, 1(1), 48-55
- McCowan, R. J., & McCowan, S. (1999). *Item analysis for criterion-referenced tests*. Center for Development of Human Services. Buffalo, New York.
- Merwe, I. F. J. (1999) Continuous Assessment in Namibia: Experience and Lessons to Be Learnt. A paper presented at the 17th Annual Conference AEAA held on September 26th–October 2nd, 1999, at Lusaka, Zambia.
- Messick, S. (1989). Validity. In R.L. Linn (Ed.). *A nation at risk: The imperative for education reform*. Washington, DC: U.S. Government Printing Office.
- Ministry of Education (2011). *National education assessment report*. Accra, Ghana. Author.

- Mohajan H.K. (2017). Two criteria for good measurements in research: Validity and reliability. *Annals of Spiru Haret University. Economics Series 17* (4), 59 – 82.
- Nitko, A. J., & Brookhart, Susan M. (2007). *Educational assessment of students*. Upper Saddle River, NJ: Pearson Education, Inc.
- Notar, C. E., Zuelke, D. C., Wilson, J. D., and Yunker, B. D. (2004). The table of specifications: Insuring accountability in teacher made tests. *Journal of Instructional Psychology*, 4(31)115-129.
- Nwana, O. C. (2007). *Textbook on educational measurement and evaluation*. Owerri: Bomaway
- Obioma. G.; and Salau. M. (2007). *The predictive validity of public examinations. A case study of Nigeria*, Abuja. Nigerian Educational Research Council (NEROC).
- Ochieng, P.A (2012). *Predictive validity of internal examinations in secondary schools in Kenya*. (Doctoral dissertation submitted) to the University of Nairobi, Nairobi, Kenya.
- Ogunkola, B. J. (2007). School-based assessment scores as predictors of students' final grades in Integrated Science. *Nigeria Journal of Educational Research and Evaluation*. 7(1), 1-7
- Okoro, O. M. (2002). *Measurement and Evaluation in Education*. Obosi: Pacific Publisher Ltd.

- Okpala, P. N., Afemikhe, O. A., & Anyanwu, E. (2012). *Educational assessment practices for quality assurance: The national examinations council's experience in Nigeria. Revolutionizing assessment and evaluation procedures in education*. Stirling-Horden Publishers Ltd., Ibadan, 15-26.
- Okyerefo, M. P. K., Fiaveh, D. Y., & Lamptey, S. N. L. (2011). Factors promoting pupils' academic performance in privately owned Junior High Schools in Accra, Ghana. *International journal of sociology and anthropology*, 3(8), 280-289.
- Oliver, P. (2010). *The Student's guide to research ethics*. McGraw-Hill Education.
- Omirin, M. S.; & Ale, V. M. (2008). Predictive validity of English and Mathematics mock examination results of senior secondary school students' performance in WASCE in Ekiti-State, Nigeria. *Pakistan Journal of Social Sciences*, 5(2), 139-141.
- Opara, I. M., Onyekuru, B. U., & Njoku, J. U. (2015). Predictive power of school based assessment scores on students' achievement in Junior Secondary Certificate Examination (JSCE) in English and Mathematics. *Journal of Education and Practice*, 6(9), 112-116.
- Orubu, M. E. N. (2013). School based assessment as predictor of students' performance in junior school certificate mathematics examination in Delta State. *Nigerian Journal of Educational Research and Education*, 12, 37-44.
- Osadebe, P.U. (2015). Evaluation techniques, DELSU, *Journal of Educational Research and Development*, 12(1), 56-63.

- Othuon, L.A. (1994). Hierarchical linear modelling of predictive validity. The case of Kenya Certificate of Primary Education Examination. *Studies in Educational Evaluation*, 181 - 190.
- Pallant, J., (2011). SPSS survival manual 4<sup>th</sup> edition: A step by step guide to data analysis using SPSS version 18. Maidenhead, Berkshire: Open University Press. Retrieved from <http://www.allenandunwin.com/spss> 13/12/2019.
- Parkes, J. (2000). Relationship between reliability and cost of performance assessment. *Education Policy Analysis Archives*, 8, 16.
- Popham, J. W. (2008). *Classroom assessment: What teachers need to know* (5th ed.). Boston: Pearson Education, Inc. Publishers.
- Quaigrain, K., & Arhin, A.K, (2017). Using reliability and Item analysis to evaluate a teacher-developed in educational measurement and evaluation. *Cogent Education*, 4 (1) 1301013.
- Quansah, M. E. (2013). *An ordinal regression analysis of students' satisfaction: A case study of the college of science in KNUTS*. (Unpublished master's dissertation), College of Science, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana.
- Ramatlala, M. S., & Nenty, H. J. (2012). Gender as a factor in the prediction of performance in Botswana general certificate of secondary education physical education examinations by coursework and forecast grades among senior secondary school students. *Scientific Research Journal*, 3, 32-37



- Ramos, V., Nunez, C., Unda, X., & Mejia, K. (2016). *Validation of a scale to measure the cultural dimension that influence the work climate in the Ecuadorian context*. Quito: National polytechnic School, Spain.
- Rothstein, J. (2004). College performance predictions and the SAT, *Journal of Econometrics*, 1(21), 297-317
- Schwarz, J. (2011). Research methodology: Tools. applied data analysis (with SPSS). Lecture 02: Measurement scales & item analysis. Lucerne University of Applied Sciences and Arts. Retrieved from [http://www.schwarzpartners.ch/Applied\\_Data\\_Analysis/Lect%2002\\_E\\_N.pdf](http://www.schwarzpartners.ch/Applied_Data_Analysis/Lect%2002_E_N.pdf)
- Sekaran, U.; & Boujje. R. (2010). *Research for business- A skill building approach*. John Wiley and Sons, New York, NY.
- Shakil, M. (2008). *Assessing student performance using test item analysis and its relevance to the state exit final exams of MAT0024 classes*. A paper presented at MDC conference day. Kendall Campus, March 6th, 2008.
- Shillingburg, W. (2016). *Understanding validity and reliability in classroom, school-wide, or district-wide assessments to be used in teacher/principal evaluations*. Unpublished manuscript.
- Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy*, 85(3), 257-268.
- Tavakol, M., & Dennick, R. (2011). Making Sense of Cronbach's Alpha. *International Journal of Medical Education*, 2, 53-55.

- The Technical White Paper Series (2007). KR20 & Coefficient Alpha; their equivalence for binary-scored items, retrieved from <http://www.pbarrett.net/techpapers/kr20.pdf>. 08/12/2019.
- Twycross, A., & Shields, L. (2004). Validity and Reliability-what's it all about? Part 2: Reliability in Quantitative Studies. *Paediatric Nursing*, 16, (10), 36.
- Ukwuije R.P.I. & Opara, I. M. (2013). School based assessment-implications for educational transformation. *Nigerian Journal of Educational Research and Evaluation*, 12(2), 9-18.
- Veloo, A., Krishnasamay, N., & Md-Ali, R. (2015). Teachers' knowledge and readiness towards implementation of school-based assessment in secondary schools. *International Education Studies*, 8(11), 194 – 195.
- Vroom, V. C. (1964). *Work and motivation*. New York: John Wiley & Sons.
- Webster, J & Watson R.T. (2002). Analysing the past to prepare for the future: Writing a literature review. *MIS quarterly*. xiii-xxiii.
- Willis, J. (2007). *Foundations of Qualitative Research: Interpretive and Critical Approaches*. SAGE Publications.
- Yip, D. Y. & Cheung, D. (2005) Teachers' concerns on school-based assessment of practical work. *Journal of Biological Education*, 39(4) 156-162
- Zahn, C. M., Saguil, A., Artino Jr, A. R., Dong, T., Ming, G., Servey, J. T., & Durning, S. J. (2012). Correlation of national board of medical examiners scores with United States Medical Licensing Examination step 1 and step 2 scores. *Academic Medicine*, 87(10), 1348-1354.

Zaida H. & Daliata R. (2007). Predicting students' academic performance. Comparing artificial neural network, decision tree and linear regression. *21st annual SAS Malaysia forum, Shangri-La. Hotel, Kuala Lumpur.*





APPENDIX B

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.gh/14/20-23

Date: 22<sup>nd</sup> June, 2020

Your Ref: .....

Dear Sir/Madam,

ETHICAL REQUIREMENTS CLEARANCE FOR RESEARCH STUDY

Chairman, CES-ERB  
Prof. J. A. Omotosho  
jamo@ucc.edu.gh  
0244784739

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

Secretary, CES-ERB  
Prof. Linda Drama Forte  
lfdr@ucc.edu.gh  
0244786680

The bearer, Edvis Agyepong....., Reg. No. ED/MEP/17/0001 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:

Quality (quantitative) and predictive relationship between internal assessment scores and students' performance in the Basic Education Certificate Examination in the KEEA Municipality.

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 Drama Forte  
(Secretary, CES-ERB)

## APPENDIX C

### INTRODUCTORY LETTER

**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  
Tele: 2552, UCC, GH.  
Telegram & Cable: University, Cape Coast  
Email: [edufound@ucc.edu.gh](mailto:edufound@ucc.edu.gh)



UNIVERSITY POST OFFICE  
CAPE COAST, GHANA

9<sup>th</sup> April, 2019

Our Ref: — —

Your Ref: **TO WHOM IT MAY CONCERN**

Dear Sir/Madam,

**THESIS WORK**  
**LETTER OF INTRODUCTION**  
**MR. ELVIS AGYEPONG**

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

Mr. Agyepong is researching on the topic: *"Quality and Predictive Power of District-Based Mock Examination on the B.E.C.E in the K.E.E.A Municipality"*.

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

Thank you.

Yours faithfully,



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

APPENDIX D

DATA COLLECTION PERMISSION

GHANA EDUCATION SERVICE

*In case of reply the  
Number and date of this  
letter should be quoted*  
Tel No: 03321-40025 /40048/40017  
Tele/Fax: 03321-40016  
E-mail: keea\_edu@yahoo.com  
Jur Ref: GES/CR/KEEA/UCCV.II/82  
Your Ref: .....



Municipal Education Office  
Komenda-Edina-Eguafo-Abrem Mun..  
P. O. Box 13  
Elmina

15<sup>th</sup> April, 2018

THE HEAD  
DEPARTMENT OF EDUCATION AND PSYCHOLOGY  
UNIVERSITY OF CAPE COAST  
FACULTY OF EDUCATIONAL FOUNDATIONS  
CAPE COAST

**RE: INTRODUCTORY LETTER  
MR. ELVIS AGYEPONG**


Reference is made to your request on the above subject dated 9<sup>th</sup> April, 2019.

Permission is granted to **Mr. Elvis Agyepong**, an Mphil student with the Department of Education and Psychology to undertake his academic project on the topic: **"Quality and Predictive Power of District Based Mock Examination on the BECE in the KEEA Municipality"**.

For any further form of assistance in his work, Mr. Agyepong is to consult **Mr. Kwaku Ackon Amoah**, the Assistant Director in charge of Supervision and Monitoring and **Mr. Ebenezer Darko Ansah**, the Municipal Examination Officer.

Management will appreciate a summary feedback on the project after completion to address lapses identified.

Count on our support in this regard.

  
PHILIP KWESI NKOOM (MR.)  
AG. MUNICIPAL DIRECTOR OF EDUC.  
K.E.E.A. - ELMINA

- cc: Mr. Elvis Agyepong, Dept. of Educ. & Psychology, UCC, Cape Coast
- Mr. Kwaku Ackon Amoah, Head of Inspectorate Unit, MEO, Elmina
  - Mr. Ebenezer D. Ansah, Examination Officer, MEO, Elmina
  - File copy

APPENDIX E

SECOND MOCK EXAMINATION (KEEA, 2019)

ENGLISH LANGUAGE 1

OBJECTIVE TEST

Duration: 35mins.

Each question is followed by **four** options lettered A to D. Find out the correct option for each question and circle it on your question paper. Select only one answer to each question. An example is given below.

..... of the sugar was wasted

A. Many

B. A few

C. Much

D. Few

The correct answer is much which lettered B and therefore answer C would be circled.

Now answer the following questions

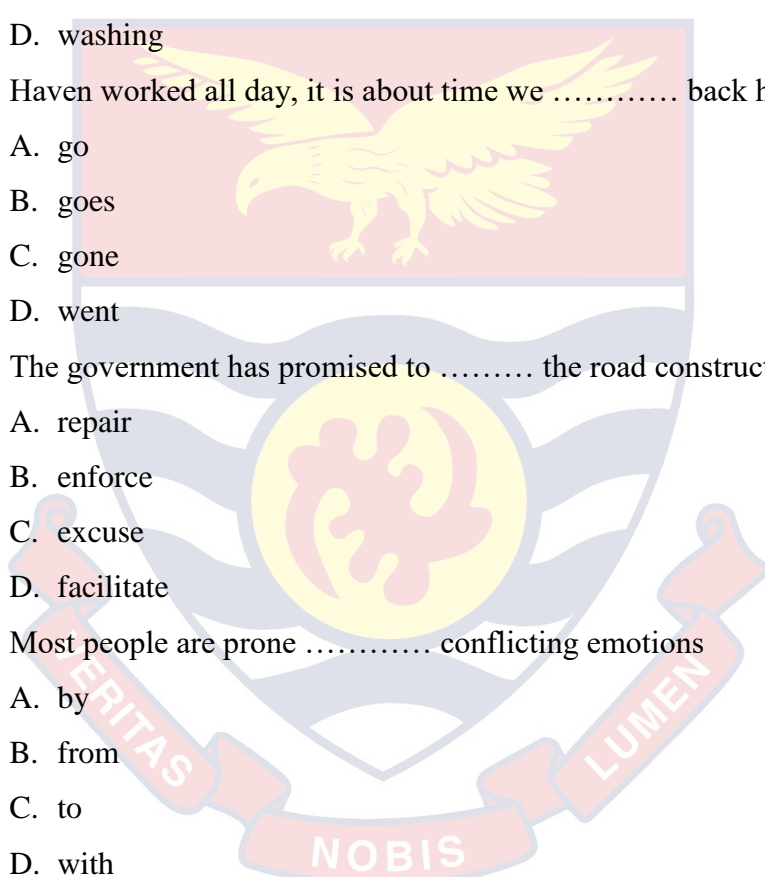
SECTION A

From the alternatives lettered A to D, choose the one which MOST SUITABLY completes each sentence.

1. Romeo and Juliet are friends, they like.....
  - A. each other
  - B. themselves
  - C. the other
  - D. one another
2. The minister of God ..... be able to solve the problem
  - A. can
  - B. has
  - C. will
  - D. ought
3. Clement, have you ever ..... in this lake?
  - A. swim
  - B. swam
  - C. swims
  - D. swum



4. Brong Ahafo region is the ..... of cocoa.
- A. nations's produce largest
  - B. largest notion's producer
  - C. producer nation's largest
  - D. nation's largest producer
5. My parents do not ..... late night parties
- A. indulge
  - B. suffer
  - C. tolerate
  - D. create
6. In some cultures, dead people are buried but in others they are .....
- A. cremated
  - B. baked
  - C. fired
  - D. flamed
7. He has since taken over the ..... of power
- A. reins
  - B. rains
  - C. reigns
  - D. realms
8. Thompson is a very good friend on ..... I can rely
- A. who
  - B. whose
  - C. which
  - D. whom
9. .... , we shall go together
- A. although you come early
  - B. if you came early
  - C. if you come early
  - D. if you had come early
10. Abuya's fever was so acute that she ..... an injection
- A. had to had
  - B. had to have
  - C. must have
  - D. ought to have

11. Everybody has to do ..... homework
- A. his
  - B. their
  - C. them
  - D. there
12. I want that spoon ....., I can still see some oil on it
- A. wash
  - B. washed
  - C. washes
  - D. washing
13. Haven worked all day, it is about time we ..... back home
- A. go
  - B. goes
  - C. gone
  - D. went
14. The government has promised to ..... the road construction project.
- A. repair
  - B. enforce
  - C. excuse
  - D. facilitate
15. Most people are prone ..... conflicting emotions
- A. by
  - B. from
  - C. to
  - D. with
- 

## SECTION B

Choose from the alternatives lettered A to D the option which nearest in meaning to the underlined word in each sentence.

16. The aftermath of the disaster is still with us

- A. cause
- B. climax
- C. conclusion
- D. result

17. The landlord has ejected the tenant from his house

- A. debar
- B. evicted
- C. expelled
- D. removed

18. Injuries keep many sportsmen from achieving their full potential

- A. capacity
- B. merit
- C. pinnacle
- D. power

19. It is possible to draw a parallel between the two positions.

- A. agreement
- B. liking
- C. line
- D. similarity

20. Writing became the focal point of his life after his retirement

- A. actual
- B. central
- C. chosen
- D. singular

### SECTION C

*In each of following sentences, a group of words have been underlined. Choose from the alternatives lettered A to D the one that best explains the underlined words.*

21. Rozzeta Benneh was admonished to paddle her canoe. This means Rozzeta Benneh should.....
- A. be in charge of the business
  - B. be more serious with life
  - C. manage her own affairs
  - D. place her interest first
22. The opportunity slipped through my fingers. This means that I .....
- A. did not use of it
  - B. expected something better
  - C. preferred something better
  - D. wanted to be more careful
23. Who to lead the group became our bone of contention. This means.....
- A. there was a quarrel
  - B. there was no appointment
  - C. there was no consensus
  - D. there was postponement
24. I did all I could to amuse him but he kept a straight face. This means that he
- A. was not convinced
  - B. remained adamant
  - C. refused to laugh
  - D. did not listen to me
25. The new salary increases were across the board. This means that the increases
- A. were very impressive
  - B. affected all board members
  - C. were approved by the board
  - D. affected all workers

**SECTION D**

*Choose from the alternatives lettered A to D the one which is nearly opposite in meaning to the underlined word in each sentence.*

26. The Louvre blades down here are opaque while those up are .....
- A. bright
  - B. distinct
  - C. plain
  - D. transparent
27. The water in that pond is muddy but that in the overhead tank is .....
- A. clear
  - B. plain
  - C. potable
  - D. stagnant
28. Alfred was admitted to the hospital, but was ..... after three days.
- A. discharged
  - B. freed
  - C. released
  - D. relieve
29. The tranquility of the rural areas contrasts with the .....in the cities.
- A. noise
  - B. panic
  - C. pollution
  - D. unease
30. There will not be any more food shortage after the ..... harvest.
- A. average
  - B. bumper
  - C. satisfactory
  - D. vast

## APPENDIX F

### SECOND MOCK EXAMINATION (KEEA, 2019) MATHEMATICS 1 OBJECTIVE TEST

Answer **all** questions in this section

Each question is followed by four options lettered A to D. Find out the correct option for each question and circle in pencil on your question paper the **answer** space which bears the same letter as the option you have chosen. Give only one answer to each question-.

An example is given below

Arrange the following whole numbers in ascending order 1101,1011,1110 and 1100

- A. 1101, 1100,1110 and 1011
- B. 1011, 1110. 1101 and 100.
- C. 1110. 1101. 1100 and 1011
- D. 1011, 1100.1101 and 1110

The correct answer is 1011, 1110, 1101 and 1100 which is lettered D and therefore answer D space would be circled

(A)            (B)            (C)            (D)

Think carefully before you shade the answer space. Erase completely any answer you wish to change.

Now answer the following questions.

1. If  $n^2 + n = 50$ , find n
  - A. 7
  - B. 24.5
  - C. 25
  - D. 49

Solve the equation  $\frac{x+2}{3} = 10$

- A. 3
  - B. 4
  - C. 5
  - D. 6
2. Write 0.00724 in standard form.
    - A.  $7.24 \times 10^2$
    - B.  $7.24 \times 10^{-2}$
    - C.  $7.24 \times 10^3$
    - D.  $7.24 \times 10^{-3}$
  3. Multiply  $(2a-b)$  by  $(2a + b)$ 
    - A.  $4a^2 - b^2$
    - B.  $4a^2 - ab + b$
    - C.  $4a - ab - b^2$
    - D.  $4a - ab - b^2$

4. K1 (2,-2) is the image of k under a translation by the vector  $T = \begin{pmatrix} -3 \\ -5 \end{pmatrix}$   
find the coordinate of K  
 A. (-1, -2)  
 B. (-5.7)  
 C. 5, -3)  
 D. (5, 3)
5. A Shop sell butteries of 68 pessewas each of GHC 2.15 for a pack of four. How much will. Osei save if he buys two packs of four instead of single batteries  
 A. GHC9,74  
 B. GHC5.44  
 C. GHC44.30  
 D. GHC1.14
6. Write 0.0963 correct to two significant numbers.  
 A. 0.096  
 B. 0.097  
 C. 0.10  
 D. 7.63
7. Calculate the size of an interior angle of regular nonagon  
 A.  $120^{\circ}$   
 B.  $185^{\circ}$   
 C.  $140^{\circ}$   
 D.  $200^{\circ}$
8. If  $V = \frac{\sqrt{\pi d^2 1 - m^2}}{8}$ , make d the subject  
 A.  $\sqrt{\frac{8v^2 - 3m}{\pi 1}}$   
 B.  $\sqrt{\frac{8v^2 - 2m}{\pi 1}}$   
 C.  $\sqrt{\frac{8v^2 + m^2}{\pi 1}}$   
 D.  $\sqrt{\frac{8m^2 v^2}{\pi 1}}$
9. If  $5^x = 125$ , find the value of x.  
 A. -1  
 B. -2  
 C. -3  
 D. -5

10. If  $q = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$  and  $r = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$  calculate  $2q - 3r$ .

- A.  $\begin{pmatrix} -5 \\ 12 \end{pmatrix}$
- B.  $\begin{pmatrix} -5 \\ -8 \end{pmatrix}$
- C.  $\begin{pmatrix} 13 \\ -4 \end{pmatrix}$
- D.  $\begin{pmatrix} 18 \\ -8 \end{pmatrix}$

One day, Asare Adu carried out a survey on the shoe sizes of 20 students, he puts the result in the table below. Use it to answer question 12 and 13

Shoe Size	1	7	8	9	10
No of Student	6	2	2	3	7

- 11. Find the median shoe size.
  - A. 7
  - B. 8
  - C. 9
  - D. 10
- 12. Calculate the mean shoe size.
  - A. 2.00
  - B. 6.88
  - C. 7.85
  - D. 8.15
- 13. How many face's has a triangular pyramid?
  - A. 3
  - B. 4
  - C. 5
  - D. 6
- 15. A point (-3,-6) is rotated through  $180^\circ$  about the origin. Find the image.
  - A. (3,-6)
  - B. (6,3)
  - C. 3,6)
  - D. 3,6)
- 16. The bearing of K from P is  $230^\circ$ : Find the bearing of P from K.
  - A.  $030^\circ$
  - B.  $040^\circ$
  - C.  $050^\circ$
  - D.  $060^\circ$



17. Solve the inequality:  $3x-7 \leq \frac{1}{2} (7x-3)$
- A.  $x \geq -11$
  - B.  $x \leq 11$
  - C.  $x \geq -17$
  - D.  $x \leq 17$
18. Simplify  $2a - 4(a+5b) - 10b$
- A.  $-2a + 30b$
  - B.  $2a - 30b$
  - C.  $-2a - 30b$
  - D.  $2a + 30b$
19. A ladder lean against a wall of height 9cm. If the foot of the ladder is 12cm away from the wall. Find the length of the ladder.
- A. 13cm
  - B. 14cm
  - C. 15m
  - D. 25m
20. If  $P = \{1, 2, 3, 4, \dots\}$  and  $Q = \{2, 4, 6, \dots\}$ . How many elements are in  $P \cap Q$ ?
- A. 2
  - B. 4
  - C. 5
  - D. 7
21. .... What is the next term in the following set of numbers 1, 3, 7, 15, 31, .....?
- A. 35
  - B. 39
  - C. 47
  - D. 63
22. If  $C = b^2 + 4r$ . find C when  $b=3$ ,  $r=4$  and  $a=5$
- A.  $\frac{19}{40}$
  - B.  $\frac{11}{20}$
  - C.  $\frac{3}{8}$
  - D.  $\frac{5}{2}$
23. Find the volume of a cuboid of sides 14cm by 12m by 3m
- A.  $36m^3$
  - B.  $45m^3$
  - C.  $180m^3$
  - D.  $504m^3$

24. A gas cooker' costing GHC820.00 has gone up by 12% in price. Find its new price.  
 A. GHC920.00  
 B. GHC.4  
 C. GHC.40  
 D. GHC908.40
25. Kofi deposited GI 10700,0,00 with a bank for-12 years at a rate of 10% per annum, find the simple interest.  
 A. GHC140.000  
 B. GHC14,000  
 C. GHC10.400  
 D. GHC104.000.
26. The following are the angles formed at the center of the circle:  $30^\circ/70^\circ$ .  $120^\circ$ .  $2x^\circ$  and  $5x^\circ$ . Find the value of x.  
 A.  $20^\circ$   
 B.  $40^\circ$   
 C.  $70^0$   
 D.  $100^\circ$
27. find the rule of the mapping for u\  
 relation.
- |   |    |   |   |   |    |    |
|---|----|---|---|---|----|----|
| x | -1 | 0 | 1 | 2 | 3  | 4  |
| ↓ | ↓  | ↓ | ↓ | ↓ | ↓  | ↓  |
| y | 3  | 5 | 7 | 9 | 11 | 13 |
28. How many integers are within the interval  $-5 < x < 7$ ?  
 A. 10  
 B. 11  
 C. 12  
 D. 13
29. Given that  $P = \left(\frac{3}{4y}\right)$  and  $Q = \left(\frac{2x-1}{4}\right)$  if  $P=Q$ . find the value if x.  
 A. 6  
 B. 3  
 C. 2  
 D. 1
30. What is the angle made by a clock when it is 4 O'clock exactly?  
 A.  $270^\circ$   
 B. 10  
 C.  $10^\circ$   
 D.  $9^\circ$

31. An angle whose measure is greater than  $90^\circ$  but less than  $180^\circ$  is called
- obtuse
  - reflex
  - straight
  - acute

32. A closed cylinder has a radius of 3cm and height 7cm. find the volume.
- $197\text{cm}^3$
  - $197.4\text{cm}^3$
  - $198.9\text{cm}^3$
  - $198\text{cm}^3$

33. Simplify  $\left(\frac{1}{4}x \cdot \frac{1}{2}\right) \div \left(12x \cdot \frac{2}{3}\right)$

- $\frac{5}{8}$
- $\frac{32}{5}$
- $\frac{3}{64}$
- $\frac{5}{25}$

34. The average of a group of six *boys* is 12 years. When one boy is withdrawn. the average becomes 11 years. Find the age of the boy who was withdrawn.
- 16 years
  - 17 years
  - 15 years
  - 14 years

35. Factorize the expression  $15x^2 - 2x - 1$ .
- $(3x - 1)(5x - 1)$
  - $(3x + 1)(5x - 1)$
  - $(3x - 1)(5x - 11)$
  - $(3x + 1)(5x + 1)$

36. Find the angle marked m in the figure below diagram
- $40^\circ$
  - $80^\circ$
  - $30^\circ$
  - $60^\circ$

37. Find the inverse of the mapping with the rule  $y = \frac{1}{2}x - 3$
- $x \mapsto 2y$
  - $x \mapsto 2y - 3$
  - $x \mapsto 2y + 6$

D.  $x \rightarrow y + 6$

38. A uniform rod of length 5m has mass 9.6kg. What is the mass per metre?

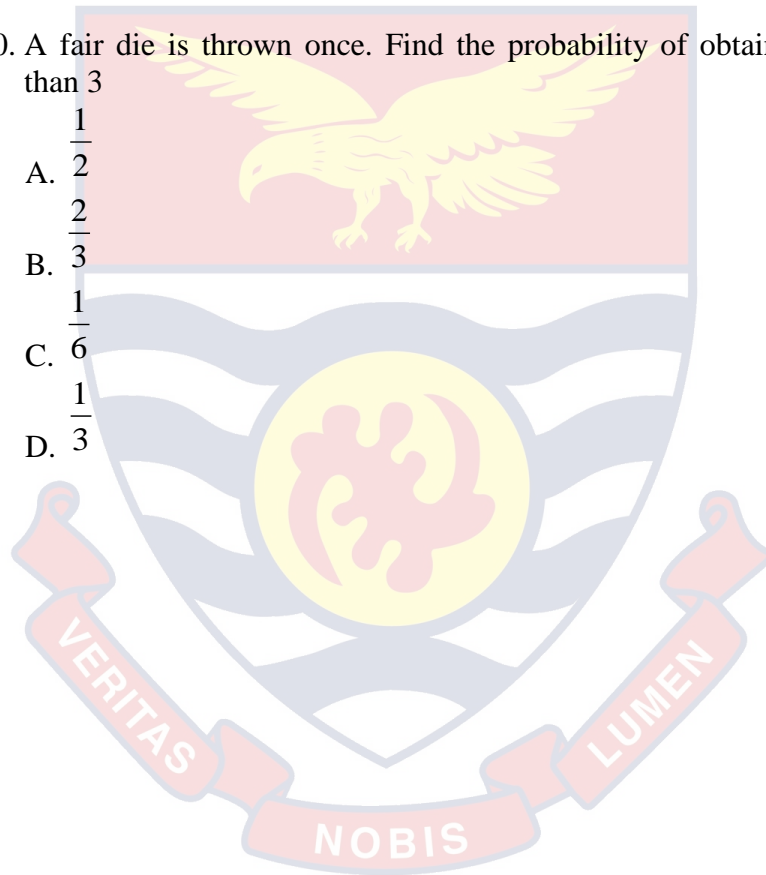
- A. 0.52kg/m
- B. 1.20kg/m
- C. 1.29kg/m
- D. 1.92kg/m.

39. Find the radius of a circle whose area is  $88\text{cm}^2$  (take  $\pi = 22/7$ )

- A. 4cm
- B. 2cm
- C. 7cm
- D. 8cm

40. A fair die is thrown once. Find the probability of obtaining a score less than 3

- A.  $\frac{1}{2}$
- B.  $\frac{2}{3}$
- C.  $\frac{1}{6}$
- D.  $\frac{1}{3}$



## APPENDIX G

### SECOND MOCK EXAMINATION (KEEA, 2019)

#### INTEGRATED SCIENCE

*Each question is followed by four options lettered A to D. Find out the correct option for each question and circle it on the question paper. Give only one answer to each question. An example is given below*

Which of these is out of place?

- A. Kidney
- B. Ureter
- C. Bladder
- D. Liver

The correct answer is *Bladder* which lettered C and therefore answer C would be circled.

*Think carefully before you circle select an answer. Erase any answer you wish to change.*

**Now answer the following question.**

1. Which of the following types of teeth is used by mammals to tear food materials?
  - A. Canines
  - B. Incisors
  - C. Molar
  - D. element Premolar
2. The with the chemical symbol S is .....
  - A. Silicon
  - B. Silver
  - C. Sodium
  - D. Sulphur
3. Which of the following statements about plant cell is correct? It.....
  - A. does not have a nucleus
  - B. contains large vacuoles
  - C. is surrounded by the cell membrane only
  - D. does not have a definite shape

4. Solid non-metals normally break into pieces when hammered because they are
  - A. brittle
  - B. ductile
  - C. Lustrous
  - D. Malleable
  
5. producers in the ecosystem are plants that.....
  - A. attract insects
  - B. feed on other plants
  - C. feed on dead materials
  - D. manufacture their own food
  
6. Which of the following elements is a semi-metal?
  - A. Carbon
  - B. Nitrogen
  - C. Silicon
  - D. Sodium
  
7. The by-products of respiration are.....
  - A. carbon dioxide and heat
  - B. carbon dioxide and water
  - C. oxygen and heat
  - D. oxygen and water
  
8. Which of the following substances is a solid-gas mixture?
  - A. Lather
  - B. Bronze
  - C. Steel
  - D. Smoke
  
9. An example of a non-living tissue used in osmosis experiment is .....
  - A. Cellophane
  - B. filter paper
  - C. potato
  - D. polythene
  
10. Which of the following processes is used to separate insoluble solids from liquids?
  - A. Crystallization
  - B. Evaporation
  - C. Filtration
  - D. Sublimation

11. Which of the following pairs of disease can be easily contracted when food is exposed to houseflies?
- Dysentery and malaria
  - Malaria and Tuberculosis
  - Dysentery and cholera
  - Cholera and Tuberculosis
12. Which of the following processes involve a change in the state of matter from liquid to solid?
- Condensation
  - Crystallization
  - Evaporation
  - Sublimation
13. An example of a disease vector is -----
- earthworm
  - liver fluke
  - tapeworm
  - tick
14. Which of the following statements about the properties of water are correct? It
- is colourless and tasteless
  - is neutral to litmus paper
  - turns red litmus paper blue
  - is a universal solvent
- I and II only
  - I and III only
  - I, II and IV only
  - I, III and IV only
15. When a green leaf is placed in a test tube containing ethanol and heated over a water bath, the leaf.....
- becomes soft
  - changes colour to brown
  - changes colour to blue-black
  - is decolourized
16. Which of the following substances is a compound?
- Hydrogen
  - Nitrogen
  - Oxygen
  - Water

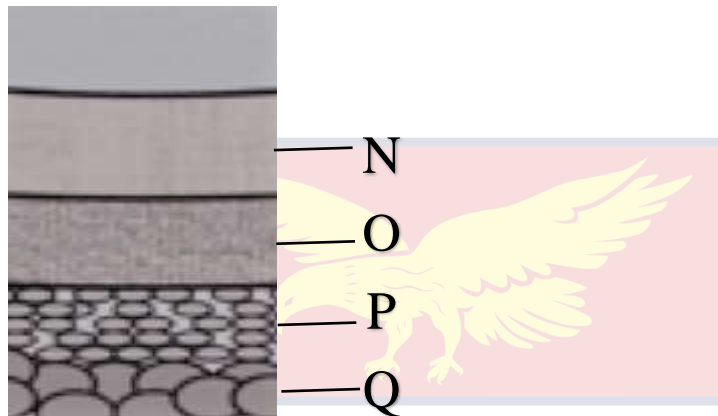
17. Which of the following organisms is/are multicellular?
- I. Amoeba
  - II. Paramecium
  - III. Onion
- A. I only
  - B. III only
  - C. I and III only
  - D. II and III only
18. Steel is an alloy of iron and .....
- A. aluminum
  - B. carbon
  - C. silicon
  - D. gold
19. Which of the pairs of structures are part of the respiratory system of humans?
- A. Fallopian tube and alveoli
  - B. Pharynx and oesophagus
  - C. Trachea and alveoli
  - D. Trachea and duodenum
20. Which of the statements about the scientific method is/are correct? It provides....
- I. Logical procedure for at knowledge
  - II. Knowledge that can be verified
  - III. Knowledge that can never change
- A. I only
  - B. I and II only
  - C. I and III only
  - D. II and III only
21. The scent of a perfume sprayed at one coner of a room fills the entire room through.....
- A. conduction
  - B. diffusion
  - C. osmosis
  - D. radiation
22. Atoms of the same elements have the same .....
- I. number of electrons
  - II. number of protons
  - III. chemical properties
- A. I only
  - B. II only
  - C. I and III only
  - D. II and III only



23. Which of the following energy transformation takes place at a hydroelectric power station?
- A. Potential energy → kinetic energy → electrical energy
  - B. Potential energy → sound energy → electrical energy
  - C. kinetic energy → electrical energy → light energy
  - D. electrical energy → light energy → light energy
24. A bottle of water removed from a refrigerator soon becomes covered with droplets because the ...
- A. bottle is porous
  - B. bottle is full of water
  - C. water in the surrounding atmosphere is colder than the water in the bottle
  - D. water in the surrounding atmosphere is warmer than the water in the bottle
25. A boy stepped on a banana peel on the ground and fell because the .....
- A. boy was not wearing shoes
  - B. ground was too hard
  - C. banana peel reduced the friction on the ground
  - D. banana peel increased the friction on the ground
26. A man did 75 work by filling a 50N load from the floor onto a shelf. Calculate the height of the shelf.
- A. 0.67m
  - B. 1.50m
  - C. 25.00m
  - D. 125.00m
27. The structure which helps a fish to breath under water is the .....
- A. Gill
  - B. Air sac
  - C. Operculum
  - D. Mouth
28. The following characteristics are features of all living things except the ability to ....
- A. Bring forth young ones
  - B. Excrete
  - C. Grow
  - D. Move from place to place
29. The greatest danger faced by mammals living on land is .....
- A. becoming too cold
  - B. becoming too hot
  - C. losing water through evaporation
  - D. Gaining too much water through absorption

30. Which of the following substances is added to the soil to reduce its acidity?
- A. charcoal
  - B. humus
  - C. lime
  - D. manure

*The figure below shows the arrangement of parts of the soil when some soil sample was shaken with water in a glass cylinder and allowed to settle. Use it to answer questions 31*



31. The layer which does *not* have the ability to retain water is .....
- A. N
  - B. O
  - C. P
  - D. Q
32. The waste products of respiration are
- A. glucose and oxygen
  - B. carbon dioxide and water
  - C. carbon dioxide and heat
  - D. oxygen and water
33. Which of the following plants reproduce(s) vegetatively?
- I. banana
  - II. mango
  - III. ginger
  - IV.
- A. II only
  - B. I and II only
  - C. I and III only
  - D. III only
34. The food substance found in yam is
- A. carbohydrate
  - B. oil
  - C. protein
  - D. vitamin

35. Which of the following factors promote germination?
- I. Adequate temperature
  - II. Moisture
  - III. Good soil
- A. I and II only
  - B. I and III only
  - C. II and III only
  - D. I, II and III
36. Which of the following substances breaks down fats into tiny droplets during digestion?
- A. Bile
  - B. Gastric juice
  - C. Pancreatic juice
  - D. Saliva
37. A free movement of the bowels is aided by .....
- A. sleeping at least 6 hours a day
  - B. washing one's body at least twice a day
  - C. eating fatty foods
  - D. eating foods containing fibre
38. How many days in a month does the moon appear?
- A. 7 days
  - B. 14 days
  - C. 21 days
  - D. 28 days
39. A smaller heavenly body which orbits a bigger one could be described as a
- A. comet
  - B. meteor
  - C. moon
  - D. sun
40. Which of the following substances is not a constituent of a balanced diet?
- A. Carbohydrate
  - B. Protein
  - C. Table salt
  - D. Vitamin

APPENDIX H

SECOND MOCK EXAMINATION (KEEA, 2019)

SOCIAL STUDIES 1  
OBJECTIVE TEST

Answer **all** questions in this section.

Duration: **45mins.**

Each question is followed by **four** options lettered A to D. Find out the correct option for each question and circle it on your question paper. Select only one answer to each question. An example is given below.

A system of government in which the views of the people are respected in the process of governance is.....

- A. dictatorship
- B. autocracy
- © democracy
- D. oligarchy

The correct answer is democracy which is lettered C and therefore answer C would be circled. Think carefully before you shade the answer spaces. Erase completely any answers you wish to change.

Now answer the following questions.

1. How many Centimetres will make one Kilometre?
  - A. 100,000cm
  - B. 10,000cm
  - C. 1000cm
  - D. 1.00cm
2. The Wli waterfall is located in which region?
  - A. Western Region
  - B. Eastern Region
  - C. Volta Region
  - D. Northern Region
3. The Dagombas are associated with.....
  - A. Nene Azumatey
  - B. Nii ayi Bontey
  - C. Naa Gbewaa
  - D. Yagbon Wura
4. The Asantes were great .....
  - A. fishermen
  - B. gatherers
  - C. warriors
  - D. miners

5. Which instrument is used in measuring atmospheric pressure?
  - A. Barometer
  - B. Thermometer
  - C. Isothermic
  - D. Isotherm
  
6. Who chairs the Regional Co-ordinating -Council (RCC)?
  - A. The President of the Regional House of Chiefs
  - B. The Regional Co-ordinating Director
  - C. The Regional Administrator
  - D. The Regional Minister
  
7. The Economic Community of West African States (ECOWAS) has its headquarters in which country?
  - A. Liberia
  - B. Nigeria
  - C. Gambia
  - D. Chad
  
8. Which of these institutions takes care of the child much?
  - A. Family
  - B. Church
  - C. School
  - D. Community
  
9. Daily temperature is got by .....
  - A. dividing maximum temperatures by two.
  - B. finding the sum of the maximum and minimum temperatures.
  - C. finding the difference between maximum and minimum temperatures.
  - D. finding the average of the maximum and minimum temperatures.
  
10. Which of these covers the largest area in Ghana?
  - A. Guinea Savannah
  - B. Sudan Savannah
  - C. Tropical Evergreen Rainforest
  - D. Moist Semi-deciduous Forest
  
11. Law and order is kept in the Nation by the .....
  - A. Immigration Service
  - B. Prison Service
  - C. Police Service
  - D. Armed Forces
  
12. The burning of lorry tyres causes.....
  - A. Heavy rains
  - B. Poor visibility
  - C. Low temperatures
  - D. High productivity

13. The .....was set up to find the causes of the 1948 Riots.
  - A. Gbedebe Committee
  - B. Watson Commission
  - C. Arden-Clarke Commission
  - D. Coussey Committee
  
14. DDT is used for illegal.....
  - A. logging
  - B. hunting
  - C. fishing
  - D. mining
  
15. The first President of the Republic of Ghana was from.....
  - A. Nzulenzu
  - B. Atuabo
  - C. Esiamam
  - D. Nkroful
  
16. Which of these is giving troubles to the District Assemblies?
  - A. Polythene
  - B. Kitchen waste
  - C. Papers
  - D. Wood
  
17. Our readiness to host others is.....
  - A. loyalty
  - B. kindness
  - C. dedication
  - D. hospitality
  
18. Friendship can be started through.....
  - A. frowning
  - B. greetings
  - C. slapping
  - D. avoiding
  
19. Which of these people has never ruled Ghana before?
  - A. Professor J. E. a. Mills
  - B. Dr. Kwasi Botchway
  - C. Dr. Hilla Limann
  - D. Dr. K. A. Busia
  
20. The wife of the President of Ghana is our .....
  - A. Queen mother
  - B. Grandmother
  - C. First – lady
  - D. Right – Honourable

21. What will be the time in a country on Long 75W if the time in Ghana (Long 0) is 2:17am?
- A. 7:17am
  - B. 9:17am
  - C. 7:17pm
  - D. 9:17pm
22. The act of not conforming to rules that govern a society is ----- behaviour.
- A. discipleship
  - B. disciples
  - C. indiscipline
  - D. discipline
23. Which of these is the shortest?
- A. Tropic of Capricorn
  - B. Tropic of Cancer
  - C. Artic Circle
  - D. Equator
24. A Ghanaian qualifies to vote when he/she is ..... years and above.
- A. 19
  - B. 18
  - C. 17
  - D. 16
25. There is ..... when the older generation teaches the younger generation their culture.
- A. socialisation
  - B. schooling
  - C. acculturation
  - D. transmittance.
26. How long does it take the earth to make a complete revolution?
- A. 365  $\frac{1}{4}$  days
  - B. 365 days
  - C. 366 days
  - D. 366  $\frac{1}{4}$  days
27. Ntim Gyakari was a King of which of these people?
- A. Fante
  - B. Asante
  - C. Denkyira
  - D. Mamprusi
28. Which of these is associates with rains?
- A. Volcanic eruption
  - B. Weathering
  - C. Tsunami
  - D. Flooding

29. The direction or position in relation to a fixed point measured in degrees is.....
- A. sketch
  - B. cardinal
  - C. position
  - D. bearing
30. The surface of the earth is made up of about .....of water.
- A. 30%
  - B. 40%
  - C. 70%
  - D. 80%
31. The Allan Burns constitution was written in which year.....
- A. 1925
  - B. 1946
  - C. 1960
  - D. 1992
32. The .....is directly elected into office.
- A. Vice President
  - B. Assembly Member
  - C. Minister of State
  - D. Chief Justice
33. The amount of goods and services produced by a worker within a working period is .....
- A. produce
  - B. production
  - C. productivity
  - D. product
34. Which of these businesses can be started with the least capital?
- A. Sole Proprietorship
  - B. Cooperation society
  - C. Partnership
  - D. Joint Stock
35. Which of these leaders ruled Ghanaian the Third Republic?
- A. Dr. K. A. Busia
  - B. Dr. Hilla Limann
  - C. Mr J. A. Kuffour
  - D. D. Gen. I. K. Acheampong
36. ....is an important feature that can be easily identified in a locality.
- A. Cardinal point
  - B. Landmark
  - C. Bearing
  - D. Compass



37. Slate is a metamorphic rock got from .....
- A. clay
  - B. granite
  - C. basalt
  - D. silica
38. The breaking down of rocks into smaller particles is .....
- A. rock formation
  - B. mass wasting
  - C. volcanicity
  - D. weathering
39. Which of these is the coldest?
- A. Europe
  - B. Australia
  - C. Antarctica
  - D. North America
40. The members of Parliament represent people from a/an.....
- A. Region
  - B. District
  - C. Constituency
  - D. Electoral area

