

**PREDICTIVE VALIDITY OF INTERNAL EXAMINATIONS  
IN SECONDARY SCHOOLS IN KENYA**

**OCHIENG PAULINE ATIENO**

**UNIVERSITY OF NAIROBI  
PSYCHOLOGY DEPARTMENT**

**UNIVERSITY OF NAIROBI  
PSYCHOLOGY DEPARTMENT  
SCHOOL OF EDUCATION**

**@ 2012**

**A RESEARCH PROJECT REPORT IN PARTIAL FULFILMENT OF THE  
REQUIREMENTS FOR THE AWARD OF THE DEGREE IN MASTERS OF  
EDUCATION IN MEASUREMENT AND EVALUATION.**

*All rights reserved. No parts of this work may be reproduced, stored in a retrieval system or transmitted in any forms or means whether electronic, photocopy, recording or otherwise without prior permission of the author or University of Nairobi.*

*@ Ochieng Pauline Atieno*

## DECLARATION

### Student's Declaration

This project is my original work and has not been presented for a degree in any other university.

**Ochieng Pauline Atieno**

**E58/74241/09**

### Supervisor's Declaration

This project has been submitted for examination with my approval as University Supervisor.

A handwritten signature in black ink, reading "Karen T. Odhiambo". The signature is stylized with a large initial 'K' and a long horizontal stroke extending to the right.

**.fofAaJma**

**Dr. Karen T. Odhiambo**

**Department of Psychology- Education Psychology**

**University of Nairobi**

## DEDICATION

I would like to thank my husband Anthony Thomas Osambo for the support, understanding during the period of my study. To, my children, Martha (Chausiku), Pendo (Ketty Ammy) Kaka (Ian Manase) and Dot P (Cherise Fassey) may God bless you for your tolerance for an absent mother.

Special tributes to my colleagues at work for their encouragement and willingness to take up my duties whenever I was away for my studies. I will never forget my parents and siblings; together they have achieved their dreams. Thanks to my brother Gabriel Ochieng, he was always there whenever i was down financially.

## ACKNOWLEDGMENT

It is not possible to express my sincere gratitude to all those who offered important advice, critique and encouragement in the development in this project. However, I wish to specifically thank my supervisor, Dr. Karen T. Odhiambo for her lovable support and insightful guidance that has not only enhanced the outcome of this study but also enriched it.

I must not forget to thank the schools which availed their examination score records for me to access my data and Pendo's role as the final reader of my project, cannot be ignored. Finally may the almighty God be praised for everything was successful due to his unconditional love. With him nothing is impossible.

## ABSTRACT

This study sought to establish the relationship between performance in internal (teacher-made) summative examinations and external summative public national examination. The objectives were to establish whether there is significant relationship between internal summative examination and external summative examination in secondary schools in Kenya, determine which subject has a greater capacity in predicting performance in external summative examination (KCSE), and the year with more weight in predicting performance in external summative examinations.

This study adopted a descriptive research design. The target population of the study comprised of all the students from four secondary schools who registered in Form 1 in 2007 and sat for their KCSE examination in 2010. Examination scores for internal summative examinations and KCSE examination scores for 60 students from form 1 in 2007 to form four in 2010 were purposively sampled from each of the four schools targeted making a sample of 240 students. The instrument used to collect data was an inventory. The inventory requested test scores derived from tests constructed, administered, and scored by teachers in secondary schools for Maths, English, Biology and Geography. Data was analysed using quantitative statistics. The relationship between the scores was determined by calculating correlation coefficient and results presented using tables and graphs with the aid of SPSS.

The study revealed that there is significant relationship between the internal summative examination scores and the external public summative examination scores with mathematics as a major predictor of student performance and that the students performance in the internal summative examination in first year cannot be used to predict the performance in the external summative examination. Fourth year is a major predictor followed by the performance in third year and then finally the performance in second year. The study recommends that the Kenya National Examination Council should standardise the internal summative examination conducted by teachers in their respective schools and that the student examination scores in the internal summative examination to be used to award students incase a student is unable to sit for external summative examination due to unavoidable situations.

## TABLE OF CONTENT

DECLARATION.....	iii
DEDICATION.....	iv
ACKNOWLEDGMENT.....	v
ABSTRACT.....	vi
TABLES OF CONTENT.....	vii
ABBREVIATIONS AND ACRONYMS.....	x
LIST OF TABLES.....	xii
LIST OF FIGURES.....	xiv
<b>CHAPTER ONE.....</b>	<b>1</b>
<b>INTRODUCTION.....</b>	<b>1</b>
1.1 Background of the Study.....	1
1.2 Statement of the Problem.....	5
1.3 Purpose of the Study.....	7
1.4 Objective of the Study.....	7
1.5 Significance of the Study.....	8
1.6 Justification of the Study.....	8
1.7 Definition of Key Terms.....	10

<b>CHAPTER TWO.....</b>	<b>12</b>
<b>LITERATURE REVIEW.....</b>	<b>12</b>
2.1 Introduction.....	12
2.2 Related Studies.....	12
2.2.1 The Concept of Predictive Validity.....	17
2.2.2 Conditions for Predictability.....	21
2.2.3 Why Predict Academic Performance.....	24
2.3 Forms of Examinations.....	26
2.3.1 Examination Practices.....	31
2.4 Examination System in Kenya.....	42
2.4.1 The Kenyan Education System.....	44
2.4.2 Examination Policy Making and Implementation.....	47
2.5 Theoretical Perspective of Assessment.....	50
2.5.1 Positivists (Scientific) View.....	50
2.5.2 Constructivism (Hermeneutic) View.....	51
2.5.3 Classical Test Theory.....	52
2.5.4 Item Response Theory.....	54
2.6 Conceptual Framework.....	56
<b>CHAPTER THREE.....</b>	<b>57</b>
<b>RESEARCH DESIGN AND METHODOLOGY.....</b>	<b>57</b>
3.1 Research Design.....	57
3.2 Target Population.....	57

3.4 Sampling Procedure.....	58
3.5 Data Collection Instruments.....	58
3.6 Data Analysis and Presentation.....	59
<b>CHAPTER FOUR.....</b>	<b>61</b>
<b>DATA ANALYSIS AND PRESENTATION.....</b>	<b>61</b>
4.0 Introduction.....	61
4.1 Biographic Information.....	61
<b>CHAPTER FIVE.....</b>	<b>85</b>
<b>SUMMARY, CONCLUSION AND RECOMMENDATION.....</b>	<b>85</b>
5.1 Introduction.....	85
5.2 Summary.....	85
5.3 Conclusions.....	87
5.4 Recommendation for practice.....	88
5.5 Recommendation for further Research.....	89
<b>REFERENCES.....</b>	<b>90</b>
<b>APPENDICES.....</b>	<b>102</b>

## ABBREVIATIONS AND ACRONYMS

<b>ANOVA</b>	Analysis of Variance
<b>CAT</b>	Cognitive Ability Tests
<b>CTT(s)</b>	Common Task Test(s)
<b>GCE</b>	General Certificate of Education
<b>GMAC</b>	Graduate Management Admission Council
<b>GMAT</b>	Graduate Management Admission Test
<b>GRE</b>	Graduate Record Examination
<b>HESI</b>	Health and Environmental Sciences Institute
<b>HLM</b>	Hierarchical Linear Models
<b>IRT</b>	Item Response Theory
<b>JCSE</b>	Junior Secondary School Certificate Examination
<b>JSC</b>	Junior Secondary School Certificate
<b>KCPE</b>	Kenya Certificate of Primary Education
<b>KCSE</b>	Kenya Certificate of Secondary Education
<b>KNEC</b>	Kenya National Examination Council
<b>LCE</b>	Leaving Certificate Examination
<b>MMPI</b>	Minnesota Multiphasic Personality Inventory
<b>NCSBN</b>	National Council of State Boards of Nursing
<b>NERDC</b>	Nigeria Educational Research and Development Council
<b>PR</b>	Percentile Rank
<b>SAT</b>	Scholastic Aptitude Test

<b>SEC</b>	Secondary Education Certificate
<b>SPSS</b>	Statistical Package for Social Sciences
<b>SSC</b>	Senior Secondary School Certificate
<b>SSCE</b>	Senior Secondary School Certificate Examination
<b>WASCE</b>	West African School Certificate Examination

**UNIVERSITY OF NAIROBI**  
•ttUVII I

## LIST OF TABLES

Table 2.1 Examination structure.....	49
Table 4.1 Gender.....	61
Table 4.2 School Category.....	62
Table 4.3 Subjects Selected for the Study.....	63
Table 4.4 Relationship between Internal summative examination and external summative examinations.....	64
Table 4.5: Mean Performance in the internal summative and external summative examination by subject.....	66
Table 4.6: Mean of medium performing students in KCSE and in internal summative examination by School category.....	70
Table 4.7: Performance of students in internal summative examination and external summative examination(KCSE) by gender.....	72
Table 4.8 Level of subjects in predicting performance in secondary schools.....	73
Table 4.9: Performance of boys and girls in internal summative examination by school category.....	76
Table 4.10: Performance of boys and girls by school category in the KCSE.....	77
Table 4.11: Mean of annual performance in internal and external summative examinations for boys and girls.....	78
Table 4.12: Performance of School Category by Gender in both internal and external examinations.....	79

Table 4.13: Performance of both boys and girls by subjects in internal and external summative examination by school category.....	80
Table 4.14: Determination of the year with more weight in predicting performance in KCSE.....	83

## LIST OF FIGURES

Figure 2.1: Structure and Organization of Education and Training in Kenya.....	46
Figure 2.2: Conceptual Framework for the Research.....	56
Figure 4.1 Relationship between Internal summative examination and external summative examinations.....	65
Figure 4.2: Mean performance of high achieving students in internal summative examination and external summative examination by School category.....	68
Figure 4.3 Mean of medium performing students in internal summative examination and external summative examination by School category.....	69
Figure 4.4: Mean of low performing students in internal summative examination and external summative examination(KCSE) by School category.....	71
Figure 4.5: Performance of boys and girls in internal summative examinations by subject	74
Figure 4.6: Performance of boys and girls in external summative examination (KCSE) by subject.....	75

# CHAPTER ONE

## INTRODUCTION

### 1.1 Background of the Study

There is a growing concern among stakeholders on whether to adapt internal assessment as this accounts for most of students assessment compared to public national examination. Also, issues that arise in cheating in exams and the cost of examinations have also called for a general concern on whether there is a relationship between student's internal summative examinations and external summative examinations. This study is designed to investigate the predictive validity of internal assessment through teacher made tests and summative assessment or public national examinations. The study is therefore intended to determine the predictive validity and success of internal examinations compared to public national examinations. The idea is to determine if there is a significant relationship between internal summative examinations and the external summative examinations. In Kenya the examination is set up such that there is internal summative examinations administered by teachers and the external summative (public national) examination constructed and administered by the examining board, (KNEC).

Predicting student's academic performance is critical for educational institutions because strategic programmes can be planned in improving or monitoring student's performance during study period in the institution, (Zaidah and Daliela, 2007). It can also be used for selection of students who will succeed in other academic endeavours. It can also be used

to reveal how successful teachers' instructions have been mastered, (Omirin and Ale, 2008).

Predictive validity is used to describe the capacity of measuring instrument to focus future performance in a related task (Graduate Management Admission Council, (GMAC), 2005.) Predictive validity can further be defined as the degree to which the operationalization can predict (or correlate) with other measures of the same construct that are measured at some time in the future.

According to Brown and Coughlin, (2007), it is the ability of one assessment tool to predict future performance either in same activity (success in college for example) or on another assessment of the same construct. The predictive validity of survey instruments and psychometrics tests is considered by scholars to be a measure of agreement between results obtained from more direct and objective measurements. The predictive validity is often qualified by the correlation coefficient between the two sets of measurements obtained by the same target population.

Internal summative examinations are teacher-made tests which are set by teachers within a school system. External summative examinations on the other hand are public examinations that are conducted by recognized examining bodies. According to Bloom, (1971) summative examinations are those given at the end of episodes of teaching (units and courses) for the purpose of grading or certifying students, or evaluating the effectiveness of a curriculum, The examining boards conducting these examinations do

not themselves organize instructional courses to prepare students for the but the national examination is designed and organized under specific terms and conditions and is based on norms that are regarded as standards (Salau, 1992).

Many situations arise that create conditions for predictions. This study will look at the predictability of internal summative examinations that are carried out at the end of every three months term (or semester). The teacher made examinations are done at the end of March and July as per term examinations. These examinations are formative in nature thus provide teachers with the information to be used as feedback to modify teaching and learning activities, Black, William and Dylan (1998). At the end of November, the internal examinations are done and they serve the summative purpose of examinations and help in making judgment about learning that has taken place. Internal summative examinations provide teachers with information they may use to draw conclusions about how well the student has attained the learning targets in the official curriculum which is the basis of the external summative examinations. The policy of pretesting students for external summative examinations is based on the assumption of those who do well in the internal summative examinations have a good chance of passing. Therefore, in preparing students for external summative examinations, future performance can be predicted from composite internal summative examinations scores.

Predicting student's academic performance is critical for educational institutions because strategic programs can be planned in improving or monitoring students' performance during their period of study in the institutions, (Zaidah and Daliela, 2007). According to

Omirin and Ale, (2008), predictions of examinations are meant to reveal how successful teachers' instructional objectives have been achieved. Predictive examinations are used for selection of students who will succeed in further academic endeavours. It also prepares students in readiness for final examinations. Universities use predictive examinations to make decisions regarding the potential of candidates' academic study (Zaidah and Daliela, 2007). For this purpose, admission examinations have been used as tools for predicting academic performances and ultimately for deciding admission into academic programs.

Daniels and Schouten, (1970) and Owoyemi, (2000) argued that a prediction of the future of an exam could be made with reasonable success on the basis of the results of previous examinations. However there are divergent views on predictability of some examinations. For example, findings made by Al-Shorayye, (1995), and Adeyemi, (1998) lent credence to this point. The findings supported the findings of other researchers that the GCE and Secondary School Certificate examination results provided the best predictor of the academic performance (1993).

Findings made by Peers and Johnston, (1994), confirmed the validity of the number and grades of passes in the Scottish Certificate of Education in predicting first year and final year university performance. Gay, (1996) also reported that high school grades could be used to predict college grades. These findings were contrary to O'Rourke, Martin and Hurley's (1989) findings that the Scholastic Aptitude Test (SAT) is unable to predict

examination performance as effectively as the Leaving Certificate Examination (LCE) point scores.

Various researchers have had divergent findings on the predictive validity of some examinations (Alonge, 1998). In other developing countries, the index of academic performance varied from one country to another. Othuon and Kishor, (1994), found that the Kenya Certificate of Primary Education scores had a moderate positive linear relationship with the Certificate of Secondary Education grades. In some other States, performance in JSC examinations has been found to be significantly related to the performance in SSC examinations, (Adeyemi, 2001). Their study established that the policy of selecting students for secondary schools education is based on the assumption that those who pass the selection exams have a good chance of succeeding in secondary schools. In their study the predictive validity did not vary from school to school. In other words future performance in the Kenya national examination is purely predictable from performance of teacher made examinations.

### **1 1.2 Statement of the Problem**

Predictive validity involves the assessment of some or all aspects of the subject ability of individuals in some context and for some set of purposes (not necessarily common to all parties). Hence, the importance of predictive validity for examinations cannot be neglected. Predictive validity of examinations enhances the value of examinations to students and other practitioners in the education sector. It is thus important to understand

the predictive validity of internal examinations in secondary schools in relation to the national examinations.

Despite the importance of the need to understand predictive ability of examinations only a few studies have been conducted in this area. The available studies which have investigated predictive ability of an examination or test done at an early stage to predict performance in subsequent examinations have been conducted in different context from the approach taken by this study. Most of the studies have been undertaken in other countries such United States (Kobbrin and Schmidt, 2006) and other developed countries (Daniel and Schouten, 1970; Gonnella, and Rothstein, 2004; Daniel, 2011). Only a few have been conducted in Africa mostly in Nigeria (Obioma and Salau, 2007; Adeyemi, 2008; Omirin and Ale, 2008). Two studies available study have been conducted in Kenya by Othuon, (1994) and Othuon and Kishor, (1994) who studied the predictive validity of KCPE in determining success of students in KCSE. No study has been conducted to assess the predictive validity of internal examinations done by students in secondary schools determining success in the KCSE examinations which are conducted by an external body (KNEC). The issue of predictive validity of internal examinations is an area of concern as it appears unfair to students to place the weight of evaluating their worth on one single examination which comes at the end of four years. Examination system in Kenya is based on internal summative examinations in schools and external summative public examinations. There has been debate on use of internal summative examination scores to award certificate to rescue students, for example, those who may fall victims of KCSE examination results cancellation. A study of this nature may be

useful in informing the debate. It would help address this issue by showing if there is a difference in performance in internal summative examinations and external summative public national examination. It is from this perspective that this study was undertaken to establish whether the internal examinations undertaken by students in secondary schools can be used to predict performance in the final KCSE examinations with a case of six schools in Nairobi.

### **1.3 Purpose of the Study**

The purpose of this study was to determine if performance in internal summative examinations can predict student performance in external summative examinations in secondary schools in Kenya.

### **1.4 Objective of the Study**

The study was guided by the following objectives;

- i) To establish whether there is significant relationship between internal summative examination and external summative examination in secondary schools in Kenya
- ii) To find out which subjects has a greater capacity to predict performance in the external examination( KCSE) in secondary schools in Kenya
- iii) To determine the year with more weight in predicting performance in KCSE, 2010

### **1.5 Significance of the Study**

This study was designed to establish predictive validity of internal examinations in secondary schools. There were only few available studies that have focussed on the predictive validity of internal examinations in Kenya. It was therefore in this line that the current study will offer teachers in secondary schools some guidance on the predictive validity of internal examinations. The findings of this study will also be of importance to educational policy makers in that they will be in a position to base their arguments based on the findings of this study. This study may also be of great importance to the Quality Assurance Standard department and the KNEC- can use the results to determine the effectiveness of the teaching and learning processes in schools. Finally, the study is expected to be of importance for it may lay the foundation for further research by various researchers and academicians who would be interested in the same area given that very few studies in Kenya have considered on predictive validity of internal examinations in secondary schools in Kenya.

### **1.6 Justification of the Study**

The findings of this study will benefit the school managers as well as the students on the importance of predictive validity. The findings of the study will also be of importance to the researchers as it can be used as a basis for further research. If the internal examination scores have no relationship with external examination scores then a research can be done on factors affecting the predictive validity of internal examinations. This can lead to improvement on internal examinations so that the scores can be integrated with the external examination scores for making judgments. Some of the studies conducted asserts

that measures of predictive validity of the results obtained in examinations at a particular level can be used with benefit for making decisions about which subjects to study or which courses to take at a higher level. For example, a study carried out by Farrugia, (2004) in the Maltese context revealed that knowledge of the predictive validity of the results obtained in examinations at Secondary Education Certificate (SEC) level can help students make the right choice of which subjects to study at Advanced Matriculation level. This study suggested that research should be carried out to find out whether it can also provide them with an indication of the probability of obtaining a good grade at this level.

## 1.7 Definition of Key Terms

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

**Test:** A test is a task, a set of tasks or questions intended to elicit a particular behaviour when presented to learners under standard conditions (Powell, 2010).

**Validity** This refers to what is assessed and how well this corresponds with the behaviour or construct that it is intended to test, Harlen (2005). It refers to the precision with which a test 'measures same particular mental ability, Othun (1994),

**Predictive Validity:** Refers to the relationship between test scores and later performance on knowledge, skill or ability (DeVellis, 2011). It is the ability of assessment tool to predict future performance either in some activity or on another assessment on the same construct, Brown and Coughlin (2007).

**Examination:** A set of questions or exercises to gauge understanding on knowledge or skill or ability (Chen et al., 2005)

**Standardized Test:** Refers to a test that is administered and scored in a consistent manner to ensure legal defensibility (Popham, 2001).

**Teacher-made tests:** These are constructed, administered and scored by teachers in schools to gauge the performance of students in tasks, abilities and knowledge (Powell, 2010).

**Internal Examinations:** These are examinations which are set administered and scored by teachers within a school system for the purpose of evaluation and grading, Adeyemi (2008).

**Summative Examination:** These are examinations given, by teachers, at the end of units, mid-term and at the end of a course, which are designed to judge the extent of

students learning of the material in a course for the purpose of grading, certification and evaluation of progress or even for researching the effectiveness of the curriculum, Bloom et al [1971].

**Formative Examination;** These examinations encompasses all those activities undertaken by teachers and by the students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged, Black, William and Dylan [1998],

**External examination:** These are examinations which are designed and conducted by recognized examination bodies (Adeyemi, 2008).

\*u\*fc»Sffy OF N  
\* U V U I. 'KffAC" L

**Public national examination:** These are external school examinations open to the general public and conducted by these examining bodies using tests that have appropriate psychometric properties, Obioma and Salau [2007]

**National assessment examination:** This is an exercise designed to describe the level of achievement, not of individual student, but a whole education system or a clearly defined part of it, Kellaghan & Greaney (2003).

**Internatinal assessment examination;** These are examinations which are administered in more than one country to provide evidence on the extent to which the treatment of common curriculum areas differ (Beaton et al, 1999; Kellaghan & Greaney, 2001b; Greaney & Kellaghan, 1999a).

**Performance;** This is the actual accomplishment showing the learner's ability as measured by a score on an achievement test. It is the level of achievement a student attains in the examination, (Thobega et al, 2008).

\*»\*UYtl //apjtflv

## **CHAPTER TWO**

### **LITERATURE REVIEW**

#### **2.1 Introduction**

This chapter summarizes the information from other researchers who have carried out their research on the subject of predictive ability of examinations. The study reviewed relevant studies related to predictive validity. The specific areas covered are related studies on predictability of examinations, the Kenyan education system, the examination system as well as theoretical perspective of assessment.

#### **2.2 Related Studies**

The predictability of examinations has been a matter of concern to many researchers, Afolabi and Faleye, (1998) carried out a research on the predictive validity of Osun State Junior Secondary Certificate Examination due to growing concern among stakeholders about the predictive validity of the state version of the JCSE for the Senior Secondary School Certificate Examination (SSCE). The study was to find out whether there is significant relationship between the overall performance of students in the JCSE and their performance in the Senior School Certificate Examinations (SSCE). The sample constituted the schools in Osun State (of Nigeria). Promotion examination scores of the students in Senior Secondary School (SSS) one (1) and SSS two (2) as well as their SSCE in the six major subjects were compared with the corresponding JCSE scores using correlation analysis procedures. The results showed that Osun State JSCE is a poor prediction of students' performance in SSCE. However, 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 overall performance in JSCE across the six subjects investigated is a poor predictor of SSCE performance (except English and Mathematics). The trend could be due to the constraints facing the MOE which perform the role of an examination body. In a study by the NERDC of 2001, public national examination scores were found to be very poor in predicting university achievements at the university level. On the contrast, continuous assessment scores obtained from cumulative point average at the university were found to have a very high predictive value on final year performance. Obioma and Salau (2007) recommended that admission of candidates into universities be based on merit, and that achievement test scores from national public examinations be replaced by other tests, such as aptitude test scores. Unlike in the above case, this study is intended to include an investigation into the predictive validity of internal examination scores to final KCSE scores in Kenya and then explore the possibility of using such scores as alternatives or supplement, to national examination scores for quality improvement.

A research on predictive validity of internal examination was also carried out by Omirin and Ale, (2008) to investigate the predictive validity of English and Mathematics Mock Examination Results of Senior Secondary School students' performance in WASCE in 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 Ekiti State, Nigeria. The study made use of the already existing data of the results of

WASCE of students and unprocessed raw scores of mock examinations from various selected schools.

Mock English and Mathematics helped significantly in predicting the success in academic performance of students in WASCE. However, English was a better predictor of success than Maths. It was then recommended that Mock examination should be made compulsory for students intending to sit for WASCE as it has been found as helpful in predicting their performance in WASCE. Adeyemi, (2008) carried out a research to examine the predictive strength of the Junior Secondary Certificate (JSC) examinations in predicting the performance of students in the Senior Secondary Certificate (SSC) examinations in Ondo State. The study revealed that JSC examinations were a good predictor of performance at SSC examinations.

According to Kobbrin and Schmidt (2006), the American Institutes to Research, in collaboration with the College Board, with their recently completed a study based on approximately 1,200 first year students from 13 colleges to examine the predictive and placement validity of new Scholastic Achievement Test (SAT) writing section. The prototype evaluated was 10 minutes shorter and had 12 fewer questions that the operational new writing section will have. The results indicated that total scores on the writing section correlated about 0.46 with first-year college grades (0.46 versus 0.43), the Sat writing scores were slightly worse than high school grades.

These results suggest that the writing section will increase predictive validity of the SAT, and will be valuable in course replacement. One of the goals of adding a writing section to the new SAT is to improve the validity of the test for predicting college success. Since the SAT test item construction, format, administration and scoring are all similar to the ones used in KCSE examinations, the usage of SAT test scores for purposes of such as placements, predicting future success and measuring performance would exhibit similar to the use of K.C.S.E, (Nsude, 1998). This argument implies that the changes that were made to improve the predictive validity of SAT by re-organizing it and increasing the test items could also apply in the K.C.S.E. examinations.

In a case of study by Nigeria Educational Research and Development Council (NERDC) the predictive validity of public examinations was investigated with an overall aim of examining the extent to which public examinations predict university student's academic achievements in Nigeria. The data was gathered from the official academic records. Correlation and ex-post factor designs were employed to investigate the relationship that exists between the performance in public examination (predictor variables) and university students' academic achievement (criterion variable). The latter design was used because both the cause and the effect had had already occurred while the data involved in the study were as they were collected from the source without any manipulation. Data was analysed using Pearson Product Moment Correlation Coefficient and forward inclusion Stepwise multiple linear regression analyses applied (Cohen and Cohen, 1983). All analyses were done using SPSS version 13 for windows. Independent Sample two-side t-tests were computed for continuous variables and t-tests as well as analysis of variance

(ANOVA) were performed to test adequacy of the postulated models. The level of significance was set at 0.01 for all statistical tests.

The results of this study revealed that even though public examinations were statistically significant, they were not of much practical importance in predicting the achievement of university students. However, it was established that the first year university examination results accounted for about 0.482 of the final year examination results. This indicates that formative evaluation plays a significant role in predicting achievement of university undergraduates (Obioma and Salau, 2007). The findings also suggest that admission of candidates into the university should be strictly based on merit. School teachers and university lecturers need to maintain the integrity of their examinations.

In a study by Othuon Lucas, (1994), the predictive validity of KCPE was investigated with an aim to determine the extent to which KCPE predicts success in secondary school. Success in secondary school was measured by the level of examinee achievement in the Kenya Certificate of Secondary Education (KCSE) examination stratified random sampling was used to select 26 secondary schools within a single district in Kenya. The 1991 KCSE data for 781 examinees in the sample were used in the analysis. The KCSE records for examinees in the sample were matched with corresponding 1987 KCPE records. The nature of the relationship was determined by use of Hierarchical Linear Models (HLM). The influence of selected moderator variables on the relationship between KCPE and KCSE was investigated as well. These variables were age gender, repetition of standard eight (8) and school size.

A moderate linear relationship between KCPE and KCSE was found. The predictive validity did not significantly vary from one school to the other. Of the three pupil level moderator variables used in the study, only age showed a significant influence on the KCPE-KCSE predictive relationship. A moderate linear relationship parallel regression slopes, and the extent to which the selected moderator variables influenced the KCPE-KCSE relationship indicates that KCPE is a moderately valid predictor of success in secondary education.

### **2.2.1 The Concept of Predictive Validity**

Predictive validity is the ability of assessment tool to predict future performances either in some activity or on another assessment on the same construct, (Brown and Coughlin, 2007). Predictive validity refers to the degree to which the operationalization measures can predict (or co-relate)with other measures of the same construct that are being measured at sometime in the future (DeVellis, 2011).

Predictive validity which refers to the degree of correlation between the measure of concept and some future measure of the same concept requires all assessments to have evidence of the reasonableness of the proposed interpretation, as test data in education have little or no intrinsic meaning. Because of the passage of time, the correlation coefficients are likely to be somewhat lower for predictive validation studies. Both types of validity are estimated with simple correlation coefficients, (Gonnella and Rothstein, (2004).

The constructs purported to be measured by assessments are important to students, faculty, or in general the school management and therefore requires solid scientific evidence of their meaning. Assessments are not valid or invalid; rather, the scores or outcomes of assessments have more or less evidence to support (or refute) a specific interpretation (such as passing or failing a course). Validity is approached as hypothesis and uses theory, logic and the scientific method to collect and assemble data to support or fail to support the proposed score interpretations, at a given point in time. Data and logic are assembled into arguments - pro and con - for some specific interpretation of assessment data.

The concept of predictive validity is used to describe the capacity of measuring instrument to focus future performance in a related task (Graduate Management Admission Council, (G.M.A.C, 2005). Gonnella, and Rothstein, (2004), define predictive validity as a measurement of how well a test predicts future performance. It is a form of criterion validity, in which the validity of a test is established by measuring it against known criteria. In order for a test to have predictive validity, there must be a statistically significant correlation between test scores and the criterion being used to measure validity. Investigation into the predictive validity of public examinations on students' future academic achievement in various contexts is well known (Gonnella, and Rothstein, 2004).

In predicting academic performance, Daniel and Schouten, (1970), emphasized the use of grades in examination and reported that grades could serve as prediction measures. They

argued that a prediction of a future examination could be made with reasonable success on the basis of a previous examination. Findings by Al-Shoree, (1995) led credence to this point thus confirming the outcomes of other researchers that GCE and secondary certificate examination and results provided the best predictor of university performance, (Uboko Bong, 1993).

In a study of predictive validity, the test scores are collected first; then at some later time the criterion measure is collected. Here the example is slightly different: Tests are administered, perhaps to the students, and then after those students do the tests for a year, the test scores are correlated with their first year job performance scores, (Al-Shoree, 1995). Another relevant example is SAT scores: These are validated by collecting the scores during the examinee's senior year and high school and then waiting a year (or more) to correlate the scores with their first year college grade point average. Thus predictive validity provides somewhat more useful data about test validity because it has greater fidelity to the real situation in which the test will be used. After all, most tests are administered to find out something about future behaviour.

In a study conducted by Daniel, (2011), on predictive validity of the graduate record examination advanced psychology test for grade performance in graduate psychology, he found out that Graduate Record Examination, (GRE) scores are often considered during the admissions process for prospective graduate students and there have been several assessments of the predictive validity of the GRE for graduate students in psychology.

The purpose of his study was to investigate the predictive validity of the GRE Advanced Psychology Test for students' grade performance in selected graduate courses. Academic records were evaluated for 236 graduate students in professional psychology programs. Higher GRE scores were significantly correlated with higher grades in several specific courses. These findings indicate that, in some instances, GRE Advanced Psychology Test scores are significantly correlated with subsequent performance in selected graduate psychology courses.

The most common use for predictive validity is inherent in the process of selecting students for university. Most universities use high-school grade point averages to decide which students to accept, in an attempt to find the brightest and most dedicated students. In this process, the basic assumption is that a high-school pupil with a high grade point average will achieve high grades at university. The most common use for predictive validity is inherent in the process of selecting students for to join other higher institutions. Most universities and colleges use high-school grade point averages to decide which students to accept, in an attempt to find the brightest and most dedicated students. In this process, the basic assumption is that a high-school pupil with a high grade point average will achieve high grades at university (Owoyomi 2000).

Quite literally, there have been hundreds of studies testing the predictive validity of this approach. To achieve this, a researcher takes the grades achieved after the first year of studies, and compares them with the high school grade point averages (Powell, 2010). A high correlation indicates that the selection procedure worked perfectly, a low correlation

signifies that there is something wrong with the approach. Most studies show that there is a strong correlation between the two, and the predictive validity of the method is high, although not perfect. Intuitively, this seems logical; previously excellent students may well struggle with homesickness or decide to spend the first year drinking beer. By contrast, underachieving college students often become dedicated, hard-working students in the relative freedom of the university environment (Moghaddam, 2010).

### **2.2.2 Conditions for Predictability**

Validity is often assessed along with reliability - the extent to which a measurement gives consistent results. An early definition of test validity identified it with the degree of correlation between the test and a criterion. Under this definition, it is evident that reliability of the test and the criterion places an upper limit on the possible correlation between them (the so-called validity coefficient). Intuitively, this reflects the fact that reliability involves freedom from random error and random errors do not correlate with one another. From this observation it holds that, the less random error in the variables, the higher the possible correlation between them (West and McCracken, 1998). Secondly, this definition, a test cannot have high validity unless it also has high reliability. However, the concept of validity has expanded substantially beyond this early definition and the classical relationship between reliability and validity need not hold for alternative conceptions of reliability and validity.

For reliability and validity to hold, Folabi and Faleye, (1998) asserts that an examination/test must produce consistent results, and not be significantly influenced by outside factors. He also asserts that validity is perhaps the most important quality of a test and that a valid test has to measure what it's intended to measure. Proper measures have to be put in place. Psychometrics is used to describe psychological tests (typically those used in educational and occupational settings). Good instruction requires a constant stream of information about students' progress or about possible reasons for students' lack of progress. Both students and teachers benefit during the teaching-learning process. Teachers after training are assumed to have the capability of teaching as well as designing valid and reliable tools of assessment to test students' academic achievement (Giacomini and White, 2006).

Questions of the adequacy of an examination as a measure of the characteristics it is interpreted to assess are answerable on scientific grounds by appraising psychometric evidence (Messic, 1980). Since all psychological measurements are subject to error. It is rare to set perfectly a reliable examination. Reliability in this context is concerned with the extent to which these errors are manifested. An examination is said to be reliable if results of individuals could be replicated upon writing the same examination again under similar conditions in an attempt to determine the degree of relationship between a predictor and criterion. It is important that errors of measurement are minimized, (Crockers and Algina, 1986). This implies that the predictor and criterion should be reasonably reliable.

Assessing the content validity of the test is undertaken to examine the match between the test items and the content, skills, and behaviours being tested. The items on the test represent a sample of content, skills, and behaviours, and the test constructor has to ensure that this sample is representative of those contents, skills, and behaviours being tested. Students' performance on the test, represented by the test scores, is used to make inferences about the students' mastery and knowledge of the materials taught in class. Therefore, it is important that the test has an acceptable level of content validity (Gansky et al., 2001).

The question of reliability pertains to the consistency of the test and the test items. The test constructors should ask themselves whether students taking the same test more than one time would obtain the same scores upon repeated testing. Most approaches to assessing test reliability involve using statistical programs or lengthy computations. However, there is one simple way to assess the test reliability that is called split half reliability. In this procedure, the whole test is divided into two halves and students' scores on the two halves are correlated using statistical tests, such as Pearson correlation. The test can be divided into two parts by placing the odd-number items in one half and the even-numbered items in the second half. The scores for each student on the two halves are correlated and the correlation coefficient is used as the index of reliability. Any other approach that would create two similar halves is acceptable because the goal is to create two similar forms of the test ( Othuon, 1994).

Test's reliability can range from 0.00 (no reliability) to 1.00 (perfect reliability). Commercial tests are likely to have reliability in the middle to high .90s for the total test, but the reliability is usually lower for subsections of the tests. Test reliability depends on several factors, not just the quality of the test items. One main factor is the length of the test. Shorter tests tend to have lower reliability than longer tests because longer tests provide a more consistent sample of students' abilities and performances, compared with shorter test where there is more room for chance and guessing (Graham and Whittaker, 2005).

### **2.2.3 Why Predict Academic Performance**

The term 'academic performance' has been described as the scholastic standing of a student at a given moment (Adeyemi, 2008). It refers to how an individual is able to demonstrate his/her intellectual abilities. The scholastic standing could be explained by the grades obtained in a course or groups of courses (Daniel and Schouten, 1970; Owoyomi, 2000). Predicting student's academic performance is critical for educational institutions because strategic programs can be planned in improving or monitoring student's performance during their period of study in the institution (Zaidah and Daliela, 2007). Prediction of examinations can be used for selection of students who will succeed in further academic endeavours (Omirin and Ale 2008). It also prepares students in readiness for final examinations.

Predictions of examinations are meant to reveal how successful teachers' instructions have been mastered, Omirin and Ale (2008). Teachers use prediction examinations (internal examinations) to know how students are progressing and where they are having trouble they can use the information to make necessary instructional adjustments such as re-teaching, trying alternative instructional approaches or offering more opportunities for practice. These activities can lead to Predicting individual student's performance attracts considerable interests in the business world when making decisions as to whom to hire and whom to promote.

The main products of universities are students. Upon graduation, the students may either continue into the post graduate programme or become manpower for the industry, government and private sector, Zaidah and Daliela (2007). Thus the student's performance is critical in ensuring the supply chain is fulfilled. Universities therefore, use predictive examinations to make decisions regarding the potential of candidates' academic study. For this purpose, admission examinations have been used as tools for predicting academic performances and ultimately for deciding admission into academic programs.

Predictive validity of examinations can be further used to provide explanations about why people do or do not perform a particular behaviour and to suggest strategies for changing that behaviour, Azjen and Fishbein (1980) and Lutz (1975). Prediction of a future examination could be made with reasonable success on the basis of the results of previous examinations, Daniel and Schouten (1970).

Predictive examinations are designed to determine each student's likelihood of meeting some criterion score on the end of year tests. A good example is the formative examinations which are conducted by teachers at the classroom level to assess learning in preparation for teachers' summative examinations at the end of the year to determine achievement of learning objectives of the course content for that particular level of study. For both of these decisions, information is used to infer future performance. Track record in previous employment, academic training, interviews, references and personality and intelligent tests are some of the information that are used for deciding (Kuncel, Hezlett, and Ones, 2001).

### **2.3 Forms of Examinations**

A test is a task, a set of tasks or questions intended to elicit a particular behaviour when presented to learners under standard conditions (Noll and Scounel, 1972). It can also be considered to be a tool or a device for obtaining a sample of an individual behaviour to measure student's progress in terms of knowledge and skills after instructions, Onocha and Opkalla (1995). Tests can be classified into two major types' namely cognitive Ability test and Non-Cognitive tests. In Cognitive Ability Tests (CAT) there are Intelligence Test (mental ability test), Aptitude Test, and Achievement Test. Cognitive ability tests try to find out how much the learner knows.

The Non-Cognitive Tests are those tests, which try to find out personal qualities and behaviour characteristics of the learner. The information obtained through the administration, collection and collection of these tests are very useful for purposes of

counselling and curriculum planning. The different types of non-cognitive tests or techniques are checklists, projective technique, personality inventory and attitude scale. They assess the affective behaviour and personality of individuals. Non-cognitive tests try to ascertain how the individual thinks, reasons, and feels, his attitudes, interests, values, preferences, personal-social adjustments and other behaviour characteristics (Powell, 2010).

To measure student's academic performance, teachers use achievement tests and learning achievement is seen from student examination performance (Ayot and Briggs, 1992). Achievement tests (examinations) are cognitive ability tests, which measure how much the individual has been able to learn from what has been taught formally, often in classroom setting. The results of achievement tests are used for promoting pupils to new classes and awarding certificates which are either for furthering education or securing a job to earn a living (Nzweze, 2009).

In a secondary school setting, teacher test refers to tests and other measures that are planned, assembled, written, or otherwise prepared by teachers for use with particular groups of students. Achievement tests include standardized and non-standardized tests. Standardized tests refer to a test that has been screened for reliability and validity on a large population and calibrated on the group and test takers for whom it is intended (National Research Council (NRC), 1999). A standardized test is one that is administered and scored in a consistent manner to ensure legal defensibility.

A standardized test is a test that is administered and scored in a consistent, or "standard", manner. Standardized tests are designed in such a way that the questions, conditions for administering, scoring procedures, and interpretations are consistent, (Popham, 1999). and are administered and scored in a predetermined, standard manner. A standardized test which has important consequences to the individual examinee is referred to as high-stakes test. Standardized tests are developed to measure skills and knowledge learned in a given grade level, usually through a planned instruction such as training or classroom instruction. Non-standardized or teacher-made tests are constructed, administered and scored by teachers in schools (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999).

There are some advantages to standardized tests. Standardized tests are used to evaluate students and schools; to help improve teaching and learning; and to generate important data from which policy decisions can be made. They are cheap, very quick to grade, and they allow analysts to look at a wide sample of individuals. For this reason, they are often used to measure the progress of a school, by comparing standardized test results with students from other schools. However, standardized tests are ultimately not a very good measure of individual student performance and intelligence, because the system is extremely simplistic. A standardized test can measure whether or not a student knows when the Magna Carta was written, for example, but it cannot determine whether or not the student has absorbed and thought about the larger issues surrounding the historical document, (Popham, 1999).

Studies on the format of standardized tests have suggested that many of them contain embedded cultural biases which make them inherently more difficult for children outside the culture of the test writers. Although most tests are analyzed for obvious bias and offensive terms, subconscious bias can never be fully eliminated. Furthermore, critics have argued that standardized tests do not allow a student to demonstrate his or her skills of reasoning, deductive logic, critical thinking, and creativity. For this reason, some tests integrate short essays. These essays are often given only short attention by graders, who frequently vary widely in opinion on how they think the essay should be scored (Popham, 1999).

Tests are often used in education professional certification, counselling psychology (MMPI), the military and other fields. The measurement which is the goal of testing is a test score and is a summary of the evidence contained in an examinee's responses to the items of a test that are related to the construct or constructs. Test scores are interpreted with regard to a norm or criterion or occasionally both. The norm may be established independently or by statistical analysis of a large number of students. Teachers use the tool-test-to evaluate students' progress. For example; to help students respond to essay items, the teacher (or test constructor) should set the task clearly and provide guidelines to help the students compose their answers. For example, the teacher might suggest to the students the desired length of the response, as well as phrase the task using words such as: explain, defend, predict, and illustrate. Completion items (also called fill-in-the-blank items) require the students to construct their own answers by filling in or completing a sentence from which a word or a phrase has been omitted. After administering the test,

the teacher has to analyze and interpret the result of the test. The teacher has to be confident that the test is valid and reliable, before making instructional decisions (e.g., assigning grades and placement of students) that are based on the results of the test. The test as a whole, as well as individual items, should be studied and assessed. This is especially important if the teacher plans to continue to use the test in the future (Kuhn et al., 2001).

From this it is clear that tests serve the functions diagnostic, motivational, self-evaluation Instructional, and improving teacher effectiveness, respectively (Kryspin, 1994). As a diagnostic tool tests indicate how prepared students are to profit from new instruction and to what degree they've mastered previous concepts and skills. Also tests serve a diagnostic role by highlighting student's strengths and weaknesses as well as trouble spots needing special remedial attention.

Standardized tests include a report of the examinee's percentile rank, where the student's score is compared to that of other similar students in the norming group or in the same school or district. The percentile rank indicates the relative position of a student and it shows the percent of students who scored at or below the student. For example, if the student's percentile rank (PR) is 75, it means that the student did as well as or better than 75% of the student in the class (Moghaddam, 2010).

On the other hand tests motivate students to study the material assigned. They want to succeed academically, they fear failure, or they want to compete when given prompt feedback they are motivated to improve, to alter their mistaken concepts and in some instances to define more deeply into the subject matter (Button, 1995).

Tests are also used for Self-evaluation as when students get feedback on tests, they learn how others appraise their efforts and abilities. In turn, they develop their own self evaluation skills. Further tests cause students to review material and consolidate and integrate ideas. In addition, preparing for tests provide them with an overall review of the lesson and its important aspects. Finally tests are used to improve teacher effectiveness as they indicate to the teacher how well he covered the material, what areas need improvement and how to organize the material for a clearer presentation to the students (Okuna, 2006).

### **2.3.1 Examination Practices**

It is always important to find out the extent the learners have acquired the knowledge and skills in the instructional objectives based on the curriculum objectives of secondary education. This process is referred to as assessment. Assessment is the process of gathering information for the purposes of making decisions about educational policy, curriculum and educational programmes or about individual student's learning, Nitko (1994). Assessment of student's academic achievements is done through examinations/or tests. An examination is an assessment often administered on paper or on the computer, intended to measure the test takers or respondents (often a student) knowledge, skills,

aptitudes or classifications in many other topics such as beliefs Gagne, Briggs and Wagner (1988). Examination is the process which comes after a period of learning and it is an organized assessment of an individual's performance, on the basis of his/her institutional procedural exposure (Adeyemi and Akindele, 2002). Nsude (1998) avers that examinations are tools intended for evaluation of the progress made by an individual in the course of acquiring skills and knowledge over a period of time. Examinations have to do with passing of value judgement on an individual's performance in a set of questions, statements or series of tasks given, with the intention of assessing how much a desired trait, skill or knowledge the individual possesses, (Nzweze, 2009).

Examinations are broad tests that may be formally written, spoken or practical while a test constitute one task an examination involves several tasks and may include a wide variety of topics measuring academic achievement. Examinations are in two categories: internal examinations and external examinations (Chen et al., 2005).

#### **a) Internal teacher made Examinations**

Internal examinations are examinations that are set by teachers within a school system. Internal assessment is set and marked by the school (teachers). Students get the CAT, and assignments marks and feedback regarding the assessment, Adeyemi (2008). They are also referred to as teacher-made tests. Internal examinations are carried out using non-standardized achievement tests. Internal examinations serve both formative and summative purposes of examinations therefore constitute both formative and summative internal examinations. Teacher's assessment on the other hand is an ongoing process in

teaching and learning situations and is used mainly for diagnostic purposes. Teacher's assessment provides feedback on challenges that should be addressed to ensure effective provision and acquisition of knowledge and skills through the school system.

According to Udo Bude, (1995), even where teachers use written tests, many of the tests they develop are wanting in quality. The tests are usually wanting in originality of style, clarity of language and abilities tested. Some teachers do not even bother to develop their own tests. Instead they simply lift questions from past national examinations or from commercial publications. Rather than help improve the learning process, these practices seem to encourage rote learning, which both the teacher and the pupils believe would improve performance in national examinations.

#### **i) Formative Teacher Made Examinations**

Formative assessment is generally carried out throughout a course or project. Formative assessment, also referred to as "educative assessment," is used to aid learning. In an educational setting, formative assessment might be a teacher or the learner, providing feedback on a student's work, and would not necessarily be used for grading purposes. Formative assessments are diagnostic. Formative examinations encompasses all those activities undertaken by teachers and by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged, Black, William and Dylan (1998). Formative examinations focus on monitoring and guiding student through the curriculum and therefore are an assessment for learning

(Nitko, 1994). Examinations become formative when the information is used to adapt to teaching and learning to meet student's needs (Carol and Boston, 2002).

According to Boston and Carol (2005), when teachers know how students are progressing and where they are having trouble, they make necessary instructional adjustments such as re-teaching, trying alternative instructional approaches or offering more opportunities for practice. These activities can lead to students' success. Formative examinations generate feedback for improvement and provide an opportunity to receive guidance for final performance. Feedback given as part of formative helps learners become aware of any gaps that exist between their current knowledge, understanding, skills and guides them through actions necessary to obtain the goal (Ramaprasad,1983); Sadler, 1989). Formative assessment provides the teacher and the student with the information that guides learning from day-to-day (Nitko, 1994).

According to Ramaprasad (1983), feedback requires that the information generated is actually used to close the gap between actual and desired level of performance. If the gap is discovered after administering an examination, but we have no idea of the nature of the discrepancy between the actual and desired performance, then this information does not help us close the gap, and therefore fails to qualify as feedback but remains for monitoring. Therefore for an examination to serve its feedback function as well as alerting us to the existence of a gap, the information must actually be useful in closing the gap between actual and desired level of performance. The information must therefore have enabled within it some degree of prescription about what must be done. The

information must be related to a development model of growth in the domain being addressed in short. It must be construct-reference (Messick, 1975). Feedback based on examinations is recognized as one of the most powerful ingredients of teaching and learning. Maximizing quality and appropriateness of feedback should be the core aim of active examinations thus informs learning.

## **ii) Summative teacher made Examinations**

Summative assessment is generally carried out at the end of a course or project. In an educational setting, summative assessments are typically used to assign students a course grade. Summative assessments are evaluative. Summative examinations are the processes by which teachers gather evidence in a planned systematic way in order to draw inferences about their student's learning based on their professional judgement and report at a particular time on their students' achievements (Harlen, 2004a, 2004b).

MJVCHtlIV *Ulr NAUttt\**  
**KUYU LIBRA\*\***

According to Bloom et al (1971), summative examinations are those examinations given at the end of units, teacher made mid-term and at the end of a term exams, which are designed to judge the extent of students learning of the material in a course for the purpose of grading, certification, evaluation of progress or even researching the effectiveness of the curriculum. An internal examination becomes summative when the information gathered after its administration is employed for recording and reporting (Wynne 2005)-<sub>u</sub>»cw®«««« «'•

Generally a summative examination result is used for internal-school tracking of students' progress, informing parents, students and the student's next teacher of what has been achieved , certification or accreditation of learning by an external body, selection for employment or higher education. It is also used with other information for monitoring the performance of teachers and schools. Internal uses of summative examinations includes using regular grading record keeping, informing decisions about courses to follow where there are options within the school, and reporting to the parents and to the students themselves. It is important that summative examinations be matched to the learning targets of the curriculum since they serve as a preparation for the external examinations which are standardized. In Kenya, internal summative examinations include and not limited to teacher made assessments and class assignments (Chen et al., 2005).

### iii) **Mock Examinations**

KCSE Mock Past Papers play a great role in students' performance in the KCSE examination. According to KNEC, (2006), KCSE mock past papers serves as a good motivation as well as revision material for the major exam the Kenya certificate of secondary examination (KCSE). Choosing the KCSE mock examination revision material saves you a lot of time spent during revision for KCSE. It is also cost effective. Mock Past papers, give an actual exam situation in readiness for the forthcoming national examination from the Kenya National Examination Council, (KNEC).

## **b) External Examinations**

External examinations are examinations that are designed and conducted by recognized examination bodies (Adeyemi, 2008). This is in the sense that the examining boards conducting these examinations did not themselves neither organize instructional courses nor prepare students for examinations. These examinations are designed and organized under specific terms and conditions based on norms that are regarded as standard. They are designed to evaluate performance at the end of a course of study programme.

According to Wayne, (2005), external examinations are summative in nature and serves the purposes of: certification by examining bodies or for vocational qualifications, selection for employment or for further or higher-education, monitoring the student's performance and school accountability, often based on the externally created tests or examinations. The various forms of external examinations are public national examinations, national assessment examinations and international assessments.

## **i) Public National Examinations**

Public examinations are viewed as external school examinations open to the general public and conducted by these examining bodies using tests that have appropriate psychometric properties (Obioma and Salau, 2007). According to Adeyegbe (2004), these tests used by various examination boards are often better developed than those prepared by teachers in schools and lecturers at the universities since they are standardized tests. It can be further be explained that it is an examination that is conducted by (an entity like KNEC) a national body which is not directly involved with instruction of learners. This

body has an important role to play in ensuring that the curriculum and learning outcomes are assessed in a vigorous and fair manner and those examinations lead to reliable and valid exit certificates.

In Kenya, the public national examinations are developed and administered by the Kenya National Examination Council (KNEC). The examination board uses the general education curriculum as a basis for the test development and is designed primarily as a measurement and evaluation system that school teachers use routinely to monitor individual student's progress and instructional effectiveness. Public examinations have been used in Kenya for decades as selection instruments for further educational training. A good example is the Kenya Certificate of Secondary Education (KCSE) which is meant to serve as a yard stick for admission into higher learning institutions. A student who is thereby admitted is assumed to possess the abilities and skills necessary to cope with the academic challenges of the university (Othuon, 1994). Public examinations are high-stakes, (Santrock, 2004). These examinations have important consequences to the parent, teacher, student and the school. High-stakes examinations put a lot of pressure on the teacher to the extent that emphasis is often shifted from teaching for acquisition of knowledge to teaching to pass examination. This come with dire consequences notably narrowing of the implemented curriculum, neglecting what is not examined, emphasising learning strategies that are superficial or short term (such as memorizing, rehearsing and rote learning) devoting a significant amount of time to test preparation activities and heavy reliance on the extrinsic rather than intrinsic motivation of student learning. Of particular significance in the Context of Education for All is the fact that teachers,

because their reputation depend on how well their students perform in examinations, may focus their efforts on students who are most likely to succeed. This in turn may be associated with high rates of grade retention and early dropout.

## **ii) National Assessment**

A national assessment is an exercise designed to describe the level of achievements, not of individual students, but a whole education system or a clearly defined part of it, Kellaghan and Greaney (2003).The centric piece of the assessment is the collection of data in schools usually students respond to assessment instruments and questionnaires in groups. Teachers may also be requested to complete questionnaires in which they provide information considered relevant to an interpretation of their achievements. National assessments were introduced in the realization of the fact that the educational data on inputs to education which had typically been collected in the past were often of little relevance or use to educational planners (Kudjoh and Mingat, 1993). National assessment would address this issue by providing information on the "products" or "outcomes" of schooling (student achievements, inequalities in the system) which it was hoped could be used in conjunction with input data to provide a sounder basis for policy and decision making. Thus national assessment would provide policy and decision-makers with relevant and reliable information about the state of the education system, its achievements and problems, which would be amenable to analysis and interpretation.

The information provided by national assessments differs from that available from public examinations which, while providing data on outcomes, do so only for those who take examinations. Information from national assessments would also differ from that provided by research and education sector studies, which are generally short-term, while national assessments, hold out possibility of integrated into the overall management and administration of the education system, of providing information relevant to education sector analysis on a continuing basis.

Kellaghan and Greanoy (2003) outlined the main elements of a national assessment as the policy needs to be addressed in the assessment are determined by the ministry. Sometimes in consultations with the key educational stakeholders (e.g. teachers representatives, curriculum specialists, business people, parents). It includes the MOE, or a steering committee nominated by it, identifies the population to be assessed; the domain to be assessed is determined; the implementing agency prepares achievement tests and supporting questionnaires and administration manuals; the tests and supporting documents are pilot-tested and subsequently reviewed to determine curricular and technical adequacy; the implementing agency selects the targeted population, arranges for printing of materials and establishes communication with selected schools; Test administrators are trained by the implementing survey instruments agency; Survey instruments (Tests and questionnaires) are administered in schools on a specified date; Survey instruments are collected, returned to the implementing agency, cleaned and prepared for analysis; Analyses are carried out and report prepared.

National assessment address issues such as: how well are students learning? Is there evidence of particular strengths or weaknesses in their knowledge and skills? Do achievements of subgroups in the population differ? To what extent is achievement associated with the characteristics of the learning environment? Do achievement of students change overtime.

The data provided by national assessment has been used for a variety of purposes. These include the need to bring to the notice of politicians policy makers and the public the need for more effective education to provide social and economic development and to justify the relocation of discretionary resources: It should be realized however that there may be little flexibility in how resources are allocated (e.g. when most resources go to teacher's salaries). Further, it is used to support policy decisions and to constrain bad policy decisions (e.g. withdrawal of funding for in-service education) which often focus on the dimension of policy by placing issues in a broader context. The other purpose include need to suggest more efficient resource allocation prompted by the findings of analyses that identifies the relationships between alternative inputs and student's achievements; to improve management efficiency through increased accountability (and in some cases competition); to use the backwash effect of the assessment (as it has been proposed in the use of public examinations) to ensure that teachers teach certain subject matter. This is keenly possible when all schools participate in an assessment and when sanctions are attached to performance (e.g.in Chile) (see Himrnel, 1996; Schiefelbein, 1993).

### **iii) International Assessment**

International assessments are examinations administered in more than one country. These examinations provide some indication of where the achievement of students in a country stands relative to the achievements of students in other countries. They also provide evidence on the extent to which the treatment of common curriculum areas differ across countries. Internal assessments have to be designed to allow administration in more than one country (Beaton et al, 1999; Greaney and Kellaghan, 1996a; Kellaghan and Greaney, 2001b). Instruments are developed to assess students' knowledge and skills but instead of representing the curriculum of only one education system, the instruments have to be considered appropriate for use in all participating systems. The age or grade at which the instruments are to be administered has to be agreed, as have procedures for selecting schools/students. International studies have been based on samples of students.

### **2.4 Examination System in Kenya**

In Kenya there is student's assessment at school level and National level school level teachers conduct internal assessment of their student periodically. At national level there are public exams offered by Kenya National Examination Council (KNEC). Assessment evaluation begins at the primary level. Assessment at primary level is done through written assignment (part of lesson activity), homework and periodic test. The student sits for end of term examination and end of year examination at every class. Most of the assessments are set by the teachers as either simple tests, case studies, essays and so on based on what was taught in class. The aim of these assessments is to measure to what extent has the learner mastered the content. At the end of 8 years student sit for national

examination which measure learner's changes cognitive domain only. The purpose of this exam is certification and selection to secondary school (Herbling, 2006).

There is formative assessment which is carried out by professional staff as they carry out teaching duties. There is continuous evaluation of students learning through written test during the term at the end of the term, and end of the year. The purpose of that test is to measure the extent that student have mastered the content and not necessary to promotion to next class. Student sits for national examination at the end of four years. The purpose of this exam is certification student are awarded Kenya Certificate of Secondary Education (KCSE) another purpose is ranking students. It is also used for selection to college, universities and other institutions of higher learning (MOE Human Resource Department). KCSE is graded on 12 point scale which is based on letter grades ranking from A (best) to E (poorest) grade: A-12points, A--11, B+-10, B~9, B-~8,C+~7, C~6, C-~5,D+~4, D~3, D—2 and E~1 point. This applies both in primary and secondary school. Test items can be objective questions or essay questions. Practical tests are also used in some subjects like Science and Technical subjects.

According to the study done by the ministry of education (1999), examination measures cumulative outcome of students academic progress over 4 years course but does not indicate progress at the end of 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> years. In grading emphasis is on ranking candidates in given subject. Grade awarded indicates how student has performed in comparison with peers rather than how much the syllabus is mastered. In the award, varying factors and conditions are not put in consideration e.g. student personal

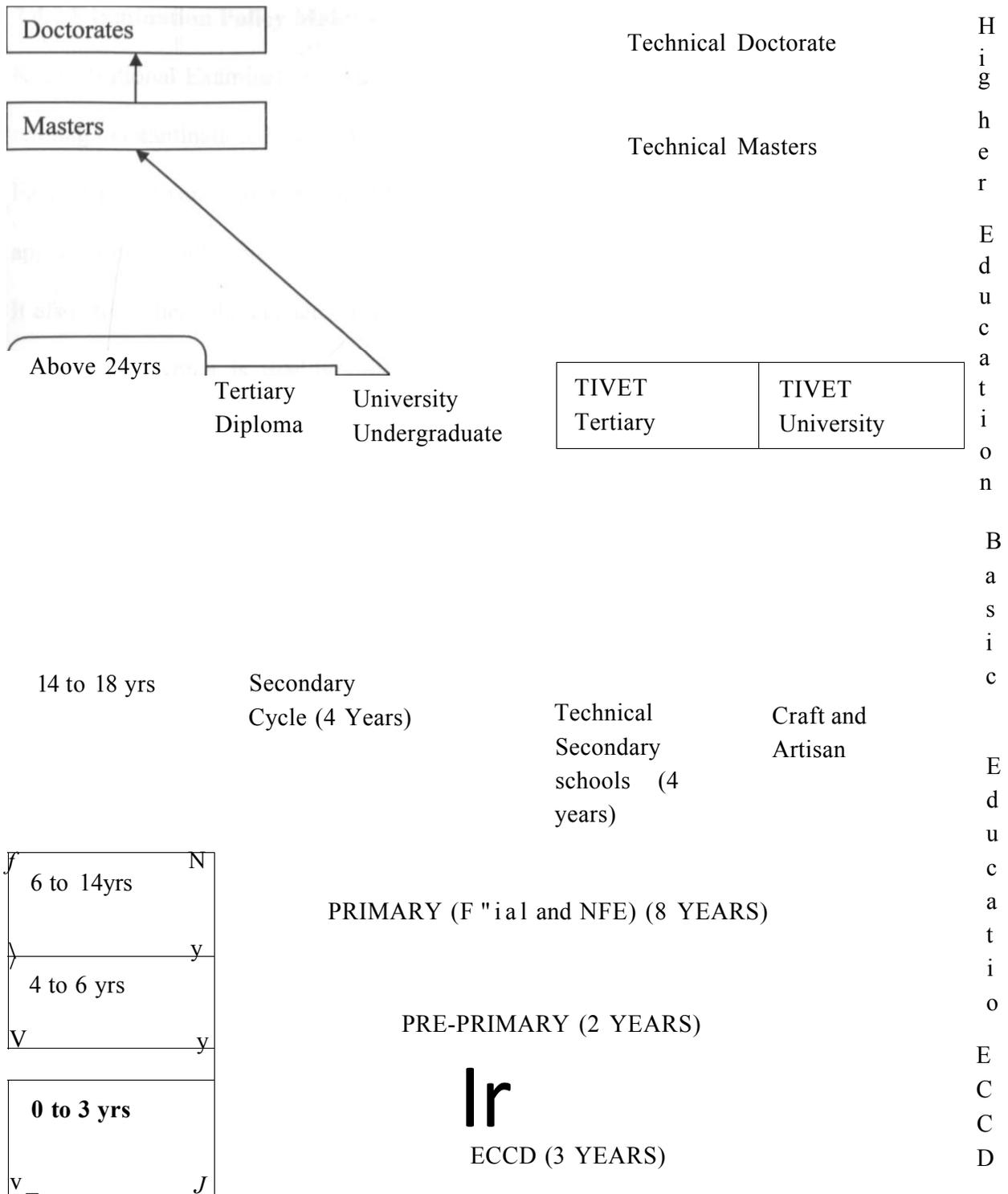
characteristics, home backgrounds, teacher characteristics and school quality which influence performance. Examination predominantly measure cognitive outcome of schooling and ignores growth values and attitudes even though nurture stated as objective of curriculum.

#### **2.4.1 The Kenyan Education System**

Kenya used to have an old education system 7.4.2.3 but in 1983, however, this basic education scheme was dropped with adoption of the 8-4-4 system; once again, secondary and tertiary education became the priority. The new system of education, known as the 8-4-4 system, was introduced in 1985. Under this system, eight years of primary schooling (leading to the Kenya Certificate of Primary Education (KCPE)) are followed by four years of secondary schooling (leading to the Kenya Certificate of Secondary Education (KCSE)) and four years of first degree studies at university. This scheme replaces one which was based on the English pattern culminating in A levels and a three-year first degree course. The introduction of the 8-4-4 education system has led to tremendous changes in the secondary school curriculum. This is in line with the need for a broad-based curriculum that prepares students for self-reliance, vocational training and further education. Primary education in Kenya begins at the age of 6 or 7 after completion of a year of kindergarten commonly known as Nursery School or pre-unit. The first class or year of primary school is known as Standard 1, the final year as Standard 8 and primary school children are known as pupils. Under the current system, students attend secondary school for four years before sitting for the school leaving exam at the end of the fourth year. In form four of secondary schools the Kenya Certificate of Secondary Examination

(K.C.S.E.) is written. Students sit examinations in eight subjects. In 1990, the first KCSE students entered university to begin four years of study for a general degree. The KCSE is administered by the Kenya National Examinations Council, (refer to figure 3.1). It is from here now that students can join tertiary institutions, universities, technical colleges and so on (Eshiwani, 1990).

Kenya has a national curriculum which is developed by the Kenya Institute of Education (KIE). The KIE develops materials for use in education at all levels below university. It is a statutory body legalized by educational act in 1968. The functions of KIE is to: Conduct research and prepare curriculum and curriculum guide for pre-school, primary, secondary, teacher, special, business, technical, adult and post education, Conduct research and provide data for curriculum development, Prepare curriculum support materials, Prepare correspondence courses for both students and teachers, It organizes in-service courses and workshops for teachers involved in experimentation of syllabus and teaching materials orientation, Guides teachers and administrators and educators on syllabus and other teaching materials and to organize and conduct curriculum evaluation.



**Figure 2.1: Structure and Organization of Education and Training in Kenya**

#### **2.4.2 Examination Policy Making and Implementation**

Kenya National Examination Council is in charge of policy making and implementation relating to examination. KNEC was established in 1980 by act of parliament and replaced East African Examination act of 1967. The council plays an advisory role by systematic appraisal of learning process through implementation of appropriate measurement tools. It also strengthens the curriculum implementation process by generating data on student performance which is readily utilized by curriculum developers in their curriculum appraisal (Shiundu and Omulando, 1992). KNEC conducts academic, technical and professional examination within Kenya is like KCPE, KCSE business examinations technical examinations and teacher examinations. It is involved in setting, moderating, administering, marking, scoring and giving feedback. It awards diploma and certificates to only successful candidates. It makes rules regulating and conduct examination, structure and organization. KNEC invites any bodies outside Kenya that wish to conduct academic technical or professional examination within Kenya or conduct these examinations jointly with council and award certificates/diplomas.

KNEC also requests other examining bodies or other institutions to conduct examination and supervise examinations sessions in Kenya National Board. It also advises anybody/bodies invited in Kenya to conduct examination on how to adopt their examinations to fit requirement of Kenya and assist such bodies to conduct such examination. Registration of candidates-only registered candidate can sit for examinations. Training of examiners only teachers that are trained by KNEC are allowed to mark examination. Discourages cheating in exam by cancelling results of candidates

who are found cheating in exams. Such candidates are not allowed to re-sit examination for a period of 2 years. Re-sitting of examination is allowed where a candidate can repeat and register all exams or can register and re-sit a few exams.

**Table 2.1 Examination structure**

<b>Level of Education</b>	<b>Examinations Conducted</b>	<b>Form of Examination</b>	<b>Purpose of examination</b>
	Mid-term exams	<b>Summative</b>	Evaluation of whether the learning objectives are being achieved.
	End of term exams	Summative	Evaluation of whether the learning objectives have been achieved.
	End of year exams	Summative	Promotion to the next class.
	Mock exams	Summative	Preparation of students for external examinations (KCPE).
	KCPE exams	Summative	Certification and selection to secondary school.
<b>Secondary</b>	Continuous assessment	Formative	Evaluates if learning objectives are being achieved.
	Mid-term exams	Formative	Evaluates if learning objectives are being achieved.
	End of term exams	Summative	Evaluates if learning objectives have been achieved.
	End of year exams	Summative	Promotion and evaluation.
	Mock exams	Summative	Prepares students for external examinations (KCSE).
	KCSE exams	Summative	Certification and selection of students into universities.
<b>University</b>	Continuous assessments	Formative	Evaluation of achievement of objectives.
	End of semester exams	Summative	Certification.

## **2.5 Theoretical Perspective of Assessment**

The study was based on three theoretical perspectives namely positivists (scientific) view, constructivism (hermeneutic) view, classical test theory and item response theory.

### **2.5.1 Positivists (Scientific) View**

Assessment is associated with writing of examinations or tests (Mckellar, 2002). According to Birtz (1991), traditionally assessment has been viewed as a means of verifying student learning and it occurs after learning has taken place. Positivists (scientific view of assessment) view an assessment is drawn on scientific concepts, beliefs and applies these to social settings such as educational and sociological contexts. Positivists assume that one can achieve objectivity and consequently uncover truths about the real world. It holds that knowledge is objective and value free (Mckellar, 2002). The goal of the curriculum is therefore to teach students these truths by employing a transmission model of instruction and in turn to assess whether students have learned these truths.

The positivist view regards knowledge as an end product, which can be measured, as well as predicted and therefore controlled (Grundy, 1987). As the purpose education in this perspective, is to prepare the youth to contribute to the well being of the society as a whole, the purpose of examination is to determine the degree to which the end product has been achieved, and to grade, rank and select according to achievement which is the purpose of public external examinations. Traditionally examination has been seen as an outside force that is imposed upon the curriculum generally and learner specifically. This

is due to the assumption that the questions which drive the curriculum must be supplied by outside experts that the majority of what is to be learned is already known, digested and organized and that there are acknowledged correct responses to the curriculum which is to be asked (Short and Burke, 1991, pg 60). This is not the truth as some instructors do not provide the learners with all that is contained in the curriculum due to the high-stakes nature of Kenya's public examinations.

The focus of examinations is thus on the products of learning—a limited range of such as facts, content knowledge and basic skills. In this perspective tests and examinations are summative. The student is regarded as a recipient of pre-existing knowledge transmitted by the expert. Student learning focuses on memorising information in order to give it back under pressure in year-end three-hour examination. Memorising without understanding and learning by rote are what (Matron and Saljo, 1988; Biggs and Entwistle, 1987; Svensson, 1977) term surface learning.

### **2.5.2 Constructivism (Hermeneutic) View**

Constructivism is a theory of knowledge (epistemology) that argues that humans generate knowledge and meaning from an interaction between their experiences and their ideas. Constructivism is a psychological theory of learning and assessment. Their argument is based on cognitive psychology, philosophy and anthropology. In this perspective reality is not regarded as 'out there' but rather seen as individually constructed and shared within a historical, social and political context (Mkellar, 2002). Constructivists define knowledge as temporary, socially and culturally mediated and thus non-objective. Since

students must actively construct their own understanding, deep learning is likely to occur, (Ramsden, 1992).

In the hermeneutic perspective learning is understood as a self regulated process of resolving inner cognitive conflicts that often become apparent through concrete experience collaborative discourse and reflection. In this perspective, institutions should provide support structures to help students in need meet institutional standards in assessment practices. The institutions are seen as gate keepers of assessment standards and practices. The students' performances therefore depend on the institutional standards and what support is provided by the institutions. This study will be guided by this theory of Constructivism.

### **2.5.3 Classical Test Theory**

Classical test theory is a body of related psychometric theory that predict outcomes of psychological testing such as the difficulty of items or the ability of test-takers. The aim of classical test theory is to understand and improve the reliability of psychological tests. Classical test theory assumes that each person has a true score-T, which would be obtained if there were no errors in measurement. A person's true score is defined as the expected number-correct score over an infinite number of independent administrations of the test. Unfortunately, test users never observe a person's true score, only an observed score, X. It is assumed that observed score equals the true score plus some error.

Reliability cannot be estimated directly since that would require one to know the true scores, which according to classical test theory is impossible. Classical test theory is an influential theory of test scores in the social sciences field (Allen and Yen, 2002).

The theory has been put in to practical use in several instances. The conceptual framework used to develop HESI exams is grounded in classical test theory and critical thinking theory. The creation, administration, and interpretation of tests are accomplished through educational and psychological measurement processes. Crocker and Algina, (2003), stated that measurement of psychological attributes occurs when quantitative values are assigned to the sample of behaviours obtained from administering a test. By observing and classifying similar behaviours, the test designer is able to draw inferences about the psychological constructs that contribute to the makeup of the test taker. The test designer may also be able to identify relationships between psychological constructs and practical consequences, therefore predicting test-taking behaviours, such as success in academic programs or nursing practice. To make such predictions, the test designer must first quantify the observations representing the constructs that define these behaviours. The nurses who design and revise the nursing exams use course syllabi from nursing programs across the United States in combination with NCLEX test blueprints provided by the National Council of State Boards of Nursing (NCSBN) to define the constructs indicative of behaviours required for entry-level practice. HESI item writers create test items for use on HESI exams that specifically measure these behaviours.

Embretson and Reise, (2000) review the ramifications ("rules") of CTTs. The first is that the standard error of measurement of a test is consistent across an entire population. That is, the standard error does not differ from person to person but is instead generated by large numbers of individuals taking the test, and it is subsequently generalized to the population of potential test takers. In addition, regardless of the raw test score (high, medium, or low), the standard error for each score is the same. Another ramification is that as tests become longer, they become increasingly reliable. Recall that in domain sampling, the sample of test items that makes up a single test comes from an infinite population of items. Also recall that larger numbers of subjects make the statistics generated by that sample more representative of the population of people than would a smaller sample. These statistics are also more stable than those based on a small sample. The same logic holds in CTT. Larger numbers of items better sample the universe of items and statistics generated by them (such as mean test scores) are more stable if they are based on more items.

#### **2.5.4 Item Response Theory**

In psychometrics, item response theory (IRT) also known as latent trait theory, strong true score theory, or modern mental test theory, is a paradigm for the design, analysis, and scoring of tests, questionnaires, and similar instruments measuring abilities, attitudes, or other variables. It is based on the application of related mathematical models to testing data. Because it is generally regarded as superior to classical test theory, it is the preferred method for the development of high-stakes tests such as the Graduate Record Examination (GRE) and Graduate Management Admission Test (GMAT) (Lord, 1980).

The name item response theory is due to the focus of the theory on the item, as opposed to the test-level focus of classical test theory, by modeling the response of an examinee of given ability to each item in the test, (Thissen and Orlando, 2001). The term item is used because many test questions are not actually questions; they might be multiple choice questions that have incorrect and correct responses, but are also commonly statements on questionnaires that allow respondents to indicate level of agreement (a rating or Likert scale), or patient symptoms scored as present/absent. IRT is based on the idea that the probability of a correct/keyed response to an item is a mathematical function of person and item parameters. According to Thissen and Orlando, (2001) the person parameter is called latent trait or ability; it may, for example, represent a person's intelligence or the strength of an attitude. Item parameters include difficulty (location), discrimination (slope or correlation), and pseudo guessing (lower asymptote).

2.6 Conceptual Framework

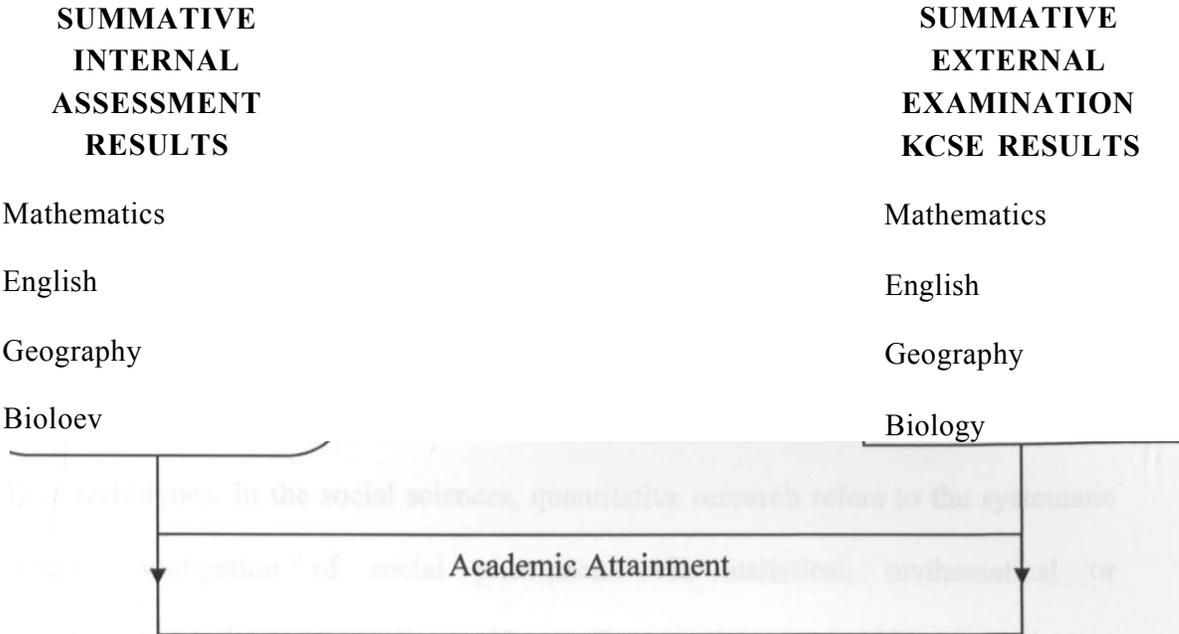


Figure 2.2: Conceptual Framework for the Research

## **CHAPTER THREE**

### **RESEARCH DESIGN AND METHODOLOGY**

#### **3.1 Research Design**

This study adopted a descriptive research design. Cooper and Schindler, (2006), described a research design as the strategy for the study and plan by which the strategy is to be carried out specifying the methods and procedures for the data collection measurements and analysis of data. The study employed quantitative and qualitative data analysis techniques. In the social sciences, quantitative research refers to the systematic empirical investigation of social phenomena via statistical, mathematical or computational techniques. According to Hunter, Laura and Leahey, (2008), the objective of quantitative research is to develop and employ mathematical models, theories and/or hypotheses pertaining to phenomena. The process of measurement of this study is central to quantitative research because it provides the fundamental connection between empirical observation and mathematical expression of quantitative relationships. This study was designed along the lines of a correlation research. Gay, (1996) described correlation research as that involving the collecting of data in order to determine whether and to what degree a relationship exists between two or more quantifiable variables.

#### **3.2 Target Population**

The target population of this study comprised of all the students from four secondary schools who sat for their KCSE examination in 2010. The data comprised of the internal summative marks results for years 2007-2010 and the summative marks results of the

2010 KCSE respectively. These were obtained from the report records in the school. A target population is a full set of cases from which a sample is taken (Saunders, Lewis and Thornhill, 1997). Cooper and Schindler, (2006), calls it a population of interest from which the individual participant or object from which the measurement is taken. Target population is a complete set of cases or objects with the same common observable characteristics from which the sample for the study is drawn, (Mugenda and Mugenda, 1999).

### **3.4 Sampling Procedure**

Summative internal test results for 576 students from form 1 (2007) to form 4 (2010) (KCSE) were purposively sampled from each of the 6 schools targeted by the study. Students' internal and external summative examination results for the 4 subjects were purposively selected from the schools based on high, middle, and low achievers in order to avoid data bias. The Summative internal test results for the targeted students were obtained from the report books dated from form one year 2007 to form four KSCE years 2010. A sample is a subject or part of the target population, (Mugenda and Mugenda, 1999). Sampling is the process of selecting the subjects or cases to be included in the study as representative of the target population (Mugenda and Mugenda, 1999).

### **3.5 Data Collection Instruments**

The instrument that was used to collect data for this study was an inventory, (Adeyemi, 2008). The inventory require data on test scores of internal and external summative examinations for students of the study year 2007- 2010 (KCSE) from each of the nine

schools. The instrument that was used to collect was an inventory as used by Adeyemi (2008). The inventory requested test scores derived from tests constructed, administered, and scored by teachers in secondary schools for Maths, English, Biology and Geography. The choice of the subjects was based on two core, one science and; one humanity subject as referred by the 8-4-4 education system in Kenya, (KNEC, 2006). The data collected was filled in the inventory sample (See Appendix 1).

### **3.6 Data Analysis and Presentation**

Data was analysed using quantitative statistics and presented using tables. Data analysis is the process of systematically applying statistical and/or logical techniques to describe and illustrate, condense and recap, and evaluate data. According to Shamo and Resnik, (2003), various analytical procedures provide away of drawing inductive inferences from data. The data was analysed using SPSS (Statistical Package for Social Sciences). Only test scores for the sampled schools were recorded. The aggregate scores, for the selected subjects for the study, were calculated in internal and KCSE examination. The two sets of examination scores were compared by calculating correlation coefficient, ( $r$ ). The purpose of the analysis was to determine the level of the relationship between performance in internal summative examination and external summative examination, (KCSE). Qualitative data analysis (content analysis) was used to interpret the results in relation to the objectives of the study. The researcher used regression analysis to determine the relationship between the summative internal examination in 1<sup>st</sup> year (2007) - 4<sup>th</sup> year (2010) and then finally to determine the year with more weight in predicting performance in KCSE, 2010.

The regression equation to be answered is:  $Y = p_0 + p_1 + p_2 + p_3 + p_4 + e$

Where 1: Internal Exams 2007 contributes

2: Internal Exams 2008 contributes

3: Internal Exams 2009 contributes

4: Internal Exams 2010 contributes

0: Constant

**Which is:**

$KCSE\ 2010 = p_0 + p_1 \text{ Internal Exams } 2007 + p_2 \text{ Internal Exams } 2008 + p_3 \text{ Internal Exams}$

$2009 + p_4 \text{ Internal Exams } 2010$

## CHAPTER FOUR

### DATA ANALYSIS AND PRESENTATION

#### 4.0 Introduction

This chapter gives the analysis of the collected data and their interpretation. The results of the findings are presented in form of tables for easy interpretation and understandings. The results are given according to the study objectives.

#### 4.1 Biographic Information

This section gives personal information of the respondents based on their gender distribution. The results indicated whether gender issue was given consideration in the study.

**Table 4.1 Gender**

Gender	F	%
Boy	576	50
Girl	576	50

Table 4.1 indicates the gender of the students under study. It illustrates that the gender response was not biased with male students (50%) being equal in number with the female students.

### **Table 4.2 School Category**

Table 4.2 indicate the student response according to school category. It illustrates how the researcher selected students based on their school category.

<b>School Category</b>	<b>F</b>	<b>%</b>
High performance schools	384	33.3
Medium performance schools	384	33.3
Low performance schools	384	33.3

The response in table 4.2 indicates that the researcher was not biased when selecting schools based on their ranking in terms of geographical coverage area. The results further indicate that respondents were taken in equal measure (33.3% each) from national level, medium school level to low school level.

**Table 4.3 Subjects Selected for the Study**

Table 4.3 indicate the study subjects under review. It gives the proportion in which the researcher selected subjects under review.

<b>Subjects</b>	<b>F</b>	<b>%</b>
Maths	<b>288</b>	<b>25.0</b>
English	<b>288</b>	<b>25.0</b>
Biology	<b>288</b>	<b>25.0</b>
Geoaphy	<b>288</b>	<b>25.0</b>
Totals	<b>1152</b>	<b>100.0</b>

The study findings illustrated in table 4.3 indicate that the researcher selected five subjects where equal response was received from all the selected subjects.

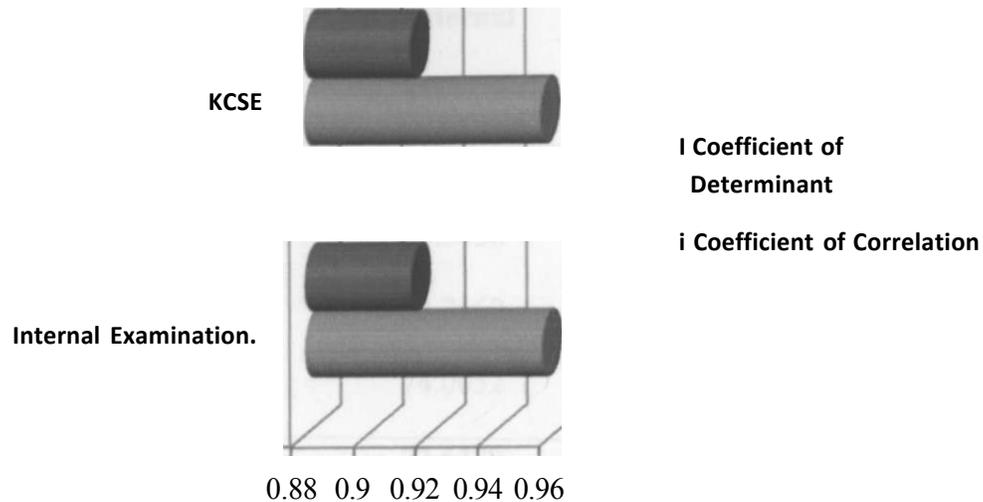
**Objective 1: To establish whether there is significant relationship between internal summative examination and external summative examination in secondary schools in Kenya.**

**Table 4.4 Relationship between Internal summative examination and external summative examinations**

Table 4.4 responds to the first objective of determining if there is significant relationship between internal examination and external examination in selected secondary schools. It gives guidance on whether internal summative examination scores can be used to predict performance in external summative examination.

<b>Form of examination.</b>	<b>Entries (exams)</b>	<b>Correlation coefficients</b>	<b>(r)</b>
Internal Examination.	576	0.956	0.9139
KCSE	576	0.956	0.9139

**Figure 4.1 Relationship between Internal summative examination and external summative examinations**



The correlation was carried out at 10% level of significance. From the data above, it shows that the correlation between internal examination and KCSE is 0.956 indicating that there is a strong positive (95.6% > 75%) relationship between internal and external examination. The coefficient of determinant ( $r^2$ ) indicated that 91% changes in external summative examination can be attributed to changes in internal summative examination. This therefore indicates the level of variability accounted for by the variables under study so in this case, internal examination account for 91.39% of the performance in the KCSE thus the selected variables can be relied upon as they account for greater variability in KCSE as indicated.

**Table 4.5: Mean Performance in the internal summative and external summative examination by subject**

<b>Subject</b>	<b>Mean in Internal examination</b>	<b>Mean in External examination</b>	<b>Difference</b>
Mathematics	72.3142	71.2986	1.01
English	70.5156	68.7847	1.78
Biology	73.3160	71.7847	1.54
Geography	74.0052	72.4444	1.56
<b>Overall mean</b>	<b>72.5378</b>	<b>9.95057</b>	<b>L47</b>

The results in table 4.5 indicate the mean performance of students in internal summative examination and external summative examination by subjects. It indicates the difference in the performance in internal summative examination performance and the performance in external summative examination.

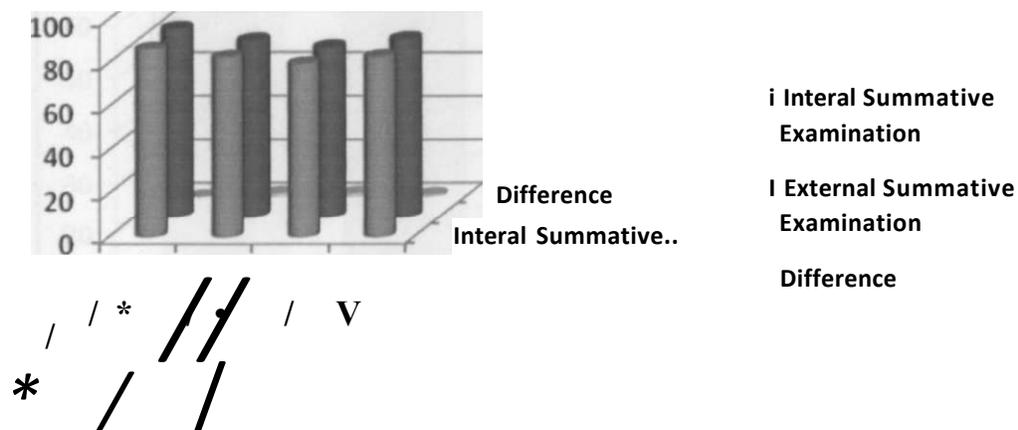
The results in table 4.5 indicate that the performance in external summative examination (KCSE) is almost equal to the performance in internal summative examination where the researcher established that the recorded performance in internal summative examination is almost equal to the recorded mean score in external summative examination. The results indicate that the mean performance in mathematics in the internal summative examination was 72.31 and the recorded performance in the external summative

examination was 71.30, a value closer to each other with an absolute difference of 1.01. In English, the recorded internal summative mean results was 70.52 and the recorded mean results for external summative examination was 68.74, a value closer to each other with an absolute difference of 1.78 meaning that the performance in the internal summative examination can be used to predict performance in the external summative results. In Biology, the mean performance in the internal summative examination was 73.32 and the performance in the external summative examination was 71.78 giving an absolute difference of 1.54.

Finally, the results indicate that the mean performance in the internal summative examination for geography was 74.00 and the performance in external summative examination in geography was 72.44 with an absolute difference of 1.47. The main findings of this study is that the mean performance in the internal summative examination can be used to predict the mean performance in KCSE just by adding + or - 1.47 to the mean in his/her prediction of the results. In these schools, the mean performance in KCSE is slightly lower than the mean performance in the internal summative examination so in predicting the mean performance of these selected schools using the given internal summative examination there has to be a subtraction of the average absolute difference value in the performances. These findings can also be adopted by the Kenya national Examination council in predicting the performance of Students in KCSE in case of a school with questionable performance or schools whose results are missing.

**Figure 4.2: Mean performance of high achieving students in internal summative examination and external summative examination by School category**

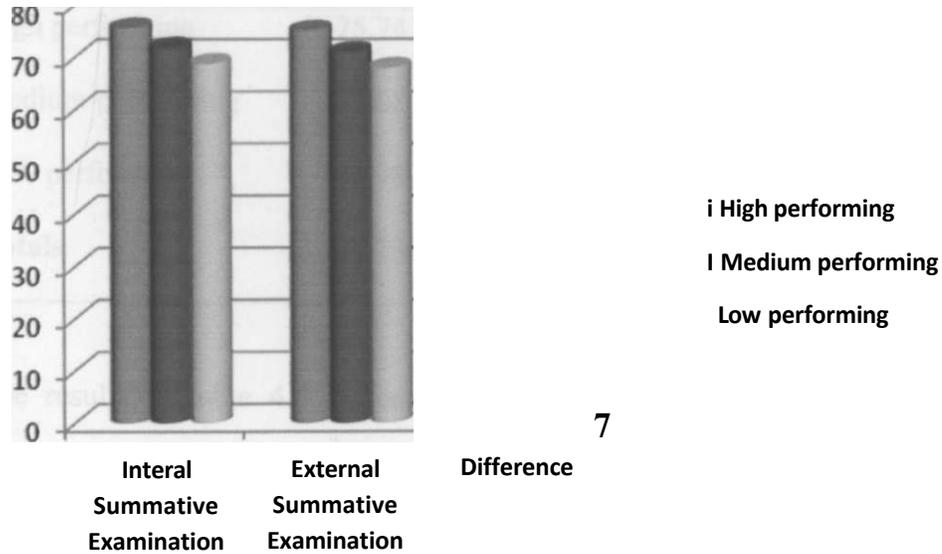
The results in figure 4.2 indicate the mean performance in internal examination of high achieving students against the mean performance in external examination (KCSE) by school category.



The results in Graph 4.6 indicate that the mean performance of internal summative examination of high performing schools in each category (83.26) (high performing, medium performing schools and low performing schools) is not much difference from the mean performance in the KCSE examination (82.04) as illustrated by a slight difference (1.10).

**Figure 4:3 Mean of medium performing students in internal summative examination and external summative examination by School category**

Figure 4.3 indicates the mean performance of medium performing students in internal summative examination against their mean performance in external summative examination (KCSE) by school category.



**Table 4.6: Mean of medium performing students in KCSE and in internal summative examination by School category**

<b>School Category</b>	<b>Mean in internal examination</b>	<b>Mean in KCSE.</b>	<b>Absolute mean difference</b>
High performing	75.74	75.30	.44
Medium performing	72.08	71.22	.86
Low performing	68.98	68.20	.78
<b>Totals</b>	<b>72.23</b>	<b>71.57</b>	<b>.69</b>

The results in table 4.6 indicate that the mean performance of internal summative examination of medium performing schools in each category (72.23) (high performing' medium performing schools and low performing schools) is not much different from the mean performance in the KCSE examinations (71.57) as illustrated by a slight difference (0.69).

**Figure 4.4: Mean of low performing students in internal summative examination and external summative examination(KCSE) by School category**

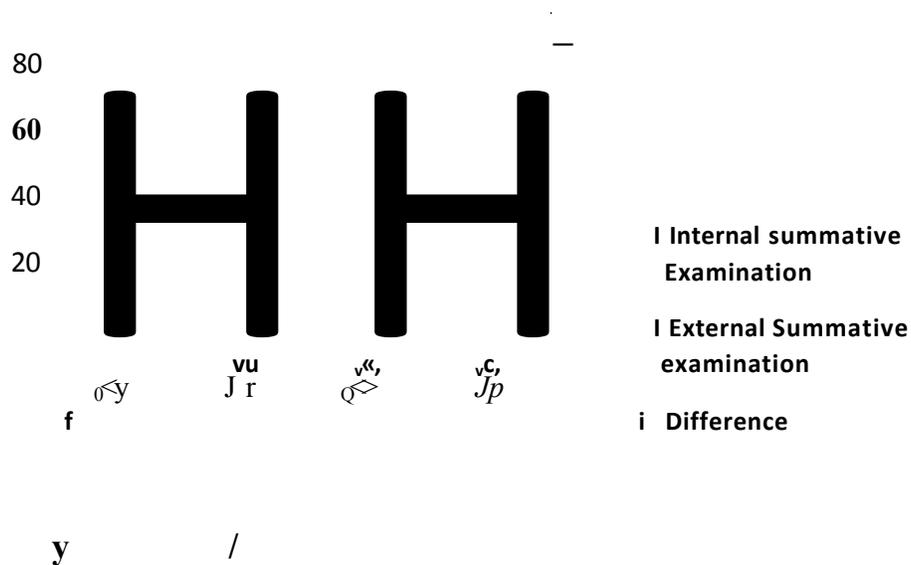


Figure 4.4 indicates the mean performance in internal summative examination of low performing schools against the mean performance in external examination (KCSE).

The results in table 4.4 indicate that the mean performance in internal summative examination in low performing schools in each category (62.23) ( high performing, medium performing schools and low performing schools) is not much different from the mean performance in the KCSE examination (59.64) as illustrated by difference (2.59). however, the mean difference in low performing schools between KCSE and internal examination is far much bigger than the ones in high and medium performance schools hence internal examination is not much a appropriate for predicting performance in external examination (KCSE) in low performing schools.

**Table 4.7: Performance of students in internal summative examination and external summative examination (KCSE) by gender**

<b>Gender</b>	<b>Internal exams</b>	<b>KCSE</b>
	<b>Mean</b>	<b>Mean</b>
Boys	72.589	71.385
Girls	72.486	70.771

The data in table 4.7 shows that boys registered a mean grade of 72.589 and girls 72.486 in the internal summative examinations. In the KCSE boys registered a mean of 71.385 and girls in the KCSE, where boys perform better than girls in both internal and external summative examination. It also shows that the student performance in both internal summative examination and the KCSE examination is consistent.

**Objective 2: To find out which subjects has a greater capacity to predict performance in secondary schools in Kenya**

**Table 4.8 Level of subjects in predicting performance in secondary schools**

The findings in table 4.8 indicate the level of subjects in predicting performance in KCSE examination. It can be used by schools to identify which subjects to put more efforts in influencing the performance in the KCSE examination.

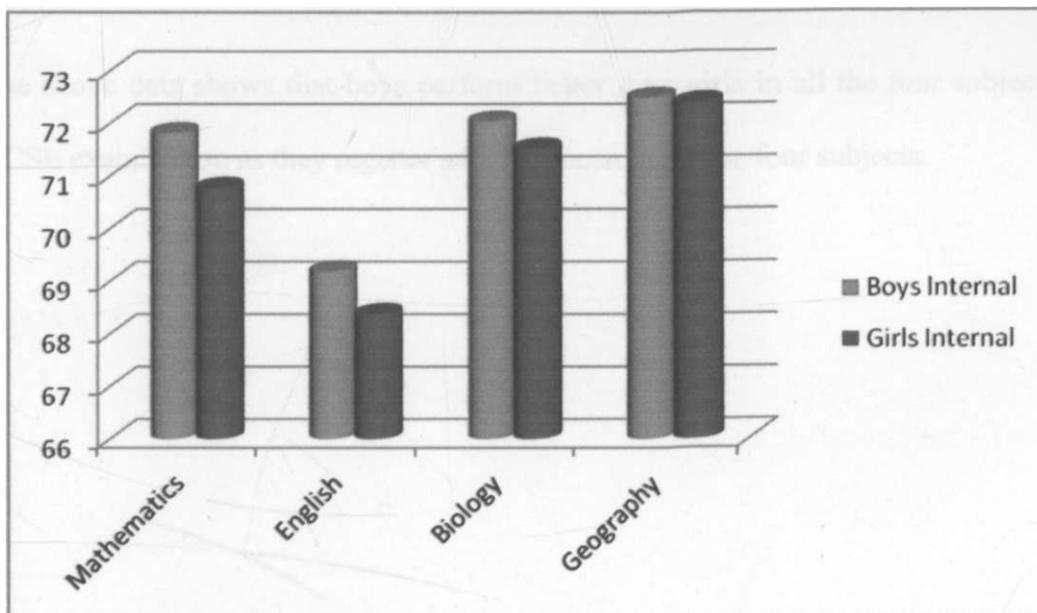
**Coefficients<sup>8</sup>**

Model	Un-standardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	-9.466	8.187		-1.156	.262
Internal maths	.671	.214	.667	3.141	.359
Internal biology	.408	.241	.294	1.697	.106
Internal geography	-.154	.164	-.132	-.939	.005
Internal English	.189	.199	.146	.951	.353

a. Dependent Variable: KCSE2010

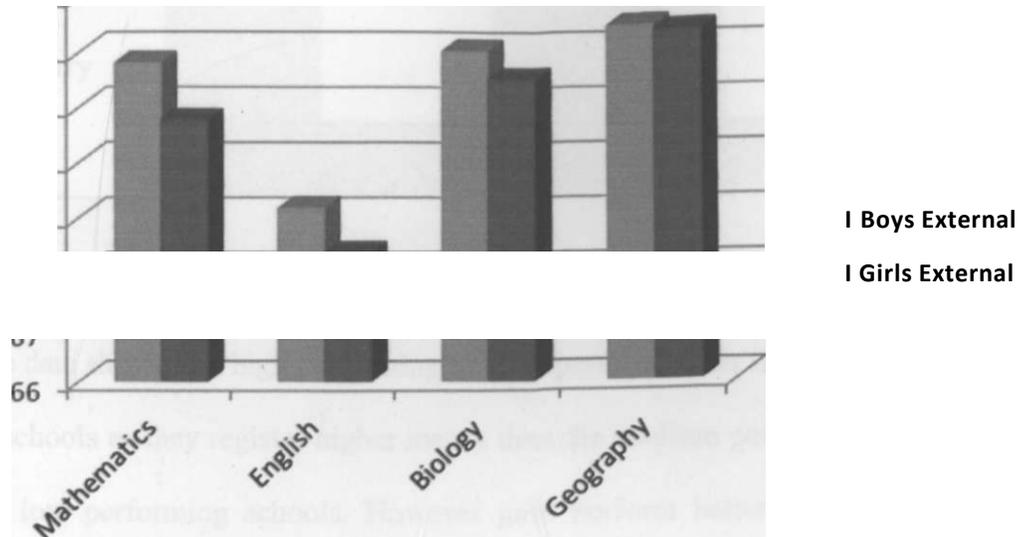
Table 4.8 indicate that internal results in mathematics count for bigger percentage (67.1%) performance level in KCSE then followed by biology (40.8) then English (18.9) and then finally geography (15.4%). The level of significance is greater than 0.05 (Sig .262) indicating that the performance in internal summative examination greatly influences the performance in external examination. The results further indicate that mathematics (Sig. 0.359) is more significant followed by English (Sig. 0.358) and finally biology (Sig. 0.106) while geography (Sig. 0.005) is not significant since its p-value is less than 0.05.

**Figure 4.5: Performance of boys and girls in internal summative examinations by subject**



The data shows that boys perform better than girls in Mathematics, English and Biology in the internal examination. However girls perform better than boys in Geography.

**Figure 4.6: Performance of boys and girls in external summative examination (KCSE) by subject**



The above data shows that boys perform better than girls in all the four subjects at the KCSE examination as they register a higher mean in all the four subjects.

**Table 4.9: Performance of boys and girls in internal summative examination by school category**

School Category	Boys		Girls		Overall mean	Overall %
	Mean	%	Mean	%		
<hr/>						

The data shows that high performing schools perform better than the rest of the categories of schools as they register higher means than the medium performing schools and the low performing schools. However girls perform better than boys in the medium school category. Overall performance shows that boys perform better than girls.

**Table 4.10: Performance of boys and girls by school category in the KCSE**

School Category	Boys		Girls		Overall mean	Overall %
	Mean	%	Mean	%		
High performing schools	76.28	17.89	74.18	17.38	75.23	35.27
Medium performing schools	69.94	16.40	71.04	16.66	70.49	33.06
Low performing schools	67.94	15.93	67.14	15.74	67.54	31.67
	71.38667	50.22	70.78667	49.78		

High achieving schools perform better than other levels of school category. Similar to table 4.6, the girls in the medium school perform better than their boy's counterparts. Boys perform better than the girls in the overall performance.

**Table 4.11: Mean of annual performance in internal and external summative examinations for boys and girls**

Table 4.11 indicates the annual mean of performance in internal and external summative examination for both boys and girls from first year to the final year where the mean for external summative examination performance is only captured in the fourth year.

year	Internal Summative		External Summative	
	Boys	Girls	Boys	Girls
2007	72.6	72.82		
2008	72.11	71.98		
2009	73.07	72.94		
2010	<b>72.58</b>	<b>72.21</b>	<b>71.39</b>	<b>70.77</b>

The results in table 4.15 further indicate that in the internal summative examination, girls (72.82) perform better than boys (72.6) while from second year, boys perform better than girls even up to the final exams that is the external summative examination where boys (71.39) had better performance than girls (70.77).

**Table 4.12: Performance of School Category by Gender in both internal and external examinations**

Table 4.12 indicates the performance of school category by gender in both internal and external summative examination. It illustrates the difference in performance of both boys and girls from different school category in internal and external examinations.

year	Internal		External	
	Boys	Girls	Boys	Girls
High performing level	76.06	75.79	76.28	74.18
Medium performing level	71.89	72.88	69.94	71.04
Low performing level	69.82	68.79	67.94	67.14

The results in table 4.12 indicate that boys in high performing schools perform better than girls both in the internal summative examination and external summative examination while in the medium performing schools, girls perform better than boys both in the internal summative examination and in external summative examination. The results further indicate that boys in the low performing schools perform better than girls in both internal summative examination and external summative examination though the difference in the performance of all the three category of schools and between boys and girls are very small.

**Table 4.13: Performance of both boys and girls by subjects in internal and external summative examination by school category**

Table 4.13 indicates the performance of both boys and girls in external and internal examination by school category on various subjects. It illustrates the extent to which boys and girls from different school category perform in different subjects.

School Category	Subjects	Internal Summative Exams		External Summative Exams	
		Boys	Girls	Boys	Girls
<b>High performing level</b>	<b>Mathematics</b>	85.38	85.42	85.38	82.92
	<b>English</b>	80.14	79.72	80.92	78.92
	<b>Biology</b>	82.36	81.97	82.29	81.00
	<b>Geography</b>	84.78	84.78	83.29	83.17
<b>Medium performing level</b>	<b>Mathematics</b>	70.95	70.59	71.21	70.50
	<b>English</b>	71.51	73.26	69.79	68.92
	<b>Biology</b>	73.56	73.26	73.29	73.29
	<b>Geography</b>	73.42	73.54	72.83	72.75
<b>Low performing level</b>	<b>Mathematics</b>	60.98	60.57	70.13	67.29
	<b>English</b>	60.25	60.19	70.58	71.75
	<b>Biology</b>	60.25	64.43	71.13	68.29
	<b>Geography</b>	63.32	63.32	72.00	72.75

g

The results in table 4.13 indicate that the performance of high performing school in mathematics, girls (85.42) perform better than boys (85.38) in the internal summative examination while at the external summative examination, boys (85.38) perform better than girls (82.92). The performance of low performing schools in mathematics, boys (70.95) perform better than girls (70.59) in internal summative examination and boys (71.21) still perform better than girls (70.50) in the external summative examination. In high performing schools, in English, boys (80.14) perform better than girls (79.72) in the internal summative examination and also in the external summative examination, boys (80.92) perform better than girls (78.92). In medium performing in English, girls (73.26) perform better than boys (71.51) in the internal examination while boys (68.92) perform better than girls (69.79) in the external examination while in low performing schools, boys (60.25) perform better than girls (60.19) in the internal examination while girls (71.75) perform better than boys (70.58) in the external examinations. The result in biology performance indicate that in the high performing schools boys (82.36) perform better than girls (79.72) at internal examination and also boys (82.29) perform better than girls (81.00) at external examination. The result also indicate that in biology performance in medium performing schools, boys (73.56) perform better than girls (73.26) in the external examination while at the external examination, boys (73.29) had equal performance with the girls (73.29) . In low performing schools, girls (64.43) had better performance than boys (60.25) in the internal examination while boys (71.13) had better performance than girls (68.29) in the external examination. Finally, the performance in geography in high performing schools had boys (84.78) having equal performance with girls (84.78) in the internal examination while in the external examination, boys (83.29)

had better performance than girls (83.17). In the medium performing schools, girls (73.54) perform better than boys (73.42) in the internal examination while boys (72.83) perform better than girls (72.75) in the external examinations. Finally, in the low performing schools, girls (63.32) and boys (63.32) have similar performance in the internal examination while girls (72.75) perform better than boys (72.00) in the external summative examination.

**Objective 3: To determine the year with more weight in predicting performance in KCSE**

The results under table 4.14 indicate the year whose scores of internal summative examinations can be used to predict performance in external summative examination (KCSE).

**Table 4.14: Determination of the year with more weight in predicting performance in KCSE**

**Coefficients<sup>8</sup>**

Model	Un-standardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	-.892	.893		-.999	.318
Internal Exams 2007	-.085	.039	-.076	-2.166	.031
Internal Exams 2008	.100	.031	.090	3.190	.001
Internal Exams 2009	.270	.046	.253	5.896	.000
Internal Exams 2010	.708	.038	.711	18.818	.000

a. Dependent Variable: KCSE2010

Table 4.14 indicate that the study results is relevant since the p values (Sig.) of the constant is 0.318 a value greater than 0.05 since we are testing at 5% significance level and the regression results under un-standardized coefficient B indicate the weight of every year in predicting KCSE results right from form 1 (2007) till for 4 (2010). The

coefficients for internal summative examination results under un-standardized coefficient

B indicate that:

$$Y = p_0 + p_1 + p_2 + p_3 + p_4 + e$$

Where 1: Internal Exams 2007 contributes

2: Internal Exams 2008 contributes

3: Internal Exams 2009 contributes

4: Internal Exams 2010 contributes

0: Constant

Internal Exams 2007 contributes -ve 8.5% to the final results

Internal Exams 2008 contributes +ve10% to the final results

Internal Exams 2009 contributes +ve27% to the final results

Internal Exams 2010 contributes +ve70.8% to the final results

**Hence KCSE 2010 = -.085 Internal Exams 2007 + .100 Internal Exams 2008 + .270**

**Internal Exams 2009 + .708 Internal Exams 2010**

## **CHAPTER FIVE**

### **SUMMARY, CONCLUSION AND RECOMMENDATION**

#### **5.1 Introduction**

This chapter gives the summary of the study findings based on the study results, it illustrates the conclusion of the study based on the researchers experience and finally it covers the recommendations of the study that should be adopted by the relevant authorities and any interested individual on the study results.

#### **5.2 Summary**

The study shares similar findings by Faleye, (1998) that there is significant relationship between the overall performance of students in external summative exams and their performance in the internal summative exams. The study also share similar findings to the one done by Carol and Boston (2002) on the effectiveness of the performance of preliminary examination to the performance in final examination at the end of the study period where the researcher established that the performance in external summative examinations (KCSE) is almost equal to the performance in internal summative examination where the recorded performance in internal summative examination is almost equal to the recorded performance in external summative examination. Based on the results of the performance per subject, the researcher established that the mean performance in mathematics in the internal summative examination and the recorded performance in the external summative examination was closer to each other with a slight absolute difference. In English, the recorded internal summative mean results and the

recorded mean results for external summative examination was also closer to each other with a slight absolute difference meaning that the performance in the internal summative examination can be used to predict performance in the external summative results. In Biology, the mean performance in the internal summative examination and the performance in the external summative examination had a slight absolute difference hence the possibility of predicting the performance in external summative examination of each of the chosen four subjects using the performance in the internal summative examinations.

Based on the study findings conducted by Nitko, (1994) on the gender performance of students, the study also had a similar findings that boys at high performing schools perform better than girls both at the internal summative examination and at the external summative examination while at the medium performing schools, girls perform better than boys both at the internal summative examination and at the internal summative examinations.

. The results further indicate that boys at the low performance schools perform better than girls at both internal summative examination and at the external summative examination though the difference in the performance of all the three category of schools and between boys and girls are very small.

Finally, the study has indicated that the results of internal summative examination results of students in first year cannot be to predict performance in the external summative examination then followed by second year performance that is relevant by 10% and then followed by the performance in third year that is relevant by 27% then finally the performance in fourth year that is relevant by 70.8% in predicting the performance in the external summative results. The study also established that the performance of subjects like geography is irrelevant in predicting performance in external examination while Mathematics performance in the internal examination is very relevant since it accounts for the 67.1% of the overall performance then followed by the performance in Biology that accounts for 40.8% while the performance in English is the least by 18.9% in predicting the student performance in the external summative exams.

### **5.3 Conclusions**

The main findings of this study is that the mean performance in the internal summative examination can be used to predict the mean performance in KCSE just by adding + or - 2 to the mean in his/her prediction of the results since there was no prediction with more than 2 as the absolute value. In these schools, the mean performance in KCSE is slightly lower than the mean performance in the internal summative examination so in predicting the mean performance of these selected schools using the given internal summative examination there has to be a subtraction of the average absolute difference value in the performances. The study also seem to agree based on all the findings in the study that irrespective of the school category or the student category, the student performance in internal summative examination can successfully predict the performance in the external

summative examination of all the subjects under study. This is useful for the external summative examination administrators that can use the internal summative examination results to award scores for missing results or detect cheating among students in the school. The study results also indicate that high performing schools perform better than the rest of the categories of schools as they register higher means than the medium performing schools and the low performing schools. However girls perform better than boys in the medium school category and also that the overall performance of boys is better than that of girls. These study findings can also be adopted by the Kenya national Examination council in predicting the performance of Students in KCSE in case of a school with questionable performance or schools whose results are missing.

#### **5.4 Recommendation for practice**

The study recommends the following based on the study results:

- i) Students and teachers should put more weight in mathematics during their study since it contributes to a greater percentage in predicting the performance in external summative examination followed by biology then English based on the study results of the four subjects under review.
- ii) Students should concentrate much in the final year (form 4) since the performance in the internal summative examination at this stage contributes more than any other year in predicting the performance in the external summative examination then followed by third year of study then finally the second year of the study

Students who perform poorly in the first year of study should not be discouraged and also students who have better performance in first year should not reduce their class concentration and their study efforts since first year performance in the internal summative examination is irrelevant in predicting the performance in the external summative exams

The Kenya National Examination Council should standardize internal summative examination given to students at various school categories so as to make it easy to administer.

Examiners of external summative examination should adopt the student performance in the internal summative examination in awarding missing results or the students who were not able to sit for the exams due to the unavoidable circumstances like heavy rains or civil wars or even acute internal security issues.

### **Recommendation for further Research**

- i) Further research should be conducted by sampling mixed day schools including rural schools.
- ii) A research can also done to find out factors affecting predictive validity of internal summative examinations in secondary schools.
- iii) Another research can be done to determine the predictive validity of internal summative examinations in primary schools in predicting performance in KCPE.

## REFERENCES

- Adeyemi T.O. (2008). Predicting students' performance in Secondary Certificate Examinations from performance in Junior Secondary Certificate Examination from performance in Junior Secondary Certificate Examinations in Ondo State Nigeria. *Humanity and Social Sciences Journal*, 3(1); 26-36.
- Adeyemi T.O. (2008). Predicting students' performance in Secondary Certificate Examinations from performance in Junior Secondary Certificate Examination from performance in Junior Secondary Certificate Examinations in Ondo State, Nigeria. *Humanity and Social Sciences Journal* 3 (1); 26-36.
- Adeyemi, T.O. (1998). School and Teachers variables associated with the performance of students in senior secondary certificate examinations in Ondo-State, Nigeria. *Unpublished PhD, Thesis*, University of Hull, UK, PP: 203-208.
- Adeyemo, G.A., 2001. *Teacher and student variables as correlates of achievement in Integrated Science in Ibadan North Local Government Area of Oyo State*. Unpublished M.Ed Thesis, Faculty of Education, University of Ado-Ekiti, Nigeria, pp: 40-45.
- Afolabi & Faleye, (1998). *Predictive validity of Osun State Junior Secondary Certificate Examination*,
- Ale, (2008), *Examination Malpractices in Nigeria Educational System: Causes, Effects and the Way Out*. *Nigerian Journal of Clinical and Counselling Psychology* 14(1). Aiken, L.R. Psychological testing and assessment.
- Allen, M.J., & Yen, W. M. (2002), *Introduction to Measurement Theory*. Long Grove, IL: Waveland Press.

- Alonge, M.F., 1998. Implications of probabilistic models of item analysis to educational evaluation. *Journal of Educational Research*, 2: 50-56.
- Al-Shorayye, S.R., (1995), *The effect of admissions policy, socio-economic factors and demographic and personal considerations on students' performance at Kuwait University. Unpublished PhD Thesis*, University of Hull UK, pp: 142-145.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). Standards for educational and psychological testing.* Washington DC: Authors.
- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84 (3): 261-271.
- Angelo, T.A., and Cross, K.P. (1993). *Classroom Assessment Techniques: A Handbook* Assessing student Achievement; A profile of classroom practices. (CSE monograph 11). Los Angeles. University of California, Centre for the study of evaluation
- Assessment Reform Group. (1999) *Assessment for Learning: beyond the black box, between formative and summative assessment' Assessment in Education*, 4, 3, pp. 365-
- Athanasou, J.A. and Lamprianou, I. (2002) *A teacher's guide to assessment*. NSW: Social Science Press.
- Baxter, J. (2007), A case study of online collaborative work in a large first year psychology class, *Paper presented at International Online Conference, Assessment for Learner Responsibility*, 29-31
- Bhatti, M. A. (1987) *Secondary Education in Pakistan: Perspective Planning*, Islamabad:

- Black P. & William D., (2003) In praise of Educational Research; formative assessment.  
*British Educational Research Journal* Vol. 29, No. 5
- Black, H. D. (1986) Assessment for learning, in: D. L. Nuttall (Ed.) *Assessing educational achievement* (London, Falmer Press).
- Black, P. (1993) Formative and summative assessment by teachers, *Studies in Science Education*, 21, 49-97
- Black, P. (1998) *Testing: friend or foe? Theory and practice of Assessment and Testing, for College Teachers*, 2nd ed. San Francisco: Jossey-Bass.
- Black; P. & William D. (1998): Assessment and Classroom learning in Education; Principles Policy and Practice, Vol (5) *Issuel L*, Pp 7, 68p.
- 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)
- Boston, C. (2002). *The concept of formative assessment. Practical Assessment, Research & Evaluation*, 8(9). Retrieved July 10, 2011 from <http://PAREonline.net/getvn.asp?v=8&n=9>.
- Boston, Carol (2002). The concept of formative assessment. *Practical Assessment Research and Evaluation* 8 (9)
- Bremme D.W. & Herman, T.L. (1986). Beijing, China: Foreign Language Teaching and Examinations Authority (Eds.) *The effects of large-scale testing and related problems*
- Brown, Richards. Coughlin, Ed (2007). The predictive validity of selected. Benchmark Assessment. Regional Educational Laboratory Mid-Atlantic. Cambridge: University of Cambridge School of Education.

- Bude U. (1995) *Strategies for using Information to Improve Learning Conditions and Instructional Practices at the School Level* DSE
- Button D. (1995) "The Rolls Royce of examinations?", *Education + Training*, 20 (9): 286-287
- Calabrese R. L., Sherwood, K., Fast, J., Womack, C., (2004) "Teachers' and principals' perceptions of the summative evaluation conference: an examination of Model I theories-in-use", *International Journal of Educational Management*, Vol. 18 (2), pp.109 -117
- Choi, C. C. (1999) Public Examinations in Hong Kong, *Assessment in Education*, 6(3), 405-418.
- Cohen, J., & Cohen, P. (1983), *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Cooper, D.R and Schindler, P.S. (2003), *Business Research Methods* (8th edn), McGraw-Hill: New York
- Crocker, L, & Algina, J. (1986). Introduction to classical and modern test theory. Fort Worth: Harcourt Brace Jovanovich College Publishers.
- Crooks, T. J. (1988) The impact of classroom evaluation practices on students, *Review of Educational Research*, 58,438-481.
- Daggett, W. R. (2005), *Achieving Academic Excellence through Rigor and Relevance*, Ed. D, pg.
- Daniel. M and 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,

- DeVellis R. F. (2011) *Scale development: Theory and applications*, Sage Publications, Inc
- Dorr - Bremme, D.W. & Herman, J.L. (1986), *Assessing student achievement; A profile of classroom practices* (CSE Monograph 11), Los Angeles. University of California, Centre for the study of evaluation *Educational Assessment*, (pp. 17-23).
- Embretson and Reise (2000), *Review the ramifications of CTTs*.
- Eshiwani, G.S. (1990). *Implementing Educational Policies in Kenya*. Africa Technical Department Series Discussion Paper. World Bank
- Eshwani, G.S. (1954), *Research in Education; the Kenya Register occasional paper number 305* p. 12, Nairobi; Bureau of Educational Research. Kenyatta University.
- Farrugia, J. (2004) *Predictive Validity of SEC Examination Results*. Report presented to the MATSEC Examinations Board. Malta: MATSEC Support Unit.
- Gagne, R.M. et al (1992) *Principles of Instructional design*. 4<sup>th</sup> Edition, pp 106 - 110. Forth Harcourt Broace Jovanovich
- Gansky SA, Pritchard H, Kahl E, Mendoza D, Bird W, & Miller AJ, (2004) *Reliability and validity of a manual dexterity test to predict preclinical grades*. *Journal of Dental Education*, 68(9): 985-94.
- Gay, L.R., (1996), *Educational research: Competencies for analysis and application*; Upper Saddle River, New Jersey: Merrill, Prentice-Hall Inc., pp: 249-305.
- Giacomini R., & White H., (2006) *Tests of Conditional Predictive Ability*, *Econometrica*, 74 (6): 1545-1578
- Gipps, C. (1994) *Beyond Testing: towards a theory of educational assessment*, Lewes:
- Gipps, C. (1996) 'Assessment for Learning' in Little, A. & Wolf, A. (eds.) *Assessment in*

- Gonnella, Erdmann, & Hojat, (2004), an empirical study of predictive validity of number grades in medical school using 3 decades of longitudinal data: Implications for a grading system, *Medical Education*, **38**(4), 425 - 434.
- Graham D. S. & Whittaker, J. M. (2005) Predicting Academic Performance and Clinical Competency for International Dental Students: Seeking the Most Efficient and Effective Measures, *Journal of Dental Education*, 69(2): 270-80.
- Gronland, (1985)). The predictive validity of Osun state junior secondary certificate examination. Department of Educational Foundations and Counselling, Faculty of Education Obafemi Awodowo University, Ife. *Electronic journal of Research in Educational Psychology*. No. 5 - 3 (1), pp. 131-144.
- Harlen W. (2005). Trusting teachers' judgement; research evidence of the reliability and validity of teaches; assessment used for summative purposes.
- Harlen, W. & James, M. (1997) *'Assessment and Learning: differences and relationships*
- Herbling, David, (2006), 'Education for all is the key to Kenya's development', *The Standard*, 13 March
- Hunter, Laura and Leahey. (2008). "Collaborative Research in Sociology: Trends and Contributing Factors". *American Sociologist* 39:290-306
- Institute of Educational Assessors (2006) *Assessment and Validity*,
- Izard, J.F. (1998a). Quality assurance in educational testing. In *National Education*
- K.S. Chen, H.T. Chen, L.L. Wu, J.T. Emanuel, (2005) "Performance assessment of health examination for freshmen", *International Journal of Quality & Reliability Management*, 22 (8), pp.849 - 859

- Kapambwe W. M. (2010). The implementation of School based continuous assessment, *Zambia Educational Research and Reviews*, 53: 99 - 107.
- Kellaghan T. & Greaney V. (2003). Monitoring performance: Assessment and Examinations in Africa. Association for the Development of Education in Africa (ADEA) International Institute for Educational Planning
- KNEC, (2006). Candidates Overall Performance Report for 2005. Nairobi: K.N.E.C.
- Kobbrin and Schmidt, (2006), *predictive and placement validity of new Scholastic Achievement Test (SAT) writing section*.
- Kryspin W. J. & Feldhusen; J. F. (1974). *Developing classroom tests. A guide for writing and evaluating test items*. Burgess Publishing Company.
- Kryspin, W.J (1974). *Developing classroom tests*. The Burgess Educational Psychology Series for the teacher. London: Falmer Press.
- Kuhn, T. M., Johnson, R. L., Arguso, S. A., & Monrad, D. M. (2001). *Put to the test: Tools & techniques for classroom assessment*. Portsmouth, NH: Heinemann.
- Kuncel, N. R., Hezlett, S. A., Ones, D. S., (2001) A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate student selection and performance, *Psychological Bulletin*, 127(1): 162-18
- Linn, R.L. and Gronlund, N.E. (2000) *Measurement and Assessment in Teaching*, 8<sup>th</sup> edn. New Jersey: Prentice-Hall.
- Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, 3, 635-694.

- Lord, F.M. (1980). *Applications of item response theory to practical testing problems*. Mahwah, NJ: Erlbaum.
- Lubisi, R. C. & Murphy, R. J. L. (2002) Assessment in south African schools, *Assessment in*
- Lynn D. (1992). The case of a national testing system. *Practical assessment Research and Evaluation. A peer review educational journal*.
- Measuring Academic Achievement, July, 2006,
- Moghaddam, S. (2010) "Cultural schemata: Iranian students' test-taking processes for cloze tests", *Education, Business and Society: Contemporary Middle Eastern*, 3 (3): 188-200
- Mugenda M. O. and Mugenda A. (1999), *Research Methods: Qualitative and Quantitative Approaches*, African Centre for Technology Studies, Nairobi, Kenya
- Mugenda, O.M and Mugenda, A.G (2003) *Research Methods, Quantitative & Qualitative Approaches*, Acts Press, Nairobi.
- Mushi, Seina L.P. (2001). Evaluating Validity and Reliability of Classroom Assessments using secondary data. Paper presented at the Annual Meeting of the American Educational Research Association (Seattle, wa, April 10-14).
- Ngechu. M. (2004), *Understanding the research process and methods. An introduction to research methods*. Acts Press, Nairobi
- Nicol, D.J. & Macfarlane-Dick, D. (2004). Rethinking formative assessment in HE: a theoretical model and seven principles of good feedback practice. In, C. Juwah, D. Macfarlane-Dick, B. Matthew, D. Nicol, D. & Smith, B. (2004) *Enhancing*

student learning through effective formative feedback, York, The Higher Education Academy.

Nitko J. A. (1994). A framework for concepts and procedures of using continuous assessment for formative and summative of student learning. *Educational resources information.*

Noll & Scaunel, (1972). Procedures of using continuous assessment for formative and summative of student learning. *Educational resources information.*

Nsude, I. (1998). Principles and Practice of Education (Education I). Enugu: Hillys Press Nigeria Ltd.

*i* Nyougo. H. (2007), use of Education Monitoring and Evaluation Information to improve classroom teaching and assessment (Kenyan case). *An unpublished independent study paper written towards attainment of M.E.D. - Measurement and Evaluation, University of Nairobi (2005)*

Nzweze. T. (2009). Management of Examinations; Ethical issues. Edo journal of counselling. Vol. No. 1.

O'Rourke, B., M.O. Martin and J.F. Hurley, 1989. The scholastic aptitude Test as a Predictor of third-level academic performance. *The Irish Journal of Education*, 23 (1): 22-35.

Obayyi (1993), Working inside the Black Box. Assessment for learning in Classroom. *The Phi. Delta Kappan*. Vol. 86, No. 1. pp 8 - 21.

Obioma. G. and Salau. M. (2007). The predictive validity of public examinations. A case study of Nigeria, Abuja. Nigerian Educational Research Council (NEROC).

- Ojerince, A. (2000), *Prediction of Academic Success in School Certificate Examination from national common Entrance Examination Scores*, M.Ed Dissertation (unpublished), Faculty of Education, University of Ife, Ile-Ife.
- Okuna, M. A. (2006) Academic goals, goal process cognition, and exam performance among college students, *Learning and Individual Differences*, 16 (3): 255-265
- Omrin. M.s. & Ale R.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.
- V Onocha and Okpala PN (1995). Tools for educational research Stirling Holden Publishers, Ibadan.
- Othuon, L. and N. Kishor, 1994. Hierarchical linear modelling of predictive validity: The case of Kenya certificate of primary education examination. *Studies in Educational Evaluation*, 20: 181-190.
- Othuon. L.A. (1994). Hierarchical linear modelling of predictive validity. The case of Kenya Certificate of Primary Education Examination. *Studies in Educational Evaluation*, 20; 181 - 190.
- Owoyemi, N., 2000. Moderation and standardisation of continuous and terminal assessment scores in junior secondary school certificate examination and primary school leaving certificate assessment. Paper delivered at the senior staff seminar, Ministry of Education, Ado-Ekiti 2nd March, pp: 2-9.
- Paris, S.G., Lawton, T.A., et al. (1991), A developmental perspective on standardised achievement testing. *Educational Researcher*. 20(4) page 40

- Peers, Ian S. and Johnston, Margaret, (1994), Influence of learning context on the relationship between A- Level attainment and final degree performance: A meta-analytic review. *British Journal of Educational Psychology*, 64: 1-15.
- Popham, W.J. (1999). Why standardized tests don't measure educational quality. *Educational Leadership*, 56(6), 8-15.
- Popham, W.J. (2001). Why Standardized Test Scores Don't Measure Educational Quality. *Educational Leadership*, 56(6) 8-15.
- Powell, J. C. (2010) Testing as Feedback to Inform Teaching, *in; Learning and Instruction in the Digital Age: Making a Difference through Cognitive Approaches*. New York: Springer.
- R.L. Linn (1992), Achievement testing. In M.C. Alkin (Ed), *Encyclopaedia of educational research* (6<sup>th</sup> edition, ppl - 12) New York; Macmillan.
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28 (1): 4-13. Research Press.
- Sadler, D.R. (1989). Formative assessment and the design of instructional systems, *Instructional Science*, 18 (2): 119-144.
- Scriven M. (1967) *The Methodology of Evaluation* (Washington, DC, American Educational Research Association).
- Thissen, D. & Orlando, M. (2001). Item response theory for items scored in two categories. In D. Thissen & Wainer, H. (Eds.), *Test Scoring* (pp. 73-140). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

- Ubokobong, H.E., (1993), Predicting educational performance at tertiary level on the basis of secondary level performance. *International Journal of Mathematical Education in Science and Technology*, 24 (2): 287.
- Uche, O.S. (2005). *National Disgrace: Examination Malpractice*. Enugu:
- West K. D. & McCracken M. W. (1998). "Regression-Based Tests of Predictive Ability," *International Economic Review*, 39(4): 817-40
- William D. & Black P. (1996). Meanings and consequences; a basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal*, 22, (5): 34-45
- Zaida H. & Daliala R. (2007). *Predicting students' academic performance. Comparing Artificial Neural Network, Decision Tree and Linear Regression*. 21<sup>st</sup> Annual SAS Malaysia Forum, Shangri-La. Hotel, Kuala Lumpur

^ S t f f ' i A i j g —

