

Sh. 1500

I

AN INVESTIGATION OF THE RELATIONSHIP BETWEEN
FORM FOUR TEACHER - MADE TESTS AND KCSE
EXAMINATION IN UASIN GISHU DISTRICT

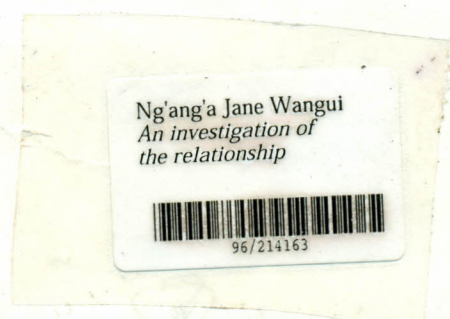
This is my original work and IN KENYA presented for a degree in any other

BY:

NG'ANG'A JANE WANGUI

21/12/20

Abstracts



Ng'ang'a Jane Wangui
*An investigation of
the relationship*



96/214163

" A Thesis submitted to the Faculty of Education in partial fulfillment for the
Degree of Master of Education at Kenyatta
University"

KENYATTA UNIVERSITY LIBRARY

1995

DECLARATION

" This Thesis is my original work and has not been presented for a degree in any other University. "

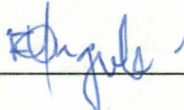


Jane W. Ng'ang'a

" This Thesis has been submitted for examination with our approval as the University Supervisors "



Dr S. K Bali



Dr F Ingule

ACKNOWLEDGEMENTS

This work has been made possible by the support of a number of people:

I dedicate this work to my dear parents Mr and Mrs Peter

Ng'ang'a Ng'ang'ira who made it possible for me to undertake the course through provision of funds.

I acknowledge the co-operation and assistance given to me by the headteachers and concerned teachers of the schools I visited.

I register my heartfelt gratitude to my parents for financing the entire course and for their encouragement all through. Much thanks to the family of Mr and Mrs Alex Ngumi who made my stay in Nairobi a success through their support. I also thank my brothers and sisters for their presence and encouragement. Special thanks to my friend Kevin K Ndubi for his prayers and support. Finally thanks to all my friends for their encouragement and

Jane Ng'ang'a

IV

ACKNOWLEDGEMENTS

This work has been made possible by the support of a number of people:

First and foremost I am grateful to my supervisors Dr Bali and Dr Ingule of the Department of Educational Psychology at Kenyatta University, for their tireless efforts in critical reading of the work and guidance all through the study. I acknowledge the co-operation and assistance given to me by the headteachers and concerned teachers of the schools I visited.

I register my heartfelt gratitude to my parents for financing the entire course and for their encouragement all through. Much thanks to the family of Mr and Mrs Alex Ngunu who made my stay in Nairobi a success through their support. I also thank my brothers and sisters for their patience and encouragement . Special thanks to my friend Kevin K Ndubi for his prayers and support. Finally thanks to all my friends for their encouragement and Catherine who typed the work.

Jane Ng'ang'a

ABSTRACT

The purpose of the study was to establish the relationship between first term, mock and KSCE examinations.

The sample used consisted of 585 form four students of the year 1990 from eight schools in Uasin Gishu District. The variables used were the seven compulsory subjects namely: English, Kiswahili, History, Geography, Mathematics, Biology and Physical Sciences, and the Average Grade.

The test scores of the three examinations in the eight variables were collected for each student for the purpose of the study.

The data was analyzed using Pearson - product moment correlations formula to help obtain correlation between the variables used and multiple regression equations were also for obtained for each examination. The results were tested for their significance using Fishers Z-transformation and F-tests at 0.05 level of significance. The result was that mock examination was highly and significantly related to the KCSE examination compared to the first term examination.

The results show that mock examination results give a measure of the level of achievement of a student and there is need to incorporate it in decision making process in addition to the KSCE examination. The mock examination can act as a supplement of the KSCE examinations.

There is need for the first term examination to be reviewed and revised in order for it to give a reliable measure of the student's level of the achievement.

LIST OF TABLES

Table 3.1	Sample used
Table 4.1	Descriptive statistics of the variables of first term examination
Table 4.2	Descriptive statistics of the variables of mock examination
Table 4.3	Descriptive statistics of KCSE examination results
Table 4.4	Intercorrelations of first term examination results
Table 4.5	Intercorrelation of mock examination results
Table 4.6	Intercorrelations of KCSE examination results
Table 4.7	Z_r -Values between first term and mock examinations
Table 4.8	Z_r -Values between first term and KCSE examinations
Table 4.9	Z_r -Values between mock and KCSE examinations
Table 4.10	Relationship between the examinations
Table 4.11	Z-Values of equality of correlation coefficients between the examinations
Table 4.12	F-Values of the examinations

TABLE OF CONTENTS

	PAGE
Chapter one 1.0 INTRODUCTION:	1
1.1 Background to the study	1
1.2 Statement of problem	11
1.3 Purpose of study	13
1.4 Significance of study	17
1.5 Definition of terms	19
Chapter two 2.0 LITERATURE REVIEW	21
2.1 Outside kenya based studies	21
2.2 Kenyan based studies	29
2.3 Statistical model	32
Chapter three 3.0 METHODOLOGY	38
3.1 Research Design	38
3.2 Population and Sample	39
3.3 Variables	40
3.4 Instruments	41
3.5 Data collection	45
3.6 Data Analysis	46
Chapter Four 4.0 RESULTS : DATA ANALYSIS	49
4.1 Descriptive results of the variables in each examination	49
4.2 Inter-correlation between the variables in each examination.	54
4.3 Analysis of the correlation results between the three examination	57

4.4	Regression analysis results	72
4.5	Summary	76
Chapter Five 5.0 DISCUSSION		79
5.1	Conclusion	86
5.2	Recommendation	88

REFERENCES	92
------------	----

APPENDICES

Appendix	1	95
Appendix	2	96
Appendix	3	97
Appendix	4	98

CHAPTER ONE

1.0 INTRODUCTION

1.1 Background to the study

In our schools, what use do we make of the tests and examinations administered by the teachers? The question can be answered if we first understand why we have these tests and examinations within the instructional process. This points out the need of tests which gives insight into some of the uses of the tests that can be applied by the teachers and those concerned with the educational process. The tests and examinations administered by the teachers are collectively referred to as teacher - made tests.

The teacher-made tests are concerned with monitoring and management of student progress. They provide continuing assessments of pupil attainment in school. According to Ebel periodic assessment of educational progress is essential to effective education and good tests afford very useful assistance to teachers in making these assessments (Ebel, 1972). These teacher-made tests are in the form of continuous assessment tests and end of term school examinations administered during the instructional process. They are constructed, administered and marked by the teachers in the schools.

A good system of assessment is of importance too in the implementation of the curriculum in the school system. Assessments give immediate feedback to both the learner and the teacher. If what is contained in the examination is not representative of the expected outcomes of education, the entire education system may be derailed.

Well-designed tests and examinations help improve the curriculum. They encourage teaching-learning process of the designated curricula in the classroom (Gándye, 1991). The tests results are therefore, in such cases of importance to the curricula developers.

The tests administered in schools that is, institutions of learning are mostly achievement tests which aim at measuring the level of mastery of a student in a given content area. These tests help in identifying how well a student has mastered a given content area hence competent in it. Such knowledge is important both to the teachers and students in the school. This is because the test results are believed to be indicators of the student level of achievement (Cronbach, 1970). This means that from the test results one can be in a position to point out how well the instructional objectives outlined prior to the teaching-learning process were achieved. They also point out how well a student has learnt and hence the success of the educational process as a whole. The test result used is dependent on the need arising during the instructional process.

A cumulative result (either a score or a grade) of many tests administered can be used to determine the rank of a student. Individual results may be used when specific content areas need to be analyzed how well they have been mastered. This implies that any test results obtained are important and can be used to meet a need at hand. At times the usefulness of test results is not only about a student but also the teacher. The test results have at least as much to tell us about the teacher as well as the student also. The teachers can learn something from test results; first they are able to tell how much a student has mastered in a given content area, and which student has mastered the content area or not. This

would prompt them to take specific measures to rectify any anomalies identified.

Secondly, teachers help students in making decisions concerning their future undertakings - such as course of study. The results would be incorporated with the knowledge by the teachers about a student's competency in a given area of study when making such decisions. This would avoid choice of subjects or study courses that one is not competent in. These functions of tests would counteract the conviction that tests within a course of study give only 'part of the picture' as compared to tests done at the end of the course and hence cannot be relied upon (Hoffman, 1983). This is because the tests give a reflection of the teaching done, study and practice the students engage in and other instructional efforts.

The need of tests also arises when during the instructional process differences among students are observed, the teacher therefore, uses the tests as a device to identify the individual differences in the area of concern such as deficiencies in competency in a given subject. The students can be given individual attention in terms of remediation work. In addition, different tests are used in testing different academic skills. The teacher may need to know how well the students can take a given test and answer questions correctly. For example, a teacher may need to know how well the students can answer such a test. Test-taking skills of the students can therefore be developed for various kinds of tests.

Different methods of instructions are used in the teaching - learning process. The tests point out to the teacher about the effectiveness of the methods of instructions employed. For example, a lecture method and question-answer

method may be used and tests given after both practices in the same content area or related content areas would aid in improvement, adoption or continuity in a given method of instruction. At the end of the teaching and learning process, teachers and educators are concerned with how well the instructional objectives have been achieved. Such objectives are best measured within the course of instruction when specific objectives guide the instructional process. The tests can be employed to measure achievement of these objectives. This would reveal any objectives not achieved and measures taken to achieve them.

The teachers may tell the effectiveness of the tests constructed if the tests achieve the objectives for which they were constructed. Otherwise, they would seek to improve the tests in an area of weakness identified. This could either be in terms of items sampled, objectives tested, or content area sampled. The testing skills a teacher has enables him/her to prepare a good test for an intended purpose.

The test scores are related to one's knowledge and educational efforts and they reward and reinforce effective study (Ebel, 1972). The student while studying needs to identify the worth of his efforts in terms of how much he has achieved or learnt. The test serves as an instrument to measure how much he has achieved or learnt. This helps the student identify areas which need attention in terms of study and the test scores would motivate and reinforce his efforts when they reveal achieved objectives.

The administrators are faced with the need of tests. The tests are needed when a student's academic progress is required by the administrators who must report to parents/guardians such progress. The schools focus on the students

progress as an ultimate criterion for acquisition of essential knowledge, understanding, skills and attitudes. The tests also serve as basis for planning and evaluating the effectiveness of the educational programs and operations. Most of the decisions made in the education field rely upon the tests administered by the teachers. With the test results a decision can be made regarding a programme - whether to continue with it, stop it or alter it. For example addition of subjects to a curriculum. There is need therefore, for careful thought and co-ordination between curriculum development, assessment and instructional methods (Mkandawire, 1991). Based on the functions outlined above it would be reasonable to conclude that the teacher-made tests would be good measures of educational achievement of a student. The teacher-made tests may even outweigh other local achievement tests in terms of educational significance (Ebel, 1972). This is because they influence how a student studies and what he learns.

Testing is a practice common in the education process in any country. This is because of its importance in determining the effectiveness of the education practice.

The Kenyan Examination System.

The Kenyan examination system is dependent upon the education system. The Education system(8-4-4) has three levels of education; primary (8 years), secondary (4 years) and university (4 years). At each given level tests are administered during the course of instruction and at the end of the instructional process.

At the secondary school level teacher-made tests are administered during the four years; continuous assessment tests as well as end of year examinations. Each academic year is divided into three terms; first, second, and third term. During a given term teachers administer tests to the students. These tests help monitor the students' progress as well as measuring their achievement level in a given content area. Using the test results the teachers would be able to identify the weak students and give them individual attention. These tests are given at a specific time determined by the teachers during the instructional process. This is because periodic assessment of the education process is essential to effective education and that good tests afford very useful assistance to the teachers in making those assessments (Ebel, 1972). These tests are referred to as before continuous assessment tests.

At the end of each term the teachers administer examinations that measure the instructional objectives for each content area. These end of term examinations are constructed, administered and marked by the teachers. The test results are used for ranking, thus, pointing out to the students their position in terms of the level of achievement against the others in his/her class. Little if any, information is drawn from these test results in terms of test-taking skills, testing skills and the like; which are of importance in the testing technology. The examination results of the second year examinations are used in decision-making. This is decision concerning the subjects to study in line with future programmes to be done by the students. Such a decision is very important and the examination results therefore need to be representative of the students' level of achievement in each content area.

The test-results may only be representative if the test and examinations are well planned and constructed. But more often than not, teachers do not take time to prepare them as they realise that there is so much to be taught and so little time to learn (Brown, 1983). The examinations thus constructed may not achieve the purpose for which they were constructed as it may lack in reliability and validity. The examinations and tests administered by the teachers ought to be for giving direction to the teaching-learning process.

The fourth year of the secondary education is the last year of the secondary cycle in education. The students therefore, take a National examination at the end of the year. During first term of this year, the teachers administer continuous assessment tests and end of term examinations. These tests and examination results help the teachers identify areas not well covered and the level of achievement of the students. More use could be made of such test results as recommended by the Kenya National Examination Council reports (KNEC). Such uses are development of the students' test-taking skills in examination questions that are of essay type, comprehension or those that require interpretation. These skills are often missing as revealed by the Kenya National Examinations (KNEC, 1991).

The teachers themselves could also benefit from such tests. The tests help point out testing skills that a teacher has in the teaching-learning process, neglect of the quality of the teacher in terms of testing skills may affect a test adversely and the performance of the students. Testing skills would be developed in areas like, construction, marking and analysis of results. Analysis of results is of importance especially in the education process as they reveal whether purposes are

achieved for which tests were constructed.

At the end of second term in the fourth year, the students take an examination known as the mock-examination. The mock examination is administered as a parallel examination to the KCSE examination. It consists of specific number of examination papers per subject as the National examination. The format of each examination paper is similar to that of the National examination. In addition, the grading of the results is done using the same grading system. The aim of the mock examination is to point out the mastery of content in different subjects and also reflect the level of achievement of the students. The district-specific mock examination is common to all schools within a district and the school specific mock examination is only done by the given individual school for which it was constructed. Apart from the aim for which it is constructed, the mock examination results could be used to identify specific areas of weakness or strengths in individual learners. The nature of the deficiencies would be ascertained and corrective measures taken (Mkandawire, 1991).

The teacher-made tests are therefore of utmost importance in the educational process. They need to be effectively used to improve the quality of education given to the student and improve its effectiveness. This would help curb wastage as a result of poor performance due to lack of necessary testing skills and test-taking skills.

Kenya Certificate of Secondary Education (KCSE)

The end of third term of the fourth year is marked by the administration of the KCSE examination. It serves both as an end of year and end of course examination. The KCSE examination is a public examination taken during the months of October and November by all students completing the secondary level of education. Therefore, it is a common examination in all schools in the country.

The KCSE examination is regarded with high respect. This is because it is a national examination and it is used as a tool for differentiating between those who will join the university and other higher learning institutions or not. Little, if any, information is drawn from the other tests a student does during his school years in determining his suitability in joining the university or other higher learning institutions or not. The importance attached to the KCSE examination is also revealed by the effort put in by the students in their studies and the teachers too.

During third term of the fourth year there is much student-teacher interaction in order to improve the student's performance in the KCSE examination. Extra time is spent to assist weak students in any given content area(s) and to cover any remaining work in the syllabi. These and many other occurrences create a tensed-up environment before and during the time of taking the KCSE examination. This indicates that relying wholly on the KCSE results in making decisions concerning a student may not be fair and just. The results may be affected by the anxiety in a student and stress too among other conditions.

The KCSE examination is administered by the Kenya National Examination Council (KNEC) which is an examination body in Kenya. The aim of the examination is selection and certification. About 29 subjects are examined through the KCSE examination. The examination is constructed by test developers within the KNEC and administered to schools with the help of supervisors and invigilators. These are teachers from various schools as well as personnel from the education department within a district. The examination is marked by teachers and other personnel in related fields of education. For example, universities and colleges. The markers undergo training to equip them with the relevant marking skills. The results are compiled and analyzed by the KNEC and released to the schools for reporting to the students.

The KCSE results reveal what was not achieved during the instructional process by the teacher and students. The information obtained would serve as a guideline to teachers while training other students who have not taken the examination and to such students as well. This has always been the case as the KNEC produces a report after a given year's KCSE results. Examples of such information are, inadequate coverage of topics included in the syllabus, concepts not learnt, lack of adequate training of students to give correct interpretation to certain words used in questions, lack of drawing skills and other specific skills required in certain subjects (KNEC, 1991,). The students taking the KCSE examination at the end of fourth year would have such weaknesses corrected if continuous assessment tests are given often by the teachers who would be able to identify them early enough during the course.

The KCSE examination results are important for they offer accessibility to various employment sectors and training institutions. Most of the decisions made about a student after school are dependent on these results. Therefore, it is the concern of all that the students achieve good grades in these examinations. This can be achieved if teacher-made tests are constructed well and used well for the different purposes during the instructional process. Good use of teacher-made tests can be motivated by using the teacher-made test results in making decisions about a students by the results being incorporated with the KCSE examination results. Such decisions are determination of entrance to a college or university or acceptance to a given job. The KCSE results obtained would be therefore related to that of the teacher-made tests. This would then give the teacher-made tests their importance and use as required. The achievement of such a goal would be realized if the teachers were aware of the uses of these tests and if the results of these tests were used even after a given level of instruction. The teacher-made tests have many uses and are of utmost importance in the instructional process.

1.2 Statement of Problem.

Testing within the course of instruction by teachers is encouraged by the Kenya National Examination Council(KNEC). This occurs through examination reports by the examination council in order to enhance the quality of education given to the students. Testing is achieved by use of teacher-made tests which serve as continuous assessment devices. The teacher-made tests measure the level

of achievement in a given content area. They are intended to be suitable indicators of mastery of content if well constructed and as a way of improving mastery of content (Ebel, 1972). This is because coverage of content areas adequately according to laid down instructional objectives would be reflected in the test results obtained. The teacher-made tests may be used for other purposes as continuous assessment instruments hence well placed than the national examinations.

This may be illustrated by considering the two main limitations of the national examinations. First, the examinations are normally administered at the end of a fairly long period of study. Therefore, improvement of student performance is not possible through such testing programmes. It would be too late to help those who have not achieved the expected outcomes of the curriculum. The teacher-made tests would help identify concepts not learnt, skills achieved and those not achieved hence improve the students performance. These are both academic and test-taking skill such as drawing, comprehension and interpretation. Secondly, the national examinations cannot adequately cover all aspects of the curriculum due to constraint of time. These examinations give results of a sample of the curriculum as a representative of the whole (Gandye, 1991). This is the concern of the study.

The results obtained from the teacher-made tests are of tremendous interest to the teachers, parents, students and those concerned with the teaching and learning process in the school. However, the use made of the test results is often different. Results are used for ranking of students in order of achievement, reporting students progress and measuring level of achievement by the teachers.

The main purposes of teacher-made tests as outlined by the KNEC may not be achieved due to lack of knowledge on how such purposes can be achieved. The teacher-made tests have important uses in the teaching - learning process. The use made of the teacher-made tests should be geared towards improving the students level of achievement and the educational process as a whole in all levels of instructions. For this to be achieved, the tests need to be related to the KCSE examination in terms of the results obtained, hence indicating that they are good achievement measures.

The study therefore, addressed itself to the following questions:

What are the uses of teacher made tests?

What is the relationship between the tests done within the course of instruction and end of course examination? ✓

1.3 Purpose of Study

Though teacher made tests have been used to determine the students' progress and level of achievement, there is much more to draw from them hence the purpose of the study. The study aimed at highlighting the possible uses of the teacher-made tests. The relationship established between the tests and the KCSE examination would help the teachers identify whether the tests they constructed were related to the KSCE Examination. This would therefore help them make use of their tests results as required. It was important to establish that the teacher-made tests are good measures of achievement otherwise they would be inadequate to form the basis of educational decisions made by the teachers.

The use of KCSE examination results alone in making various decisions about a student is questionable because of unknown reliability and validity. Examples of these decisions are, selection decision of a course of study in a college or university, the job to take or career to pursue, placement and classification decisions. The teacher-made tests results could complement the information about the observations made for the purposes which cannot be accomplished by the limitations of the national examinations in the testing process. This is of importance so that decisions made about a student are fair and just. In such cases therefore, the Grade point averages would be an added piece of information. The use of results from teacher made tests would itself be a motivation for the teachers and an encouragement in competency in the testing practice in the instructional process.

Since good tests are required to achieve the continuous motivation of students and ensure learning, the teachers would follow the requisite procedures of test construction. This would help them achieve the intended purposes and enhance exercising their professional freedom as teachers in the testing process. The tests given would help teachers check whether the tests focus on all aspects that is, knowledge, comprehension, application, analysis, synthesis and evaluation abilities. The kind of test administered would help achieve this objective thus pointing clearly that a test constructed should be for an intended purpose. The result would be improvement of the instructional validity of the assessment.

Students' mastery of content and coverage of given content would be determined through testing. This would help in giving individualized instruction through remediation practice; thus improving the students' performance at the end

of the course of instruction. The tests would serve as a motivation factor in the learning process by encouraging the well achievers who show mastery in content. Adequate testing of the entire curriculum content in each content area would also be achieved.

Adequate coverage of content would be achieved through continuous assessment practice during the instructional process. The students would be encouraged in their competency in each content area. In some cases failure may not necessarily be a reflection of inadequate coverage of content, poor mastery of content but poor test-taking skills. This would be reflected in the national examinations if not dealt with earlier. Teachers would be of great help in such cases. Through teacher-made tests administered, test-taking skills of the students would be tested and training/guidance given where need is identified. Test-taking skills need to be achieved by the students during the instructional process, hence a duty of all teachers. The training given is dependent on the type of test administered and the type of questions to be answered. The results of such training would be reflected in the test result.

The teacher-made tests would be used in guiding students if they are good measures of achievement. Guidance and counselling in the choice of courses of training after a given level of education or for further studies is of utmost importance. The teachers would be of great assistance in such cases and the tests would give them greater confidence in using the test results to make such decisions. These decisions are made during the course of instruction and not after. Therefore, it is important that the test results used are representative of a student's level of achievement. Students too would be able to make sound

decisions in areas of vocation and education.

The investigation of the relationship between first-term and mock examinations would assess whether the two are different. This is in terms of their measurement of the student's level of achievement. This may provide empirical evidence in support of banning of the mock examination as an external examination, in that, it was duplicating what was achieved by the first-term examination, hence not necessary. The fact that the two are different in terms of construction, and marking whereby an individual teacher does the work for first-term examination while a group of teachers do the work for the mock examination. It is worth finding out whether this difference is reflected in the test results obtained. The correlations obtained helped make the conclusions.

The relationship between the results in the various subjects would be used to identify an individual's weak subjects and strong ones. This is because dependence on ones overall test result may not really result in a confident differentiation between the subjects. This is as a result of the many factors that influence the ultimate performance of an individual on a given test. In order to conclude that one is good in a given subject and not the other, reliable evidence such as test results can be used instead of relying on opinion or other wise. The aggregate score or grade has been used for a long time in making various decisions regarding the students. The study set out to show out that the aggregate score is a reliable measure of the student's level of achievement. This is because it is a representative score or grade of the information received from the individual subjects considered. In some certain cases individual subject grades may be used to make more specific decisions. This is because some subjects give higher

contribution than others to the aggregate score or grade. The subjects with high contributions may be considered in making some decisions such as which course (s) would benefit most in a college or universities. In a case where they are already in operation this move would be strengthened by the results obtained.

1.4 Significance of study

Educational measurement emphasizes use of tests that are both reliable and valid. This gives meaning to any test used and the result obtained. A test is only useful if it serves the purpose for which it was constructed. This minimizes overrating of the test results. For example, examinations given by the teachers will have a defined purpose and will assume their status as measures of student achievement in the education process.

Therefore, using the information from this study, examinations administered during instructional courses would be considered with high regard by the students, teachers and the public. Rarely ever do the students know the purpose of a test because the teacher normally does not identify nor does he communicate this information to his students. This is because tests given frequently are constructed hastily without a defined objective. The tests therefore tend to be of questionable quality.

The test results are not evaluated and are not used for evaluation. The correlation established between teacher-made tests and KCSE examination would indicate the worth of the tests and should therefore be taken more seriously.

The students need to be notified of the purpose and aim/objective of a test administered to them. This would enable them interpret their scores correctly.

For example, a test on mastery thus taking corrective measures accordingly.

The teachers who are central in the testing program within the school, need to be well-equipped with the necessary testing techniques. There is need to improve testing skills of the existing teachers through in-service courses, seminars and short period training when necessary.

A well-skilled teacher in testing would be confident to undertake the exercise as results obtained are to be used and meaningful. This would serve as a motivation in the testing profession. The teachers would be encouraged to evaluate their test results drawing meaning from them in line with the purpose of the test. The evaluation would call for analysis of test results. This requires that guidelines need to be provided on the techniques of scaling and weighting to be followed by the teachers/schools in obtaining the scores from the tests. The result would be proper use and interpretation of the test results.

Both the teachers and the students would regard tests not as a threat but a technique to help them achieve the desired goals and objectives in education. Tests would also serve as a motivation to the teachers and students during the instructional process.

Since the testing process is of importance, it would be the concern also of the KNEC . The KNEC need to provide guidelines on techniques of assessments especially construction of tests and analysis of results. This would help monitor testing exercises in all schools and give importance to teacher-made tests. The KNEC need to venture into the use of teacher-made test to measure the affective

and psychomotor domains which are less catered for by current examination exercise. This would encourage testing other skills by the teachers other than cognitive skills.

The teacher-made tests given as a whole should function as feedback mechanisms to facilitate teachers' ability to adapt instructions to individual students. The tests should also meet the feedback condition which will facilitate learning in the entire learning process.

1.5 Definition of terms

1. A-Level - KACE examination done at the end of two years of high school education.
2. Achievement - performance in school in a series of educational tests.
3. Achievements test - A test measuring the effects of a specific programme of instruction or training in a given area of the curriculum.
4. Certification - To certify an individual's competence in a particular field.
5. Content - The subject matter taught in school.
6. Correlation coefficient - A statistical summary of the relation between two variables.
7. O-level - KCE examination done at the end of four years in secondary education.
8. School Examination - Tests done at the end of the first and the second term of the fourth year in secondary education.

- 9 Selection - A process through testing of the individuals who would and who should not go on with a programme of instruction while others do.
- 10 Significant - A correlation that differs by more than a chance amount from a postulated 'true' population value of zero.
- 11 Teacher made tests. - Continuous assessment tests and school examinations administered during the instructional process by the teacher.
- 12 Table of Specifications - A table drawn in order to obtain a balance of items included in a test covering some content.
- 13 Test battery - A group of tests yielding several comparable scores on the individuals or group.
- 14 Variable - A property whereby the member of a group or set differ from one another.

CHAPTER TWO

2.0 LITERATURE REVIEW

It has been in the interest of researchers to study the common concurrent validity of various tests designed to measure a common trait. These traits are for example, achievement, intelligence and anxiety. A number of researches carried out have revealed that tests designed to measure a common trait have concurrent validity for they have yielded positive and significant correlations. The tests can therefore be used in conjunction with one another yielding advantageous results.

Concurrent validity is an empirical technique of studying the relationship between the test scores and the criterion (an independent external measure). Theoretically, tests constructed to measure a common trait, are highly related and should therefore yield high and significant correlations empirically. These would indicate that they have concurrent validity. In addition to concurrent validity established of tests, the importance of tests such as teacher-made tests has been researched on and various writings have been put down regarding this. In order to illustrate this, a few studies and papers written will be cited below. The studies done outside Kenya will be reviewed first followed by Kenyan studies.

2.1 Outside Kenya based Studies

A study was carried out by Ansley and Forsyth 1983, in a university in united states (US) to investigate the relationship of elementary and secondary school achievement test scores to college performance. The test scores they used were of the Iowa Tests of Basic Skills(ITBS) and Iowa Test of Educational

Development (ITED). A population of about 2,000 freshmen who had entered Iowa University in 1977 was used. They obtained test scores in grades 4,6,8,9,10,11 and 12 for some of the students. Criterion scores were obtained from three criterion measures in college. These were first semester 1977 Grade Point Average (GPA), 1988 GPA and Graduation status (graduated/not graduated).

Using correlation analysis technique they established that the correlations obtained were similar to those reported for tests of the college entrance type. This supported the fact that ITED and ITBS are significantly related to first semester college GPA, final GPA and graduation status. The conclusion was that any of the achievements could be reasonably used to help in predicting future academic performances.

Another study by Bruno 1981, in Canada investigated the concurrent validity of learning ability profile against college GPA. The purpose of the study was to determine the concurrent validity of the Learning Ability Profile (LAP) by correlating the results obtained on the LAP with GPA. The sample used consisted of 80 French - speaking students in the Bachelors degree program in Business Administration. Scores on LAP, both weighted and non-weighted were correlated with correlations of .30 ($P < .01$) and .31 ($P < .01$) respectively with GPA. It was concluded that learning ability measures obtained with this test could present an acceptable validity in academic contexts in which selection is concerned. They employed the pearson-product moment correlation formula in their analysis.

Eaves and Simpson 1983, in Auburn university carried out a study to determine the concurrent validity of the Woodcock Reading Mastery Tests (WRMT) relative to the Peabody Individual Achievement Tests (PIAT) among

retarded Adolescent. Both the instruments, WRMT and PIAT, were administered to 56 adolescents. Scores on the PIAT reading and spelling subtests were correlated with scores on the WRMT subtests which were designed to measure similar skills. The presence of relatively high and positive partial correlations between appropriate subtests supported the concurrent validity of WRMT relative to the PIAT and to the criterion variables.

Partial correlation co-efficient were computed which controlled for the inflationary effect that age has on inter-correlation among Achievements test scores. These were for each of the three PIAT subtests and for three WRMT subtests. The correlations were positive and significant at $\alpha = 0.01$. The partial correlation ranged from .49 to .96. A strong relationship between variables used was revealed by the correlation obtained ($r > .87$). The subjects yielded high correlation (.94 to .90) indicating that they share a great deal of common variance (88% and 81% respectively) and they can be used in conjunction with one another advantageously with the samples used.

In another study by Eaves and Simpson 1984, in Auburn University the same was revealed. The study was designed to investigate the concurrent validity of the Peabody Individual Achievement Test (PIAT) relative to the Keymath Diagnostic Arithmetic Test (KDAT) among adolescents. The two tests, PIAT and KDAT, were administered to 121 adolescents of mean age 12.8 years. Mathematics scores on the PIAT were correlated with all scores on the KDAT. Partial correlation co-efficient were computed according to age, sex, race, IQ and total group. The presence of relatively high and positive partial correlation between the scores support the concurrent validity of the instruments.

Esezobor (1991), in Nigeria did a research on the procedure of carrying out school based continuous assessments (CASS) and its use in public examination at the secondary schools in Nigeria. He observed that Nigeria has been relying on School certificate/ General certificate of education (SC/GCE) ordinary and advanced level as the major examination for evaluating pupils attainment at the completion of secondary school. There was a need to implement continuous assessment in order to handle the formidable task of using information about the child in the three domains of educational objective that is cognitive, affective, psycho-motor domains. This was for the purpose of certifying the pupils attainment at the end of the secondary school course.

The result was the introduction of CASS which is a component of the candidate's overall assessments. CASS imposes greater responsibilities on the classroom teacher and involves massive data from various sources. To facilitate adequate assessment of pupils, detailed guidelines have been provided on the techniques of assessment especially on the construction of these tests. Guideline on scaling and weighting to be followed by teachers/schools in obtaining internal assessment scores are also provided. It is the concern of the West African Examination Council (WACE) how CASS is handled at the school level.

Gandye (1991), in his paper pointed out that studies on school based assessment in Tanzania show that, where the program has appropriately been carried out there has been success concerning the issue of quality, variety and relevance of assessment of knowledge, abilities, skills and attitude including character formation. Through terminal tests, for example, some teachers have been able to follow the requisite procedures of item preparation, paper assembly

and administration. This has enhanced the instructional validity of the assessment and exercise by teachers of their professional freedom as teachers.

The school-based assessment cannot be used on its own for various purposes due to some limitation observed. School-based assessment compared to external examinations have poor discrimination index and test item focus less on higher level abilities such as analysis, synthesis or evaluation. Both external and school-based assessment should be used for better achievement of the objectives and goals of the curriculum. The limitation of school-based assessment, stem mainly from lack of well trained and motivated teachers. This can be achieved by giving short course and seminars of assessment procedures to existing teachers. Also specific guidelines must be given to schools so that there is uniformity as far as possible in the way the assessments are carried out. This is what the National Examination Council of Tanzania (NECTA) has been doing to curb the situation.

Hattie, et al 1984, in Australia investigated the relationship between the two measures of self - actualization. Personal Orientation Dimensions (POD) and Personal Orientation Inventory (POI) were administered to 94 high school students. The POD had been introduced as a refinement and extension of concepts of self - actualization which were first measured by POI. A canonical correlation analysis was carried between the two inventories. The results obtained revealed that there were relations between the two major POD and POI scales. These scales were Time-oriented and Time-competence and Core centredness and Inner directedness. Both (POD and POI) inventories had almost equal overlap with the other. It was concluded that there is little theoretical justification for development of the POD because it was not demonstrably better than the instrument it

intended to replace.

Hughes 1983, in Australia also carried out a study to determine the comparison of error-based and time-based learning measures as predictors of general intelligence. He examined two methodological issues involved in determining the relationship between learning and general intelligence. These issues were, the use of student learning strategies in a paired-associates learning task and the importance of time in the measurement of learning efficiency. The subjects used were 34 male and 46 female undergraduates who were divided in two groups (strategy instruction and no-strategy instruction). They were matched for intelligence on the Raven Standard Progressive Matrices. For the strategy-instruction group the use of an appropriate learning strategy correlated significantly with general intelligence scores, but not for the no-strategy instruction group thus the conclusion was that, the two error-based and time-based learning measures are related in prediction of general intelligence.

In another study, in Birmingham, Jackson and Brooks, 1986, investigated the relationship among the Medical College Admission Test (MCAT) reading subtest, Nelson - Denny Reading Test (NDRT) and Medical School Achievement test. They used a sample of 292 medical students. On comparing the predictive validities of the tests (MCAT and NDRT) for medical school achievement, they established that NDRT was related to MCAT reading subtest; though NDRT predicted well or was highly related to the medical school achievement test compared to MCAT; both being considered independently.

McCausland 1985, in New Zealand found out that there was a correlation between estimated and achieved marks of pupils in external examinations. He looked at the accuracy of teachers' predictions of pupils marks in school certificates using a sample of 1,440 secondary school students. This was a six year study. He established that the teachers' predictions correlated highly with the achieved marks of the pupils.

Mkandawire 1991, in Namibia, in his research paper pointed out the importance of continuous assessments. This was based on the various decisions made as a result of information obtained from them. Continuous assessments can be used for instructional decisions. Periodic testing helps monitor students' progress and identify areas where change maybe needed to modify instruction or otherwise. This facilitates student learning. Diagnostic tests can be developed for skill subjects like reading, arithmetic and spelling for they are used to identify specific areas of weaknesses or strengths in individual learners. The nature of the deficiencies would be ascertained and corrective measures taken such as provision of remedial work. The students, through continuous assessments would make sound decisions in areas of vocation and education. This was because they would have accurate self-concepts of themselves based on their own strengths and weaknesses obtained from the assessment data. The use made of the test would therefore help improve the quality of education given to the student and improve its effectiveness. He observed also that there was too much reliance on external one-short examination for final selection and certification which would not be adequate.

The study by Phelps and Ensor 1986, in Columbia investigated the concurrent validity of the WISC-R using deaf norms and the Hiskey - Nebraska. Test results of the Hiskey - Nebraska Test of Learning - Aptitude (H-NTLA) and an adaption of Weschler Intelligence Scale for Children - Revised (WISC-R) were obtained. Performance scales for the deaf children developed by R.J. Anderson and F.H. Sisco among 50 prelingually deaf children aged 6.8 - 16.3 years were used as the deaf norms. The H-NTLA was administered using the standard demonstration and/or Pantomime instructions, while a total communication directional set was used with the WISC-R. A Pearson - Product - Moment correlation of .913 between the two instruments demonstrated overwhelming commonality and supported the hypothesis that the WISC-R performance IQ may provide a convenient alternative to the H - NTLA when the score is computed using the Anderson and Sisco deaf norms and a total communication approach for administration.

Ryan 1986, in Leavework looked at the relationship between subtests. He investigated the validity of the combination of the Wechsler Adult Intelligent Scale (WAIS), Information, Arithmetic, Picture Completion and Block Design subtests (short form) as an estimate of WAIS - R (Revised) fullscale IQ (Intelligence Quotient). He found that the correlation between the short form and full scale IQ was highly significant ($P < .001$), that is, at .001 level of significance. A pairwise t-test between mean IQ's for the forms of the scales was significant at .005 level of significance. The average difference between the two was small. Therefore the short form WAIS was related to the WAIS full scale IQ.

According to a Study by Saigh and Khouri, 1983, in Beirut, scales measuring a common trait can be included in a battery in assessment of student's ability. He carried out a study to determine the concurrent validity of the Maths Anxiety Rating Scales for Adolescents (MARS-A) in relation to the Academic Achievement of Lebanese students. The MARS-A was administered to five classes (n=133) of Lebanese junior and senior high school students. The scores of the students were correlated with their semester grades in Physics, Chemistry and Maths as well as with their overall GPA. As the correlations accounted for moderate portions of the shared variance between the MARS-A and the aforementioned achievement measures it was suggested that the MARS-A should be included in the assessment batteries of Lebanese practitioners. The Pearson - product moment correlation formula was employed in the analysis and the correlations obtained ranged between $-.21$ to $-.60$.

According to the studies outlined above tests that are designed to measure a given trait bear a strong relationship with each other and can supplement each other.

2.2 Kenya Based studies

In Kenya the researches done are of the predictive nature. No studies thus far have been done to investigate concurrent validity of tests measuring a given common trait especially achievement. In addition the studies done have used scores from public examinations like Certificate of Primary Education, (CPE), Kenya Certificate of Education (KCE), and Kenya Advanced Certificate of Education (KACE). None of the studies have made use of the current public examinations like Kenya Certificate of Secondary Education (KCSE).

In the light of this, a few of the studies that have been carried out to investigate the predictive validity of tests will be briefly reviewed. A study by Bali, et al 1984, investigated the contribution of aptitude tests IDEA (Intelligence and Development Tests for East Africa) in predicting school performance. The result obtained showed that the aptitude tests had predictive validity, the correlations ranged between .54 to .74. In the same study CPE examination was found to be a good predictor of KCE examination. The correlation between the two was .62 which was significant.

Gatumu 1976, looked at the predictive validity of O-Level (KCE), A-Level(KACE) and Aptitude tests in relation to performance at the University of Nairobi in five faculties (Education, Architecture, Agriculture, Arts and Medicine). He found that A-Level had the highest correlation with university performance followed by O-Level and then Aptitude tests. The correlation coefficients were .31, .27 and .25 respectively with first year examinations. Correlations with third year examinations yielded correlations of .21, .25, and .19 respectively.

Using an example of primary level of education Ogula, 1991, in Kenya noted that there was need for continuous assessments to determine students' capabilities. Continuous assessment is an integral and essential part of the teaching-learning process. It makes possible for teachers to take non-cognitive aspects of performances into account. They can be standardized by KNEC on the basis of the final writing examinations.

He observed that with the importance associated with KCPE the education process in the upper primary classes was dominated by examinations which to a very large extent determine the methods of teaching. Many teachers did not use essay and short answer test formats because they did prepare children for KCPE multiple choice examinations. He noted that the KNEC pays little attention to affective and psychomotor objectives and does not measure the whole of what students have learned. There was need therefore to use continuous assessments to meet this need.

Another study was by Opinya 1980, who tried to determine the relationship between History in Mock and East African Certificate of Education (EACE) examinations administered in form four (Mock examination is common to all schools within a district). He used a sample of 8 secondary schools in Nairobi and test scores of the year 1976 - 1979. He found out that there was a high correlation between Mock and EACE history examinations. Some correlation coefficients were as high as .885 and he concluded that history EACE results could be predicted from History Mock results.

Ssalli 1986, also carried out a study using an aptitude test. The study was designed to measure educational potential with the Draw-A-Person (DAP) tests. She used a sample of 360 children in classes 1, 2 and 3 in four primary schools in Nairobi. The results revealed a high correlation that were significant at .05 level of significance. For example, DAP (man) test correlated with DAP (woman) test with correlation coefficient of .62, that between DAP (M) and DAP (S) was .63, where DAP (S) is DAP total score, and also for DAP (S) and DAP (W) ($r=0.63$). Therefore, the DAP test scores were good predictors of educational potential

among the children.

The studies in Kenya that have used achievement tests have investigated their predictive validity. Therefore, the present study is designed to investigate the concurrent validity of teacher-made tests at the secondary school level relative to the KCSE examination. Since from the studies discussed, tests measuring a common trait yield high correlations, indicating their relationship, the same was investigated in this study. The teacher-made tests and the KCSE examination are achievement tests designed to measure the students level of achievement. If this is so, it is theoretically true that they will yield high correlations with one another.

The subtests ought to share a great deal of common variance as indicated by the correlations obtained between them (Eaves and Simpson, 1983). This was also investigated between subtests of one another. This was to give empirical support of the concurrent validity of the test under investigation. The contribution of the subtests (subjects) towards the average grade was determined. The use of GPA results is a common practice instead of the individual subject grades. The assumption is that subjects make an equal contribution towards the aggregate score or grade, which is considered as a representative of the whole.

2.3 Statistical model

From the researches, Pearson-product moment correlation formula (r) has been used to investigate the concurrent validity of tests. This technique is used in data analysis in order to obtain the degree of association between the measurements (tests) and some other observable measurements (Lord and Novick,

1968). The formula r ; is

$$r = \frac{N\sum XY - \sum X \sum Y}{\sqrt{(N\sum X^2 - (N\sum X)^2)(N\sum Y^2 - (\sum Y)^2)}}$$

where,

\sum = indicates summation

X and Y - the two variables being correlated

r - the correlation coefficient.

(source: Lemke and Wiersma, 1976)

The use of this technique is an application of the correlation theory.

Correlation being the degree of relationship between the variables. A number of other techniques may be used to determine the degree of relationship between one variable and another. For example, Canonical Correlation analysis (Hattie, et al, 1984). In such an analysis two linear combinations are constructed and their correlations obtained. This technique is an extension of multiple regression. The first linear combination may be denoted by X_1, X_2, \dots, X_p and the second Y_1, Y_2, \dots, Y_q for p and q number of cases in each.

An extension is also made of the Pearson Product Moment analysis whereby a number of variables are compared instead of two. This yields multiple correlation analysis. In such a case a correlation is obtained between the test scores and the criterion in their best weighted combination.

Appropriate weights are found through pivotal condensation which are calculated in form of partial regression coefficient known as beta () coefficients. Other special cases of the Product Moment correlation are Point Biserial Formula and Biserial correlation. The Point - Biserial formula assumes that one variable is continuous while the other is dichotomized. While Biserial correlation is correlation between two variables which are continuous but one is artificially dichotomised by putting a cut-off point.

These correlation techniques yield correlations between variables that have been defined according to the purposes of a given study. Though they are widely used in validity researches they have one major disadvantage. The correlation techniques tend to break down complex behaviour into simpler components (Borg, 1967). The success in many of the complex behaviours that interest us can be attained in a number of different ways. For example, a study attempting to relate success of high school principals to specific individual variables might fail because of lack of any set of characteristics common to all successful principals.

Despite the disadvantage observed, the correlation technique has a number of advantages which outweigh the disadvantage hence its use in research.

1. The correlation technique permits one to measure a great number of variables and their inter-relationships simultaneously. In experimental technique this is not possible and it permits the manipulation of only a single variable resulting in a casual relationship being found. This is because a number of variables may be contributing causes of a particular behaviour pattern.

2. It provides information concerning the degree of relationship between the variables being studied. Therefore, one is able to understand the way in which variables are operating which cannot be found in other designs.
3. The correlation technique minimizes the high level of artificiality into research situations encountered in behavioral sciences when experimental technique is used. A far more realistic setting is permitted by the correlation technique.
4. The partial correlation allow for the statistical control of variables that one wishes to hold constant and does so without changing the field situation as for an experimental technique.

The correlation technique has been in use in most researches especially prediction and relationship studies. The study carried out used the correlation technique in order to establish the relationship between the teacher-made tests and the KCSE examinations. This was with an aim of testing the hypotheses outlined in the study.

In addition to the correlation analysis, regression analysis was also employed. This was because the contribution of the subjects towards the aggregate grade needed to be established. The result of which would help understand why GPA'S are more in operation than individual subject grades when making certain decisions regarding the individual or student.

The regression analysis technique was used to determine the contribution of the subject towards the aggregate grade. The linear regression function is of the form:-

$$Y^1 = a + b_{yx} X \text{ where}$$

\hat{Y}^1 - predicted score

a - an intercept constant

b - regression constant

X - score on test score (predictor)

This function can be extended to n variables such that the general form for n cases is

$$Y^1 = a + b_1 X + b_2 X + \dots + b_n X$$

where $X = (x_i - \bar{x}), \dots, (x_n - \bar{x})$ for

i = 1, 2, 3, ..., n variables

The regression technique for purpose of the study is modified whereby Y^1 represents the aggregate score and X the subtests. The b's give the contribution of each X and Y^1 .

The assumption made was that the variables followed a multivariate normal distribution.

The methods used to obtain the predictive validity of the tests were the same ones employed in obtaining the concurrent validity of the tests. Therefore, the study at hand used the regression analysis technique and Pearson - Product Moment correlation technique in analysis of the data. This was to help test the null hypotheses identified in the study.

The present study was therefore set to test the following null hypotheses.

1. There was no significant relationship between teacher-made tests and KCSE examination.
2. There was no significant relationship between subjects in each examination.

3. There was no significant contribution of each subject towards the average grade in a given examination.

This report is divided into various sections and is a description of early work for grades. The report begins with an overall review of the study, the population and sample, and outlines the population used, the sample selected and the variables used in the study. The section on variables gives a brief description of the variables used in the study. A detailed description of the instrument used is given, followed by an outline of the data collection method. The next section describes the methods of data analysis employed in the study and the research hypotheses tested.

1.1. Hypotheses

This was a correlational study which aimed at explaining the relationship between the mean correlation between teacher-made tests and the national examination (KCSI) was studied. The significant values of the teacher-made tests was established in order to answer further the hypotheses.

The number of 20 items in the questions used in the sample was 245 which were obtained from 1980-1981. The selected schools in Uasin Gishu district. The grades of these students in 1980 were compared in a set variables namely English, Mathematics, Kiswahili, History, Geography, Biology, Physical Science and Overall grade for the first term, since the KCSI examination. Data collection was done in four periods of March, April and May 1991 with the aim of a data

CHAPTER THREE

3.0 METHODOLOGY

This chapter is divided into various sections and a description of each will be given. The research design gives an overall review of the study. The population and sample section outlines the population used, the sample obtained and the sampling technique employed. The section on variables gives a brief description of the variables used in the study. A detailed description of the instruments used is given followed by an outline of the data collection method. The last section deals with the methods of data analysis employed in the study and the research hypotheses tested.

3.1 Research Design

This was a correlation study which aimed at exploring the uses of teacher made tests. The inter-correlation between teacher-made tests and the national examination (KCSE) was studied. The concurrent validity of the teacher-made tests was established using grades achieved in these examinations.

The number of the form four students used in the sample was 585 who were obtained from eight randomly selected schools in Uasin Gishu district. The grades of these students in 1990 were collected in eight variables namely English, Mathematics, Kiswahili, History, Geography, Biology, Physical Science and Overall grade for the first term, mock and KCSE examinations. Data collection was done in the months of March, April and May 1992 with the help of a data

collection sheet (appendix 1). At the end of the exercise data analysis was done using relevant techniques to the study.

3.2 Population and Sample

The population considered in the study was the form four students in the year 1990 in 48 secondary schools in Uasin Gishu district. The students were from both mixed and single sex schools (boys and girls). Since some of the schools were boarding, some students were from various parts of the country but for the day schools the majority of the students came from the district or surrounding regions.

The sample was obtained by stratified sampling technique - the strata were identified in terms of the categories considered in this study in which each school belonged as indicated in the table below.

Table 3.1 Sample used

School	Categories	Number of School	Number of Form Four Students
Secondary Schools	Boys	2	136
	Girls	2	147
	Mixed	4	302
Total		8	585

The table indicates the categories considered in the study, the number of schools considered obtained from each category and the number of form four students obtained from each category. A total of eight schools (Appendix 4) were selected with a total of 585 students. The eight schools were randomly selected from which all form four students were considered to make up for the sample. All the girls' and boys' schools were boarding and two of the mixed schools were day schools. The number of schools selected was in such a way that a balance was obtained across the categories considered. The students were of an average age range of 16 to 18 years.

The data collected for the sample was grade for all the eight variables that were common to all. The set up was in a field situation in order to obtain data that would be of practical use.

3.3 Variables

(a) School examination Variables

These were two sets of school examination results obtained from first term and mock examinations. For each examination seven subjects which were compulsory in the school syllabus in 1990 were considered and the overall grade. These seven subjects and the overall grade were referred to as variables for purposes of the study therefore, there were eight variables in total. The eight variables were the same for each examination. These were namely English, Mathematics, Kiswahili, History, Geography, Biology, Physical Sciences and Overall grade. Physical Science was a combination of the Physics and Chemistry

results in schools where the two subjects were done separately. For each variable the corresponding score or grade awarded by the teachers was obtained from the students' records files.

(b) Criterion Variables

The criterion variables were the KCSE examination results in seven subjects and the overall grade. These were eight in number and the same as the school examination variables. The variables therefore were English, Mathematics, Kiswahili, History, Geography, Biology, Physical Science and the Overall grade. The grade for each subject was collected which reflected the criterion for success after four years of study. The data needed for each student was obtained in the school's headteachers office.

3.4 Instruments

(a) School Examinations

These were examinations done within the school without supervision from the examination council. It was mainly a concern of the teachers and headteachers of the schools involved. Two kinds of school examination were used in these study - first term and Mock examinations.

(i) First term examinations

This was an internal examination constructed by individual subject teachers in each school. The examinations in all the subjects (Mathematics, English,

Kiswahili, Geography, Biology, History and Physical Science) were used. These examinations were supervised by the subjects' teachers or headteachers of each subject's department in each school. The aim of the examination was to measure the students level of achievement at the time of the examination in relation to what they had learnt during the term. Thus it measured the achievement of the objectives outlined by the teachers at the beginning of the term.

The examination consisted of one theory paper per subject except for science subjects which included a practical paper. It was administered at the end of first term during the fourth year of education and it covered all the subjects taught in school. The examination did not have the same number of test papers as the KCSE examination. The type of questions were mainly objective and essay type except for languages and sciences which included comprehension and interpretation question. They were similar to the KCSE examination questions. Marking and scoring was done by the subject teachers. The marks assigned to each student were in percentage and only for a few schools had teachers accorded the corresponding grades according to the scale used by KNEC (Appendix 3). The marks were then reported to the students and parents. A record of the results was retained in the school's student record file. Therefore, the records of this examination were available in the school headteacher's office.

(ii) Mock examination

This was an external examination done at the end of second term in the fourth year during the months of July/August by all schools in the district. The examination was constructed by a number of teachers from different schools who

formed subject panels. Each panel was responsible for constructing an examination in its subject area. The constructing exercise was coordinated by the District Education Officer (DEO) with the assistance of headteachers in the schools.

The examination consisted of two examination papers in each subject administered at different times like the KCSE examination. For example, Mathematics had two papers, I and II. The science subjects had the second paper as a practical paper. The test items paralleled those of the KCSE examination because they included essay, objective and comprehension type questions. The examination was administered during the month of July and the exercise lasted for about three weeks.

The marking of the examination was done by teachers from different schools within the district. Initial scoring was in percentages which was then converted to grades using a twelve point grading system (Appendix 3) in all the schools in the district. The marking process was continuously checked using a sample of papers marked by the heads of the examination subject panels to ensure that marks were allocated objectively to every paper marked. The results were all reported in grade form.

The aim of the examination was to measure the level of achievement of the students in second term during the fourth year of study. It was intended to be a reflection of one's possible performance in the KCSE examination. The examination also helped point out loopholes in the learning process or content covered to be filled.' The questions were set therefore in line with the course objectives and students were given time to prepare for the examination. Much of

the testing exercise was similar to that of the KCSE testing practise. The results were then released to all schools. The records, therefore, were available in the headteacher's office in each school sampled.

b) National Examination (KCSE)

This was a public examination done at the end of the fourth year of secondary education in the 8:4:4 system of education. Its prime purpose was selection and certification. This was selection into the university and other institutions of higher learning. Selection into the relevant courses was dictated by ones performance in the required subjects for each course of study. The subjects a student studied were chosen from four groups indicated in the syllabi (Appendix 2). The average grade received reflected a student's overall level of achievement in these subjects. The KCSE examination did not categorise candidates as having passed or failed in the examination and therefore every candidate received a certificate of achievement based on personal performance.

The study used the records of the KCSE examination done in 1990. In 1990 the minimum number of subjects one could sit for was ten but from 1993, the students would sit for a minimum of eight subjects and a maximum of ten subjects (Muya, 1991). The criterion for school success was achievement of grade C+ as the final average grade. The results were therefore reported in grade form. The grade C+ made one eligible for admission into a public university.

The 1990 KCSE examination was administered to all schools in the country. It was the second public examination administered in the secondary cycle of the 8:4:4 system of education. It was conducted by the Kenya National

Examination Council (KNEC). The construction of the examination was aided by the use of table of specifications in each subject to ensure that items included in the examination measured a given objective and a number of subject experts were involved. The examination was conducted between 22nd October and 19th November 1990. For subjects which involved project work like Agriculture, Woodwork and the like, the examinations were started earlier (Ministry of Education, 1990). Two papers were done in each subject, for the sciences, the second being a practical paper. During the administration process, supervisors, invigilators and security personnel from within the district ensured that it was conducted smoothly. Officers and assessors were sent by the examination council to monitor how the examination was being conducted. This was to avoid cheating during the taking of the examination. The examination papers were then sent back to the examination council for marking.

Marking of the examination was done by examiners for a period of about five weeks. Intermaker (rater) reliability was ensured so that marks obtained were valid and reliable. The examination was then processed for another one and a half months (2nd January to 11th February 1991). The processing was necessary as the marks were originally assigned as percentages and conversion to grades had to be done by the KNEC. This was done using the scaling system given (Appendix 3). After processing, the results were publicly announced and released to all schools. Therefore, the records of the results were obtained in the headteachers' office.

3.5 Data Collection

The data that was needed by the researcher was obtained in the schools' academic record files that were found in the schools visited. These records were kept as permanent academic record of students for future reference. The data collected was for the year 1990 in both the school examinations and the KCSE examinations in eight schools.

The data collection exercise lasted for a period of seven weeks. This was because the schools were far apart and only five schools had been visited by the end of first term. Therefore, for the remaining schools data was collected during the first three weeks of second term. Data collection was aided by use of a data collection sheet (Appendix 1). This helped the researcher keep record of data for each student in each subject considered for the three examinations (instruments). Use of the data collection sheet ensured that no omissions in each subject was done for each student. Only students with full records were considered in the study. This ensured that the same number of grades were collected for each subject and for each examinations.

The data obtained for the first term examinations from some of the schools was in percentage form. Therefore, using the format for grading used in each school, the percentages were converted to corresponding grades. The grades were then converted to the corresponding grade point (Appendix 3) for all the three examinations and across all eight variables. For easier management and handling of data the schools were assigned letters (A,-----,H) and the variables numbers (1,-----,8)). The examination were designated as Exam A, Exam B and Exam C

for first term, mock and KCSE examinations respectively.

3.6 Data Analysis

Using SPSS (Statistical Package for Social Sciences) Package descriptive analysis of data was done. The analysis included computation of:

- 1). The means of each school examination variable and each criterion variable.
- 2). The standard deviation for each variable in the three examinations. The standard deviation gives an overall measure of variability for a set of test scores hence characterizes a distribution of scores. That is, it is a measure of how the marks are spread around the mean. A small standard deviation indicates that marks are clustered around the mean and a large one indicates that the marks are widely spread from the mean.
- 3) Each variables' kurtosis and skewness were calculated. This was in order to obtain the distribution of each variable in relation to the normal distribution model.
- 4) The ranges of the values were computed for each variable and their variances. This enhanced the knowledge obtained in each of the variables for each examination.
- 5) The intercorrelation of each examination. These were the correlations between variables within a given examination. In addition correlations between the three examinations were obtained, that is between, first term, mock and KCSE examinations. These correlations and inter correlations coefficients obtained gave a knowledge of the degree of association between

the variables and the examinations. The correlations gave further knowledge on the concurrent validities of the examinations.

- 6) Finally regression equations were obtained for all the examinations. These were eleven in number. The independent variables were $X_1, X_2, X_3, X_4, X_5, X_6$ and X_7 namely English, Kiswahili, History, Geography, Mathematics, Biology and Physical Science in each examination. The dependent variable was X_8 (average grade) in each examination. These gave a knowledge of the contributions of each subject to the average grade.

The statistical significance of the results was obtained in order to help in the interpretation of the meaning of the correlation coefficients obtained. The interpretation was at 0.05 level of significance. This means that five in a hundred the statistical results will not be correct due to chance errors in sampling. The statistical significance of a correlation coefficient usually is different from zero at a given level of a significance. When a correlation is said to be significant it means that the coefficient is sufficiently high. This is such that we may be reasonably confident that some relation exists between the variables we have correlated. This helps determine the variable under the given situation.

The statistical significance of the correlation coefficients obtained between the variables was tested using the correlation coefficients statistical table. This was a two - tailed test.

CHAPTER FOUR

4.0 RESULTS : DATA ANALYSIS

The study-at hand was carried out to investigate the relationship between teacher-made tests (school examinations) and KCSE (National) examination. The results obtained were of the form four students of the year 1990. The null hypotheses tested were as follows:-

- (1) There was no significant correlation between subjects (variables) in each examination.
- (2) There was no significant correlation between the teacher-made tests and the KCSE examination.
- (3) There was no significant contribution of each subject towards the average grade in a given examination.

In order to test the above hypotheses the following results were obtained during the analysis of the data: Descriptive results, correlation coefficients and regression results of the examinations considered. Each of the results obtained will be described below.

4.1 Descriptive results of the variables in each examination

The initial analysis done was to obtain the descriptive statistics for each variable across all the three examinations. The results obtained indicate that the mean values across the examination did not vary. The mean values range between 2.07 to 6.42. The variance results obtained explain this further. At first, the descriptive results for the first term examination were obtained in all the eight

schools. The descriptive results were summed up together resulting in composite scores in each variable. Descriptive results of the composite scores were then obtained and reported. Analysis of the mock examination results of the combined scores in all the eight variables was also done. The same was done for the KCSE examination results.

The descriptive analysis involved obtaining the means of all the variables and the variances which gave information about the variability within the variables. The skewness and kurtosis obtained gave a description of the distribution of the scores in relation to the normal distribution curve. Finally the standard error of measurements for each variable were obtained. The results of the descriptive analysis for each examination will be considered separately as given below:

Table 4.1: Descriptive Statistics of the variables of first term examination
(N=585).

Variables	Mean	Variance	Skewness	Kurtosis	Std. Error
English	5.17	5.61	.05	-.63	.10
Kiswahili	5.14	7.13	.13	-.87	.11
History	6.03	6.62	-.20	-.46	.11
Geography	5.47	6.15	-.19	-.80	.10
Mathematics	2.07	4.21	2.21	4.47	.09
Biology	4.10	7.08	.65	-0.32	.11
Physical Science	3.07	5.11	1.09	.41	.09
Average Grade	4.47	3.87	.09	-.51	.08

Table 4.1 above gives a descriptive analysis of the first term examination results. These were results from eight schools. History had the highest mean (6.03) while Mathematics had the lowest mean (2.07). All the variables except the Average Grade had a high variance which was greater than four. Therefore, the variables had a high variability within them in the grade points. History and Geography had a negatively skewed distribution while all the rest had a positive skewness. English had the least skewed distribution (0.05). The range of the standard error of measurement of 0.08 to 0.11 was obtained for the variables.

Table 4.2: Descriptive statistics of the variables of mock examination

(N=585).

Variables	Mean	Variance	Skewness	Kurtosis	Std. Error
English	4.62	4.26	-.24	-.11	.09
Kiswahili	5.01	3.94	-.32	.22	.08
History	5.47	3.51	-.40	.28	.08
Geography	5.63	3.57	-.06	.34	.08
Mathematics	2.08	3.24	1.78	2.87	.07
Biology	4.05	4.94	.40	-.17	.09
Physical Science	3.10	3.80	.74	-.26	.08
Average Grade	4.61	2.67	-.42	-.04	.07

The descriptive results of the mock examination results were presented in Table 4.2 above. None of the variables had a mean above the expected value (6.00) on the twelve - point scale. The highest mean was that of Geography (5.63) and the lowest mean was obtained in Mathematics (2.08). All the variables had a high variability in the results obtained as indicated by the variances, the highest variance was that of Biology (4.94) while the Average Grade had the lowest variance (2.67). Mathematics had the highest positive skewness (1.78) and Geography had the lowest positive skewness (0.06). The range of the the standard error of measurement of 0.08 to 0.12 in all the variables was obtained.

Table 4.3 : Descriptive statistics of KCSE examination results(N=585).

Variables	Mean	Variance	Skewness	Kurtosis	Std. Error
English	4.46	4.82	1.05	.26	.09
Kiswahili	6.42	6.56	.16	-.79	.11
History	5.88	4.00	-.06	-.74	.08
Geography	5.48	5.10	.28	-.89	.09
Mathematics	3.73	7.90	.94	-.09	.12
Biology	3.74	3.94	1.27	1.32	.08
Physical Science	3.87	4.07	1.16	.69	.08
Average Grade	4.97	3.46	.40	-.45	.08

The KCSE examination descriptive results were presented in the above

Table 4.3. As observed, only Kiswahili had a mean grade point (6.42) above the expected value (6.00) on the twelve - point scale. The lowest mean value was that of Mathematics (3.73). The variability of all the variables in KCSE examination results was high even as compared to first term and mock examination results.

The highest value was that of Mathematics (7.90) while the lowest was that of the Average Grade (3.46). The distribution of all the variables differed as indicated by the skewness of each variable. History had a negative skewness while the rest had a positive skewness. The range of the standard error of measurement obtained was .09 to .12. The standard error of measurement gives the accuracy with which the examination measures a given trait.

4.2 Inter-correlation between the variables in each examination

The relationship between the variables within an examination was investigated. This was in order to obtain more information about the relationship between the subjects in a given examination. The relationships between the eight variables were obtained for the first term, mock and KCSE examinations. For each of these examinations, the results have been tabulated below. Tables 4.4, 4.5 and 4.6 give the results of the inter-correlations of the eight subjects of the first term, mock and KCSE examination results respectively.

The subjects were positively related and all the correlation coefficients obtained for the results in the three examinations were statistically significant.

Table 4.4: Inter-correlation of first term examination results(N=585).

Variables	1	2	3	4	5	6	7	8
1.English	-							
2.Kiswahili	.41*	-						
3. History	.41*	.36*	-					
4. Geography	.27*	.42*	.27*	-				
5.Mathematics	.09*	.09*	.24*	.08	-			
6. Biology	.16*	.24*	.42*	.38*	.33*	-		
7. Physical Science	.24*	.30*	.42*	.46*	.48*	.61*	-	
8. Average Grade	.48*	.60*	.62*	.63*	.37*	.63*	.74*	-

Note: *p < .05

As can be expected, there was a strong and positive relationship between the school subjects. Low positive relationships were obtained between Mathematics and the rest. The relationship between Mathematics and Geography (0.08) only is not statistically significant.

Table 4.5 : Intercorrelation of Mock examination results(N=585).

Variables	1	2	3	4	5	6	7	8
1. English	-							
2. Kiswahili	.62*	-						
3. History	.58*	.51*	-					
4. Geography	.48*	.49*	.62*	-				
5. Mathematics	.41*	.33*	.37*	.31*	-			
6. Biology	.54*	.51*	.56*	.62*	.52*	-		
7. Physical Science	.56*	.45*	.55*	.55*	.69*	.72*	-	
8. Average Grade	.61*	.65*	.62*	.72*	.53*	.81*	.62*	-

Note: The variables have a positive and significant correlation coefficients as indicated by the asterisk (*).

Mathematics has a low relationship with the other variables though the observed correlation coefficients were statistically significant. None of the variables had a relationship that was not statistically significant. The correlation coefficients range between .31 to .81. High correlation coefficients were observed between the Average Grade and all the other variables. Therefore, the Average

Grade was highly related to all the other variables.

Table 4.6: Intercorrelations of KCSE examination results

Variables	1	2	3	4	5	6	7	8
1. English	-							
2. Kiswahili	.58*	-						
3. History	.53*	.49*	-					
4. Geography	.55*	.55*	.67*	-				
5. Mathematics	.53*	.38*	.45*	.46*	-			
6. Biology	.49*	.46*	.44*	.61*	.62*	-		
7. Physical Science	.57*	.46*	.50*	.54*	.77*	.71*	-	
8. Average Grade	.65*	.64*	.60*	.74*	.68*	.82*	.75*	-

Note: The asterisk (*) indicates that the correlation coefficients are statistically significant.

As can be observed from the Table 4.6 above, all the variables had positive relationships. The observed correlation coefficients range between .44 and .82.

The Average Grade had high and positive relationships with all the other variables compared to the observed correlation coefficients of the others. The highest relationship being that between the Average Grade and Biology (.82)

The Intercorrelation results obtained indicate that the results of mock examinations and KCSE examinations were comparable. This was because all the correlation coefficients obtained in both cases were high and statistically

significant. In both cases also the Average Grade had the highest relationship with the other subjects.

The three examinations considered gave statistically significant Intercorrelations between the subjects and the Average Grades. The first term examination Intercorrelations were lower as compared to those of the mock and KCSE examinations.

4.3 Analysis of correlation results between the three examinations

This section gives the results obtained when the interrelations between the three examinations were investigated. The results were all tabulated. A correlation table reports how much a particular test/examination had in common with each other in turn. The correlation coefficients therefore, give an indication of the degree of relation between tests in terms of a given attribute (achievement). The degree of the relationships between first term, mock and KCSE examinations' results were obtained by determining the correlation coefficient between the seven school subjects and the Average Grade considered in the study.

In order to test for the null hypotheses, one and two, that is, to test for significance between the inter correlations between the variables and that between the examinations, the correlation coefficients were transformed using Fisher's Z-transformation. The results will be dealt with in different sections that follow:

Relationship between the variables in the examinations.

This is a section that deals with the relationships between the various

school subjects and the Average Grades in different examinations. The relationship between each variable in first term examination and each variable in both mock examination and KCSE examination results was determined.

Correlation coefficients were obtained for each variable in mock examination and KCSE examinations. The results obtained are all presented in table form for easier interpretations (Tables 4.7, 4.8, 4.9). To test the hypothesis $H_0: r_{xy} = 0$ where x and y are variables (subjects) used, all the values were transformed to Z_r values through Fisher's Z - transformation of r. The level of significance was .05. The results of Fisher's Z-transformation of r - values for the various variables were as follows:

(a) The table below gives the transformed values of the correlation coefficients between first term and mock examinations results:-

Table 4.7 Z_r values between first term and mock examination

mock examination results	1.Eng	2.Kisw	3.Hist	4.Geog.	5.Math	4.Geog.	6.Bio.	8.Av.Grade
First term/Mock exam results								
Variables	1	2	3	4	5	6	7	8
English	.266*	.288*	.213*	.182*	.121*	.245*	.244*	.255
Kiswahili	.182*	.172*	.172*	.245*	.070*	.266*	.161*	.255*
History	.255*	.141*	.224*	.141*	.172*	.192*	.224*	.223*
Geography	-.100*	-.040	.020	.213*	.110*	-.110*	-.100	.070
Mathematics	.010	.030	.030	-.100*	.020	-.040	.020	-.020
Biology	-.080	-.040	-.010	.080	.010	-.030	.040	.010
Physical Science	.060	.080	.070	.080	.020	.030	.080	.070
Average Grade	.060	.080	.090*	.131*	.000	.121*	.100*	.131*

* - Indicates that the correlation coefficients obtained are statistically significant.

There was a great difference between the variables in terms of their relationship with one another. The first term results in English, Kiswahili and History only had significant relationships with all the variables in the mock examination results. Biology and Physical science first term results did not have significant relationships with any of the mock examination results in all the variables. The overall Average Grade of first term examination was significantly

related to History, Geography and Biology only with Z_r values of .090, .131 and .121 respectively. In general, few variables were significantly related with one another.

Except for the science subjects (Mathematics, Biology and physical Science) some subjects yielded high and significant correlations with each other in the two examinations compared to the correlations between different subjects. The values reveal that the science subjects in first term examination were not well related with both the arts and science subjects in the mock examinations.

This could be due to sampling of items and testing skills used. The mock examination also had different number of papers per subject compared to first term examination which had only one paper per subject. Little training in test-taking skills would possibly have taken place by the end of the first term compared to second term. Hence the low and significant correlations obtained in the science subjects as observed earlier.

(b) The table below gives the transformed values of the correlation coefficients between first term and KCSE examination results.

Table 4.8: Z_r Values between first term and KCSE examinations(N=585).

KCSE examination results	1.Eng.	2.Kisw.	3.Hist.	4.Geog.	5.Maths	6.Bio	7.Phy/Sc.	8.Av.Grade
First term/KCSE results								
Variables	1	2	3	4	5	6	7	8
English	.245	.172*	.182*	.234*	.255*	.245*	.299*	.266*
Kiswahili	.070	.245*	.182*	.234*	.172*	.245*	.255*	.234*
History	.266*	.255*	.310*	.343*	.192*	.213*	.224*	.299*
Geography	-.161*	-.000	.100*	.131*	.050	-.121*	.060	.050
Mathematics	.010	.030	.090*	-.080*	.040	-.020	.060	-.030
Biology	-.030	-.040	-.131*	.121*	.060	-.000	.000	.030
Physical Science	.020	.030	.072	.161	.070	.110	.141*	.121*
Average Grade	.050	.121*	.213*	.245*	.110*	.203*	.192*	.172*

*- Indicates that the correlation coefficients obtained are statistically significant.

There was a weak relationship between the first term and KCSE examination results. The correlations coefficient range between .002 and .290 whose Z_r values were .000 and .310 respectively. The significant values were those between first term English and History results and all KCSE variables only. All first term examination results in the variables were significantly related to KCSE History and Geography variables only. First term Mathematics was only significantly related to KCSE History results. The Average Grade in first term examination was significantly related to all the KCSE variables except English. Therefore, English,

History and Average Grade results in first term examination were well related with the KCSE examination results in all the variables.

The same trend was observed in the relationship between the first term examination and KCSE examination as in the relationship between first term and mock examinations. Low and insignificant correlations were obtained between first term Mathematics and Biology examination and KCSE Mathematics, Biology, Physical Science and Average Grade. Low but significant correlations were obtained between first term Geography and Physical Science with all the KCSE examination subjects. The trend may have resulted from the differences in the number of examination papers per subject in the first term and KCSE examinations. This may also result from the examination being set by individual teachers who may not have been well equipped with the appropriate testing skills. The students test - taking skills may also be questionable at this stage.

(c) The results of the transformed correlation coefficient values between mock and KCSE examinations are given in the table below:

Table 4.9 Z_r Values between mock and KCSE examination.

KCSE examination results	1.Eng.	2.Kisw.	3.Hist.	4.Geog.	5.Math.	6.Bio.	7.Phy/Sc.	8.Av. Grade
Mock term/KCSE results								
Variables	1	2	3	4	5	6	7	8
English	.497*	.424*	.424*	.436*	.310*	.377*	.365*	.436*
Kiswahili	.310*	.400*	.299*	.332*	.310*	.255*	.245*	.400*
History	.424*	.400*	.424*	.400*	.343*	.377*	.343*	.412*
Geography	-.234*	-.343*	.354*	.472*	.388*	-.321*	.234*	.472*
Mathematics	.377*	.245*	.266*	.234*	.266*	-.299*	.299*	-.354*
Biology	-.299*	.343*	.299*	.377*	.563*	.472*	.472*	.618*
Physical Science	.400*	.332*	.332*	.332*	.400*	.436*	.436*	.497*
Average Grade	.365*	.424*	.343*	.523*	.590*	.354*	.354*	.758*

* - Indicates significant correlation coefficients.

The results presented in Table 4.9 above support the fact that examinations testing the same attribute and whose items are drawn from the same content areas to measure similar objectives are related. The transformed Z_r -values were all statistically significant and they range between .234 and .758. This was for all the variables in both the examinations. The highest relationship was that between the Average Grades in both the examinations (Z_r - .758). The lowest relationship was obtained between Mathematics in mock examination and the variables in KCSE

examination results. The mock examination was therefore well related with the KCSE examination results for the variables considered in the study.

The results obtained in Table 4.9 reveal further that the mock examination was well related to the KCSE examination. This points out that the mock examination was better placed in measuring the level of achievement of the students compared to the first term examination. Sampling of test items by different teachers results in a good examination that is both reliable and valid. The two examinations had the same number of examination papers per subject which helped in adequate sampling of the content.

The results of the correlations between the examinations in Tables 4.7, 4.8 and 4.9 reveal that the same subjects in different examinations in general had high and significant correlations compared to the correlations observed in different subjects. This was with just a few exceptions. Such high relationship may result from the fact that the same content area was examined. Also the format of the examinations may be the same. The exceptional ones may point out to differences in testing skills among other testing techniques required at the face value.

Relationship between the examinations

The following null hypotheses will be tested in this section:-

H_0 :

- (i) The product - moment correlation coefficient between first term (T₁) and KCSE (N) examination results is equal to zero ($r_{T_1N} = 0$).
- (ii) The product - moment correlation coefficient between mock (T₂) and KCSE (N) examination results is equal to zero ($r_{T_2N} = 0$).

- (ii) The product - moment correlation coefficient between mock (T_2) and KCSE (N) examination results is equal to zero ($r_{T_2N} = 0$).
- (iii) The product - moment correlation coefficient between first term and mock examinations is equal to zero ($r_{T_1T_2} = 0$)

The relationship between similar subjects in each examination will be considered.

To test the above null hypotheses, the correlation coefficient (r) values were transformed by means of Fisher's Z-transformation of r into Z_r . The test statistic for testing the null hypotheses is:-

$$Z = \frac{Z_r - 0}{1/n - 3}$$

Where,

Z_r - the transformed value

n - sample size

At $\alpha = .05$, the critical values of Z were -1.96 and $+1.96$. If a Z -value exceeded 1.96 ; H_0 was rejected hence the given Z -value was significant and also if it was below -1.96 .

The results obtained after the Fisher's Z - transformation of r - values are presented in the table below.

Table 4.10: Relationship between the examinations(N=585).

Variables	Z_{rT1N}	Z_{rT1T2}	Z_{rT2N}
English	.245*	.266*	.497*
Kiswahili	.245*	.174*	.400*
History	.304*	.224*	.424*
Geography	.131*	.213*	.472*
Mathematics	.038	.020	.299*
Biology	-.005	-.030	.563*
Physical Science	.141*	.080	.485*
Average Grade	.182*	.131*	.758*

T1 - First term examination results

T2 - MOCK examination results

N - KCSE examination results

Note (*), indicates significant Z_r values at $\alpha=0.05$ level of significance.

From Table 4.10 above the following can be observed.

- (i) The first term examinations in English, Kiswahili, History, Geography, Physical Science and the Average Grade were significantly correlated to the same in the KCSE examination results. Hence, these subjects in first term examination were well related with the same subjects in the KCSE examination results. No significantly different correlation coefficients from zero were obtained for Maths and Biology. Thus the first term

examinations in Maths and Biology were not related with the same in KCSE examinations. Only Biology results were negatively correlated in both first term and KCSE examinations.

- (ii) The correlation coefficients obtained between mock and KCSE examinations reveal that all the subjects results in KCSE examination were well related with the mock examination results in the same subjects. The correlation coefficients were all significantly related. The highest correlation coefficient was that between the Average Grades with a Z_r - value of .758 and the lowest was that between the Mathematics results with a Z_r - value of .299. This implied that the overall results in mock examination was related to the overall results in KCSE examination.
- (iii) The first term examination results in English, Kiswahili, History and Geography were well related with the same subjects in mock examination. The correlation coefficients obtained for Mathematics, Biology and Physical Science were not significantly different from zero. This implied that Mathematics, Biology and Physical Science results in mock examination were not well significantly related with the results of the same subject in first term examinations.

The first term examination results correlated well with the KCSE examination. This was to a larger extent than it did for the mock examination. This was inferred from the number of subjects correlated with significantly in each case and the size of the correlation coefficients obtained also. Physical Science was not significantly correlated with the Physical Science in first term examination

in the case of mock examination, yet it was well correlated for in the case of KCSE examination.

The significance of the relationship between the examinations

The section above, had indicated that the correlation coefficients obtained by relating any two of the three examinations were different. In this section the statistical significance of this difference was investigated. For this to be established the following null hypotheses were tested:-

H_0 :

- (i) The correlation coefficients between first term and KCSE examinations were equal to those between first term and mock examinations.
- (ii) The correlation coefficients between mock and KCSE examinations were equal to those between the first and the KCSE examinations.
- (iii) The correlation coefficients between mock examinations and KCSE examinations were equal to those between first term and mock examinations.

The test statistic for the testing the null hypotheses is:

$$Z = \frac{\sqrt{n}(r_{xy} - r_{xz})}{\sqrt{(1-r_{xy}^2)^2 + (1-r_{xz}^2)^2 - 2r_{yz}^3 - [(2r_{yz} - r_{xy}r_{xz})(1-r_{xy}^2 - r_{xz}^2 - r_{yz}^2)]}}$$

Where; X - first term examination (T1)

Y - mock examination (T2)

Z - KCSE examination (N)

n - sample size.

The α - level of significance was 0.05.

The hypotheses were rejected if Z-value exceeded 1.96 and if lower than -1.96; hence the correlation coefficients would be considered statistically insignificant.

The Z-values obtained for the test statistic were presented in the table below:-

Table 4.11: Z-values of equality of correlation coefficients between the examinations(N=585).

Hypothesis variables	$r_{NT2} = r_{T1T2}$	$r_{NT2} = r_{NT1}$	$r_{NT1} = r_{T1T2}$
English	4.372*	4.851*	0.485
Kiswahili	4.408*	2.843*	1.566
History	3.977*	2.130*	1.853
Geography	4.644*	6.509*	1.867
Mathematics	4.908*	4.495*	0.406
Biology	10.480*	9.839*	0.376
Physical Science	7.485*	6.079*	1.396
Average Grade	10.870*	10.120*	1.450

Key

r_{NT2} - Correlation coefficient between mock and KCSE examinations.

r_{NT1} - Correlation coefficient between first term and KCSE examinations

r_{T1T2} - Correlation coefficient between first term and mock examinations.

* - Indicates statistically significant correlation coefficients.

The results obtained reveal the following :-

(i) The correlation coefficients obtained between mock and KCSE examinations in all the subjects (variables) considered in the study were not equal to the correlation coefficients obtained between first term and mock examinations in all the subjects. This was because the Z-values exceed the required critical value of 1.96. Thus the hypothesis that the correlation coefficients between the examinations being equal with respect to being related to mock examination was rejected. Therefore, the accuracy with which KCSE examination can correlate with mock examination was not the same as the accuracy with which first term examination can correlate with mock examination. This was strongly so for Biology ($Z = 10.48$) and Average Grade ($Z = 10.87$). The Z -values range between 3.977 and 10.87.

(ii) The Z - values obtained for the $H_0: r_{NT2} = r_{NT1}$; exceeded the required critical value of 1.96. This implied that the hypothesis could be rejected. The hypothesis that the first term and mock examinations were alike with respect to being related with the KCSE examination was rejected. Therefore, the examinations were not alike. This was in terms of the accuracy with which first term examination in all the subjects considered were related with KCSE examination in the same subject. This was not the same as the accuracy with which mock examination results were related KCSE examinations results. The highest Z - values were for Biology (9.839) and the Average Grade (10.120). The Z - values range between 2.13 to 10.12.

(iii) The $H_0 : r_{NT1} = r_{T1T2}$ was supported by the Z -values obtained. The Z -values fall below the required critical value of 1.96. They range between 0.406 and 1.867. Therefore, the hypothesis that KCSE examinations and mock examinations were alike in the degree of relationship with first term examination could not be rejected. The accuracy with which KCSE and mock examinations were related with first term examinations or vice versa was the same.

The three hypothesis tested point out that the first term and mock examinations were related with the KCSE examinations results with different degrees of accuracy. The correlation coefficients from the first term and mock examination may not give the same information about the KCSE examination results. The mock examination was well related with the KCSE examination to a greater degree than the first examination.

Since the accuracy of the relationship between KCSE and first term and mock and first term examinations was the same. The mock and the KCSE examinations were well related and measure the level of achievement of a student to the same degree. First term examination may not represent a student's mastery of content or level of achievement due to the fact that they were set by individual teachers who may not have had adequate testing skills. The mock examination was comparable to the KCSE examination since the format of the examination and the number of examination papers was the same as the National Examination which was not the case for the first term examination. Not over-looking the fact that less content was covered by the end of the first term as compared to the time the mock examination was done.

4.4 Regression analysis results.

In this section, the results of multiple regression analysis will be given. Each examination was considered separately with an aim of determining the contribution of the seven subjects to the Average Grade in each. This analysis helped investigate whether the same variable(s) made a higher or lower contribution in the three examinations considered in the study.

Regression analysis yielded the following regression equations:-

1) Regression equations for the first term examination results.

where,

X_1 Eng. X_2 Kisw. X_3 Hist. X_4 Geog.

X_5 Math. X_6 Bio. X_7 Phy/Sc X_8 Av. Grade.

Independent variables are X_1 X_7

Dependent Variable is X_8

$$(i) \quad X_8 = 0.14 + 0.008X_1 + 0.23X_2 + 0.18X_3 + 0.08X_4 + 0.05X_5 + 0.13X_6 + 0.32X_7.$$

The coefficient of determination R^2 was 0.80. The regression equation above was for all the subjects in the study. The highest contribution was that of Physical Science with a coefficient of 0.32. The lowest being that of English(.008).

$$(ii) \quad X_8 = 0.51 - 0.027X_1 + 0.28X_2 + 0.33X_3 + 0.13X_4.$$

The value of R^2 was 0.60. The regression equation above was for arts subjects namely English, Kiswahili, History and Geography. History had a high

weight (0.33). Among the arts subjects, History had a higher contribution towards the Average Grade.

$$(iii) \quad X_8 = 2.105 + 0.011X_5 + 0.21X_6 + 0.48X_7$$

The value of R^2 was 0.59. The regression equation above gives the contribution of the science subjects namely Mathematics, Biology and Physical Science towards the Average Grade. The highest coefficient observed was for Physical Science subject (0.48)

2). Regression equations for mock examination results.

$$(i) \quad X_8 = 0.17 + 0.13X_1 + 0.16X_2 + 0.17X_3 + 0.15X_4 + 0.09X_5 + 0.14X_6 + 0.15X_7.$$

The value of R^2 is 0.88. All the seven subjects were considered. Their contribution towards the Average Grade was obtained as indicated by the beta weights determined. Physical Science had the highest contribution with a coefficient of 0.15. This observation was in agreement with the high correlation between the Average Grade and Physical Science

$$(ii) \quad X_8 = -0.15 + 0.23X_1 + 0.19X_2 + 0.26X_3 + 0.23X_4.$$

The value of R^2 was 0.79. The subjects considered were English, Kiswahili, History and Geography which were art subjects. The one with the highest contribution towards the Average Grade was History with a coefficient of 0.26.

$$(iii) \quad X_8 = 2.05 + 0.08X_5 + 0.34X_6 + 0.33X_7$$

R^2 was 0.72. The subjects included in the regression equation were Mathematics, Biology, and Physical Science. Biology had the highest contribution towards the Average Grade with a coefficient of 0.34 associated with

it.

3) Regression equations for KCSE examination results.

$$(i) \quad X_8 = 0.20 + 0.16X_1 + 0.12X_2 + 0.12X_3 + 0.12X_4 + 0.12X_5 + 0.15X_6 + 0.20X_7.$$

R^2 is 0.91. The regression equation for the seven variables reveals that physical science has the highest contribution (0.20).

$$(ii) \quad X_8 = 0.18 + 0.30X_1 + 0.15X_2 + 0.22X_3 + 0.22X_4$$

The value of R^2 is 0.79. The subjects included were the art subjects namely English, Kiswahili, History and Geography. Amongst these English had the highest contribution with a coefficient of 0.30.

$$(iii) \quad X_8 = 1.84 + -.15X_5 + 0.32X_6 + 0.35X_7$$

R^2 is 0.75. The subjects included in the regression equation were Mathematics, Biology, and Physical Science. Physical Science with a coefficient of 0.35 had the contribution towards the Average Grade amongst the three.

Statistical Significance of R^2 .

In order to determine the statistical significance of the coefficient of determination (R^2) obtained in the previous regression equations an F-test was employed. The F-test served as the statistical test for R^2 . The F-tables were used for rejection or acceptance of the hypothesis at .05 level of significance. The null hypothesis tested was:-

H_0 : The R^2 values were not significantly different from zero.

The test statistic used was:-

$$F = \frac{R^2/K}{(1-R^2)/N-K-1}$$

Where;

K - number of variables

N- sample size

R2 - coefficient of determination for each regression considered.

The test statistical yielded the following results of F-values.

Table 4.12: F-Values for the Examination

Variables	First term	Second term	K C S E
X ₁ , X ₂ , X ₃ , X ₄	216.7*	552.0*	488.5*
X ₅ , X ₆ , X ₇	280.9*	493.8*	589.2*
X ₁ , X ₂ , X ₃ , X ₄ , X ₅ , X ₆ , X ₇	324.8*	589.9*	810.9*
X ₁ - Eng. X ₃ - Hist. X ₅ - Math. X ₇ Phy/Sc X ₂ - Kisw. X ₄ - Geog. X ₆ - Bio.			

The F-values obtained reveal that the R² values that resulted from regression analysis were all statistically significant.

The coefficients obtained for each variable were significant hence indicate the degree of contribution of each school subject towards the Average Grade.

4.5 Summary

Data analysis of the three examinations revealed the following;

1. The intercorrelations of first term examination were statistically significant. Low but significant correlations were observed between Mathematics and the other variables. The correlations ranged between 0.08 to 0.74. The inter correlations of mock examination revealed statistically significant relationship. The correlations were high and ranged between .31 to .81. All the correlations were significant. High and significant correlations were also observed among the subjects in the KCSE examinations. All the correlations were significant and ranged between .38 to .82. In all the three examinations the correlations between Mathematics and the other subjects were low though significant and the correlations between the Average Grade and the others were significant and were the highest. The highest being that between the Average Grade and Physical Science.
2. The inter-examination correlations of the three examinations obtained revealed that; the relationship between first term and mock and KCSE examinations was low but significant in the art subjects. The science subjects in first term examination did not yield significant correlations with KCSE examinations in all the subjects. Mathematics and Biology yielded non-significant results. The relationship between first term and mock examination yielded similar results. Mathematics, Biology and Physical Science yielded non-significant correlations.

This information was further verified by the test of significance of the correlations obtained. The Z_r values obtained between first term variables and those of mock and KCSE variables were low. Only those of Mathematics and Biology were not significant in both cases.

The relationship between mock examination and KCSE examination was highly significant in all the variables. The highest correlation being that between the Average Grades (0.758) in the two examinations. This high relationship was further supported by the Z_r values obtained between the two examinations (Table 4.9). The results indicate that there was much in common between the mock and KCSE examinations both in terms of sampling of content and the format of the examination paper.

The low relationship between the first term and mock examination may reflect the importance attached to a given examination.

3. The contribution of each subject to the average grade was studied and the results revealed that in the first term examination, the Physical Science subject had the highest contribution and the lowest was English. In the mock examination, Physical Science also had the highest contribution while the Mathematics had the lowest. The highest contribution was that of Physical Science in the KCSE examination and Kiswahili, History, Geography and Mathematics had the lowest. Each subject significantly contributed to the average grade. The correlations and regression analysis reveal that Physical Science in general was highly related to the Average Grade. Physical Science results may therefore give as much information

about a student compared to the other subjects. At the same time, the mock examination may tell no more than first term examination in relation to the KCSE examination.

The study was a correlational study which aimed at exploring the uses of teacher-made tests. The main objective of the study was to determine the relationship between teacher-made tests and the national examination. A sample of 525 form four students was used of the year 1991. The examination was conducted very late into the year of KCSE examinations.

The analysis was done using Pearson-Product Moment correlation coefficient. The results are as follows:

1. There was a significant correlation between students' first term examinations and the mock examination.

2. There was no significant correlation between teacher-made tests and the KCSE examinations.

3. There was a significant distribution of the subjects towards the mock examination in each examination.

The inter-relationships which obtained were tested using a Z test to determine the significance of the results. The results were written out in Chapter Five and their interpretation and the distribution of the beta weights are also given in the Appendix. The results are in the following discussions.

1. Inter-relationship between the examinations

The inter-correlations between the variables in the first term examination were high, the range was between 0.46 to 0.61 except for Mathematics (0.09 to 0.24). The Average Grade had the highest inter-correlation with the school

CHAPTER FIVE

5.0 DISCUSSION

This study was a correlation study which aimed at exploring the uses of teacher-made tests. The main objective of the study was to determine the relationship between teacher-made tests and the national examination. A sample of 585 form four students was used of the year 1990. The examinations considered were first term, mock and KCSE examinations.

The analysis of data was done using Pearson-Product Moment correlation coefficient formula. The hypotheses tested were:-

1. There was no significant correlation between subjects i in the three examinations.
2. There was no significant correlation between teacher-made tests and the KCSE examinations.
3. There was no significant contribution of the subjects towards the Average Grade in each examination.

The correlation coefficients obtained were tested using a Z-test for their statistical significance. The results were written out in Chapter Four and their interpretations. The significance of the beta weights obtained were tested using an F-test. The results led to the following discussions.

1. Interrelationship between the variables.

The inter-correlations between the variables in the first term examination were high, the range was between 0.16 to 0.61 except for Mathematics (0.09 to 0.24). The Average Grade had the highest inter-correlation with the school

subjects (0.37 to 0.74). This indicated that the subjects were well related and hence measured the content areas concerned. The highest correlation obtained was that between the Average Grade and Physical Science subjects (.74). The lowest were the correlations between Mathematics and English (0.09), Kiswahili (0.09) and Geography (0.08). Mathematics had higher correlations with the Science subjects namely Biology (0.33) and Physical Science (0.48) hence highly related with them as compared to the Art subjects.

The results of the mock examination obtained were higher than those of the first term examination. The correlations had a range between 0.31 to 0.81. The correlations were high and significant. The same trend was observed whereby the Average Grade had the highest inter-correlations with the subjects (range 0.53 to 0.81). The highest correlation 0.81 was that between Average Grade and Biology. Mathematics had significant correlations with the Art subjects though low (range 0.31 to 0.41) namely English, Kiswahili, History and Geography. Higher correlations were observed between Mathematics and the Science subjects; Biology (0.52) and Physical Science (0.69).

The subjects within the mock examination were well related, hence measure educationally relevant concepts. They are achievement measures and the examination for each school subject measured the desired objective. A higher relationship within the Mock examination as compared to the first term examination may seem to suggest that there was a difference between the examinations maybe due to the way the two examinations were prepared and scored. First term examination was prepared and scored by an individual subject teacher while a number of teachers prepared and scored the mock examination

within the area of study.

The inter-correlations obtained between the subjects in the KCSE examinations ranged between 0.38 to 0.82. The subjects were highly related with one another as indicated by these correlations. The Average Grade had the highest correlations with the subjects (range 0.60 to 0.82). The highest correlation was that between Biology and the Average Grade (0.82). The KCSE examination compares with the mock examination as revealed by the correlations; First, the highest correlations in both was that between the Average Grade and Biology (0.81 and 0.82 respectively). The score in a Biology examination would suggest the Average Grade to be obtained by a student hence can be used in the absence of an Average grade.

High and significant correlations were obtained in both cases and Mathematics was highly related with the Science subject compared to the Arts subjects. The high relationship within the subjects in the KCSE examination (the criterion) gives confidence in the use of the mock and first term examinations which have subjects that were well related. Therefore, the mock examination could serve in certain purposes in the absence of the KCSE examination.

2. The Relationship between the three examinations

The correlations obtained between first term and mock examinations were not all statistically significant. Some were high, others substantially lower and for some (Mathematics, Biology and Physical Science) negligible. The correlations between first term English, Kiswahili and History and all the variables in the mock examinations (range 0.070 to 0.288) were significant revealing that there was a

relationship between them. The first term Mathematics, Biology and Physical Science yielded non-significant correlations (0.010 to 0.080) with all the variables in the mock examinations. These meant that there was no relationship between these variables. The objectives tested could have been different and content areas tested too. The results further support that the first term examination was not well related to the Mock examination. Though both were teacher-made tests, the first term was an internal examination while the mock examination was an external examination. Despite this fact and that they were both achievement measures, the kind of the examination influenced the relationship obtained. In that case the first term may need some improvements so that it can be well related with the mock examination and hence supplement the mock examination.

The first term and KCSE examinations were related in some subjects but not in others. As in the relationship with the mock examination only first term English, Kiswahili and History resulted in significant correlations with all the variables in the KCSE examination (range 0.070 to 0.299). These correlations were comparable with those of the mock examination (0.070 to .288). The three subjects were well related with the KCSE examination. Mathematics, Biology and Physical Science registered very low correlations (range 0.000 to 0.172) with the KCSE examination.

The first term examination was not well related with the criterion (KCSE examinations). The results may cause one to question the first term examination. There is definite need for its improvement as a measure of achievement. This could be in terms of testing skills and test-taking skills. Objectives outlined in the content areas should be addressed by the first term examination to improve its

validity and reliability.

The first term examination was not a good achievement measure and at that moment not much information could be drawn from it in terms of a students' level of achievement.

The results obtained between the mock and KCSE examinations emphasis earlier discussed results. All the correlations obtained were significant ranging between 0.234 to 0.758. The highest correlation (0.758) was that between the Average Grade. These indicated that the two examinations were highly related. The relationship was between the school subjects and the Average Grades too.

The high relationship could be attributed to the nature of the examinations which was the same. The mock examination had the same number of papers in each subject as the KCSE examination. This was not the case for the first term examination. The two (mock and KCSE) examinations were external examination. This meant that a number of people were involved in their construction and scoring. The scores obtained from each were therefore reliable.

The relationship determined between the mock and KCSE examinations suggested that the mock examination would be a good supplement of the KCSE examination. Any weakness identified when answering questions in KCSE examination could well be identified by the use of mock examination results. The mock examination results could also be used when KCSE examination results are not available. It would therefore not be a valid move to do away or replace mock examination with school specific examination. This would be the case if the school specific teacher made tests are improved so as to match the KCSE examination. The school-specific examination results could only be used if well

related with the KCSE examination.

Apart from the use of mock examination as a supplement of the KCSE examination, it could be used to improve the students' performance in the KCSE examination. Since the mock examination was related to the KCSE examination, the weakness noted in students could be corrected in good time before the taking of the national examination. This would mean that the effect of factors such as lack of test-taking skills, lack of adequate content coverage would not influence the KCSE examination result of a student. The results of the KCSE examination would reflect the achievement level of a student and not otherwise.

On comparing the results between first term and KCSE examination and that between mock and KCSE examination the marked difference suggested that there was need to have the mock examination as a measure of the students level of achievement. The first term examination may need improvement in various areas such as item sampling for its results to be helpful in the instructional process. The teachers needed to regard this tests as important and take time to construct and mark them. Analysis of their results would need to be done so that improvement in revealed weak areas would be done. This would raise the quality of the tests and the results too.

Further analysis of the correlations between results in the same content areas (for example first term English and mock English results) revealed the same results as discussed. The correlations between first term and mock examinations (range 0.020 to 0.266) indicated that only Art subjects (English, Kiswahili, History and Geography) were related. The same was the case for the correlations between first term and KCSE examinations (range 0.038 to 0.304). But for the

correlation between second term and KCSE examination (range 0.299 to 0.758) all the subjects were related. The Average Grades were related in all the three cases. In general, the results implied that the mock examination was a good predictor of the KCSE examination.

3. Contribution of the school subjects towards the Average Grade:

The Average Grade which had been in use in making various decisions concerning a student serves its purpose as a representative grade of the one's level of achievement. The contribution of each subject towards the Average Grade was determined in terms of beta weights (regression coefficients). The results obtained revealed that all the subjects had a significant contribution towards the Average Grade. In all the three examinations, Physical Science had the highest coefficient. First term (0.32), mock (0.15) and KCSE (0.20) examinations. These results were obtained when all the subjects were considered.

When only the art (English, Kiswahili, History, and Geography) subjects were considered the History had the highest coefficient in the first term (0.33) and mock (0.26) examinations. English had the highest coefficient in KCSE examination (0.30). On considering the science (Mathematics, Biology and Physical) subjects only Physical Science had the highest coefficient in first term (.48) and KCSE (.35) examinations. Biology had the highest coefficient (.34) in mock examination.

The results indicated that for each examination considered the contribution of each subject was significant and that the examinations bore a relationship with one another. The use of the Average Grade was valid in making decisions

concerning a student but when specific subjects are to be considered, one subject could be chosen in each category (art or science). Physical Science in terms of science subjects and History in terms of art subjects to serve the a given purpose.

5.1 Conclusion

The teacher-made tests are important within the teaching-learning process as a measure of the level of achievement of the students. The mock examination turned out to be a better instrument as a measure of the student level of achievement than the first term examination. This was revealed on obtaining the coefficients of relationship of the two (first term and mock) with the KCSE examination. The correlations of all the variables were significant in the relationship between mock and KCSE examinations. Only the art subjects and the Average Grade had significant correlations in the relationship between first term and KCSE.

A definite difference was revealed between the teacher-made tests (first term and mock examinations) by the correlations obtained between the two. Therefore, all teacher-made tests may not measure the student's level of achievement to the same degree. A lot of things need to be considered. For example, testing-skills in the teachers, test-taking skills in the students, defined objectives of the tests and well sampled test items. These seemed to be inadequate in the first term examinations.

The relationship established between the teacher-made tests and the KCSE examination indicated that; some of the major problems incurred in answering the

KCSE examination would be dealt with earlier if teacher-made tests were well used. Analysis of teacher-made tests results would reveal these weaknesses. Such weaknesses are lack of test-taking skills, inadequate coverage of content, concepts not learnt and knowledge not acquired. The result would be that teachers would take time to not only help students acquire knowledge but also train them how to answer specific questions in different content areas.

Besides the use of teacher-made tests in monitoring and diagnostic decisions, they can be used for administrative purposes. An institution needs to maximize the objectives of the education system in terms of decisions made. These decisions entail selection, classification and placement of the students. Analysis of results obtained from these tests would help select weak students from strong ones academically in specific content areas. The teachers would then give remedial work to improve the performance of such students. In classification, the administrators would be able to classify students correctly and have confidence in the classes obtained even without the KCSE results. This would help students at an earlier stage to identify where they belong with reference to the other students and hence be able to adjust accordingly where possible.

The mock examination was an important teacher-made test that needed to be encouraged in order to help the students in the various areas of concern. The mock examination results could be used whenever the KCSE results are not available in making decisions regarding a student. This would be the case unless the school specific mock examinations are found to be correlated with the KCSE examinations. Otherwise, the district specific mock examination would be a necessary examination in the education process.

The average grade was a representative grade of ones performance in all the three examinations. Therefore, the use of the Average Grade in making decisions concerning a student was valid. In addition to the use of the Average Grade, individual-subject grades are used in some cases. This helps in achieving specific information as concerns certain fields of operations. As revealed from results obtained there is need at times for such considerations. Since the subjects give significant contribution towards the Average Grade . The Average Grade is well related to the subjects considered and hence its use instead of the use of all subject grades. Therefore, for a given set of examinations, an overall Average Grade would reliably give valid information as would the individual sets of examinations. This supports the use of the Average Grade in decisions concerning a student.

5.2 Recommendations

The teachers need to make use of tests in testing mastery of content and monitoring students' progress. The tests should function as feed-back mechanisms to facilitate teachers' ability to adapt instructions to individual students.

Depending on the aim of a test constructed either objective can be met.

Therefore, the purpose for which the tests are given need to be articulated by both the teacher and the students. The tests given must serve the purpose for which they were constructed.

In order to monitor a students progress, the teachers need to use the results of periodic tests given on a continuous basis to identify areas where change may

be needed to modify instruction or otherwise. This would also help students determine how well they have mastered a given content area and to identify areas not covered hence take corrective measures. This should be encouraged by all concerned with the instructional progress.

Tests given by teachers are rarely used for diagnostic purposes. This is an important aim of a test since the tests aimed at achieving diagnostic purposes help identify specific areas of weaknesses or strengths in individual learners and then ascertain the nature of deficiencies. The teachers need to aim at achieving this purpose through testing for such tests would help point out the specific corrective guideline measures to be undertaken to improve a students achievement. The areas of improvement could be in answering questions, learning difficulties test-taking skills and the like. The teachers need to equip the students in these areas apart from equipping them with the necessary academic skills.

In order for the teachers to be able to analyse the test results and address the information obtained, the teachers need to be well equipped with the testing teehnology. This may occur through pre-service courses. These pre-service courses ought to emphasis all that is involved in the testing process. Analysis of test results ought to be emphasised such that the knowledge acquired would help the teachers maximize the use of test. The teacher also need to continually attend in-service courses, seminars and workshops to improve in their skills. Such organised courses could equip the teachers with any new modes of testing to be adapted to improve the testing process and use of the tests. These trainings are often given to personnel involved in marking the KCSE examination and need to be extend to persons within the schools involved in achieving outlined educational

objectives.

The testing process within the schools in the area studied should be a concern of all persons involved with the instructional process. The KNEC could provide annual reports of information obtained from analysis of the KCSE examination. This would serve as a guideline to the teachers when carrying out their teaching and testing. The teachers would be able to use the tests they give to identify and rectify weaknesses revealed by the national examination. The result would be improvement of the students performance if these weaknesses are corrected within the course of instruction. The teachers would also not be satisfied with the score of a student in a test. But would be concerned with how the students answer the questions thus giving necessary guidance where needed.

The Heads of departments would also contribute in improvement of testing within the school. With the assistance of the Heads of Schools, the Heads of Departments would guide the teachers in the testing process and use of test results. Through close supervision, the teachers would be helped to achieve the best through the tests they give.

Improvement done on the tests given by the teachers would enable the teachers use the results from these tests with confidence. The administrators and others involved with the instructional process would also regard the teacher-made tests with high respect and confidence as measures of the level of achievement of the students. The optimal results about a student's future training or undertakings.

The teacher-made tests need to be used to improve a students level of achievement. The teachers can achieve this if well equipped with the testing skills. If this be the case, the teacher-made tests results would serve as predictors

of the national examination results, thus improving their uses within the instructional process.

Further research need to be done in a number of areas of concern; First , the mock examination used in the study was a district-specific one. As mentioned earlier, there are school specific Mock examinations. There is need therefore for a study to look into the difference between the school specific and district - specific mock examinations as far as their relationship with the KCSE is concerned. Other than factors mentioned it would be worth investigating into factors that would result in the low relationship between first term and mock examinations. The low relationship between the school subjects in the first term examination need be investigated also. This would give more information on the need to improve the teacher-made tests and emphasis their use.

For a generalization of the findings, studies in other districts within the country need to be done. This would give information that would be applicable within the entire educational field concerning the use of teacher made tests. In addition, studies need to be carried out which would consider all the subjects contributing to an average grade. This would entail taking into account individual optional subjects.

Goodrich, J. J. (1971). *Test and Measurement in Education*. New York: Educational Measurement, Washington, DC, American Council on Education.

Evans, R. C. and Simpson, R. G. (1981). The Construct Validity of the Wechsler's Reading Mastery Test relative to the Prason's Individual Achievement Test among K-12 level Adolescents. *Journal of Psychological Measurement*, Vol. 4.

Evans, R. C. and Simpson, R. G. (1981). The Construct Validity of the Peabody Individual Achievement Test relative to the Key Math Computer Arithmetic Test among Adolescents. *Journal of Psychological Measurement*, Vol. 4, No. 8.

REFERENCES

1. Allen, M J and Yen, W M (1979), Introduction to Measurement Theory, California Brooks/Cole Co.
2. Anastasi, A. (1982), Psychological Testing, New York, Macmillan.
3. Ansley T N and Forsyth R A; (1983): Relationship of Elementary and Secondary School achievement test scores to college performance in Educational and Psychological measurement; Vol. 43.
4. Bali, S K , Drenth, P J D, Flier, H and Young, W C E (1984), Contribution of Aptitude Tests to the prediction of School Performance in Kenya : A Longitudinal Study, Lisse, Swets and Zeitlinger.
5. Bokhorst, F D, (1986), Confidence - Weighting and the Validity of achievement tests in Psychological reports, 56, pp 383-386
6. Borg, W R (1967), Educational Research : An Introduction. New York, David Mckay.
7. Bruno, F. (1981), A concurrent Validity Study of the learning profile against college GPA in Educational and Psychological measurement; Vol. 43.
8. Brown, F G (1983), Principles of Educational and Psychological testing : New York. Hole, Rinehardtand Winston.
9. Christie, T and Forrest, G M (1981), Defining Public Examination Standard, London, Macmillan Education.
10. Cronbach, L J (1970) Essentials of Psycholological Testing, New York, Harper and Row.
11. Cronbach, L J (1971), Test Validation. In Thorndike R L (Ed), Educational Measurement, Washington, DC, American Council on Education.
12. Eaves, R C and Simpson, R G (1983). The Concurrent Validity of the Woodcock Reading Mastery Tests relative to the Peabody Individual Achievement Test among Retarded Adolescents in Educational and Psychological measurement; Vol. 43.
13. Eaves, R C and Simpson, R G (1984), The Concurrent Validity of the Peabody Individual Achievement Test relative to the Key Math Diagnostic Arithmetic Test among Adolescents in Psychological Abstract, Vol. 71, No.8.

14. Ebel, R L (1972), Essentials of Educational Measurement, Englewood Cliffs, New Jersey, Prentice - Hall.
15. Eisner, E W (1986), The Art of Educational Evaluation : A Personal View, London, The Falmer Press.
16. Fergusson, G A (1981), Statistical Analysis in Psychology and Education, New York, McGraw - Hill.
17. Flexer, B K (1984), Predicting 8th grade Algebra achievement. In Journal for Research in Maths Education, 15, pp 352 - 360.
18. Gandye, P P (1991), Assessment of National Educational goals and objectives: the case for continuous assessment. Unpublished paper presented at IAEA 17th conference.
19. Gatumu, H J N , (1976), A study of predictive validity of O-Level and an aptitude test in relation to the performance at the University of Nairobi, Unpublished master thesis, University of Nairobi.
20. Glass, G V and Stanley, J C (1970), Statistical Method in Education and Psychology, New Jersey, Prentice Hall.
21. Gronlund, N E (1976), Measurement and Evaluation in Teaching, New York, Macmillan.
22. Hattie, et al, (1984); The relationship between two measures of self-actualisation in Psychological Abstract, Vol 7. No. 7.
23. Hughes, O L (1983), A comparison of error-based and time-based learning measures as predictors of general intelligence in Psychological Abstract. Vol. 7. No. 7.
24. Kline, P (1979), Psychometrics and Psychology, London, Academic Press.
25. KNEC, 1991, 1989 KCSE Examination Report; Unpublished Report.
26. Lemke, E and Wiersma, W (1976), Principles of Psychological Measurement, Boston, Houghton Mifflin Co.
27. Lewis, D G (1974), Assessment in Education, London. University of London Press. Press.
28. Lord, F M and Novick, M R (1968), Statistical Theories of Mental Test scores, Reading Mass, Addison - Wesley.
29. Maundu, J N (1987), Family background and student achievement in Kenyan National Examinations, paper submitted for the Bureau of Education Research Seminar, Nairobi.

30. McCausland, (1985), The accuracy of teachers predictions of pupils' marks in school certificate, Psychological abstract, Vol. 73 No.11
31. Mehrens W A and Ebel, R L (1976), Principles of Educational and Psychological Measurements, Chicago, Rand McNally and Co.
32. Mkandawire, D S J (1991), Role of criterion, norm-referenced and continuous assessment in post -independence curriculum and examination reforms in Namibia. Unpublished paper presented at IAEA 17th conference.
33. Muya, W (1991), Subjects reduced in 8 - 4 - 4 reform, Daily Nation, Nairobi, October 8, Tuesday Page 1 Col.4 - Page 2 Col 1.
34. Nunnally, J (1967), Psychometric Theory, New York, McGraw - Hill.
35. Opinya, J O (1980), Mock History as a Predictor of success in the East African Certificate of Education (EACE). Unpublished Master Thesis, Kenyatta University.
36. Phelps, A and Ensor, A (1986), Concurrent Validity of the WISC-R using deaf norms and the Hiskey-Nebreska in Psychological Abstract; Vol 73 No. 8.
37. Popham, W J (1981), Modern Educational Measurement, Englewood Cliffs, New Jersey, Prentice -Hall.
38. Ryan, J J (1986), Application of a WAIS-R Short form with neurological patients: Validity and correlational findings in Psychological Abstract; Vol 73. No. 2.
39. Saigh, P A and Khouri, A (1983), The Concurrent validity of the Maths anxiety Rating Scales for Adolescents (MARS-A) in relation to the Academic Achievements of Lebanese students in Educational and Psychological Measurement; Vol. 43.
40. Satterly, D (1981), Assessment in School, Oxford, Basil Blackwell.
41. Sharpley, C F and Edgar, E (1986), An empirical investigation of agreement between two indices of achievement. Psychological abstract. Vol 73 No 6
42. Ssali, R. E. T. (1986), measuring educational potential with Draw - A - Person test. Unpublished master thesis, Kenyatta University.
43. Stanley J. C. , B. Hopkin K., ,D. (1981) Educational and Psychological Englewood Cliff, New Jersey, Prentice - Hall.
44. Svec, H (1986), Overestimation of academic competence by High School dropouts. Psychological Reports, Vol 59 pp 669 - 670.

APPENDIX 2

<p>Group 1</p> <p>English Kiswahili History and Government Geography Mathematics Biology Physics Chemistry Physical Sciences Biological Sciences</p>	<p>Group 2</p> <p>Christian Religious Education Islamic Religious Education Social Education and Ethics Hindu Religious Education</p>
<p>Group 3</p> <p>Home Science Art and Design Agriculture Woodwork Metalwork Building Construction Power Mechanics Electricity Drawing and Design</p>	<p>Group 4</p> <p>French German Arabic Music Accounting Commerce Economics Typewriting with Office Practice</p>

NOTE: The superior performance is awarded grade A and the weakest performance grade E.

APPENDIX 3

GRADING SYSTEM ANALYSIS	
GRADE	NUMBER OF POINTS
A	12
A-	11
B+	10
B	9
B-	8
C+	7
C	6
C-	5
D+	4
D	3
D-	2
E	1

NOTE: The superior performance is awarded grade A and the weakest performance grade E.

APPENDIX 4

SECONDARY SCHOOLS STUDIED

1. Arnesen's Boys school
2. Dry's Girls school
3. Harambee Mixed school
4. Hill Mixed school
5. Kapkong Mixed school
6. Kapng'etuny Boys school
7. Kipkabus Mixed school
8. Moi Girls school

KENYATTA UNIVERSITY LIBRARY