A STUDY OF THE PREDICTIVE VALIDITY OF THE KENYA CERTIFICATE OF PRIMARY EDUCATION EXAMINATION: APPLICATION OF HIERARCHICAL LINEAR MODELS

By

LUCAS A. OTHUON

B.Ed., University of Nairobi, 1980

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS OF THE DEGREE OF MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY AND

SPECIAL EDUCATION

We accept this thesis as conforming to

the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

July, 1993

ABSTRACT

Public examinations have been used in Kenya for decades as selection instruments for further education and training. The Kenya Certificate of Primary Education examination (KCPE) is the first of such selection examinations. A basic assumption is that those who pass the examination and are selected to join secondary school have a good chance of succeeding in secondary school. However, evidence that may verify such an assumption, that is, a study of the predictive validity of KCPE, has received little attention. The purpose of this study was to determine the extent to which KCPE predicts success in secondary school. Success in secondary school was measured by the level of examinee achievement in the Kenya Certificate of Secondary Education examination (KCSE).

Stratified random sampling was used to select 26 secondary schools within a single district in Kenya. The 1991 KCSE data for 781 examinees in the sample were used in the analysis. The KCSE records for examinees in the sample were matched with corresponding 1987 KCPE records. The nature of the relationship between KCPE and KCSE was determined by use of Hierarchical Linear Models (HLM). The influence of selected moderator variables on the relationship between KCPE and KCSE was investigated as well. These variables were age, gender, repetition of Standard 8 (i.e., writing KCPE more than once), and school size.

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

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ACKNOWLEDGEMENTS

It is not possible to express my sincere gratitude to all those who offered important advice, critique and encouragement in the development of this thesis. However, I wish to specifically thank my supervisor, Dr. Nand Kishor, and members of my thesis committee, Dr. R. Conry and Dr. W. McKee, for their being available and ever willing to offer guidance whenever such a need was deemed necessary. Dr. D. Allison's role as an external reader of the manuscript was highly appreciated.

I must not forget to thank the Canadian International Development Agency and the Kenya Government for their kindness in ensuring that my stay in Canada was rewarding and full of happiness. It was such a wonderful experience!

Last but not least, I cannot forget to thank my wife, Clarice; and my children, Vera, Harry, and Rowney, for their understanding and perseverance during my two years' stay away from them.

Chapter 1

INTRODUCTION

Public examinations have been used in Kenya for decades as selection instruments for further education and training. The Kenya Certificate of Primary Education examination (KCPE) is the first of such examinations. One of the aims of having an examination at primary school level was to provide evidence of a well rounded education. It was also expected that secondary schools would continue to give such training (Kenya Education Commission Report, 1964).

Eisemon and Schwille (1992) stated that the object of primary schools' national examination was to reduce the number of students eligible for secondary schooling. They further suggested that one of the factors affecting student performance is the characteristics of critical examinations used to select students for further education. For example, a good primary school examination should bear evidence of having a significant positive linear relationship with a secondary school examination in the same system. If mass failure gets reported at secondary school level, as is often the case in Kenya (see The Weekly Review, April 16, 1993, p. 10), then there is need to find out the nature of the relationship between the examination that was used for selecting candidates into secondary school, and student achievement in secondary school. Such predictive validity studies of performance on KCPE has received little attention (Kellaghan & Greaney, 1988). This concern is crucial if proper selection decisions are to be made based on performance in KCPE.

This study is an attempt to examine the predictive validity of KCPE. The study also attempts to show how the predictive relationship is influenced by selected pupil-level and school-level variables. Because the predictor and criterion variables are examinations set and scored in Kenya, it is important to give some background information on education in Kenya.

Education in Kenya

On attainment of independence in 1963, the system of education in Kenya was predominantly examination oriented. Pupils who sat for examination during their final year in primary school were being branded as failures if their performance did not meet the required standards. In 1975, out of 220,000 primary school leavers, about 32% of the group were offered secondary school or vocational training places (Report of The National Committee on Educational Objectives and Policies, 1976). Of the 341,000 primary school leavers who wrote KCPE in 1987, 51% were successful in obtaining places in secondary school (Kellaghan and Greaney, 1992). These figures illustrate how competitive it was for students to join secondary school in Kenya.

Over the years, there has been concern about the impact of external examinations on curriculum implementation. The former 7-6-3 education system, which stands for seven years of primary education, six years of secondary education, and three years of university education, was much criticized. Parents and the public in general expressed their dislike for the examination at the end of primary school. They saw it as the traumatic event which determined the fate of the majority of children who did not manage to get places in secondary schools (Kenya Education Commission Report, 1964). The education system was further criticized as not being practical-oriented and that it offered a narrow based curriculum. Its graduates were thought to be ill-equipped to fit into the changing job market, and hence the rise in unemployment. It was also claimed that the system did not cater for the development of the most appropriate attitudes that children needed, particularly at primary school level.

In 1985, a major change of curriculum was effected in Kenya. The 7-6-3 system of education was replaced with 8-4-4 system of education, with a view to alleviating the problems that beleaguered the old system. The 8-4-4 stands for eight years of primary education, four years of secondary education and four years of university education. It was apparent, however, that even with the 8-4-4 system, basic issues and concerns related to curriculum implementation and evaluation still persisted. For example, whereas the new system called for much more than the old one ever had in terms of physical resources, notably laboratories and workshops, not many schools could be said to have been prepared in this area. The responsibility for equipping them was transferred to parents under the new cost-sharing system. In addition, a few of the teachers at that time were trained to handle the new 8-4-4 curriculum. It was small wonder, therefore, that critics of the system claimed that pupils were being used as guinea pigs to test the system, and their worst fears appeared to be confirmed when mass failure in the 1989 Kenya Certificate of Secondary Education examination was reported.

One area where the negative aspects of the 8-4-4 system have come out glaringly is at secondary school level. It became apparent that other than equipping its graduates with the necessary skills to enable them to survive in the competitive job market upon graduating from school, the system left them worse off. Employers were not ready to give them jobs, preferring instead the graduates of the old system. Something seemed to be inherently wrong with the new system, calling for a closer look into the system with a view to offering guidelines that may lead to its restructuring.

Examinations in Kenya

The Kenya National Examinations Council Act, (Cap 225A, Laws of Kenya), enacted in 1980, made provision for the establishment, constitution, control and administration of the Kenya National Examinations Council (KNEC). Amongst the important examinations set and scored by KNEC are the Kenya Certificate of Primary Education examination (KCPE) and the

Kenya Certificate of Secondary Education examination (KCSE). KCPE is an external examination taken by pupils at the end of their eighth year in primary schools in Kenya. A certificate is issued to each candidate irrespective of level of achievement. The certificate is meant to certify that the bearer has undergone eight years of primary education and has attained a level of achievement as shown by the grades contained therein. KCPE examination under the current 8-4-4 system was done for the first time in 1985. In 1987, six papers were written in the examination and each was scored out of 100 points. The six papers were English (ENG1), Kiswahili (KIS1), mathematics (MAT1), science and agriculture (SC&A), geography, history, civics and religious education (GHCR), and art & craft, home science and music (ACHM). All items in the examination were of multiple choice type, except for English and Kiswahili, each of which had a short composition section. The examination was computer scored except for the composition sections of English and Kiswahili papers.

Other than being used for examinee certification, KCPE is also used for selection of candidates into secondary schools. In 1989, out of over 600,000 KCPE candidates, only 166,748 pupils were selected to join secondary schools the following year (Statistical Abstract, Kenya, 1990). The selected group formed a proportion of about 30% of the registered candidates who wrote the examination in 1989.

The use of external examinations as a selection tool in Kenya has quite often received criticisms. The Kenya Education Commission Report (1964), in a section dealing with the primary school examination that used to be offered at that time, recommended that:

> The issue of KPE [Kenya Primary Education examination] certificates to successful pupils should be replaced by the issue of school leaving certificates to all pupils. Cramming for KPE should be discouraged. Alternative selection procedures should be the subject of research.

Members of the public also expressed their concern about possibilities of a correlation between achievement in primary school and age of the examinees. They thought that this realization could give rise to the extensive rates of repeating before children were submitted to write the primary school examination (Report of The National Committee on Educational Objectives and Policies, 1976). However, these feelings have rarely been supported with concrete research evidence and should therefore be seen as purely impressionistic.

The Kenya Certificate of Secondary Education examination (KCSE) is also an external examination written by pupils at the end of their fourth year in secondary school. Those who wrote KCPE examination in 1987, and joined secondary school, wrote KCSE in 1991. In the 1991 KCSE examination, each candidate registered for a minimum of ten subjects. Three of the compulsory subjects in the examination were English (ENG2), Mathematics (MAT2) and Kiswahili (KIS2). The 1991 KCSE examination was scored by subject specialists who had received training as examiners by the Kenya National Examinations Council. Other than being used for selection into tertiary institutions, KCSE grades are also used by some employers to recruit workers. It is therefore worth stressing that, to date, examinations remain the basic procedure for selection of students for further education and training in Kenya.

The fact that an examination system can be used for a variety of purposes should not be taken to imply that such a system can readily serve all purposes equally well. For example, an examination system that efficiently selects the pupils most likely to benefit from further education might contribute to the identification of a technical elite and over time might even have important effects on the economic performance of a nation (Heyneman, 1987). However, such a system could have serious and damaging effects on the educational experiences of many students if it ignores the fact that for many students, learning has to have utility beyond that of qualifying individuals for the next level of education (Kellaghan & Greaney, 1992). Thus, a procedure that most efficiently selects students may be inadequate for certification purposes. Similarly, a procedure that is adequate for certification is unlikely to be the most appropriate one for monitoring the quality of performance of a school or of the educational system in general. This calls for the evaluation of the validity of examinations at each stage within the educational system. However, only ad hoc evidence has so far been provided on the predictive validity of KCPE examination since its inception in 1985.

Purpose of the Study

The policy of selecting students for secondary school education is based on the assumption that those who pass the selection examination have a good chance of succeeding in secondary school. In other words, in selecting pupils for secondary school admission, future performance is predictable from composite KCPE scores. One problem that requires analysis is whether KCPE examination, the predictor variable that was used for selection, has significant relationship with success in secondary school. Thus, this is a study of the validity of KCPE examination. The term validity, when applied to a test, refers to the precision with which the test measures some particular mental ability. Three main types of validation studies exist; criterion-related validity, construct validity, and content validity. Criterion-related validity may be either concurrent or predictive depending on whether the scores predict a criterion at the time a test is administered or at some point in future. Since KCPE scores were used to predict success in secondary school at a future date, this validation study is a predictive one. Success in secondary school was measured by KCSE examination data. The study also examined the extent to which the relationship between KCPE scores and success in secondary school is moderated by selected pupil-level and school-level factors. The major research question is whether KCPE is a valid predictor of KCSE. This question was addressed by the following sub-questions:

1. What is the correlation between KCPE and KCSE?
2. How much do secondary schools in a single district in Kenya vary in their mean KCSE achievement levels?
3. What is the nature of the pupil-level relationship between KCPE and KCSE?
4. What is the impact of selected pupil-level and school-level variables on the relationship between KCPE and KCSE?

Significance of the Study

As in other countries, education continues to be a dominant sector in Kenya's economy. According to the Development Plan (1984-88), education accounted for about 7.2% of Kenya's GDP in 1981. By 1983, it was already receiving 30% of the national recurrent budget. There is, therefore, a need for identifying variables that govern the quality of our educational output as well as the inter-relationships involved between them. It is only through such means that we can propose important policy decisions regarding our system of education. This would ensure that meagre resources are directed where they are mostly needed and where the benefits would be optimal, thus increasing efficiency.

The number of studies that focus on the predictive validity of KCPE has been low. It is therefore hoped that this study will offer teachers, educational researchers and policy makers some guidance on the predictive validity of KCPE. The study is also an addition to similar research done elsewhere in the world using hierarchical linear models (see, for example, Willms (1985) & Willms and Jacobsen (1990)). Hierarchical linear models capture hierarchical data structures in a manner that was not possible in most of the previous research.

Chapter 2

REVIEW OF LITERATURE

This chapter discusses some important factors that may affect predictive validity of an examination in relationship to a criterion. The chapter highlights a number of important statistical approaches to prediction studies. A conceptual framework of multilevel modelling is also presented.

Factors that may affect Predictive Relationships

Several factors may influence predictor-criterion relationship. The factors include psychometric characteristics of examinations, characteristics of examinees, and characteristics of schools.

Characteristics of Examinations

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

Characteristics of Examinees

Examinee characteristics may influence the predictor-criterion relationship in a predictive validity study. The selected examinee characteristics are age, gender, and repetition.

Age

The relationship between age and achievement is made more and more complex by the fact that grade retention, maturation, learning, and nature of instruction are confounding variables in the relationship. In a study involving some twelve countries, each with a distinctive educational system, there was support for the idea that older children generally performed better academically than younger children (see Husen, 1967). Choppin (1969) found that the effect of age on achievement, after controlling for social class, differed across countries. Walsh (1988) also found that children who were youngest in their class showed the highest chances of failure. However, Smith and Shepard (1987) noted that it was not age alone but a combination of young age and low ability that had an effect on performance.

Harnisch and Archer (1986) gathered data from Japan, India, and Illinois (U.S.) on high school students who had completed the High School Mathematics Test. Results not only showed differences in achievement across the three countries, but also differences among countries in the relative influence exerted by age on students' achievement in mathematics. For example, there appeared to be a positive linear trend in the Japanese sample, with scores increasing as students' age increased. However, in India the younger students had higher scores than their older counterparts. The correlations between mathematics achievement and age in Japan and Illinois were .32 and -.08, respectively. It seems that the relationship between age and achievement may vary from one setting to another.

Gender

There is a substantial body of evidence to suggest that from the beginning of secondary schooling, males frequently outperform females in mathematics (Fennema & Leder, 1990). Fox & Cohn (1980) gave school variables as one of the reasons. These school variables include organization procedures as well as behavior, expectations, and beliefs. Another possible explanation is that teachers often interact differently with their male and female students, with males attracting more and qualitatively different interactions (Brophy, 1985). Peer group influence is yet another important dimension in explaining differential achievement by gender. It acts as an important reference for childhood and adolescent socialization and further perpetuates sex role differentiation through leisure-activities, friendship patterns, subject preferences, and career intentions.

To date, there is no conclusive evidence as to whether gender has an effect on achievement. Several studies reveal an inconsistency of findings, with males performing better in some studies and females in others. Stroud and Lindquist (1942) gave an exhaustive review of literature on experiments about gender differences and achievement in elementary and secondary schools. They reported that judging from past experiments, correlations between IQ and achievement data were higher for girls than for boys. However, this superiority cannot be generalized for it is more pronounced only in certain subjects. They also reported that earlier research showed no significant differences in reading, whereas there existed a significant difference in language and literature in favour of girls. On the other hand, boys were deemed to excel in mathematics, history, and science over girls.

Aiken (1971) and Werdelin (1961) found that on the average, girls tended to score higher than boys on tests of verbal fluency, arithmetic fundamentals, and rote memory. They also found that boys were superior in spatial ability, arithmetic reasoning and problem solving. However, these sex differences were found to be less pronounced in the early grades. Dwyer (1973) stated that it has been a common research finding that girls are generally better readers than boys.

Simon & Danna (1990) conducted a study that evaluated the accuracy of Law School Admission Test scores in predicting student law school performance. Male and female scores and White, and Black, or Hispanic scores were compared. Data were drawn from 1987 and 1988 graduating classes of five geographically diverse law schools. No significant differences between gender groups were found.

Over the years, girls achievement in mathematics and science in Kenyan secondary schools has been lower than that of boys (Ndunda, 1990). This outcome may be attributed to the girls' experiences in mathematics and science, and of socio-cultural forces that interact to influence their conception of mathematics and science. Eshiwani (1985) noted that whereas female enrollment patterns improved significantly during the 1975-1984 decade, there were regional disparities and significant differences in attrition rates and the number of girls attending school. When confronted with constraints of limited opportunities or resources for primary schooling, parents have generally favoured the education of male children. Eshiwani (1985) found that the only subjects in which secondary school girls in Kenya performed better at than boys were English and Christian Religious Education.

Repetition

Repetition or grade retention is the practice of requiring a student who has been in a given grade level for a full year to remain at that level for a subsequent year. Two main reasons behind retention are to remedy inadequate academic progress and to aid in the development of students who are considered to be emotionally immature. In 1986, 13% of the total primary school enrollment in Kenya were repeaters (Kellaghan & Greaney, 1992). One consequence of the difficulty and selectivity of KCPE examination is high repetition rates for Standard 8 in order to rewrite the KCPE examination (Eisemon and Schwille, 1991). These cannot be estimated with precision since ministry regulations forbid repeating. Most of the research on retention effects have been inadequate because the researchers defined the treatments poorly, used small sample sizes, and did not consider the hierarchical nature of most organizational structures. Jackson (1975) identified three such designs commonly used in the study of the effects of grade retention on achievement. The first type of design used, for example, in the studies of Coeffield and Bloomers (1956) and Scott and Ames (1969), involved comparing the condition of retained students after promotion with their condition before promotion. Such a design is biased toward indicating that pupils gain from grade retention because of lack of control for possible improvement resulting from causes other than the retention experience itself.

The second type of design, used by researchers like Chansky (1964), Briggs (1966), and Reinherz and Griffin (1971), employed an analysis in which students retained under normal school policies were compared with students promoted under normal policies. Such comparison is biased toward indicating that grade promotion has more benefits than grade retention, because it compares retained students who are having difficulties with promoted students who usually are not having as severe difficulties. Although matching retained and promoted students on indicators of classroom achievement like age, IQ, or SES is likely to improve such designs, results from such studies could still be flawed because the criteria for promoting students do vary. Additionally, student performance also varies among schools, thereby making any inferences drawn from such studies spurious. Thus, even when retained and promoted pupils have been matched on age, IQ, and SES, there is inadequate assurance that the pupils were initially similar in respect to the actual conditions which precede grade retention.

The third type of design, used by Cook (1941), for example, compared pupils with difficulties who had experimentally been assigned to promotion or grade retention. This type of design is superior and more reliable than the first two designs. However, such experimentally controlled designs are few in number as to allow for broad generalizations about the effects of grade retention on students' academic achievement. It appears that a design that takes care of the hierarchical nature of student performance would be most suitable in this case.

Holmes and Matthews (1984) did a meta-analysis on methods used for examining the effects of nonpromotion on elementary and junior high school pupils. They defined effect size as the difference between the mean of the retained group and the mean of the promoted group, divided by the standard deviation of the promoted group. This procedure results in a measure of the difference between the two groups expressed in quantitative units. From their study, they found that the effect of nonpromotion on pupils academic achievement was measured in 31 out of 44 studies. Altogether, 367 effect sizes were calculated. When the mean of the effect sizes was calculated a value of -.44 was obtained, indicating that the promoted group on the average had achieved .44 standard deviation units higher than the retained group, ${}_{.999}t_{366} = 12.57$.

Niklason (1987) measured retention effects for specific groups of children under four classification variables: group (retained vs. promoted), district, ability, and grade level. Retention was not found to benefit the children academically or in personal or social adjustment. There was further evidence that the arithmetic scores declined the following year for the younger retained children, but not for the younger promoted children.

Lenarduzzi and McLaughlin (1990) used a quasi-experimental design to examine the effects of nonpromotion on academic achievement and scholastic effort in junior high school. Results indicated that the students who were retained improved significantly with respect to academic achievement and scholastic effort when compared to those who were promoted. However, when they later did a follow up study on the same students, results showed no significant differences between the two groups in academic achievement.

Characteristics of Schools

School characteristics may affect predictive relationships. Such characteristics are, in general, called school climate (see, for example, Tagiuri, (1968), Anderson (1982), and Bryk & Driscoll (1988)). Tagiuri (1968) stated that school climate is a summary concept dealing with the total environmental quality within a school system. These may be geographical, organizational, or functional. Specific examples include the quality of teachers, type of supervision, location of the school, and curriculum organization. Anderson (1982) found that each school has a unique climate, and that climate affects many student outcomes. Anderson (1982) also found that understanding the influence of climate will improve the understanding and prediction of student behavior. However, climates in different schools are often elusive, and difficult to describe and measure. It is for this reason that school size was selected as a characteristic of the school that may affect predictive relationships. This was with the belief that school size is a potential climate mediator.

School Size

A number of studies have reported on the effect of school size on achievement. Some of these reports seem to be conflicting. Kimble (1976) administered the Stanford Achievement Test to 2,186 high school students in an attempt to find out whether or not school size was important in determining student achievement. The sample included 1,311 sophomores and 875 seniors. Results showed that, at the senior level, no significant differences existed in the mean test scores based on school size. However, sophomores of the larger schools scored better than did those from smaller schools. The New York State Department of Education (1976) did a study on how school size affects academic outcomes and found that the larger the school size, the lower the academic outcomes. Rutter, Maughan, Mortimore, Ouston and Smith (1979) found that school size had no effect on educational outcomes. The findings of the New York State Department of Education (1976) were similar to the findings of a study by Coladarci (1983) who carried a meta-analysis on the effects of institution size on pupil progress: smaller schools showed a definite superiority to larger schools. Coladarci (1983) further found that much of the research on the relationship between school size and achievement had evidentiary and inferential errors, intellectual puritanism, and rational extravagance. Wyatt and Gay (1984), in a review of research and case studies concerning how the size of an institution affects academic achievement, showed that size should not be seen as an independent variable having any direct impact on achievement. This suggests that size of institution should be used as a moderator variable rather than a predictor in studies involving predictive relationships because of its low correlation with the criterion.

Methodological Limitations in Previous Research

The foregoing review reveals that a number of results in studies on factors that affect achievement do not concur. The inconclusive findings about the impact of age, gender, repetition, and school size on achievement, suggest that more studies are required in this direction. Additionally, most of the previous studies made use of single level linear models. Such studies neglected the hierarchical nature of organizational structures. In an education system, for example, students are nested within schools, and schools within districts. Using a linear model that does not take care of this type of hierarchical structure often results in aggregation bias and misestimated precision. A more appropriate model that resolves the problems inherent to single level analyses is therefore necessary if sensible inferences are to made.

Statistical Analyses in Predictive Validity Studies

Several statistical approaches to prediction research exist, and the type chosen for use will quite often depend on the design of the study. This does not necessarily imply that only one method should be used in a single study. In quite a number of cases, more than one method is usually used in analyzing data. The following are some of the important statistical approaches to prediction studies.

Correlation Analysis

Correlation analysis involves determination of the strength and direction of relationship between variables. For situations involving more than two variables, partial correlation, multiple correlation, canonical analysis, factor analysis or discriminant analysis may be used. Many researchers report correlation coefficients not as their fundamental analytical approach, but as a component of their analysis (e.g., Jacobsen, 1990). The presence of a correlation between two variables does not necessarily mean there exists a causal link between them. However, correlation between variables can be useful in identifying causal relationships when coupled with other methodological approaches (Glass & Hopkins, 1984).

Ordinary Least Squares Regression Analysis

Ordinary least squares regression is a statistical technique that allows one to attribute amount of change in one variable to amount of change in other variables. One purpose of a linear regression equation is to make predictions on a new sample of observations from the findings on a previous sample of observations. For example, given individual scores on an independent variable X, it is possible to predict scores on a dependent variable Y by use of regression analysis. Obviously, the predicted scores quite often differ from the actual scores due to error of estimate.

Multiple regression is one of the most widely used statistical techniques in educational research. Borg and Gall (1983) attribute this to its versatility to yield information about relationships between several variables. It has, therefore, been used extensively in situations that require prediction. For example, universities usually admit and reject candidates mainly on the basis of predictions about their probable future performance made from achievement tests in high school. Similarly, insurance agencies heavily rely on actuarial studies to predict future events that may affect their operations. This in turn helps them to adjust policy premiums accordingly so that their business remains self-sustaining.

Hierarchical Linear Modelling

Raudenbush and Bryk (1986) suggested that a common weakness with ordinary least squares regression technique is that all explanatory variables are considered to be at the same level, yet most organizational systems do not operate as single level structures. They asserted that in the past, most researchers used single-level regression models only because of lack of viable alternatives. However, these alternatives are now available.

In ordinary least squares regression analysis of data sets from groups, aggregation bias is a common problem. Aggregation bias can be described as the difference between a slope obtained in a regression of means on means and the slope obtained in an individual level analysis. Raudenbush (1988) defined aggregation bias as a situation in which a variable takes on different meanings and has different effects at different levels of aggregation. Longford (1989) suggested that the use of hierarchical linear models may help reduce aggregation bias because such models take care of all sources of variation simultaneously.

Goldstein (1986) proposed multilevel linear modelling, also known as Hierarchical Linear Modelling (HLM) estimation by use of iterative generalized least squares, and discussed its application to nested and longitudinal data structures. The advantage of this approach is that, unlike single level ordinary least squares regression technique, it applies comparatively to a variety of mixed models with two or more levels of aggregation, with the result of reduced aggregation bias.

Raudenbush (1988) noted that multilevel models have parameters at a lower level of aggregation (microparameters) that are presumed to vary as a function of the parameters at the next higher level (macroparameters). As an example, a two-level HLM may have pupils at the micro-level (e.g.

examination performance data) nested within the school at the macro-level. This gives a linear model with pupil-level regressors and with a pupil-level dependent variable. Each school has its own model. The macro-model relates the parameters of the micro-models, which are the regression coefficients and the error variances, to macro-level regressors. Thus, the advantages of HLM over single level linear modelling arise from the former's more realistic portrayal of the effects of grouping. HLM incorporates the fact that individuals within groups share common features; they are not the completely independent entities assumed in ordinary least squares regression analysis.

The iterative generalized least squares approach used in HLM is an efficient method of fitting multilevel models. It uses all the information available in the data by differentially weighting each school's contribution. Schools whose coefficients would be poorly estimated if a series of ordinary least squares regressions were conducted benefit from the other schools' data. This occurs, for example, when some schools have far fewer students than others or when there is little inter-pupil variation. Additionally, in HLM covariances among coefficients are exploited to optimize precision of the estimates of the individual schools' slopes and intercepts (Rasbash, Prosser, & Goldstein, 1989).

With the foregoing considerations, HLM seems to offer a superior alternative method of data analysis with the type of data used in this study. The following section discusses the methodological framework of HLM as used in this study.

Methodological Framework of HLM

Previous studies on achievement relationships gave little consideration to the fact that most organizations operate as hierarchical structures. For example, in an attempt to determine how achievement relates with age and school size using multiple linear regression analysis, one ought to consider that age is a pupil-level variable and school size is a school-level variable. The use of a model that does not utilize the distinction between multilevel units leads to smaller estimates of standard errors of regression coefficients (Rasbash, Prosser, & Goldstein, 1989). Inferences made from such an analysis may be spurious.

Hierarchical linear modelling (HLM) utilizes the multilevel units in hierarchical data structures. Estimates of standard errors of regression coefficients using HLM are larger than the corresponding values from an ordinary least squares regression analysis, because the intraclass correlation among the measurements is taken into account. As an improvement over most of the previous research, the following hierarchical linear model was used in this study:

If data are collected in J schools, each of which contains n_j students (j = 1,...,J) and the interest is in determining the relationship between an examinee's KCPE examination scores (predictor) and KCSE examination grades (criterion), then for school j, a possible linear relationship between KCSE and KCPE is

$$Y_{ij} = b_{0j}X_0 + b_{1j}(X_{1ij} - \bar{X}_{..}) + e_{ij} \quad (1)$$

at pupil level. This is an ordinary least squares regression equation where Y_{ij} represents KCSE grade for individual i in school j, b_{0j} is a within-school intercept, X_0(=1) is a constant term, b_{1j} is the average change in KCSE grade for each unit change in KCPE score (i.e., the slope in the KCSE-KCPE relationship), X_{1ij} is the KCPE score for individual i in school j, $\bar{X}_{..}$ is the grand-mean for KCPE, and e_{ij} is a random residual variable, assumed to have an expectation of zero. This model permits each school to have its own slope and intercept (Rasbash, Prosser, & Goldstein, 1989).

Generally, b_{0j} and b_{1j} may vary across schools. Therefore, they are treated as random variables at level 2. If Z represents the size of secondary school, for example, then its impact may be analyzed by use of the following between-school model:

$$b_{0j} = c_{00} + c_{01}Z_j + u_{0j} \quad (2)$$

$$b_{1j} = c_{10} + c_{11}Z_j + u_{1j} \quad (3)$$

In equation 2, within-school intercept is given as a function of school size. Generally, c_{01} is the average effect of school size on mean KCSE performance for candidates having average KCPE scores. In other words, it represents the benefit of secondary school size on KCSE mean achievement levels. Similarly, c_{11} in equation (3) represents the average increment to the slope b_{1j} explained by differences due to school size. The random variables u_{0j} and u_{1j} represent the effects on the b's not explained by school size and are assumed to have a joint distribution with mean zero and a covariance matrix in level 2. It should be noted that equations (2) and (3) need not necessarily be modelled as functions of Z (See Rasbash et al., 1989).

Equations (1), (2), and (3) can be represented by the following general matrix expressions for a model where the random term is associated with the intercept:

$$\mathbf{Y}_j = \mathbf{X}_j \mathbf{b}_j + \mathbf{e}_j \tag{4}$$

for within-unit model for the jth level 2 unit. $\mathbf{Y}_j$ is the response vector values for group j, $\mathbf{X}_j$ is a matrix of group members' values on a set of explanatory variables (including $\mathbf{X}_0$), $\mathbf{b}_j$ is a vector of the coefficients for the group , and $\mathbf{e}_j$ $(=[e_{1j},...,e_{nj}]')$ is a vector of level 1 random terms. The between-unit model for the coefficients can be written as

$$\mathbf{b}_j = \mathbf{Z}_j \mathbf{\Gamma} + \mathbf{u}_j \tag{5}$$

where $\mathbf{Z}_j$ is a between-unit matrix, $\mathbf{\Gamma}$ is a vector of the fixed coefficients, and $\mathbf{u}_j$ is a vector of random terms. All the matrices given have conformable dimensions.
Combining (4) and (5) gives

$$\mathbf{Y}_j = \mathbf{X}_j\mathbf{Z}_j\mathbf{\Gamma} + (\mathbf{X}_j\mathbf{u}_j + \mathbf{e}_j) \qquad (6)$$

where $\mathbf{X}_j\mathbf{Z}_j\mathbf{\Gamma}$ is called the fixed part and $(\mathbf{X}_j\mathbf{u}_j + \mathbf{e}_j)$ the random part.

Summary

This chapter has described some pupil-level and school-level factors that may affect predictive relationships. Such factors include psychometric characteristics of examinations, characteristics of examinees, and characteristics of schools.

One such psychometric characteristic is reliability of both the predictor and the criterion. Reliability is a measure of how consistent the scores of individuals are over repeated administration of the same examination or its parallel form under same conditions. An examination is said to be valid if it measures what it purports to measure.

Characteristics of examinees may also affect the predictor-criterion relationship. The selected characteristics are age, gender, and repetition.

School climate may also affect achievement relationships. However, climates in different schools are often elusive and difficult to describe and measure. For this reason, school size was selected as a school level variable that may affect achievement relationships. This was on the basis that school size is a potential climate mediator.

The chapter has also provided an account of some important approaches to prediction studies, some of which are correlation analysis, regression analysis, factor analysis, canonical analysis, and discriminant analysis. Most of the previous research was limited to the extent that they did not consider the hierarchical nature of organizational structures. HLM was selected as a more appropriate choice for data analysis because it accounts for the hierarchical nature of organizational structures. A description of methodological framework of a simple two level hierarchical linear model was also provided. Chapter 3 describes the methodology employed in this study.

Chapter 3

METHODOLOGY

This chapter discusses the criterion, predictor, and moderator variables used in this study. The population is defined, and methods of obtaining a sample of schools considered representative of the population are also discussed. This is followed by a description of the procedure employed in data collection. Finally, data analytic techniques used in the study are described.

The Criterion Variable

The 1991 KCSE grades were used as the criterion variable. KCSE examination grades were in the form of letter grades ranging from the lowest grade, E, to the highest grade, A. For purposes of statistical analysis, the letter grades were converted to an equivalent twelve point scale with 1 corresponding to grade E, the lowest grade, and 12 corresponding to grade A, the highest grade. This scale gives E = 1, D- = 2, D = 3, D+ = 4, C- = 5, C = 6, C+ = 7, B- = 8, B = 9, B+ = 10, A- = 11, and A = 12. Grades on three secondary school subjects: English (ENG2), Kiswahili (KIS2), and mathematics (MAT2), were recorded for each examinee in the sample. Similarly, the overall KCSE grade for each examinee was also recorded.

The Predictor Variable

1987 KCPE scores were used as the predictor simply because those who wrote the examination were the same candidates who wrote KCSE examination (the criterion) in 1991. The Kenya National Examinations Council scored each of the six KCPE examination subjects as a percentage. These subjects were English (ENG1), Kiswahili (KIS1), mathematics (MAT1), science and agriculture (SC&A), geography, history, civics and religious education (GHCR), and art and craft, home science and music (ACHM). The six subjects offered in the 1987 KCPE examination therefore formed a composite test with a possible maximum of 600 points. Selection to secondary school was done on the basis of an examinee's KCPE composite score, calculated as a sum of the six subject scores. Selection was done on merit and the cut-off point depended on the number of available Form 1 places.

Moderator Variables

According to Cohen and Cohen (1983), a moderator variable refers to an independent variable that potentially enters into interaction with a predictor variable, while having a negligible correlation with the criterion itself. Baron and Kenny (1986) defined a moderator variable as a qualitative or quantitative variable that affects the direction or strength of the relation between an independent and a dependent variable. It is a third variable that affects the zero order correlation between two other variables. The role of pupil-level moderator variables, namely, age, sex, and repetition, on the relationship between KCSE and KCPE examinations, were examined in this study. REPETITION refers to an examinee having sat for KCPE examination more than once. AGE refers to how old an examinee was in 1987 when he/she took the KCPE examination. GENDER refers to whether an examinee is male or female. GENDER was dummy coded 0 for males and 1 for females.

Population

The population in this study comprised of secondary schools that presented candidates for the 1991 KCSE examination in South Nyanza district within the Republic of Kenya. The district was chosen on the basis of convenience, for it was the most accessible one to the researcher at the time of the study. Of the 45 districts in the country, the one that had the highest mean score in 1987 KCPE had 347.15 points and the one with the lowest mean score in the same examination had 236.34 points (KCPE Newsletter, 1988). The national mean score in 1987 KCPE examination, computed from 45 district means, was 294 points with a standard deviation of 25. South Nyanza district from which the sample in this study was drawn had a mean score of 273.48 points in the 1987 KCPE examination.

Whereas records kept at the District Education Office, South Nyanza district, showed that the district had 82 secondary schools that presented candidates for the 1991

KCSE examination, 1991 KCSE examination data were available for only 51 secondary schools. Distribution of the population of 51 secondary schools by gender was 20 boys' schools, 12 girls' schools, and 19 mixed schools.

Sample

Stratified random sampling was used to select about 50% of the available 1991 KCSE data. First, three lists were prepared: one for boys' schools, one for girls' schools, and one for mixed schools. One school was selected at random from the first two schools on each list. Thereafter, every second school was picked from each list, giving a sample of 26 schools. The sample consisted of 10 boys' schools, 6 girls' schools, and 10 mixed schools with a total of 781 cases. This sample was considered to be representative of secondary schools in South Nyanza district. Results of this study are therefore generalizable to the population of 1991 KCSE examinees in South Nyanza district.

Data Collection

A year prior to presentation of candidates for KCSE examination, each secondary school in Kenya, through their respective district education offices, submits student returns to the Kenya National Examinations Council. The returns give details of examinee background information. This information includes codes that each of the candidates used in KCPE examination, KCPE examination scores, codes to be used by each of the examinees in the forthcoming KCSE examination, sex, and year of birth. Those examinees who wrote KCPE more than once had unique KCPE codes that made their identification possible.

For each secondary school, the Kenya National Examinations Council compiled and forwarded copies of 1991 KCSE examination data to the District Education Offices for filing and distribution. Each of the schools' computer printed KCSE examination data contained information on examinees' KCSE subject codes and grades, sex, and examinees' composite KCSE grades.

Two people, whose duty was to maintain examination records in the district education office, were involved in the data collection exercise. After photocopying the district's 1991 KCSE examination data, they were briefed by the researcher on how to retrieve and record the required information. They then embarked on the exercise under the guidance and supervision of the researcher who cross-checked 10% of their work. First, an examinee's full name on the 1991 KCSE examination data list was read aloud. This name was then located on the KCSE returns list that was prepared by schools one year earlier. Using the name on the returns list, the examinee's code in 1987 KCPE examination was read and recorded next to his/her name on the 1991 KCSE examination data list. This KCPE code was used to locate the examinees in the 1987 KCPE examination data list. Once located in the KCPE data list, an examinee's KCPE examination data were copied next to his/her name in the KCSE examination data list. This sequence was repeated until all the examinees' KCSE and KCPE data were matched.

Data Analysis

In this study, data analysis was done by use of ML 2 computer software. This program, unlike similar ones in the market, has extensive data manipulation facilities and provides high resolution graphics. The program is imbedded in another software package called NANOSTAT, which allows data preparation and manipulation before, during, and after modelling. Because of its integration with NANOSTAT, ML2 is one of the most flexible of the packages in the market today (Arnold, 1992). However, the different packages for multilevel analysis produce similar results (Kreft, De Leew, & Kim, 1990).

The following data analytic techniques and models were used:

a) <u>Inter-Correlation Matrix</u>

Inter-correlations between variables were calculated to help in the investigation of the degree of relationship between the pupil-level variables being studied.

b) Distributions

One of the basic requirements in regression analysis is that variables be normally distributed. A serious violation of normality may lead to spurious results. Normal plots of variables and stem-and-leaf plots of mean scores were used to check the normality assumption. Skewness and kurtosis were also calculated to find out the nature of the tails and 'peakedness' of the distributions, respectively.

c) Intra-School Correlation

In order to justify the use of multilevel analysis, the intra-school correlation over the sample of schools was computed from the formula

$$\delta = \sigma^2_0/(\sigma^2_0 + \sigma^2_e) \qquad (7)$$

where σ^2_0 is the variance of the schools' intercepts after the predictor variable has been fitted, and σ^2_e the within-school residual variation. There would be no need to proceed with multilevel analysis if the measurements within a school, but not between schools, were highly correlated (Goldstein, 1987, p.13). The magnitude of this correlation would indicate whether a portion of the variation in achievement is explained by the differences between schools.

d) Model Fitting

Five models were fitted to help in drawing validity inferences from the data. Table 1 shows the models and their uses.

Table 1

Models and their Uses

MODEL	USE
1. Variance Components (VC)	To partition the variance in KCSE grades into within- and between-school components in order to determine the extent to which mean KCSE achievement levels differ across schools.
2. Random Coefficients (RC)	To determine the relationship between KCPE and KCSE.
3. Moderator (MOD)	To determine the impact of age, gender, and repetition on the KCPE -KCSE relationship.
4. Omnibus (OB)	To determine the order of importance of moderator variables in their impact on the KCPE-KCSE relationship.
5. School Size (SS)	To account for the variation in intercepts and slopes across schools.

The following is a detailed description of how the models were fitted:

Variance Components Model

The variance components model was used to determine whether the mean KCSE achievement levels differed significantly across schools. This is a model with no independent variables, at either pupil or school level. It was used to partition the variance in the criterion variable (KCSE) into within- and between-school components. The model is:

$$Y_{ij} = b_{0j}X_0 + e_{ij} \quad (8)$$

and

$$b_{0j} = c_{00} + u_{0j} \quad (9)$$

In this model, the intercept variable X_0 has the value of 1 for every examinee. Each school has its own mean level of achievement, b_{0j}, and these school means vary about the overall mean c_{00}.

Random Coefficients Model

The random coefficients model has two parts, a within-school part and a between-school part. The within-school part relates an examinee's KCSE grade Y_{ij}, to his/her KCPE deviation score from the sample grand mean. The within-school model (i.e., pupil-level) is

$$Y_{ij} = b_{0j} + b_{1j}(X_{1ij} - \overline{X}_{1..}) + e_{ij} \quad (10)$$

where Y_{ij} is the 1991 KCSE grade for student *i* in school *j*, b_{0j} is the mean KCSE grade for school *j*, b_{1j} is the KCSE-KCPE achievement relationship (i.e., slope) in school *j*, X_{1ij} is the 1987 composite KCPE score of student *i* in school *j*, $\bar{X}_{1..}$ is the KCPE grand mean, and e_{ij} is the error of estimate for student *i* in school *j*. Deviating KCPE scores from the sample grand mean leaves each of the schools' slopes in the KCSE-KCPE relationship invariant. However, the transformation changes each of the schools' intercepts in the KCSE-KCPE relationship. The between-school part (i.e., school-level) is

$$b_{0j} = c_{00} + u_{0j} \tag{11}$$

and

$$b_{1j} = c_{10} + u_{1j} \tag{12}$$

where c_{00} is the grand mean for KCSE scores across all schools, c_{10} is the mean slope for the KCSE-KCPE relationship pooled within all schools, u_{0j} is the residual of school *j* on the KCSE mean achievement, and u_{1j} is the residual of school *j* on the KCSE-KCPE slope. The parameters of fundamental concern here are:

i) the grand mean for KCSE scores across all schools (c_{00})

ii) the variance of the schools' intercepts ($\sigma^2{}_0$)

iii) the variance of the schools' slopes ($\sigma^2{}_1$)

iv) the average KCSE performance boost contributed by the KCPE scores, i.e. the mean slope for the KCSE-KCPE relationship pooled across all schools (c_{10}).

In setting up this model, the predictor variable was added to the fixed part. The estimated mean within-school regression equation was used to test the hypothesis that $c_{10}=0$ (i.e., that the mean KCSE-KCPE slope is zero). Considering that under the null hypothesis, $c_{10}/[\text{s.e.}(c_{10})]$ has a *t* distribution (Bryk and Raudenbush, 1992), it was possible to determine whether the coefficient of KCPE in the KCSE-KCPE linear relationship could have occurred by chance alone.

Moderator Models

Selected pupil-level variables were added to the random coefficients regression model as explanatory variables, one at a time, to find out the impact of each one of them on the relationship between KCSE and KCPE. The moderator variables were GENDER, AGE, and REPETITION.

Omnibus Model

An omnibus model involving all the selected pupil-level variables was fitted, first with raw scores of explanatory variables; then with standardized scores for the same explanatory variables, giving a prediction equation of standardized weights. The prediction equation with standardized weights was used to determine the order of importance of the explanatory variables in their impact on the KCSE-KCPE relationship.

School Size Model

In an attempt to account for the variation in intercepts and slopes across schools, school size was used as an explanatory variable together with KCPE. The magnitude of estimates of parameter variance were compared with those in the random coefficients regression model to determine whether school size partly accounted for the variation in schools' intercepts and slopes.

Summary

Chapter 3 specified variables used in this study. The criterion was 1991 KCSE examination. The predictor was 1987 KCPE examination. Moderator variables that may have an impact on the relationship between KCSE and KCPE examinations were also discussed. These were age, gender, and repetition. School size was discussed as a variable that may help account for the differences in schools' intercepts and/or slopes across schools.

The population and sample used in the study were outlined. The population consisted of secondary schools that presented candidates in the 1991 KCSE examination in a single district in Kenya. Stratified random sampling was used to select 26 secondary schools whose examination data were used in the analysis. 1991 KCSE data were collected from computer printouts kept at the district education office. The data were matched with corresponding 1987 KCPE examination data to allow for analysis. The data analytic techniques employed in the study were outlined as well. The techniques included checking the assumptions and fitting the variance components model, the random coefficients regression model, the moderator model, the omnibus model, and the school size model.

The next chapter presents the results obtained from the analysis of data.

Chapter 4

RESULTS

The results given in this chapter were from an analysis of 1991 KCSE and 1987 KCPE examination data from a single district using ML 2 computer software. The distributions, correlations between variables, scatterplots, normality plots, and variability of scores are provided. The variance components model and the random coefficients regression model were used to examine KCSE-KCPE relationship, thus providing the prediction equations relating KCPE and KCSE examinations. These models were further extended to include other pupil-level and school-level variables in an attempt to find out the extent to which these variables affect the KCSE-KCPE relationship.

Distribution of the Sample by Schools and Gender

Table 2 shows the distribution of the sample by schools and gender. The sample contained ten boys' schools, six girls' schools, and ten mixed schools. However, some of the mixed schools presented candidates of the same gender. There were 190 girls and 591 boys in the sample, giving a total of 781 cases, with no missing data.. Girls formed 24% of the total sample. The proportion of girls enrolled in secondary schools in Kenya in 1990 was about 40%, which means girls were underrepresented. Out of the 26 schools in the sample, the smallest school formed 0.4% of the total sample and the largest school formed 9.0% of the total sample.

Table 2

Sample Distribution by Schools and Gender

SCHOOL ID	TYPE	BOYS	GIRLS	TOTAL	% BY TYPE
1	Boys	70	0	70	
11	Boys	46	0	46	
12	Boys	57	0	57	
19	Boys	41	0	41	
2	Boys	70	0	70	60.2
26	Boys	25	0	25	
3	Boys	42	0	42	
4	Boys	49	0	49	
6	Boys	55	0	55	
9	Boys	15	0	15	
13	Girls	0	30	30	
16	Girls	0	16	16	
23	Girls	0	21	21	20.9
25	Girls	0	34	34	
5	Girls	0	38	38	
8	Girls	0	24	24	
10	Mixed	9	6	15	
14	Mixed	16	2	18	
15	Mixed	14	1	15	
17	Mixed	7	2	9	
18	Mixed	14	5	19	18.9
20	Mixed	3	0	3	
21	Mixed	20	3	23	
22	Mixed	9	0	9	
24	Mixed	24	6	30	
7	Mixed	5	2	7	
TOTAL		591	190	781	100.0

Range of Scores

Table 3 shows minimum, maximum, and range of scores for all the examinees in the sample in three KCSE subjects, overall KCSE scores, six KCPE subjects and KCPE composite scores.

Table 3

Minimum, Maximum and Range of Scores

EXAM	MIN.	MAX.	RANGE
KCSE*			
English	1	9	8
Kiswahili	1	12	11
Mathematics	1	12	11
KCSE TOTAL TEST	2	9	7
KCPE			
English	20	95	75
Kiswahili	1	82	81
Mathematics	18	97	79
SC&A	1	88	87
GHCR	1	94	93
ACHM	1	91	90
KCPE TOTAL TEST	172	533	361

* KCSE scores are based on letter grades.
SC&A = Science and Agriculture;
GHCR = Geography, History, Civics, and Religious Education;
ACHM = Primary Art and Craft, Home Science, and Music.

The results in Table 3 indicate that there were no obvious errors in scoring or recording of data because the distributions are within expected range.

Means and Standard Deviations of Scores

Table 4 gives means and standard deviations of KCPE examination scores and three KCSE subject grades. The KCPE subject with the highest mean score was mathematics (MAT1), followed by English (ENG1) and then Kiswahili (KIS1). However, this order was reversed in KCSE, with the highest mean score in Kiswahili (KIS2), followed by English (ENG2) and then mathematics (MAT2). The overall average mark at national level in the 1987 KCPE mathematics, English, Kiswahili, and science and agriculture papers were 41%, 43%, 54% and 58% respectively (1988 KCPE Newsletter, KNEC).

Amongst the three KCPE subjects, mathematics had the highest variability followed by Kiswahili and then English. This order was maintained in the three KCSE subjects, with mathematics having the highest variability followed by Kiswahili and then English.

Table 4

Means and Standard Deviations of Scores

	PREDICTOR (KCPE)		CRITERION (KCSE)	
	Mean	(SD)	Mean*	(SD)
English	59.7%	(12.1)	4.23	(1.69)
Kiswahili	45.8%	(15.2)	4.36	(2.26)
Mathematics	61.8%	(15.4)	3.59	(2.68)
SC&A	59.4%	(12.2)	-	-
GHCR	60.8%	(12.6)	-	-
ACHM	60.2%	(13.4)	-	-
TOTAL EXAM	58.0%	(57.9)	-	-

*KCSE means are based on letter grades
SC&A = Science and Agriculture
GHCR = Geography, History, Civics, and Religious Education
ACHM = Art and Craft, Home Science, and Music.

Reliability Estimates

The reliability of the predictor and criterion were not available from the Kenya National Examinations Council.

However, an alternative was to estimate the reliabilities by use of Kuder-Richardson (KR21) formula, which is

$$KR21 = k/(k - 1)\{1 - [\mu(k - \mu)/k\sigma^2]\} \qquad (13)$$

upon making certain assumptions. In the KR21 formula, k is the number of items on the test, μ is the mean total score, and σ^2 is the total score variance. Only the 1987 KCPE mathematics examination was used in the estimation of the reliability of the composite KCPE examination. KCPE mathematics examination had 50 multiple choice items, assumed to be of equal difficulty. The sample grand mean for KCPE mathematics was 61.8%. If each item was worth a point, the grand mean of 61.8% would be equivalent to 30.9 points (i.e., half of 61.8), with an adjusted total test score variance of 59.3 (i.e., one-quarter of 15.4^2). With the crude estimates, the reliability of the 1987 KCPE mathematics examination using KR-21 was found to be 0.82. This suggests that the entire 1987 KCPE examination was likely to have reliability greater than 0.82.

The mean and variance of the 1991 KCSE mathematics examination were 4.36 and 7.17 respectively (N = 781). Considering that KCSE mathematics examination had two papers having 24 items each, and assuming that the items were of equal difficulty, KR-21 gave a reliability estimate of 0.46. The reliability of the entire 1991 KCSE examination would therefore be higher than 0.46 because the total test was longer than the mathematics sub-test, and the longer the test, the higher the reliability. This reliability estimate is large enough to allow for the use of 1991 KCSE examination data as a criterion variable.

Crocker and Algina (1986) suggested that the correlation between a predictor and a criterion is, in most cases less than or equal to the square root of the product of the reliability estimates of the predictor and criterion. Using .82 and .46 as reliabilities of the predictor and criterion respectively, an upper bound estimate of the correlation between 1987 KCPE examination and the 1991 KCSE examination is 0.61. In other words, the correlation between KCPE and KCSE examinations is not likely to exceed 0.61. This suggests that if a linear relationship exists between KCPE and KCSE examinations, then KCPE scores may account for not more than approximately 37.2% of the variability in KCSE grades.

Inter-Correlation Matrix of Variables

Table 5 shows inter-correlations between pupil-level variables. The correlation between KCPE composite scores and KCSE composite grades was 0.56. All correlations involving KCSE subjects and KCPE subjects were positive, with a range of 0.62 and a median of 0.45. These correlations are large enough to allow for meaningful regression analysis.

Table 5
Inter-Correlation Matrix of Student-Level Variables

	ENG2	KIS2	MAT2	KCSE	ENG1	KIS1	MAT1	SC&A	GHCR	ACHM	KCPE	GNDR	AGE	REP
ENG2	1.00													
KIS2	.45	1.00												
MAT2	.43	.22	1.00											
KCSE	.69	.47	.68	1.00										
ENG1	.52	.20	.34	.45	1.00									
KIS1	.26	.55	.10	.25	.27	1.00								
MAT1	.34	.02	.57	.45	.51	.05	1.00							
SC&A	.40	.12	.41	.46	.56	.19	.58	1.00						
GHCR	.37	.09	.33	.42	.54	.18	.53	.62	1.00					
ACHM	.31	.04	.38	.39	.49	.15	.54	.64	.57	1.00				
KCPE	.50	.25	.49	.56	.77	.45	.74	.81	.78	.77	1.00			
GNDR	-.08	.03	-.33	-.25	-.14	-.11	-.26	-.31	-.28	-.25	-.31	1.00		
AGE	-.30	-.10	.20	-.25	-.30	-.01	-.22	-.18	-.18	-.18	-.24	-.16	1.00	
REPT	-.03	-.05	-.05	-.04	-.04	-.06	.03	.05	.03	.04	.01	-.04	-.23	1.00

ENG2 = KCSE English;
KIS2 = KCSE Kiswahili;
MAT2 = KCSE Mathematics;
ENG1 = KCPE English;
KIS1 = KCPE Kiswahili;
MAT1 = KCPE Mathematics;
SC&A = KCPE Science and Agriculture;
GHCR = KCPE Geography, History, Civics, and Religious Education;
ACHM = KCPE Art and Craft, Home Science, and Music;
GNDR = Gender (coded 1 for girls & 0 for boys);
REPT = Repetition (coded 1 for repeaters & 0 for non repeaters).

Scatterplots

Appendix B contains scatterplots of KCSE grades against KCPE scores. The scatterplots showed a linear relationship between KCSE and KCPE scores, thus justifying the use of linear regression analysis.

Normality of Distributions

One of the basic assumptions in regression analysis is that the dependent variable be normally distributed. The stem and leaf plot of school KCSE means as well as the normal and box plots of the dependent variable reveals the degree of violation of the normality assumption. Stem and leaf plots, and normal and box plots for 1991 KCSE mean grades and 1987 KCPE mean composite scores for the 26 sample schools are in Appendix C.

The stem and leaf plots, together with the normal and box plots, indicate that the distribution of sample school means for KCPE examination was near normal. Fisher's measures of skewness and kurtosis for KCPE composite scores were -.43 and -.18 respectively, implying that the scores had a tolerable negative skewness and a tolerable heavier tail than the normal distribution.

KCSE English, Kiswahili and mathematics showed a tolerable positive skewness. The median grade for composite KCSE grades was in the centre of the box. The tails in the box plot were almost of equal length. The normal plots showed a linear trend. These suggest that the distribution of KCSE composite grades was near normal, with a tolerable positive skewness of 0.41 and Fisher's measure of kurtosis of 0.05.

School Means and Standard Deviations

Table 6 shows KCSE and KCPE means and standard deviations for each of the 26 secondary schools in the sample. A scatterplot of the school means given in Table 6 is shown in Figure 1. No influential data points (outliers) were detected.

Table 6

School Means and Standard Deviations

	SCHOOL												
	1	2	3	4	5	6	7	8	9	10	11	12	13
KCSE Mean Grade	5.16	5.10	5.38	5.53	4.58	4.42	4.57	3.83	4.40	3.60	6.52	5.23	3.73
KCSE S.D.	1.4	1.2	1.1	1.6	1.0	1.2	1.2	0.9	2.0	1.0	1.3	1.0	1.0
KCPE Mean Score	387	359	361	402	375	347	369	310	310	300	429	359	297
KCPE S.D.	42	43	29	38	32	42	53	41	62	36	33	28	43

	SCHOOL												
	14	15	16	17	18	19	20	21	22	23	24	25	26
KCSE Mean Grade	4.28	3.27	3.56	4.22	3.79	4.90	2.00	3.17	4.22	3.76	4.53	4.65	4.88
KCSE S.D.	1.1	1.1	0.8	0.6	1.0	1.3	2.0	1.0	1.7	1.5	1.1	1.4	0.8
KCPE Mean Score	316	291	284	281	255	362	269	311	324	268	337	337	323
KCPE S.D.	56	61	34	51	39	28	18	50	63	53	44	26	45

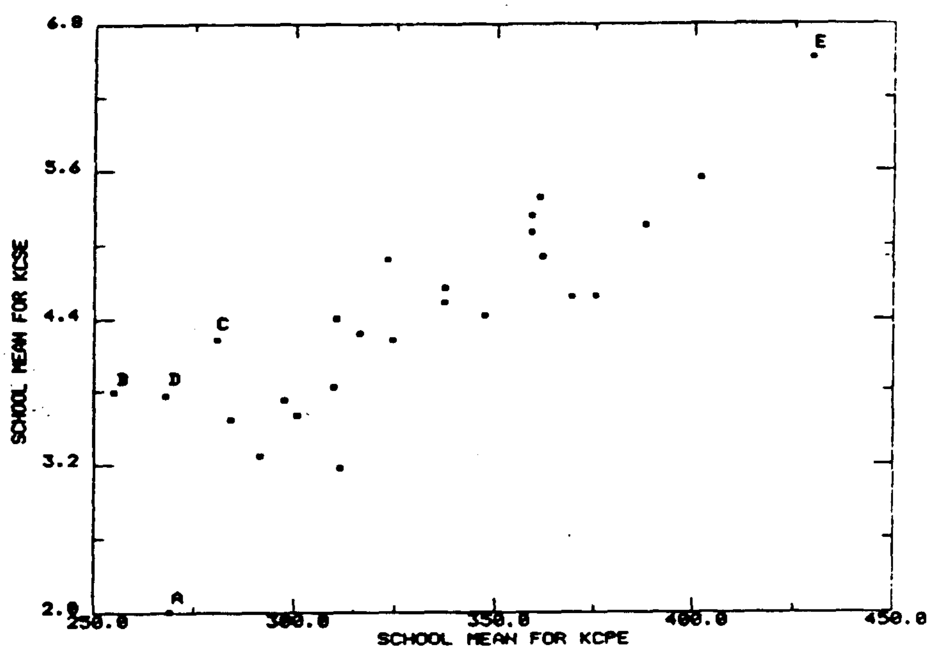


Fig. 1: Scatterplot of School Means

It is evident from the scatter-plot in Figure 1 that the school means for KCSE and KCPE have a strong positive linear correlation. This suggests that on the average, secondary schools with low KCPE achievement levels tended to have low KCSE achievement levels and those with high KCPE achievement levels tended to have high KCSE achievement levels.

Whereas there were schools like A (Kendu Muslim) and E (Kanga) whose performances were as expected (i.e. school A had a low KCPE mean (269 points) and a low KCSE mean (2.00 points) and school E had a high KCPE mean (429.39 points) and a high KCSE mean (6.52 points)), some schools like B (Kegonga), C (Arambe) and D (St. Lucy's, Raruowa) had low mean intake scores but ended up recording reasonably high mean KCSE scores as compared with other schools with low intake scores.

Predictive Models

In an attempt to address the research questions, five predictive models were fitted; the Variance Components Model (VC Model), the Random Coefficients Model (RC Model), the Moderator Model (MOD Model), the Omnibus Model (OB Model) for pupil-level variables, and the School Size Model (SS Model). These models mainly differ in the number and level of explanatory variables that are considered in a sequential order. Iterative generalized least squares convergence criterion was used in each of the models. Convergence was achieved in less than ten iterations in all cases.

Model 1: The VC Model

The VC model was used to determine whether the variation in KCSE achievement levels varied significantly across schools. The model is

$$Y_{ij} = b_{0j}X_0 + e_{ij} \qquad (14)$$

and

$$b_{0j} = c_{00} + u_{0j} \qquad (15)$$

In this model, the intercept variable X_0 has the value of 1 for every examinee. Each school has its own mean level of achievement, b_{0j}, and these school means vary about the overall mean c_{00}. The school-level residuals, u_{0j}, are the deviations of each school's mean from the district average. Y_{ij} is the response variable of KCSE scores for pupil *i* in school *j*. Table 7 shows the results of the VC Model.

Table 7

Variance Estimates in the VC Model

PARAMETER	ESTIMATE	(S.E.)	*t*	*p*
		SCHOOL LEVEL		
INTERCEPT	0.626	(0.195)	3.2	.002*
		PUPIL LEVEL		
INTERCEPT	1.535	(0.790E-01)#	19.4	.000*

0.790E-01 = 0.0790
* Significant at .01 level

The between-school intercept variance is more than three times its standard error, suggesting that the average level of achievement in KCSE differed significantly across schools ($t = 3.2$, $p < .01$). It is also clear from Table 7 that most of the variation in KCSE grades was between pupils. The maximum likelihood point estimate for the grand-mean KCSE achievement was 4.42 with a standard error of .17, indicating a 95% confidence interval of

$$4.42 \pm 1.96(.17) = (4.09, 4.75).$$

To determine if the use of multilevel regression analysis was in order, the intra-school correlation over the sample of schools was computed from the formula

$$\delta = \sigma^2_0/(\sigma^2_0 + \sigma^2_e) \quad (16)$$

where σ^2_0 is the variance of the schools' intercepts, and σ^2_e is the within-school residual variation. The intra-school correlation, whose value is $\delta = 0.29$, represents the proportion of variance in KCSE between secondary schools. This indicates that about 29% of the variance in KCSE is between schools. This is a justification for the use of multilevel analysis (Goldstein, 1987, p.13).

Model 2: The RC Model

The RC Model serves two purposes. First, it provides an estimate of the mean equation for the regression of KCSE scores on KCPE scores pooled within schools. Second, it tests the null hypothesis that the KCPE-KCSE relationship, (i.e., the KCPE-KCSE slope), does not vary across schools.

According to Raudenbush and Bryk (1986), the RC Model for determining the required KCPE-KCSE relationship may be written as

$$Y_{ij} = b_{0j} + b_{1j}(X_{1ij} - \bar{X}_{1..}) + e_{ij} \quad (17)$$

at pupil level and

$$b_{0j} = c_{00} + u_{0j} \quad (18)$$

$$b_{1j} = c_{10} + u_{1j} \quad (19)$$

at school level. Equation (17) is a within-school model that relates students' KCSE achievement to their respective KCPE scores when the KCPE scores are deviated from the grand mean. At pupil level, Y_{ij} is the response variable, the KCSE achievement for pupil *i* in school *j*, X_{1ij} is the predictor variable i.e., the KCPE achievement for pupil *i* in school *j*, $\bar{X}_{1..}$ is the KCPE grand mean, b_{0j} is the KCSE mean achievement for school *j*, b_{1j} is the KCPE-KCSE achievement relationship (i.e., KCPE-KCSE slope) for school *j*, and e_{ij} is a residual for pupil *i* in school *j*. At school level (Equations (18) and (19)), c_{00} is the grand mean for KCSE achievement across schools, c_{10} is the mean slope for the KCSE-KCPE relationship pooled across schools, u_{0j} is the effect of school *j* on the mean KCSE achievement level, and u_{1j} is the effect of school *j* on the KCSE-KCPE slope. It is assumed that u_{0j} and u_{1j} are multivariate normally distributed, both with expected values of 0 (Bryk and Raudenbush, 1992).

The model was set by having the constant vector of 1's and the deviation scores of KCPE as explanatory variables. In this regression model, both intercepts and slopes are allowed to vary across schools. Table 8 shows the coefficient estimates in the RC Model.

Table 8

Coefficient Estimates in the RC Model

PARAMETER	ESTIMATE	(S.E.)	*t*	*p*
INTERCEPT	4.647	(0.884E-01)	52.6	.000*
SLOPE	0.121E-01	(0.104E-02)	11.6	.000*

* Significant at .01 level

Using the parameter estimates in Table 8, an average within-school linear regression equation relating KCPE and KCSE scores is

$$KCSE_{ij} = 4.647 + 0.0121(KCPE_{ij} - \overline{KCPE}_{..}) \qquad (20)$$

In Equation (20), a change of 1 unit in the predictor variable (KCPE) is associated with a change of 0.0121 unit in the criterion variable (KCSE), and the higher the KCPE examination score, the higher the KCSE examination score. In other words, a change of 83 points in KCPE (based on a 600 point interval scale) is associated with a change of 1 grade point in KCSE (based on a 12 point interval scale). From Table 8, the ratio of the observed intercept to its standard error gives a significant $t = 52.6$, $p < .01$. The ratio of the observed slope to its standard error also gives a significant $t = 11.6$, $p < .01$. These results suggest that the observed intercept and slope in the KCPE-KCSE relationship could hardly have occurred by chance alone.

Therefore, on the average, there is a significant positive linear relationship between KCPE and KCSE within schools. The relationship is substantial, equivalent to a correlation of approximately 0.56.

Equation (20) can be used to predict KCSE grades from given KCPE scores. For example, an examinee who scored 350 points in KCPE was likely to score:

$4.65 + 0.012(350 - 347.9) = 4.7$ points or C- in KCSE, if all other factors were held constant. However, a student with a KCPE score of 280 points was likely to score:

$4.65 + 0.012(280 - 347.9) = 3.8$ points or D+ in KCSE, if all other factors were held constant.

Table 9 shows between-school and between-pupil parameter variances in the KCPE-KCSE relationship. The estimated parameter variance for b_{0j} (intercepts) was used to test whether the observed differences among schools in mean achievement levels could have occurred by chance alone.

Table 9

Variance/Covariance Estimates in the RC Model

PARAMETER	ESTIMATE	(S.E.)	*t*	*p*
		SCHOOL LEVEL		
INTERCEPT	0.883	(0.917)	0.96	.173
INTERCEPT/SLOPE	-.208E-02	(0.257E-02)	-.81	.213
SLOPE	0.576E-05	(0.735E-05)	0.78	.221
		PUPIL LEVEL		
INTERCEPT	1.329	(0.691E-01)	19.2	.000*

* Significant at .01 level

The ratio of the between-school intercept variance to its standard error gives $t = 0.96$. This is a non-significant t at .01 level, suggesting that the between-school variation in intercepts (KCSE mean achievements) could have occurred just by chance. The ratio of between-school slope variance to its standard error gives $t = 0.78$ which is non-significant at .01 level, implying that the relationship between KCPE and KCSE within schools (i.e., slope) does not vary significantly across the population of schools. Thus, the regression lines for the 26 schools may, for all practical purposes, be considered parallel. This suggests that the strength of the relationship between KCPE and KCSE is similar across schools. Any observed between-school differences in mean KCSE achievement levels after considering the schools' mean KCPE scores could be due to sampling error or other random factors. Similarly, any observed between-school differences in slope could be due to sampling error of other random factors. Based on the negative value of the intercept/slope covariance in Table 9, it can be inferred that the average slope in the KCPE-KCSE relationship decreases with increasing KCSE mean achievement levels.

The fact that the variation across schools in mean KCSE achievement levels is significant, other things being equal, and that the variation becomes non-significant when mean KCPE scores are considered, has a far reaching impact on policy decisions that relate to ranking of secondary schools based entirely on KCSE mean achievement levels. This issue is elaborated in Chapter 5.

Any linear relationship between two variables in a two dimensional space is completely determined if both the intercept and slope in the relationship are known. Figure 2 is a plot of slopes against intercepts for the 26 secondary schools in the sample. Since the abscissa represents intercepts and the ordinate represents slopes, it is possible to write linear regression equations for each of the 26 secondary schools.

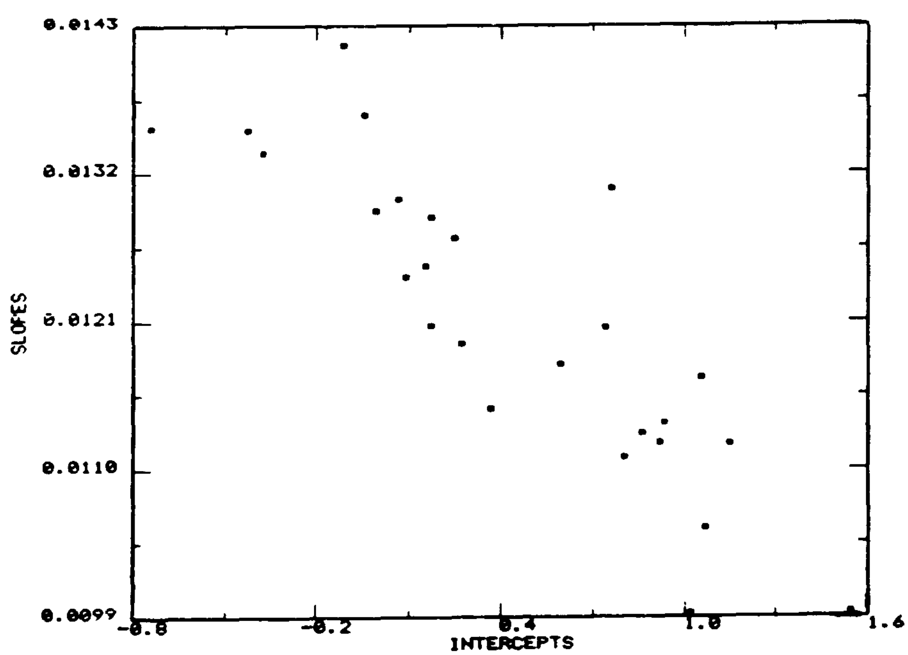


Fig. 2: Plot of Slopes against Intercepts

Individual school coordinates were identified from Figure 2 by setting the cursor on the required point, upon which the coordinates were displayed on the computer screen along with the serial number of the point. Using these coordinates, 26 regression lines representing the KCPE-KCSE relationship were drawn, one for each school. The regression lines are given in Figure 3.

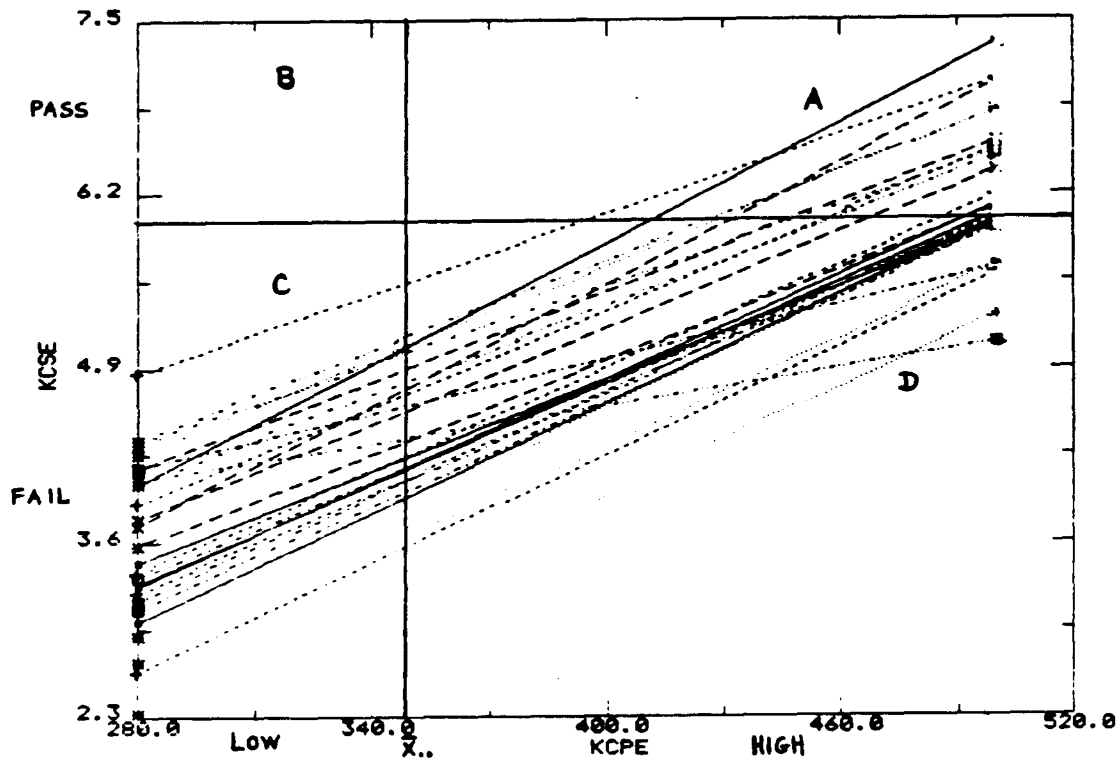


Fig. 3: Predicted KCSE grades vs KCPE scores

Quadrant A in Figure 3 represents examinees with KCPE scores above the grand mean and who succeeded in high school. Quadrant B represents examinees with KCPE scores below the KCPE grand mean who succeeded in high school. Quadrant C represents examinees with KCPE scores below the KCPE grand mean who did not succeed in high school. Quadrant D represents examinees with KCPE scores above the KCPE grand mean who did not succeed in high school.

It can be deduced from quadrant C that those examinees who had KCPE scores below the KCPE grand mean were unlikely to succeed in high school. However, a number of examinees who had KCPE scores above the KCPE grand mean did not succeed in high school as evidenced in quadrant D.

Model 3: The MOD Model

Having concluded from the VC Model that KCSE mean achievement levels differed significantly across schools, and having concluded further from the RC Model that these differences became non-significant if mean achievement levels in KCPE for secondary schools were considered, an attempt was made to determine the separate roles of three pupil-level variables as moderators in the KCPE-KCSE relationship. The three pupil-level variables are the age of examinees at the time they took their last KCPE examination, the sex of examinees, and whether an examinee wrote KCPE once or more than once. The MOD Model may be written as

$$Y_{ij} = b_{0j} + b_{1j}(X_{1ij} - \bar{X}_{1..}) + b_{2j}X_{2ij} + e_{ij} \quad (21)$$

at level 1 and

$$b_{0j} = c_{00} + u_{0j} \tag{22}$$

$$b_{1j} = c_{10} + u_{1j} \tag{23}$$

$$b_{2j} = c_{20} + u_{2j} \tag{24}$$

at level 2. At Level 1, Y_{ij} represents KCSE grades, X_{1ij} represents KCPE scores, $\overline{X}_{1..}$ represents the schools' KCPE grand mean, X_{2ij} represents the pupil-level variable which acts as a moderator in the KCPE-KCSE relationship (i.e., age, gender, or repetition), b_{0j} is the mean KCSE achievement for school j, b_{1j} is the KCPE-KCSE achievement relationship in school j, and b_{2j} represents the mean coefficient of the moderator variable for school j. At Level 2, c_{00} represents the overall mean KCSE achievement for all schools, u_{0j} is the residual of school j on the mean KCSE achievement, u_{1j} is the residual of school j on the mean slope in the KCPE-KCSE relationship, c_{10} is the mean slope for the KCPE-KCSE relationship pooled for all schools, and c_{20} represents the mean coefficient of the moderator variable pooled across all schools.

Age as a Moderator

The MOD Model was set by adding AGE as an explanatory variable in the within-school KCPE-KCSE relationship. Therefore, there were three explanatory variables in the model; the constant vector of 1's, KCPE scores, and age. The response variable was KCSE grades. Table 10 shows the parameter estimates for the MOD Model with X_2 = AGE.

Table 10

Coefficient Estimates in the MOD Model with X_2 = AGE

PARAMETER	ESTIMATE	(S.E.)	*t*	*p*
INTERCEPT	7.110	(0.428)	16.6	.000*
KCPE	0.109E-01	(0.112E-02)	9.73	.000*
AGE	-.163	(0.274E-01)	-6.0	.000*

* Significant at .01 level

Using Table 10, the equation that represents within-school KCSE-KCPE relationship when age is taken into account is:

$$KCSE_{ij} = 7.11 + 0.011(KCPE_{ij} - \overline{KCPE}_{..}) - 0.163(AGE)_{ij} \quad (25)$$

Under this model, the ratio of the coefficient of AGE to its standard error gives a significant $t = -6.0$, $p < .01$. The coefficient of KCPE under this model also gives a significant $t = 9.73$, $p < .01$, implying that the coefficient of KCPE could not have occurred by chance alone. Consideration of age has, however, lowered the KCSE-KCPE slope from 0.012 to 0.011, a drop of approximately 8.3%. In other words, consideration of AGE as a moderator variable has changed the KCPE-KCSE relationship. As an illustration, consider two pupils aged 15 and 18 respectively who had the same KCPE score, and who were attending the same secondary school. Equation (25) tells us that, on the average, the pupil aged 15 was likely to score 0.5 points more in KCSE than the one aged 18. It is clear from the equation that in general, older pupils with the same KCPE grades as their younger counterparts attending the same secondary school tended to perform poorer than their younger counterparts in the KCSE examination.

Gender as a Moderator

In an attempt to establish how gender moderates the KCPE-KCSE relationship, the MOD Model was fitted with X_2 = GENDER. GENDER was dummy coded 0 for boys and 1 for girls. Table 11 shows the parameter estimates for this model at convergence using the IGLS approach.

Table 11

Coefficient Estimates in the MOD Model with X_2 = GENDER

PARAMETER	ESTIMATE	(S.E.)	*t*	*p*
INTERCEPT	4.712	(0.938E-01)	50.2	.000*
KCPE	0.120E-01	(0.106E-02)	11.3	.000*
GENDER	-.245	(0.168)	-1.5	.078

* Significant at .01 level

Using Table 11, an equation that represents the within-school KCPE-KCSE relationship when the sex of examinees is considered is:

$$KCSE_{ij}=4.71+0.012(KCPE_{ij} - \overline{KCPE}_{..})-0.245(GENDER)_{ij} \qquad (26)$$

In this model, the coefficient of KCPE has not changed much from what it was under the RC Model. The ratio of the coefficient of GENDER to its standard error gives $t = -1.46$ which is non-significant at .01 level. It can therefore be inferred from Equation (26) that if you compare the KCSE mean achievement levels for boys and girls having the same KCPE grades, and who attended secondary schools having the same KCPE mean achievement levels, boys tended to outperform girls in KCSE by approximately 0.25 points, although this could have occurred by chance alone.

Repetition as a Moderator

In an attempt to determine how repetition moderates the KCPE-KCSE relationship, the pupil-level variable REPETITION was substituted for X_2 in the MOD Model. This gives an equation in which the response variable is a linear function of KCPE and REP. The variable REP was coded 0 for non-repeaters and 1 for repeaters. Table 12 shows fixed parameter estimates for this model.

Table 12

Coefficient Estimates in the MOD Model with X_2 = REP

PARAMETER	ESTIMATE	(S.E.)	*t*	*p*
INTERCEPT	4.682	(0.920E-01)	50.9	.000*
KCPE	0.121E-01	(0.107E-02)	11.3	.000*
REPETITION	-.243	(0.117)	-2.1	.025

* Significant at .01 level

Using Table 12, an equation that represents the linear relationship involving KCSE as a criterion variable and KCPE as a predictor, after taking into account repetition of examinees is:

$$KCSE_{ij} = 4.68 + 0.012(KCPE_{ij} - \overline{KCPE}_{..}) - 0.243(REP)_{ij} \quad (27)$$

Equation (27) shows that the inclusion of REP i.e. REPETITION, as a moderator variable hardly changed the KCPE-KCSE relationship. The coefficient of REP gives $t = -2.1$ which is non-significant at .01 level. Holding other factors constant, repeaters of Class 8 were associated with lower KCSE mean achievement levels than non-repeaters, although this could have occurred by chance alone.

Model 4: The OB Model

Under the OB Model, all the selected pupil-level variables were fitted simultaneously. Two OB Models were fitted for the explanatory variables KCPE, AGE, GENDER, and REPETITION. The first employed raw scores of explanatory variables. This model may be written as:

$$Y_{ij} = b_{0j} + b_{1j}X_{1ij} + b_{2j}X_{2ij} + b_{3j}X_{3ij} + b_{4j}X_{4ij} + e_{ij} \quad (28)$$

at pupil-level, with

$$b_{pj} = c_{p0} + u_{pj}, \; p = 0,\ldots,4, \; j = 1,\ldots,26 \quad (29)$$

at school-level. In Equation (28), Y_{ij} is the KCSE grade for pupil i in school j, $b_1,\ldots,b_4$ are coefficients of $X_1,\ldots,X_4$ respectively, $X_1,\ldots,X_4$ are the four pupil-level variables KCPE, GENDER, AGE, and REPETITION, respectively, and e_{ij} is the error term. b_{0j} is a constant representing the intercept. At the school-level (Equation 29), each of the b-coefficients has a grand mean of c_{p0}, and a residual u_{pj}. Table 13 shows the results of the OB Model for raw scores.

Table 13

Coefficient Estimates in the OB Model with Raw Scores

PARAMETER	ESTIMATE	(S.E.)	*t*	*p*
INTERCEPT	3.53	(0.650)	5.43	.000*
KCPE	0.107E-01	(0.114E-02)	9.39	.000*
AGE	-.164	(0.284E-01)	-5.8	.000*
GENDER	-.389	(0.175)	-2.2	.018
REP	-.862E-01	(0.118)	-.73	.236

* Significant at .01 level

Using the results in Table 13, a prediction equation for KCSE grades from raw scores of explanatory variables is:

$$KCSE_{ij} = 3.53 + 0.011(KCPE_{ij}) - 0.164(AGE_{ij}) - 0.389(GENDER_{ij}) - 0.086(REP_{ij}) \quad (30)$$

The second OB Model was fitted in order to determine the order of importance of the selected pupil-level explanatory variables in predicting KCSE grades. In this case, the variables KCSE, KCPE, GENDER, AGE, and REPETITION were standardized before running the model. The OB Model of standardized regression coefficients is given by:

$$KCSE_{ij}^{*} = \beta_{0j} + \beta_{1j}(KCPE)_{ij} + \beta_{2j}(GENDER)_{ij} + \beta_{3j}(AGE)_{ij} + \beta_{4j}(REPETITION)_{ij} + e_{ij} \quad (31)$$

at pupil-level, with

$$\beta_{pj} = c_{p0} + u_{pj}, \; p = 0,\ldots,4, \; j = 1,\ldots,26 \quad (32)$$

at school-level.

In Equation (31), $KCSE_{ij}^{*}$ represents the standardized KCSE grades, $\beta_{1j},\ldots,\beta_{4j}$ represent standardized coefficients of KCPE, GENDER, AGE, and REPETITION, respectively, and e_{ij} is the error term. At the school-level (Equation 32), each of the β-coefficients has a grand mean c_{p0}, and a residual u_{pj}. Table 14 shows the results of this analysis.

Table 14

Coefficient Estimates in the OB Model using Standardized Scores

PARAMETER	ESTIMATE	(S.E.)	*t*	*p*
INTERCEPT	-.043	(0.069)	-0.62	.270
KCPE	0.427	(0.046)	9.28	.000*
GENDER	-.114	(0.051)	-2.24	.017
AGE	-.181	(0.031)	-5.84	.000*
REPETITION	-.021	(0.029)	-0.72	.239

Note: The coefficient estimates are standardized.
* Significant at .01 level

Using the results in Table 14, a prediction equation for standardized KCSE grades using standardized scores of the predictor and three pupil-level explanatory variables is:

$$Y_{ij}^{*} = -.043 + 0.427(KCPE_{ij}^{*}) - 0.114(GENDER_{ij}^{*}) - 0.181(AGE_{ij}^{*}) - 0.021(REPETITION_{ij}^{*}) \quad (33)$$

where * signifies a standardized variable. In equation (33), the pattern of standardized weights indicates that of the three moderator variables considered in the KCPE-KCSE relationship, AGE had the greatest impact followed by GENDER and REPETITION, in this order.

Model 5: The SS Model

An attempt was made to account for the observed variation in mean achievement levels (intercepts) and slopes across schools. School size, a school level variable, was used in this attempt. Small schools were coded 1, medium schools were coded 2, and large schools were coded 3. In order to set up the model, a cross product of KCPE scores and secondary school size was created. This product was named KCPE*SIZE. Table 15 shows the coefficient estimates for the SS Model.

Table 15

Variance/Covariance Estimates in the SS Model

PARAMETER	ESTIMATE	(S.E.)	*t*	*p*
	SCHOOL LEVEL			
INTERCEPT	0.699	(0.855)	0.82	.210
INTERCEPT/SLOPE	-.167E-02	(0.243E-02)	-.69	.248
SLOPE	0.484E-05	(0.702E-05)	0.69	.248
	PUPIL LEVEL			
INTERCEPT	1.33	(0.689E-01)	19.3	.000*

* Significant at .01 level

In Table 15, the parameter variance of schools' intercepts dropped from 0.883 as given under the RC Model (see Table 8) to 0.699 in the SS Model, a drop of 16%. The slope variance also dropped from 0.576E-05 as given under the RC Model to 0.484E-05 in the SS Model, a drop of 16%. These reductions suggest that variability in schools' KCSE achievement means for students with low/average KCPE scores is partly accounted for by secondary school size.

Table 16 shows the coefficient estimates in the SS Model.

Table 16

Coefficient Estimates in the SS Model

PARAMETER	ESTIMATE	(S.E.)	*t*	*p*
INTERCEPT	4.372	(0.238)	18.4	.000*
KCPE	0.143E-01	(0.262E-02)	5.46	.000*
SIZE	0.664	(0.453)	1.47	.077
KCPE*SIZE	-.143E-02	(0.130E-02)	-1.10	.141

* Significant at .01 level

The coefficient of SIZE in Table 16 suggests that an increase in secondary school size by a single stream is associated with an increase in KCSE mean achievement levels of approximately 0.66 points. However, based on the negative coefficient of KCPE*SIZE, an increase in secondary school size by a single stream is associated with a drop of approximately 0.001 units in the KCSE-KCPE slope, implying a drop in the benefit from KCPE.

KCPE MATHEMATICS AS A PREDICTOR OF KCSE MATHEMATICS

In this section, two models were fitted to help determine the extent to which achievement in KCPE mathematics examination predicts achievement in KCSE mathematics examination. The two models are the variance components model and the random coefficients regression model. The role of three pupil-level variables as moderators in the KCPE-KCSE relationship in mathematics was also analyzed.

Variance Components Model

A VC Model was fitted for KCSE Mathematics (MAT2) as a response variable. This model helps in finding out whether achievement levels in KCSE mathematics vary across secondary schools. The model is given by

$$\text{MAT2}_{ij} = b_{0j} + e_{ij} \qquad (34)$$

$$b_{0j} = c_{00} + u_{0j} \qquad (35)$$

Under this model, each school has its own mean level of KCSE mathematics achievement, b_{0j}, and these school means vary about the overall mean in mathematics achievement, c_{00}. u_{0j} is the residual at school level. Table 17 shows the parameter variances in this model.

Using Table 17, the ratio of the between-school intercept variance to its standard error gives $t = 3.2$, suggesting that mean achievement levels in KCSE mathematics differed significantly from school to school. The maximum likelihood point estimate for the grand-mean KCSE mathematics achievement was 3.03 with a standard error of .03, indicating a 95% confidence interval of

$$3.03 \pm 1.96(.03) = (2.97, 3.09).$$

Table 17

Variance Estimates in the VC Model for KCSE Mathematics

PARAMETER	ESTIMATE	(S.E.)	*t*	*p*
		SCHOOL LEVEL		
INTERCEPT	2.104	(0.653)	3.22	.002*
		PUPIL LEVEL		
INTERCEPT	4.888	(0.252)	19.4	.000*

* Significant at .01 level

The RC Model

A random coefficients regression model was fitted for mathematics scores. This model is given by:

$$MAT2_{ij} = b_{0j} + b_{1j}(MAT1_{ij} - \overline{MAT1}_{..}) + e_{ij} \qquad (36)$$

at pupil level and

$$b_{0j} = c_{00} + u_{0j} \qquad (37)$$

$$b_{1j} = c_{10} + u_{1j} \qquad (38)$$

at school level.

Table 18 shows parameter estimates in the RC Model for achievement in mathematics.

Table 18

Parameter Estimates in the RC Model

PARAMETER	ESTIMATE	(S.E.)	t	p
INTERCEPT	3.200	(.203)	15.8	.000*
MAT1	.720E-01	(.828E-02)	8.70	.000*

* Significant at .01 level

Based on the parameter estimates in Table 18, the relationship between KCPE mathematics and KCSE mathematics is:

$$MAT2_{ij} = 3.20 + 0.07(MAT1_{ij} - \overline{MAT1}_{..}) \qquad (39)$$

The ratio of the slope to its standard error is t = 8.7, suggesting that the slope is significantly different from zero at .01 level. It may be concluded that KCPE Mathematics is a moderate predictor of KCSE Mathematics.

The parameter variances in the RC Model are given in Table 19. The between-school intercept variance in KCSE mathematics after considering the effect of KCPE mathematics was 0.584 (S.E = 0.914), giving a t-value of 0.64 which is non-significant at .01 level. This implies that whereas schools differed significantly in KCSE mean mathematics grades, these differences ceased to be significant when the schools' KCPE mean mathematics scores were considered, and any observed differences could have been due to chance.

Table 19

Variance/Covariance Estimates

PARAMETER	ESTIMATE	(S.E.)	*t*	*p*
	SCHOOL LEVEL			
INTERCEPT	0.584	(0.914)	0.64	.264
INTERCEPT/SLOPE	-.241E-01	(0.202E-01)	-1.19	.123
SLOPE	0.837E-03	(0.467E-03)	1.79	.043
	PUPIL LEVEL			
INTERCEPT	3.925	(.2041)	19.2	.000*

* Significant at .01 level

Gender as a Moderator

The equation relating KCSE mathematics scores and KCPE mathematics scores when GENDER is considered is:

$$MAT2_{ij} = 3.33 + 0.07(MAT1_{ij} - \overline{MAT1}_{..}) - 0.35(GENDER_{ij}) \quad (40)$$

GENDER was coded 0 for boys and 1 for girls. Equation (40) therefore suggests that for pupils with the same KCPE mathematics scores and attending schools with the same KCPE mean mathematics score, boys tended to outperform girls in the KCSE mathematics examination by approximately 0.35 points.

Age as a Moderator

The contribution of age of an examinee at the time of taking KCPE in the KCSE-KCPE Mathematics relationship can be deduced from the ML2 regression equation:

$$MAT2_{ij} = 6.10 + 0.07(MAT1_{ij} - \overline{MAT1}_{..}) - 0.19(AGE)_{ij} \quad (41)$$

Equation (41) suggests that for pupils with the same KCPE score attending secondary schools with the same KCPE mean mathematics score, younger pupils tended to outperform their older counterparts in KCSE mathematics.

Repetition as a Moderator

Equation (42) is a regression equation involving KCSE Mathematics scores, KCPE Mathematics scores, and whether a candidate sat for KCPE once or more than once.

$$MAT2_{ij} = 3.34 + 0.07(MAT1_{ij} - \overline{MAT1}_{..}) - 0.55(REP)_{ij} \quad (42)$$

Considering that REPETITION was dummy coded 1 for repeaters and 0 for non-repeaters, it is clear from Equation 42 that for constant KCPE mathematics scores, KCSE mathematics scores were higher for non-repeaters than for repeaters by approximately 0.55 of a point.

KCPE KISWAHILI AS A PREDICTOR OF KCSE KISWAHILI

The two basic models used in the prediction of KCSE mathematics scores from KCPE mathematics scores were used in the prediction of KCSE Kiswahili scores from KCPE Kiswahili scores.

The VC Model

A variance components model was fitted for Secondary School Kiswahili (KIS2). This model is given by:

$$KIS2_{ij} = b_{0j} + e_{ij} \qquad (43)$$

and

$$b_{0j} = c_{00} + u_{0j} \qquad (44)$$

where $KIS2_{ij}$ represents KCSE Kiswahili grades, b_{0j} is the KCSE Kiswahili mean grade for school j, c_{00} is the overall mean for KCSE Kiswahili, e_{ij} is the residual at pupil level and u_{0j} is the residual at school level. This model helps in finding out whether achievement levels in KCSE Kiswahili varied significantly across schools. Table 20 shows the results for the variance components model for KCSE Kiswahili.

Table 20

Variance Estimates in the VC Model

PARAMETER	ESTIMATE	(S.E.)	t	p
		SCHOOL LEVEL		
INTERCEPT	1.281	(0.408)	3.14	.002*
		PUPIL LEVEL		
INTERCEPT	3.856	(0.198)	19.47	.000*

* Significant at .01 level

The between-school parameter variance of 1.281 (S.E.=0.408) gives $t = 3.14$, $p < .01$. This indicates that there was significant variation in KCSE Kiswahili achievement levels across schools. The maximum likelihood point estimate for the grand-mean KCSE Kiswahili achievement was 4.22 points with a standard error of .24, indicating a 95% confidence interval of

$$4.22 \pm 1.96(.24) = (3.75, 4.69).$$

The RC Model for Kiswahili

A random coefficients regression model was fitted for KCSE Kiswahili. This model is given by:

$$\text{KIS2}_{ij} = b_{0j} + b_{1j}(\text{KIS1}_{ij} - \overline{\text{KIS1}}_{..}) + e_{ij} \qquad (45)$$

at pupil level and

$$b_{0j} = c_{00} + u_{0j} \qquad (46)$$

$$b_{1j} = c_{1j} + u_{1j} \qquad (47)$$

at school level. Table 21 shows the fixed coefficients for KCSE Kiswahili as a function of KCPE Kiswahili.

Table 21

Coefficient Estimates for Kiswahili

PARAMETER	ESTIMATE	(S.E.)	*t*	*p*
INTERCEPT	4.28	(0.159)	26.92	.000*
SLOPE	0.715E-01	(0.749E-02)	9.55	.000*

* Significant at .01 level

The coefficient estimates under this model enable us to determine a linear relationship between KCPE Kiswahili (KIS1) and KCSE Kiswahili (KIS2). The equation is:

$$KIS2_{ij} = 4.28 + 0.07(KIS1_{ij} - \overline{KIS1}_{..}) \qquad (48)$$

The ratio of the slope to its standard error in the KCPE-KCSE Kiswahili relationship gives a significant $t = 9.55$, $p < .01$. KCPE Kiswahili is therefore a moderate predictor of KCSE Kiswahili.

Table 22 shows variance/covariance estimates for the RC Model for Kiswahili. The ratio of the intercept variance to its standard error gives a non-significant $t = 1.45$ at .01 level, suggesting that the variation of intercepts in the relationship of KCPE-KCSE Kiswahili could have occurred by chance.

Table 22

Variance Estimates in the RC Model for Kiswahili

PARAMETER	ESTIMATE	(S.E.)	*t*	*p*
		SCHOOL LEVEL		
INTERCEPT	0.773	(0.532)	1.45	.080
INTERCEPT/SLOPE	-.215E-01	(0.136E-01)	-1.58	.063
SLOPE	0.803E-03	(0.386E-03)	2.08	.024
		PUPIL LEVEL		
INTERCEPT	.905	(0.152)	19.11	.000*

* Significant at .01 level

Similarly, the ratio of the slope variance to its standard error gives a non-significant t = 2.08 at .01 level.

Gender as a Moderator in the Kiswahili Relationship

Parameter estimates of coefficients were used in the determination of the effect of gender on the KCSE-KCPE relationship in Kiswahili. Equation (49) shows this relationship.

$$KIS2_{ij} = 4.15 + 0.07(KIS1_{ij} - \overline{KIS1}_{..}) + 0.46(GENDER_{ij}) \quad (49)$$

With GENDER coded 0 for boys and 1 for girls, it can be deduced from equation (49) that for examinees with the same KCPE grades in Kiswahili, girls tended to outperform boys in KCSE Kiswahili by approximately 0.46 points.

Age as a Moderator

The impact of AGE in the KCSE-KCPE Kiswahili relationship can be observed from the computed regression equation given by:

$$KIS2_{ij} = 7.29 + 0.07(KIS1_{ij} - \overline{KIS1}_{..}) - 0.20(AGE)_{ij} \quad (50)$$

This equation suggests that for examinees with the same KCPE Kiswahili score attending secondary schools having the same KCPE Kiswahili mean grade, younger examinees tended to perform much better than older examinees in KCSE Kiswahili.

Repetition as a Moderator

Parameter estimates for fixed coefficients were used to write equation (51). This equation represents the relationship between KCSE Kiswahili (KIS2) and KCPE Kiswahili (KIS1) after considering whether an examinee sat for KCPE once or more than once before joining secondary school.

$$KIS2_{ij} = 4.30 + (KIS1_{ij} - \overline{KIS1}_{..}) - 0.12(REP_{ij}) \quad (51)$$

REPETITION was coded 0 for non-repeaters and 1 for repeaters, implying that for pupils with the same KCPE Kiswahili score attending secondary schools with the same KCSE mean grade in Kiswahili, non-repeaters outperformed repeaters in the KCSE Kiswahili examination.

OB Model for Composite Exam, Mathematics and Kiswahili

Table 23 shows the coefficient estimates in the OB Model for composite scores, mathematics scores, and Kiswahili scores. The predictive validity of the composite examination was moderate. Similarly the predictive validities of mathematics and Kiswahili sub-tests were also moderate.

Table 23

Coefficient Estimates in the OB Model

	KCPE	AGE	GENDER	REP	VALIDITY
COMPOSITE	.011	-.164	-.389	-.086	MODERATE
MATHS	.026	-.190	-.706	-.072	MODERATE
KISWAHILI	.071	-.200	0.365	-.116	MODERATE

Summary of Results

The data used in this study were for 781 examinees in 26 secondary schools drawn from one district in Kenya whose 1987 KCPE results were most representative in the province where this district is located. Scatterplots showed that the relationship between achievement in KCPE and achievement in KCSE could be approximated by a linear function. Histograms and normal plots for the predictor and criterion variables suggested that the distribution of scores were appropriate for regression analysis. The inter-correlation between variables were moderate to large. A plot of KCSE school means against KCPE school means revealed no influential data points. Thus, the data were such that multiple regression analysis could be meaningfully used to study the prediction relationships.

Several models were fitted to find out the variation of achievement levels across schools as well as the relationship between KCPE and KCSE examinations. The models revealed that secondary schools differed significantly in KCSE achievement levels but this difference became statistically non-significant when the schools' mean achievement levels in KCPE were considered. The KCSE-KCPE slope varied from school to school although this variation remained non-significant. This implies that schools' regression lines for predicting KCSE from KCPE are, for all practical purposes, parallel.

The role of pupil-level and school-level variables on the KCSE-KCPE relationship was also examined. From the results of KCSE-KCPE relationship, younger examinees had higher achievement levels than their older counterparts, boys outperformed girls, and repetition had a negative effect on the predictive relationship of KCPE and KCSE. Whereas the variation across schools in mean achievement and in regression slopes were non-significant, secondary school size helped in explaining the observed differences across schools. It was found that the bigger the school, the higher the achievement mean and the lower the slope in the KCSE-KCPE relationship.

In mathematics, boys had a higher achievement level than girls, younger examinees outperformed their older counterparts, and repeaters performed below non-repeaters. In Kiswahili, girls had a higher achievement level than boys, younger examinees outperformed their older counterparts, and repeaters of Class 8 performed below non-repeaters. Composite KCPE was found to be a valid predictor of composite KCSE. However, the validity was only moderate, with a correlation of .56 between the two composite examinations. Similarly, KCPE mathematics and Kiswahili were found to be valid predictors of KCSE mathematics and Kiswahili respectively. The next chapter discusses the principal findings of the study.

Chapter 5

DISCUSSION

This final chapter presents important findings that address the research questions in this study. The implications of these findings in the context of educational policy that may be useful to decision makers are discussed. The findings are considered important in prediction research because the study involved the use of hierarchical linear modelling, a linear regression approach that takes care of differences between units at one level (pupils in this case) and differences between units at another level (schools in this case).

Validity of KCPE

It was found that KCPE scores had a moderate positive linear relationship with KCSE grades, with a correlation of 0.56 between them. The proportion of variance in KCSE grades explained by KCPE scores was, therefore, 31%. In other words, 31% of the variance in pupils' KCSE grades may be attributed to differences among the pupils' KCPE scores.

The variance components model revealed a significant variation in KCSE grades across schools. However, when the schools' mean achievement levels in KCPE were considered, the variation across schools in KCSE means became non-significant. The slope in the KCSE-KCPE relationship of 0.012 was found to be significant. A change of 83 points in KCPE (measured on a 600 equal interval scale) was found to be associated with a change of 1 grade in KCSE (measured on a 12 point equal interval scale). Secondary schools with low KCSE mean grades were generally associated with low KCPE mean scores; those with high KCSE mean grades were generally associated with high KCPE mean scores.

The between-school variation in the KCSE-KCPE relationship (slope) was non-significant, suggesting that other than differences due to chance, the KCPE-KCSE relationship (slope) was more or less the same for the population of schools from which the sample was drawn. A statistically significant KCSE-KCPE slope, a non-significant variation in KCSE-KCPE relationship (slope) across schools, and a moderate positive linear correlation between KCPE and KCSE, are indicators that KCPE is a valid predictor of KCSE, although the validity is only moderate.

Whereas there were no significant differences in KCSE-KCPE slopes across secondary schools, schools with low KCSE means tended to have higher slopes than schools with high KCSE means. The finding that schools with higher KCSE means benefitted less from KCPE than schools with lower KCSE means could be attributed to the regression effect. Unless r=1.0 or -1.0, all predictions of a criterion variable from a predictor variable involve a regression toward the mean. In other words, pupils in secondary schools whose KCPE means are lower tend to have more room for improvement than their counterparts in schools with higher KCPE means. Another possible explanation is that pupils in schools with lower

KCPE means are usually aware that they are academically weak. For this reason, they tend to work much harder in order to succeed.

Reliability affects validity coefficients. Since the predictor and criterion were not perfectly reliable, the moderate validity found in the KCPE-KCSE relationship is a lower bound.

Influence of Age

The results showed that the influence of age at which KCPE was taken on the KCPE-KCSE relationship was significant. As a moderator variable, it was associated with a drop in the KCSE-KCPE slope from 0.012 to 0.011, a drop of 8.3%. This means that the use of age as a moderator variable tends to weaken the relationship between KCPE and KCSE, thereby making prediction poorer. Results further showed that younger pupils generally had higher achievement levels than older pupils.

Influence of Gender

Gender had a non-significant influence on the KCPE-KCSE relationship (slope). However, boys generally showed higher mean achievement levels in KCSE than girls, with boys scoring an average of 0.245 points above girls. Boys also outperformed girls in KCSE mathematics, scoring an average of 0.35 points above girls. However, girls turned out to be better in KCSE Kiswahili than boys, scoring an average of

0.46 points above boys. The finding that boys excelled in KCSE mathematics over girls is similar to that of Fennema and Leder (1990) who found that from the beginning of secondary schooling, boys frequently outperformed girls in mathematics.

Influence of Repetition

Repetition of Class 8 had a non-significant influence on the KCPE-KCSE relationship (slope). However, non-repeaters had higher mean achievement levels (intercepts) than repeaters. Non-repeaters scored an average of 0.24 points above repeaters in KCSE. This trend was maintained in KCPE-KCSE mathematics and KCPE-KCSE Kiswahili relationships. Non-repeaters of Class 8 scored an average of 0.55 points above repeaters in KCSE mathematics and 0.11 points above repeaters in KCSE Kiswahili.

The result that repetition does not help in improving achievement levels is similar to the findings of Niklason (1987).

Influence of Secondary School Size

The variability in schools' KCSE achievement means for average KCPE students was found to be partly accounted for by secondary school size. Similarly, the variability in schools' slopes was found to be accounted for in part by school size. Larger schools had higher mean achievement levels in KCSE (intercepts) than smaller schools, although the difference was not significant. This could be attributed to the fact that larger schools tend to have more established physical and human resources as compared to what is normally found in smaller schools. The larger established schools do receive more financial aid from the government in the form of grants than the smaller, less established ones. Additionally, such "maintained" schools select some of the best candidates, leaving only mediocre candidates for the "non-maintained" schools. "Maintained" schools are schools that receive financial aid from the central government in the form of grants. This further adds to the observed differences in mean achievement levels between the larger and smaller schools. Therefore, such secondary schools are not comparable to other schools because selection is already biased.

Another possible explanation for the differences in mean achievement levels between larger and smaller schools is that larger schools are normally popular choices for pupils who expect to score higher in KCPE. These pupils usually have the notion, which turns out to be true, that larger schools are more established, and that this may increase their chances of succeeding in secondary school. Most of the larger schools therefore tend to select pupils with higher KCPE scores than smaller schools.

The finding that larger secondary schools tend to have higher mean achievement levels than smaller schools is similar to the finding of Kimble (1976). In summary, the predictive validity of KCPE was found to be moderate, with a correlation of 0.56 between them.

Implications for Educational Policy

This study has revealed that KCPE is a valid predictor of KCSE, although the validity is only moderate. This is a desirable outcome because secondary education is supposed to cater not only for the interests of those who may want to go for further education, but also for the majority of candidates who opt to join the job market after completion of secondary school. For this reason, the use of KCPE as a selection instrument for secondary school admission may be continued. However, it is worth noting that out of the three pupil-level variables that were considered as possible moderators in the KCPE-KCSE relationship, only age had a significant influence on the relationship. It may therefore be necessary to ensure that pupils who are in a particular grade level, on the average, do not differ significantly in chronological age. This may improve the predictive validity of KCPE.

KCPE mathematics and KCPE Kiswahili were also found to have moderate predictive validity with respect to KCSE mathematics and KCSE Kiswahili, respectively.

The fact that secondary schools showed significant variations in KCSE means, and that these variations became non-significant when the schools' KCPE means were considered, has important policy implications in the manner in which secondary schools are normally ranked based only on their KCSE means. Such ranking may be spurious, and it is recommended that this be discontinued or be used with caution. Rather, each secondary school's KCPE mean should be considered as a base line before ranking of secondary schools is done. For example, in the event that two secondary schools have the same KCSE mean but different KCPE means, the one with the higher KCPE mean is likely to be of lower rank than the one with the lower KCPE mean. In such a situation, the one with a lower KCPE mean may be deemed as having attained a higher achievement gain than the one with a higher KCPE mean.

The fact that boys generally showed higher achievement levels than girls could be an indicator that the curriculum or the examinations are biased in favour of boys. This in effect may jeopardize government efforts in its struggle towards providing equal educational opportunities to both males and females. Alternatively, it may be that education for females is simply lagging behind that for males, particularly in science education. There seems to be a belief, very strongly embedded in the system, that mathematics and science are not suitable for girls and are not reconcilable with the demands of family life (Ndunda, 1990). Such negative attitudes towards girls ability by teachers and society in general may hinder girls progress in school.

Repetition of Grade 8 seemed not to improve achievement in secondary school. This could be explained by the fact that those who repeated grade 8 were, in most cases, academically weak. Repetition of Grade 8 may therefore have helped such students to gain entry into secondary school, but with the consequences that the same students, on the average, may end up failing in secondary school. These outcomes suggest that the policy that prohibits repetition at primary school level should be further enforced in order to make good use of the scarce resources, because as found in the present study, repetition of Class 8 hardly improved mean achievement levels in secondary school.

Based on the outcome that larger schools tended to outperform smaller schools, it is recommended that rather than build more smaller schools that are ill equipped, priority attention may be given to the expansion of existing facilities in smaller schools, with a view to increasing their respective enrollments to a capacity of three streams. This is because in this study, schools with three streams outperformed those with either one or two streams. Such a move will be economically prudent for Kenya.

Weaknesses and Strengths of the Study

Researchers in the social sciences are rarely able to obtain information on all the factors that may influence predictor-criterion relationship. Using predictors as explanatory variables may not account for all the variance in KCSE. Whenever use of nonexperimental designs is made, and the predictor variables fail to account for all of the variance in the criterion, inferences about the effects of the predictors on that criterion may be biased. The present study is no exception.

A number of psychometric characteristics could have affected the results in this study. Of major concern are the effects of restriction of range and criterion contamination. The effect of restriction of range arises from the fact that those examinees who failed the predictor were not part of the sample in the study. This is a problem that is difficult to eliminate in a predictive validity study of this nature. Restriction of range tends to lower the correlation coefficient between the predictor and the criterion variables. Criterion contamination was seemingly inevitable because most teachers knew how their students performed in KCPE. This might have influenced their teaching.

Another important limitation of the study, and which turns out to be a limitation of most predictive studies, is the possibility of decreasing predictive validity across time (Hulin, Henry & Noon, 1990). This is so because student abilities are not fixed, but dynamic. If predictive validities vary randomly across time, and in excess of random fluctuations, then there may be no linear factor involved in the variability.

Socioeconomic background is a useful variable in most educational studies of this nature. However, this study could not incorporate students socioeconomic background due to lack of appropriate records.

Other weaknesses of the study include inability to fit a prediction model for English due to failure of convergence, inability to get other school level and pupil level variables that might have had an influence on the relationship, and inability to calculate the actual reliability of the predictor and criterion.

The major strength of the present study is the use of HLM which considers the hierarchical nature of educational data.

Future Research

Future research on predictive validity of KCPE should be a large-scale national study, making use of three level hierarchical linear modelling. The levels to be considered may be pupil-level, school-level, and district-level. Such a study may provide comprehensive and valuable information about the predictive validity of KCPE as well as the relative achievement levels between pupils, schools, and districts. Such a study should incorporate more school level variables.

In conclusion, the findings of this study showed that KCPE is a valid predictor of KCSE, although the validity is only moderate. Whenever examinations are used as selection instruments, their predictive validity should be empirically studied from time to time in order to make possible improvements if necessary.

REFERENCES

Aiken, L.R. (1971). Intellective variables and mathematics achievement: Direction for research. Journal of School Psychology, 9, 201-212.

Anderson, C. S. (1982). The search for school climate: A review of the research. Review of Educational Research, 52(3), 368-420.

Arnold, C. L. (1992). Methods, Plainly Speaking. An Introduction to Hierarchical Linear Models. Measurement and Evaluation in Counseling and Development, 25, 58-85.

Baron, M. R. and Kenny, D. A. (1986). The Moderator-Mediator Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. Journal of Personality and Social Psychology, 51(6), 1173-1182.

Borg, W.R. and Gall, M.D. (1983). Educational Research. An Introduction: Longman.

Briggs, L. D. (1968). The impact of failure on elementary school pupils (Doctoral dissertation, North Texas State University, 1966). Dissertation Abstracts, 27, 2719-A. (University Microfilms, No. 67-2571) [O]

Brophy, J. (1985). Interactions of male and female students with male and female teachers. In gender influences in classroom interaction (pp 115-142). N.Y. Academic Press.

Bryk, A. S. & Driscoll, M. E. (1988). The school as community: Theoretical foundations, contextual influences, and consequences for students and teachers. Madison, WI: National Center on Effective Secondary Schools.

Bryk, A. S. and Raudenbush, S. W. (1992). Hierarchical Linear Models: Applications and data analysis methods. Sage Publications.

Chansky, N. M. (1964). Progresses of promoted and repeating grade 1 failures. Journal of Experimental Education, 32, 225-237.

Choppin, B. H. (1969). The relationship between achievement and age. Educational Research, 12(1), 22-29.

Coeffield, W. H., & Bloomers, P. (1956). Effects of non-promotion on educational achievement in the elementary school. Journal of Educational Psychology, 47, 235-250.

Cohen, J. & Cohen, P. (1983). Applied Multiple Regression Analysis for Behavioral Sciences (2nd Ed.). Lawrence Erlbaum Associates, Inc., Publishers.

Coladarci, A. (1983). Paradise regained: An apodictic analysis of the relationship between school size and public achievement. Research in Rural Education, 2(2), 79-82.

Cook, W. W. (1940). Some effects of the practice of non-promotion of pupils of low achievement. Washington DC.: American Educational Research Association.

Corbett, H. D. & Wilson, B. Raising the stakes in state wide mandatory minimum competency testing. Research for Better Schools Inc., Philadelphia, Pa, 1989. (ERIC Document Reproduction Service No. Ed. 338641.

Dwyer, C. A. (1973). Sex differences in Reading: An evaluation and a critique of current theories. Review of Educational Research, 43, 455-467.

Ebel, R. L. (1982). The practical validation of tests of ability. Paper presented at the NAEP Conference on Large-Scale Assessment, Boulder, Colorado: Educational Research Association.

Eisemon, T. O. & Schwille, J. (1991). Primary schooling in Burundi and Kenya: Preparation for secondary education or self-employment? The Elementary School Journal, 92(1), 23-39.

Eshiwani, G. S. (1985). The education of Women in Kenya, 1975-1984. Kenyatta University College, Nairobi.

Fennema, E. and Leder, G. C. (1990). Mathematics and gender. Teachers College Press.

Ferguson, G. A. (1981). Statistical analysis in psychology and education. McGraw-Hill, Inc.

Fox, L. H. and Cohn, S.J.(1980). Sex differences in the development of precocious mathematical talent. Women and the mathematical mystique (pp. 94-111) Baltimore: Johns Hopkins University Press.

Goldstein, H. (1987). Multilevel models in educational and social research. Oxford University Press.

Government Printer, Nairobi (1976). Report of the National Committee on Educational Objectives and Policies.

Government Printer, Nairobi (1989). Development Plan (1984-88).

Government Printer, Nairobi (1990). Statistical Abstract.

Haislett, J. and Hafer, A. (1990). Predicting success in engineering students during the freshman year. Career Development Quarterly, 39(1), 86-95.

Harnisch, D. & Archer, J. (1986). Mathematics productivity and educational influences for secondary students in Japan, India and the United States. Paper presented at the annual meeting of the American Educational Research Association (70th San Francisco, CA, 1986).

Hedges, L. V. (1982). Fitting continuous models to effect size data. Journal of Educational Statistics, 7, 245-270.

Heyneman, S. P. (1987). Uses of examinations in developing countries: Selection, research, and educational sector management. International Journal of Educational Development, 7, 251-263.

Holmes, C. T. and Matthews, K. M. (1984). The effects of nonpromotion on elementary and junior high pupils: A meta-analysis. Review of Educational Research, 54(2), 225-236.

Hulin, C. L., Henry, R. A. & Noon, S. L. (1990). Adding a Dimension: Time as a factor in the generalizability of predictive relationships. Psychological Bulletin, 107(3), 328-340.

Husen, T. (1967). International Study of Achievement in Mathematics. New York: Wiley.

Jackson, G. B. (1975). The research evidence on the effects of grade retention. Review of Educational Research, 45(4), 613-635.

Jacobsen, S. S. (1990). Identifying children at risk: The predictive validity of kindergarten screening measures. Ph.D Thesis, University of British Columbia.

Kenya Education Commission Report. (1964). Government Printer, Nairobi.

Kenya National Examinations Council, (1988). K.C.P.E. Newsletter.

Kellaghan, T. & Greaney, V. (1992). Using examinations to improve education. A study in fourteen African countries. World Bank Technical Paper Number 165. Africa Technical Department Series.

Kimble, J. W. (1976). Basic quality of secondary education in rural Montana. Bulletin 685. (ERIC Document Reproduction Service No. Ed. 212418).

Kreft, I. G. G., De Leew, J., & Kim, K. (1990). Comparing four different statistical packages for hierarchical linear regression: GENMOD, HLM, ML2, and VARCL (CSE Technical Report 311, 70-71. Los Angeles, CA: UCLA Center for Research and Evaluation, Standards, and Student Testing.

Lenarduzzi, G. P. & McLaughlin, T. F. (1990). The effects of nonpromotion in junior high school on academic achievement and scholastic effort. Unpublished manuscript.

Longford, N. T. (1989). Fisher scoring algorithm. In R. D. Bock (Ed.), Multilevel analysis of educational data (pp. 297-310). Academic Press, Inc. San Diego, California 92101.

McAfee, J. K. (1981). Toward a theory of promotion: does retaining students really work? Paper presented at the meeting of the American Educational Research Association, Los Angeles, CA.

McMillan, J. H. & Schumacher, S.(1984). Research in education. A conceptual introduction. Little, Brown and Company (Canada) Limited.

Messick, S. (1980). Test validity and ethics of assessment. American Psychologist, 35(11), 1012-1027.

Ndunda, M. M. K. (1990). Because I am a woman: Young women's resistance to science careers in Kenya. M.Ed. Thesis, Queen's University, Canada.

New York State Department, Bureau of School Programs Evaluation (1976). Which school factors relate to learning? Summary of findings of three sets of studies. Albany, N.Y.

Niklason, L. B. (1987). Do certain groups of children profit from a grade retention? Psychology in the Schools, 24, 339-345.

Paterson, L. & Goldstein, H. (October, 1991). New statistical methods for analysis of social structures: an introduction to multilevel models. Multilevel Modelling Newsletter, p. 6.

Rasbash, J., Prosser, B. & Goldstein, H. (1989). ML 2 software for two-level analysis; Users' Guide. Institute of Education, University of London.

Raudenbush, S. W. (1988). Educational applications of hierarchical linear models: A review. Journal of Educational Statistics, 13, 85-116.

Raudenbush, S. W. & Bryk, A. S. (1986). A hierarchical model for studying school effects. Sociology of Education, 59, 1-17.

Reinherz, H. & Griffin, C. L (1971). The treadmill of failure. Quincy, Mass.

Rutter, M., Maughan, B., Mortimore, P., Ouston, J., & Smith, A. (1979). Fifteen thousand hours. London, England: Open Books.

Simon, R. J. & Danna, M. J. E. (1990). Gender, race, and the predictive validity of the LSAT. Journal of Legal Education, 40(4), 525-529.

Smith, M. L. & Shepard, L. A. (1987). Effects of kindergarten retention at the end of first grade. Psychology in the Schools, 24, 346-357.

Stanley, J. C. & Hopkins, K. D. (1972). Educational and psychological measurement and evaluation: Prentice-Hall, inc., englewood cliffs, n.j.

Stroud, J. B. & Lindquist, E. F. (1942). Sex differences in achievement in the elementary and secondary schools. Journal of Educational Psychology, 33, 657-667.

Tagiuri, R. (1968). The concept of organizational climate. In R. Tagiuri & G. H. Litwin (Eds.). Organizational climate: Exploration of a concept. Boston: Harvard University, Division of Research, Graduate School of Business Administration.

System dogged by controversy. (1993, April 16). The Weekly Review, p. 10.

Walsh, D. J. (1988). The two-year route to first grade: The use of testing to validate decisions based on class and age. Paper presented at the Annual Meeting of the American Educational Research Association in New Orleans.

Werdelin, I. (1961). The geometrical ability and the space factor in boys and girls. Lund, Sweden.

Willms, J. D. (1985). Catholic-school effects on academic achievement: New evidence from the High School and Beyond Follow-up study. Sociology of Education, 58, 98-114.

Wyatt, J. F. & Gay, J. D. (1984). The educational effects of different sizes and types of academic organization. Oxford Review of Education, 10(2), 211-223.

Population Studies No. 106 (1989). World Population Prospects, 1988. Department of International Economic and Social Affairs, United Nations, New York, 1989.

Yates, B. J. (1991). A comparison of effectiveness ratings of selected principals and NASSP assessment center ratings. Paper presented at the annual meeting of the American Educational Research Association.

APPENDICES

APPENDIX A

Mathematics Grade Statistics by Gender

	A	A-	B+	B	B-	C+	C
BOYS	675	442	710	974	1327	1867	576
GIRLS	136	108	136	250	332	547	828
TOTAL	811	550	846	1224	1659	2414	1404
%	0.62	0.42	0.65	0.94	1.28	1.86	1.08

	C-	D+	D	D-	E	TOTAL	%
BOYS	3375	4776	12982	19541	29318	76563	58.9
GIRLS	1188	1766	6133	12475	29582	53481	41.1
TOTAL	4563	6542	19115	32016	58900	130044	100
%	3.51	5.03	14.70	24.62	45.29	100	-

APPENDIX B

Scatterplots of KCSE grades against KCPE scores.

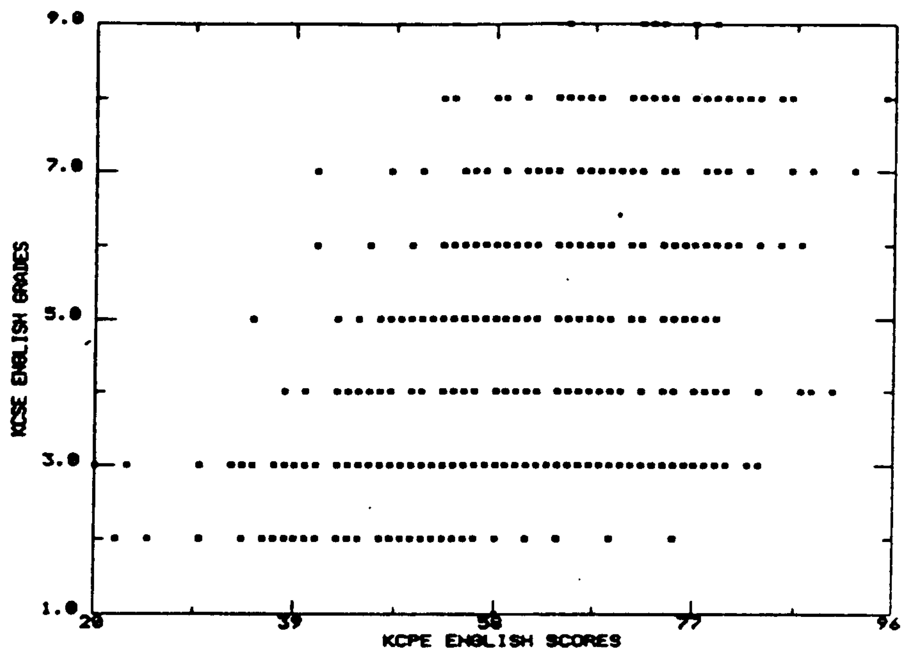


Figure B-1: KCSE English against KCPE English

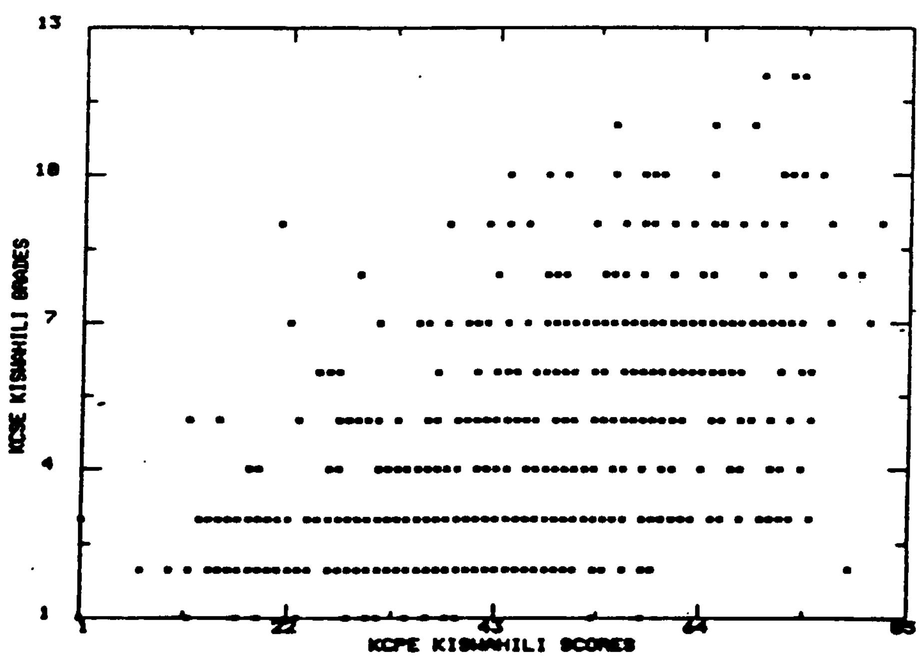
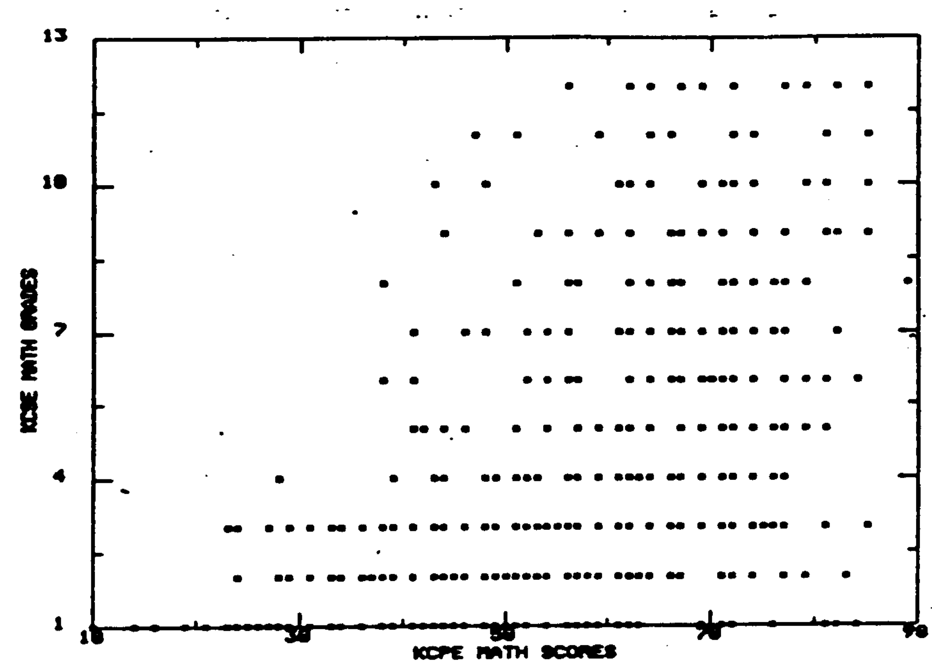
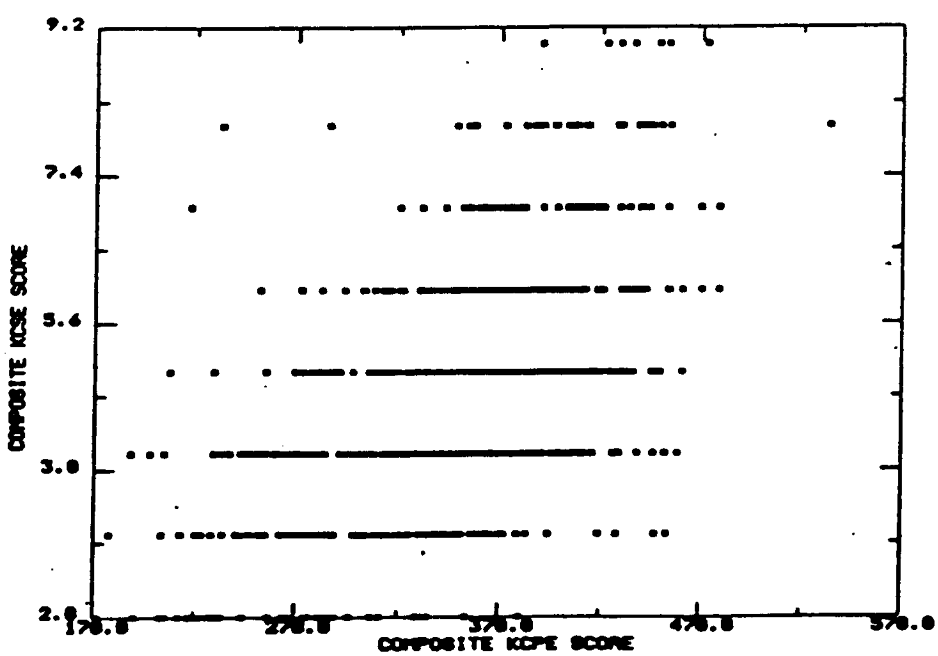


Figure B-2: KCSE Kiswahili against KCPE Kiswahili



<u>Figure B-3</u>: KCSE Math Scores against KCPE Math Scores



<u>Figure B-4</u>: Composite KCSE against Composite KCPE

APPENDIX C

KCSE and KCPE Stem and Leaf, Box, and Normal Plots.

Leaf unit is 0.100

Lower limit	N		
2.000	1	:	0
2.500	0	:	
3.000	2	:	12
3.500	6	:	567778
4.000	5	:	22244
4.500	6	:	555689
5.000	4	:	1123
5.500	1	:	5
6.000	0	:	
6.500	1	:	5

Figure C-1: Stem and Leaf Plot for KCSE School Means

Leaf unit is 10.0

Lower limit	N		
240.0	1	:	5
260.0	2	:	66
280.0	4	:	8899
300.0	5	:	00111
320.0	4	:	2233
340.0	3	:	455
360.0	4	:	6667
380.0	1	:	8
400.0	1	:	0
420.0	1	:	2

Figure C-2: Stem and Leaf Plot for KCPE School Means

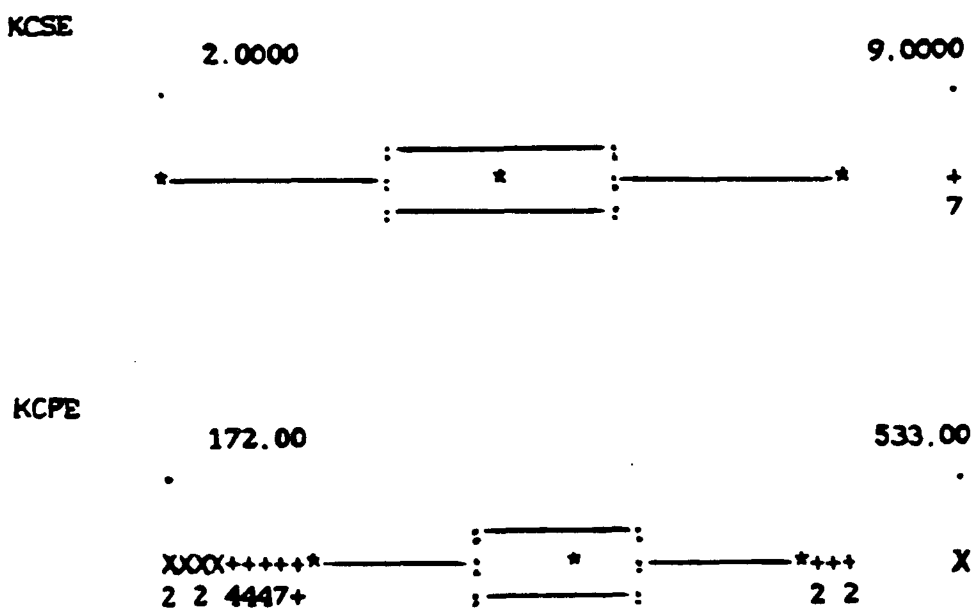


Figure C-3: Box Plots for KCSE and KCPE

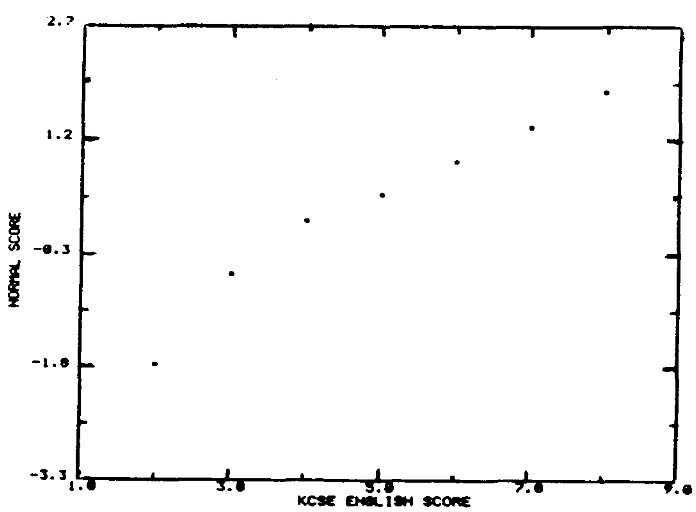


Figure C-4: Normal Plot for KCSE English

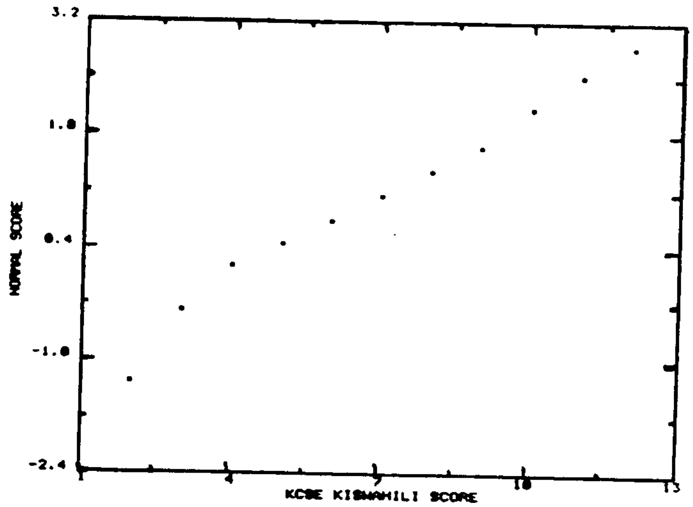


Figure C-5: Normal Plot for KCSE Kiswahili

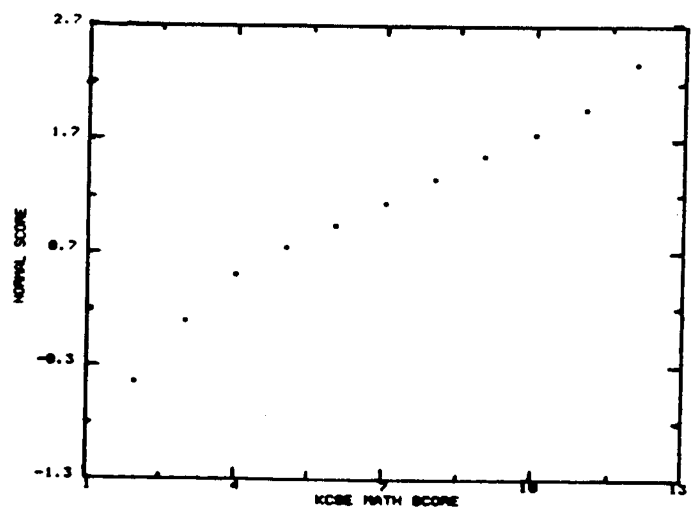


Figure C-6: Normal Plot for KCSE Mathematics

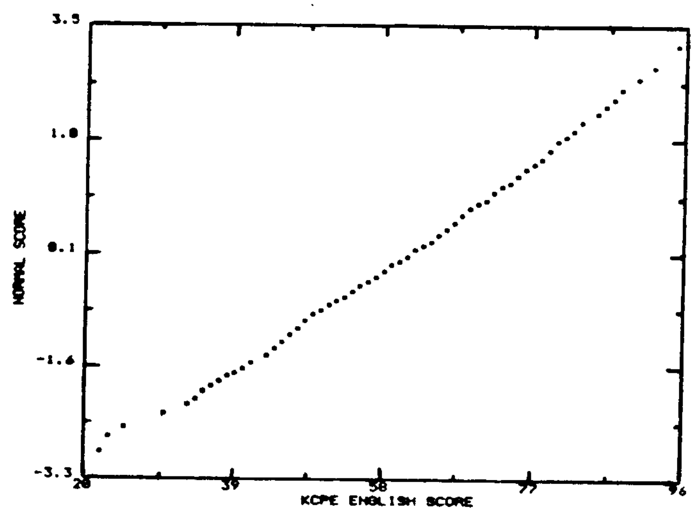


Figure C-7: Normal Plot for KCPE English

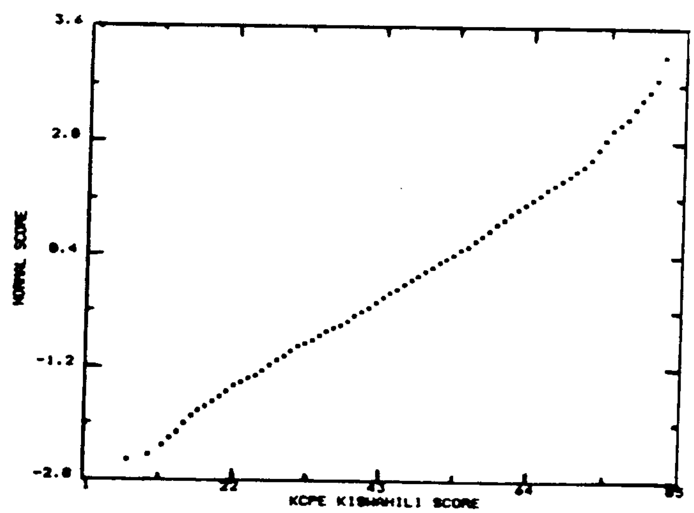


Figure C-8: Normal Plot for KCPE Kiswahili

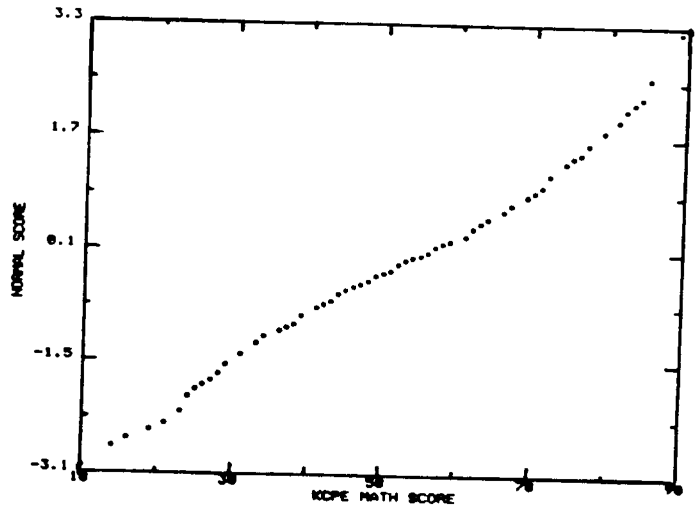


Figure C-9: Normal Plot for KCPE Mathematics

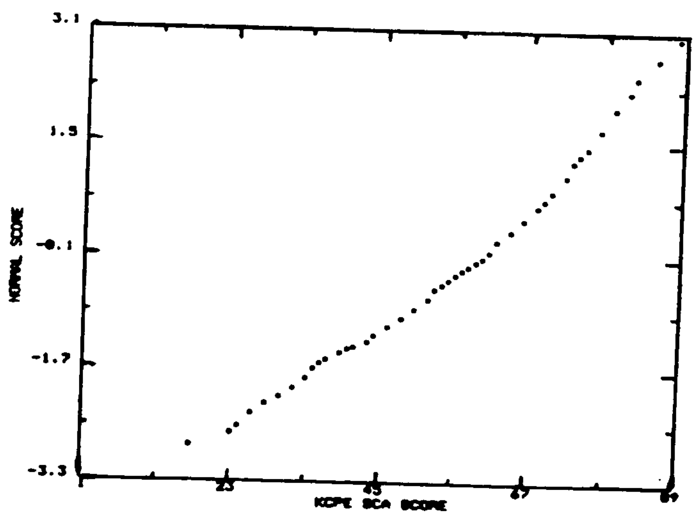


Figure C-10: Normal Plot for KCPE Science and Agriculture

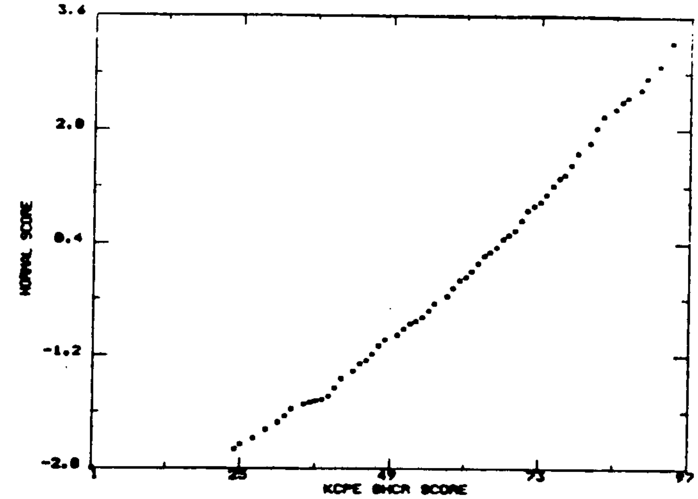


Figure C-11: Normal Plot for KCPE GHCR

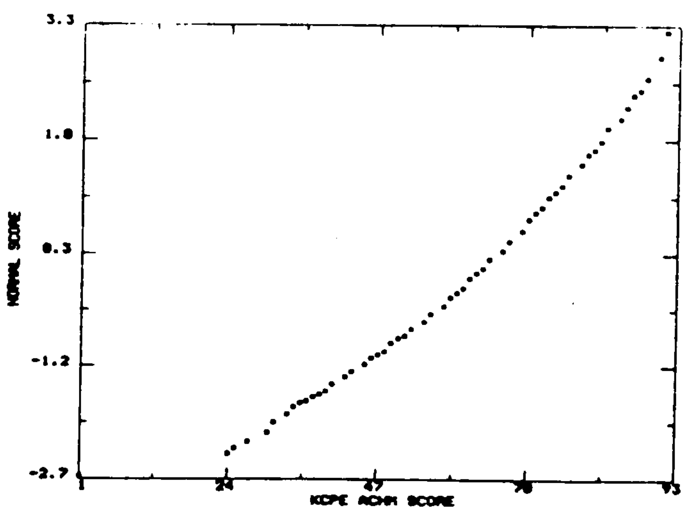


Figure C-12: Normal Plot for KCPE ACHM

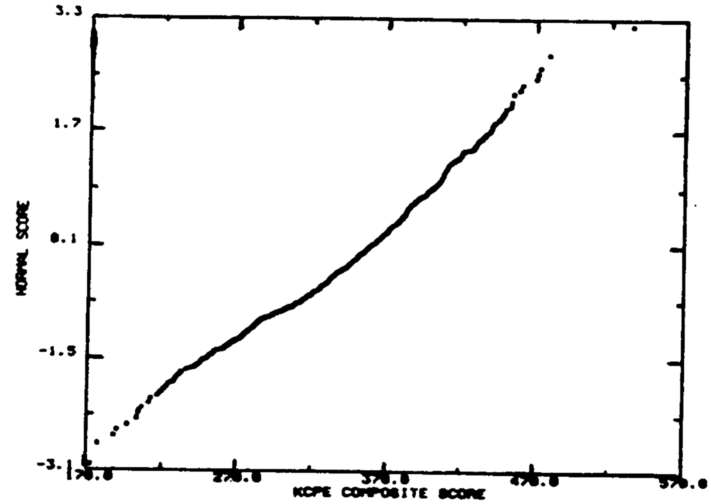


Figure C-13: Normal Plot for KCPE Composite score

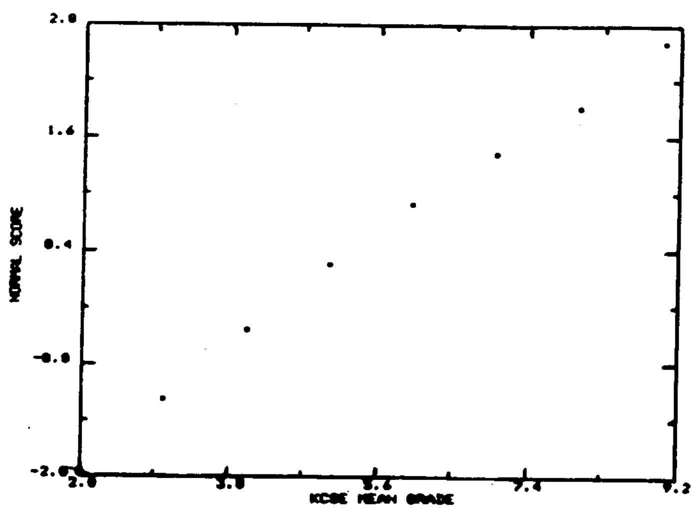


Figure C-14: Normal Plot for KCSE Composite Score