Global Journal of Human-Social Science: G
Linguistics & Education
Volume 16 Issue 2 Version 1.0 Year 2016
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 2249-460X & Print ISSN: 0975-587X

Multivariate Analysis of Factors Influencing Achievement of Students in Selected Subjects at Secondary School Level: A Case Study of Grade 10 Students at Hawassa City, Ethiopia

By Endris Assen Ebrahim

Debre Tabor University, Ethiopia

Abstract- This study has been an attempt to determine factors influencing academic achievements of grade 10 students (normally under age 18 years) on specific subjects. A cross-sectional survey was conducted on a total of 719 sample students of grade 10 from 11 different government and non-government secondary schools using multistage sampling technique. A designed questionnaire was used to obtain data from the respondents. The secondary data on students EGSECE scores were obtained from the Education Department as achievements of students in the five selected subjects: Mathematics, Biology, Physics, Chemistry and English. Descriptive analysis, factor analysis and multivariate multiple linear regression analyses were used to analyze the data. From the descriptive results both governmental and non-governmental school students were achieved poorest in physics and best in English. However, on average, non-governmental school students' achievements were better than governmental school students.

Keywords: *factor analysis; multivariate multiple linear regression analysis; school subjects; achievement; grade ten.*

GJHSS-G Classification : *FOR Code: 930101p*

MULTIVARIATEANALYSISOFFACTORSINFLUENCINGACHIEVEMENTOFSTUDENTSINSELECTEDSUBJECTSATSECONDARYSCHOOLLEVELACASESTUDYOFGRADE10STUDENTSATHAWASSACITYETHIOPIA

Multivariate Analysis of Factors Influencing Achievement of Students in Selected Subjects at Secondary School Level: A Case Study of Grade 10 Students at Hawassa City, Ethiopia

Endris Assen Ebrahim

Abstract- This study has been an attempt to determine factors influencing academic achievements of grade 10 students (normally under age 18 years) on specific subjects. A cross-sectional survey was conducted on a total of 719 sample students of grade 10 from 11 different government and non-government secondary schools using multistage sampling technique. A designed questionnaire was used to obtain data from the respondents. The secondary data on students EGSECE scores were obtained from the Education Department as achievements of students in the five selected subjects: Mathematics, Biology, Physics, Chemistry and English. Descriptive analysis, factor analysis and multivariate multiple linear regression analyses were used to analyze the data. From the descriptive results both governmental and non-governmental school students were achieved poorest in physics and best in English. However, on average, non-governmental school students' achievements were better than governmental school students. In factor analysis, self-concept, motivation to the subjects and teaching-learning process explained most of the variations. Multivariate regression results revealed that, the factors, sex, school type, school facilities, family status, school volume, interest to the subject, motivation to the subject, self-concept, safe reading and trouble (anxiety) to the subjects, had significant influence on achievements of students with respect to most of the subjects. Factors like sex, school facility, family status, motivation to the subject, interest to the subject had a significance positive impact on achievements. However, trouble of the subject and school volume had a significant negative influence on students' achievements on Biology, Physics and English subjects. It is suggested that academic facilities and managements at schools, beside home and students' personal efforts need to be promoted for better academic achievements of students in subjects.

Keywords: *factor analysis; multivariate multiple linear regression analysis; school subjects; achievement; grade ten.*

I. Introduction

The current educational system in Ethiopia is organized in cycles or levels of formal schooling that includes ten years of general education. General education is completed at the end of the first cycle of general secondary school education (Grade 9 and 10). Moreover, this cycle is intended to enable students to identify area of interest (Natural Sciences Stream and Social Sciences Stream) for further training in the second cycle of general secondary education (Grades 11and 12) to prepare students for continuing their studies at higher education level (University or collages) or selecting their own vocations. Students appear for the New National Examination at the end of grade 10 (normally under 18 years old) which is known us the Ethiopian General Secondary Education of Certificate Examination (EGSECE). This is after the students have successfully achieved school examinations in all school subjects. However, students should score a minimum of 2.00 on a scale of 4.00 in EGSECE or a minimum of 50 out of 100 in standard school exams at least in five or seven subjects: English, Mathematics (both compulsory) and any other three or five science (Natural or Social) subjects in order to appear in EGSECE [14].

Girls' education is one of the fundamental pillars for ensuring sustainable economic development, democratic participation and poverty reduction. As a result, gender discrimination affects not only women but also the overall growth of the economy. In this connection, the Ethiopian government has given more attention to girl's education. In 2003-2004, due to the favorable policy environment, the gross enrollment of female students at general secondary first cycle (9-10) was about 37.0% and at the preparatory level (11-12), it was 29.0%. Moreover, in technical, vocational and training institutions (colleges), it was 49.0%, whereas it was 25.2% in higher education/Universities. Nevertheless, there was a great variation of students' achievement at different school type (non-governmental and governmental) based on their gender. Without controlling for student background differences, non-governmental schools scored higher than government (public) schools ([4]; [13]).

It is obvious that students at schools can be classified as clever (high achievers), medium (average achievers) and lazy (low achievers) with respect to individual's achievements in specific school subjects based on exam scores or general test results of

Author: Debre Tabor University, Faculty of Natural and Computional Sciences, Department of Statistics, Debre Tabor, Ethiopia, Addis Ababa. e-mails: end384@gmail.com, assen_endris@yahoo.com

subjects. The general belief is that, if the student is intelligent or clever, he/she is expected to perform well at school in compulsory and science school subjects and is well fitted for national and regional exams. But intelligence is not the only influential factor of academic achievement in school subjects. In addition to intelligence, there are various factors influencing academic achievement of students at school in each school subjects ([2]; [19]).

This study has been undertaken to investigate multivariate evaluation of the impacts of family with student and school characteristics variables on academic achievement of students on five selected subjects at secondary schools, specifically in grade 10. The presence of all or some of the factors identified above may have resulted in the poor academic achievement of students on each school subjects in some areas of our country. However, evidence of the availability of these factors as well as other factors need to be obtained or checked. The purpose of this study, therefore, is to obtain the factors that are responsible for the poor academic achievement of students with school type and gender gap on school subjects among secondary schools of grade 10 students at Hawassa city, in SNNPR state.

a) *Statement of the Problem*

In 2007/08 the number of students who sat for grade 10 national exams, at SNNPR state, was 92,836 (male 61,742 and female 31,094). Out of these who get CGPA of 2.00 and above out of 4.00 were 33,211 (25,085 males and 8,126 females). The percent of promoted students in a successive three years, 2005/06, 2006/07 and 2007/08, were 45.8%, 44.2% and 35.8%. Specifically, the percentage of promoters (scored 2.00 and above) at Hawassa City Administration in 2007/08 were 46.7 %.

[12] Reported that the test items (exam questions) of the EGSECE for English were not relatively content valid. Hence, test items did not match with the syllabus contents.

Students might pass from one class level to the other as they evaluated on CGPA result of all subjects. But due to achievement variation with respect to each school subject, students get difficulty and being unsuccessful in higher level education which leads directionless. The current education system of Ethiopia gives a great attention, about 70%, on natural sciences subjects, to enhance sciences and technology. Therefore, it is better to find solutions to the problems and factors one faced in his/her academic achievements in selected subjects: Mathematics, Biology, Physics, Chemistry and English at secondary schools in grade 10 distinctly but dependably.

Many reasons have been attributed for the high failure rate and poor academic achievements in secondary schools. Some researchers traced that the high failure rate of students was due to student's inability to comprehend and balanced the principles of some subjects such as Mathematics, Physics and others. Others are of the view that the abysmal school achievement is due to loaded curriculum (there is too much to be taught within a short time) ([8]; [12]).

Again some people suggest it on lack of proper supervision on the part of school administration and family control in student's self-carelessness ([7]; [10]).

Likewise, [13] claimed that gender stereotype and student's interest to the subjects have also great influential effect. Peculiar nature of some factors and the students low and unbalanced success rate have led to this study on the multivariate analysis of the determinants of students' academic achievement measured in five selected subjects at general secondary school completion level, first cycle, grade 10.

The following research questions have been developed to guide this study:

- What are the key factors that influence students' academic achievements in Mathematics, English, Biology, Chemistry and Physics at secondary school level in grade 10?
- What relationships (correlations) are there among the selected subjects at student and school levels and what gender and school type gap is observed in terms of the five school subjects?
- How much of the variations (level differences) of the academic achievements are accounted for at school with respect to each response measurement scores of the school subjects?
- How much variations are explained with the interrelationships of general home-school characteristic variables and as students' opinions over a group of items about each separate school subjects?

b) *Objectives of the Study*

The general objective of the study has been to determine the key factors influencing academic achievements of students measured in exam scores of five subjects in grade 10 (Mathematics, Biology, Physics, Chemistry and English), and to assess the variations accounted at school and individual (student) level for each response (school subjects). The Specific Objectives are

- To identify the most important factors (covariates) influencing academic achievements of student's in each component of selected subjects in grade 10.
- To determine the relationship among the school subjects at both school and student level; and whether there is gender and school type differences in this relationship.
- To quantify and determine the within and between schools variation for each components of selected subject at secondary schools.

- To determine the groups or clusters of interrelated observed variables or items as component factors that explain the variation of achievement indictor variables.

II. Materials and Methods

a) Description of the Study Area and Population

The study was conducted in Hawassa, the capital city for SNNPR state, which was established in 1960. It is located at about 275 km South West of Addis Ababa, and near to Hawassa Lake. Geographically it lies between 07^0 $05'$ Latitude North and 38^0 $29'$ Longitude East. According to the report of [6], the estimated population size of the city (urban) in 2007 was 159,013 out of which 81,984 were males and 77,029 females. There are 4 governmental colleges and one university, 8 non-government (private) colleges, 5 governmental high (secondary) schools, about 15 non-governmental high (secondary) schools. The gross enrollment rate of secondary school students at Hawassa Town Administration has been 62.1%.

The target population for this study was grade 10 students of both government and non-government schools registered in 2010-2011 academic year at Hawassa City secondary schools. The total population of students in all high schools of the city was 6,384 in 2010-2011 academic year.

Exclusion criteria were made on the students who were transferred to other schools or those dropped out, only completed enrolment procedures at the school but did not yet attend the national exam or left the school or had been absent for more than four continuous weeks (excluding school vacations) and had no examination results in 2 of the most targeted school subjects (compulsory subjects). This was because full information about those students was not available.

b) Sampling Design and Procedure

A cross-sectional study with stratification sampling designed to take independent samples for different sub-populations was conducted. The stratums were governmental and non-governmental secondary schools as school type.

Sampling methods are scientific procedures of selecting those sampling units which would provide the required estimator with associated margins of uncertainty arising from examining only a part not the whole of the population. The main purpose of stratification is to reduce sampling error. Moreover, stratified sampling is a technique which uses any relevant information that might be available in order to increase efficiency. It involves the division or stratification of a population by partitioning the sampling frame in to non-overlapping and relatively homogeneous groups [5].

A list of grade 10 students was obtained from Hawassa City Administration Education and Capacity Building Department. The population of grade 10 students was stratified into governmental and non-governmental school and the required sample size for the study was determined from each stratum. The multistage sampling procedure was employed as:

- Stage one: Stratification by school type

All secondary schools except those with number of students in class less than 15 and far away from the city center were considered.

Stratum 1: Grade 10 students in government schools with population size N_1 and sample size n_1.

Stratum 2: Grade 10 students in non-governmental school with population size N_2 and sample size n_2.

- Stage two: proportional allocation or proportion by sample size method

Sample of students was taken from sampled schools by proportional allocation, to enrollment size of grade 10 students at selected schools, of total sample size n in to sample sizes of governmental and non-governmental schools, n_1 and n_1.The selection of a simple random sample was usually carried out according to a set of mechanical instructions which guarantees the random nature of the selection procedure. This is an equal probability of selecting individual units for all elements in the population of the school.

- Stage three: simple random sampling of students from class

Taking a list of students with their registration number in each school, then refer to a table of random numbers; the required sample students were selected. In simple random sampling, the selection of one individual was independent of the selection of another individual.

i. *Sample Size Determination*

In the planning of a sample survey or researches, a stage at which a decision must be made about the size of the sample is always required. However, too large a sample implies wastage of resources, and too small a sample diminishes the utility of the results. Therefore the decision should be made with a minimum cost but the estimate will explain the population characteristics with a high probability. However, several formulas developed for sample size calculations that conform to different research situations [5].

The sample size for this study was determined based on stratified sampling with proportional allocation at 95% confidence level using the general formula for sample size determination adopted as:

$$n = \frac{\sum_{h=1}^{2} \frac{W_h^2 S_h^2}{W_h}}{V + \frac{1}{N}\sum_{h=1}^{2} W_h S_h^2}$$, where h = the stratum,

$W_h = N_h / N$ = stratum weight, n_h = number of units in strata h, $V = \left(\frac{E}{Z_{\alpha/2}}\right)^2 = Var\left(\hat{Y}_{str}\right)$ = desired variance for estimate of population mean which is

$\hat{Y}_{str} = \sum_{h=1}^{2} W_h \bar{y}_h$. $\bar{y}_h = \frac{\sum_{i=1}^{n_h} y_{hi}}{n_h}$ = sample mean,

$\bar{Y}_h = \frac{\sum_{i=1}^{N_h} y_{hi}}{N_h}$ = true mean (mean for the population measurements) and $S_h^2 = \frac{\sum_{i=1}^{N_h}\left(y_{hi} - \bar{Y}_h\right)^2}{N_h - 1}$ = true variance (variance of the population measurements). S_1^2 = government school sample variance of students' academic achievement

S_2^2 = non-government school sample variance of students' academic achievement

$\bar{y}_1$ = government school students sample mean of students' academic achievement and

$\bar{y}_2$ = non-government school students sample mean of students' academic achievement,

$z_{\alpha/2} = z_{0.025} = 1.96$ is the critical value for 95% confidence level with standard normal distribution.

The known methods of estimating s^2 for calculating sample size of any survey were by taking the sample in two steps; one by the results of a pilot survey and another by previous studies sampling of the same or similar population and guesswork about the structure of the population [5].

But for the present study, s^2 and the margin of (absolute) error E were determined from the results of previous studies of similar population. The sample variance s_1^2 = 0.20885 and mean $\bar{y}_1$ = 2.62 were taken for government school from the study which assessed the determinants of students' academic performance in government schools of grade 10 at Hawassa town taking a sample of 920 students (Hanna; 2010).

Then, E for this study was calculated as:

$$E = Z_{\alpha/2}\sqrt{\frac{S^2}{n}} = 1.96\left(\sqrt{\frac{0.208849}{920}}\right) = .0295$$

On the other hand, the sample variance s_2^2 = 0.13421 was taken for non-government schools from the previous study at the same area [11]. The total population was (number of students in 11 selected secondary schools of grade 10) $N = 5006$ from 5 governmental and 6 non-governmental selected secondary schools, which contained total number of grade 10 students in governmental schools $N_1 = 3755$ and total number of grade 10 students in non-governmental schools $N_2 = 1251$.

After all, using the weight $W_h = N_h / N$ and S_h^2 were more convenient for computing the sample size n from the estimated sample size, n_0.

$$n_0 = \frac{\sum_{h=1}^{2} W_h S_h^2}{V} = \left(\frac{Z_{\alpha/2}}{E}\right)^2 \sum_{h=1}^{2} W_h S_h^2 = \left(\frac{1.96}{0.0295}\right)^2 \left(\frac{3755 \times 0.20885}{5006} + \frac{1251 \times 0.13421}{5006}\right) = 840.$$

Then $n = \frac{n_0}{1 + \frac{n_0}{N}}$ = 719 was the total sample for this study. Thus, using the above results, the following sample sizes for both school types (Governmental and Non-governmental) as proportional allocation by school type as a factor is: $n_h = \left(\frac{N_h}{N}\right)n$ for $h = 1,2$; $n_1 = \left(\frac{3755}{5006}\right) \times 719 = 539$ (Sample for governmental schools), $n_2 = \left(\frac{1251}{5006}\right) \times 719 = 180$ (Sample for non-governmental schools).

c) *Methods of Data Collection*

In assessing the academic achievement of students' measured by exam results scored in school subjects, Mathematics, Biology, Physics, Chemistry and English at both government and non-government sample secondary schools, both primary and secondary data were used. The primary data was collected using questionnaire method. The questionnaire consisted the student's, family background and school characteristic variables on the student's academic achievements evaluated in selected 5 subjects. Individuals sampled for this study were asked to complete the determinants of students' outcome (in five school subjects) study questionnaire. The secondary data on academic achievements of respondents was measured by their EGSECE results (scores) in each of the five selected subject (Mathematics, Biology, Physics, Chemistry and English). Besides, school records with regard to students' exam registration number and some profiles of teachers and schools were taken from record offices of the schools. Sampled grade 10 students were taken with their exam scores of all five school subjects and the student's results were standardized and scaled to be 4.00.

d) Variables of Interest in the Research

The outcome variables used in this study were the five selected school subjects as individual's achievement measures using EGSCEE results or scores on the five school subjects (Mathematics, Biology, Physics, Chemistry and English). All achievement scores were taken as standardized and transformed to assure that all scores were scaled in the same metric. This also allowed us to interpret the between school variances as the percentage of variation in student achievement accounted for by schools in PCFA, MVML and multivariate multiple linear regression analysis with respect to each response. The set of explanatory variables included were the composite common factors of students, family, teachers and schools characteristic variables.

i. *Students and Family Characteristic Variables*

These were: Age, gender, religion of student, parents' employment status, natural talent, students' job aspiration, time spent on study, peer(group) effect, student class attendance(absence), skipped class, student's satisfaction with school administration, satisfaction with school rules and regulations, academic confidence, preferred study time, preferred study place, distance of the school from students' home, availability of text and reference books at home, home location, parental involvement, fathers'/guardians' level of education, comfort of study place at home, mothers' education level, average family expenditure, other expenses related to education, satisfactions in food type available in home, pervious grade scores, students attitude and perception on school subjects (difficulty, boringness, preference, etc.).

ii. *School Characteristic Variables*

These were: teachers average workload, average year of experience, teachers average educational level, teacher preparation, class size, teaching method, standard of examination, parent to teacher communication, teacher absence, teacher late, average size of school, school fee, completion of the syllabus, school type, student-teacher ratio, teacher efficient and skills, school location/environment, current curriculum, human resources (teachers per subjects, principals, supervisors), infrastructure (buildings, classrooms, sport facilities), library facility, equipment (desks, blackboard, telephone, duplicating computers), amentias (toilets, electricity, water), and availability instructional materials (text and reference books, maps and charts), laboratory facilities, academic counseling service, health service (first aids).

III. Methods of Data Analysis

a) Factor Analysis Model

This analysis describes the covariance relationships among many variables (items) in terms of a few underlying and unobservable random quantities.

The observable random vector X with P components has mean μ and covariance Σ. The factor model postulates that X is linearly dependent upon a few unobservable random variables $f_1, f_2, \ldots, f_m$ called common factors, ($m < p$) and p additional source of variation $\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots, \varepsilon_p$ called specific factors.

The factor analysis model is given by: X = LF+ ε, where L_{PxM} is a matrix of unknown constants called factor loadings.

$$L_{pxm} = \begin{pmatrix} l_{11} & l_{12} \ldots & l_{1m} \\ l_{21} & l_{22} \ldots & l_{2m} \\ \ldots\ldots & & \\ l_{p1} & l_{p2} \ldots & l_{pm} \end{pmatrix} \quad F = \begin{pmatrix} f_2 \\ f_1 \\ \ldots \\ f_m \end{pmatrix} \quad \text{and} \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \ldots\ldots \\ \varepsilon_p \end{pmatrix}$$

The coefficient l_{ij} is the loading of the i^{th} variable on the j^{th} factor.

i. *Assumptions of Factor Model*

1. $E(F) = \mathbf{0} = (0, 0, \ldots, 0)^T$
2. $cov(F) = E(FF^T) = I_m$
3. $E(\varepsilon) = \mathbf{0} = (0, 0, \ldots, 0)$
4. $Cov(\varepsilon) = E(\varepsilon\varepsilon^T) = \Psi_{pxp}$, Ψ is a diagonal matrix
5. $Cov(\varepsilon, F) = E(\varepsilon, F^T) = \mathbf{0} = (0,0,\ldots,0)^T$

ii. *Covariance Structure for Orthogonal Factor Model*

1. $Cov(X) = LL^T + \psi$
2. $Var(X_i) = l_{i1}^2 + l_{i2}^2 + \ldots + l_{im}^2 + \psi_i$, where ψ_i is the i^{th} specific factor.

3. $E(X_i, X_k) = l_{i1}l_{k1} + l_{i2}l_{k2} + \ldots + l_{im}l_{km}$
4. $Cov(X_i, F_j) = l_{ij}$
5. $Cov(X, F) = L$, loading matrix.

Communality is defined by: $h_i^2 = l_{i1}^2 + l_{i2}^2 + \ldots + l_{im}^2$

The factor model assumes that $p + \frac{p(p-1)}{2} = \frac{p(p+1)}{2}$ variables and covariance for X can be reproduced from pm factor loadings l_{ij} and p specific variables ε_i.

The factor model provides a simple explanation of the covariation in X with parameters $(p + pm)$ which are fewer than $p(p+1)/2$ parameters in Σ.

iii. *Methods of Estimation of Loading*

If the off diagonal elements of sample covariance S are small or those of the sample correlation matrix R essentially zero (identity matrix), the variables are not related. This implies that a factor analysis will not prove useful and in these circumstances, the specific factor plays a dominant role. If covariance matrix appears to deviate significantly from a diagonal matrix, then a factor model can be entertained and the initial problem is one of estimating the factor loading l_{ij} and specific variance ψ_i. There are two popular methods of parameter estimation, Maximum Likelihood (ML) Method and Principal Component Method. However, for this study, the principal component method was used.

iv. *The Principal Component Method*

The spectral decomposition of covariance Σ having eigenvalues-eigenvector pairs (λ_i, e_i) with $\lambda_1 > \lambda_2 > \ldots > \lambda_m > 0$ is given as $\Sigma = \lambda_1 e_1 e_1^T + \lambda_2 e_2 e_2^T + \ldots + \lambda_p e_p e_p^T$

From the above equation, we can obtain the loading, $L = \left(\sqrt{\lambda_1}e_1, \sqrt{\lambda_2}e_2, \ldots, \sqrt{\lambda_p}e_i \right)$.

v. *The Contribution to the Total Sample Variances*

In applying the principal component to perform factor analysis, we have use, the sample covariance matrix S was used. Observe that $S_{11} + S_{22} + \ldots + S_{pp} = tr(S)$ =trace of the sample covariance matrix and $\hat{\lambda}_1 + \hat{\lambda}_2 + \ldots + \hat{\lambda}_p = \rho$ = trace of sample correlation matrix, where, $\hat{\lambda}_i$'s, $i = 1,2,\ldots\ldots, p$ were estimated eigenvalues of S.

$$\begin{pmatrix} \text{The proportion of total sample} \\ \text{variance due to } j^{th} \text{ factor} \end{pmatrix} = \begin{cases} \dfrac{\hat{\lambda}_j}{tr(S)} & \text{for factor analysis of sample covariance.} \\ \dfrac{\hat{\lambda}_j}{\rho} & \text{for factor analysis of correlation.} \end{cases}$$

Researchers have no single agreement about selecting the required number of principal components. However, the best choices for researchers to fix the number of factors retained have been the proportion variance explained being at least 50-60% and the Scree plot test examining the graph of the eigenvalues by looking for the natural bend or break point in the data where the curve flattens out. The number of data points above the "break" is usually the number of factors to retain, although it can be unclear if there are data points clustered together near the bend ([16;[21]]).

vi. *Rule of Thumb (Convention)*

- Choose the number of positive eigenvalues of sample covariance matrix *S* $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \ldots \geq \hat{\lambda}_m \geq 0$ and
- Choose the number of eigenvalues of sample correlation matrix R which are larger than 1.

vii. *Factor Rotation and Factor Scores*

Factor rotations are an orthogonal transformation of the factor loadings, as well as the implied orthogonal transformations of the factors. If $\hat{L}$ is the pxm matrix of estimated factor loadings obtained by any method, then $\hat{L}^* = \hat{L}T$, where $TT' = T'T = I$, was a pxm matrix of 'rotated' loadings, where I is the identity matrix. This shows that the estimated covariance (correlations) matrix remains unchanged since $\hat{L}\hat{L}' + \hat{\Psi} = \hat{L}TT'\hat{L}' + \hat{\Psi} = \hat{L}^*\hat{L}^{*'} + \hat{\Psi}$.

A useful byproduct of factor analysis was factor scores. Factor scores were composite measures that can be computed for each individual on each common factor. They are standardized measures with a mean = 0.00 and a standard deviation of 1.00, computed from the factor score coefficient matrix. For the given original data x_{ij} ($i = 1,2,3,\ldots n$ and $j = 1,2,3,\ldots, p$) the factor score of the i^{th} individual student on the k^{th} principal component retained can be calculated as: $\hat{f}_{ik} = \hat{l}_1 x_{1i} + \hat{l}_2 x_{2i} + \hat{l}_3 x_{3i} + \ldots + \hat{l}_p x_{pi}$, where

$\hat{f}_{ik}$ =factor score of the i^{th} respondent/student for the k^{th} factor retained,

x_{ij} = observation of the i^{th} on the j^{th},

$\hat{l}_j$ = the principal component (factor) loading of variable j [15].

b) Multivariate Multiple Linear Regression Model

The multivariate extension of multiple linear regression was used to model the relationship between m responses $Y_1, Y_2, \ldots, Y_M$ and a single set of r predictor variables $z_1, z_2, \ldots, z_r$. Each of the m response was assumed to follow its own regression model, so that

$$\mathrm{Y}_i = \beta_{0i} + \beta_{1i} z_i + \beta_{2i} z_i + \ldots + \beta_{ri} z_r + \varepsilon_i \text{ for all } i = 1,2,3,\ldots,m.$$

The vector of error term ε has $\mathrm{E}(\varepsilon) = E\left(\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdots \\ \varepsilon_m \end{bmatrix}\right) = 0$, $Var(\varepsilon) = \Sigma$

Thus, the error terms associated with different responses may be correlated.

Conceptually, we can let $(z_{j0}, z_{j1}, \ldots, z_{jr})$ denote the values of the predictor variables for the j^{th} trial (individual student) and $\mathrm{Y}_j = \begin{bmatrix} \mathrm{Y}_{j1} \\ \mathrm{Y}_{j2} \\ \cdots \\ \mathrm{Y}_{jm} \end{bmatrix}$,

$\varepsilon_j = \begin{bmatrix} \varepsilon_{j1} \\ \varepsilon_{j2} \\ \cdots \\ \varepsilon_{jm} \end{bmatrix}$ be the responses and errors for the j[th] (trial) individual student.

Thus we have a $n \times (r+1)$ design matrix of explanatory (predictor) variables or factors

$$\boldsymbol{Z} = \begin{pmatrix} Z_{10} & Z_{11} \cdots & Z_{1r} \\ Z_{20} & Z_{21} \cdots & Z_{2r} \\ & \cdots\cdots\cdots\cdots & \\ Z_{n0} & Z_{n1} \cdots & Z_{nr} \end{pmatrix}$$

Setting the matrix of response (dependent) variables, $\mathbf{Y}$ and a matrix of fixed unknown parameter, β and matrix of errors ε.

$$\mathrm{Y}_{(n\times m)} = \begin{pmatrix} \mathrm{Y}_{11} & \mathrm{Y}_{12} \cdots & \mathrm{Y}_{1m} \\ \mathrm{Y}_{21} & \mathrm{Y}_{22} \cdots & \mathrm{Y}_{2m} \\ & \cdots\cdots\cdots & \\ \mathrm{Y} & \mathrm{Y} \cdots & \mathrm{Y} \end{pmatrix} = \left(\mathrm{Y}_{(1)} \mid \mathrm{Y}_{(2)} \mid \ldots \mid \mathrm{Y}_{(m)} \right)$$

$$\beta_{((r+1)\times m)} = \begin{pmatrix} \beta_{01} & \beta_{02} \cdots & \beta_{0m} \\ \beta_{11} & \beta_{12} \cdots & \beta_{1m} \\ & \cdots\cdots\cdots & \\ \beta_{r1} & \beta_{r2} \cdots & \beta_{rm} \end{pmatrix} = \left(\beta_{(1)} \mid \beta_{(2)} \mid \ldots \mid \beta_{(m)} \right)$$

$$\varepsilon_{(n\times m)} = \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} \cdots & \varepsilon_{1m} \\ \varepsilon_{21} & \varepsilon_{22} \cdots & \varepsilon_{2m} \\ & \cdots\cdots & \\ \varepsilon_{n1} & \varepsilon_{n2} \cdots & \varepsilon_{nm} \end{pmatrix} = \left(\boldsymbol{\varepsilon}_{(1)} \mid \boldsymbol{\varepsilon}_{(2)} \mid \ldots \mid \boldsymbol{\varepsilon}_{(m)} \right)$$

The multivariate linear regression model is: $Y_{(nxm)} = Z_{(nx(r+1))} \beta_{((r+1)xm)} + \varepsilon_{(nxm)}$ with $E(\varepsilon) = 0$ and $Cov(\varepsilon_i, \varepsilon_k) = \sigma_{ik} \times I$ for $i,k = 1,2,...,m$. The '*m*' observed responses on the j[th] trial (student) have covariance matrix $\Sigma = (\sigma_{ik})$, but observations from different trials (individual students) are uncorrelated ([9]; [21]).

i. *Method of Parameter Estimation*

In the model above β and $\sigma_{i,k}$, i,k =1, 2, 3,…, *m*, are unknown parameters. The ordinary least squares (OLS) estimates $\hat{\underset{\sim}{B}}$ are found in a manner analogous to the uni-variate case. We begin by taking a single response solution as: $\hat{\beta}_{(i)} = (Z'Z)^{-1} Z'Y_{(i)}$. Then collecting the uni-variate least squares estimates yields:

$$\hat{\beta} = (Z'Z)^{-1} Z' \left(Y_{(1)} \mid Y_{(2)} \mid \ldots \mid Y_{(m)} \right) = (Z'Z)\hat{\beta}^{-1} Z'Y.$$

Using a matrix $\hat{\beta}$, one can easily ascertain that the matrices of predicted values: $\hat{Y} = Z\hat{\beta} = Z(Z'Z)^{-1} Z'Y$ and residuals: $\hat{\varepsilon} = Y - \hat{Y} = \left[I - Z(Z'Z)^{-1} Z' \right] Y$. If the model is of full rank, rank (Z) = r + 1 < n; and ε and β are also uncorrelated. Furthermore, because $Y = \hat{Y} + \hat{\varepsilon}$, then one have

$$\underset{\text{Total sums of squares and cross products (SSCP)}}{Y'Y} = \underset{\text{Predicted sums of squares and cross products (PSSCP)}}{\hat{Y}'\hat{Y}} + \underset{\text{Residual sums of squares and cross products (RSSCP)}}{\hat{\varepsilon}'\hat{\varepsilon}}$$

Residual SSCP= $\hat{\varepsilon}'\hat{\varepsilon} = Y'Y - \hat{\beta} Z'Z\hat{\beta}$ and the unbiased estimator of $\boldsymbol{\Sigma}$ is $\hat{\Sigma} = \hat{\varepsilon}'\hat{\varepsilon} / (n - r - 1)$.

ii. *Test of Hypothesis*

The hypotheses of all explanatory have no effect on academic achievements of students jointly on the responses, i.e. the i^{th} school subject doesn't depend on the 'r' explanatory variables: $H_0 : \beta_{(s,i)} = 0$ Vs. $H_a : \beta_{(s,i)} \neq 0$ for all $s = 1,2,3,...,r$ and $i = 1,2,3,...,m$.

A test statistic: $t_{cal} = \hat{\beta}_{(s,i)} \Big/ S.E\left(\hat{\beta}_{(s,i)}\right) \sim t(n-(r+1))$, where $S.E\left(\hat{\beta}_{(s,i)}\right) = \sqrt{\mathbf{var}\left(\hat{\beta}_{(s,i)}\right)}$

$$Var\left(\hat{\beta}_{(i)}\right) = \hat{\sigma}_{i,i}(Z'Z)^{-1} = Diag\left[\hat{Var}(\beta_{0i}), \hat{Var}(\beta_{1i}), \hat{Var}(\beta_{2i}), ..., \hat{Var}(\beta_{r,i})\right], \quad \hat{\sigma}_{i,i} = \hat{\varepsilon}_{(i)}'\hat{\varepsilon}_{(i)} \Big/ n-r-1.$$

Decision Rule: if $t_{cal} > t_{(n-r-1)tab}(\alpha/2)$ or p-value less than $\alpha = 0.05$, we reject the null hypothesis. On the other hand, the confidence ellipsoid for β can be easily contracted with the one-at- a-time *t* value $t_{n-r-1}(\alpha/2)$ and using intervals $\hat{\beta}_i \pm t_{n-r-1}(\alpha/2) \times SE(\hat{\beta}_i)$. Here if the confidence interval includes $\beta_i = 0$, the variable z_i might be dropped out from the regression model [9].

iii. *Checking the Goodness of Fit of the Model*

It is imperative to examine the adequacy of the model before the estimated function becomes a permanent part of the decision making apparatus [9]. All the sample information on lack of fit is contained in the residuals.

iv. *Residuals*

The residuals are defined as:

$$\hat{\varepsilon} = Y - \hat{Y} = \left[I - Z\left(Z'Z\right)^{-1} Z' \right] Y$$

Since a residual may be viewed as the deviation between the data and the fit, it is also a measure of the variability in the response variable not explained by the regression model. Plotting residuals is a very effective way to investigate how well the regression model fits the data and to check the assumptions.

v. *Normal Probability Plot*

The most commonly used methods of checking normality of an individual variable are the Quantile-Quantile plot (Q-Q plot), P-P plot and Normal Curve Histogram. The P-P plotted as expected cumulated probability against observed cumulated probability of standardized residuals – line should be at 45 degrees. The variable is normality distributed if this plot illustrates a linear relationship. In case of the assumption that says the combinations of variables follow a multivariate normal distribution, one can generally test each variable individually and assume that they are multivariate normal if they are individually normal [3]; [1]).

vi. *Ethical Issue/ Considerations*

Ethical approval was obtained from research ethics committee of Hawassa University, Postgraduate school of Computional sciences. Following the endorsement by the research ethics committee and acceptance of the postgraduate school and statistics department, Hawassa City Administration Education and Capacity Building Department was informed about the study through a support letter from Hawassa University research Postgraduate research office. Then verbal permission had been obtained from respective department of the city administration.

Following the endorsement by Hawassa City Administration Education and Capacity Building Department, the selected schools were informed about the objective of the study through a support letter from Hawassa City Administration Education and Capacity Building Department and oral permission and supports were obtained from the respected school principals, teachers and students. As the study was conducted through review of academic records, the individual person was not subjected to any harm as far as the confidentiality is kept. Consent was obtained from individual person or student who was selected to fill the study questionnaire. To preserve the confidentiality, data recorders or file keepers, in the City Administration Education and Capacity Building Department extracted the data from the academic records. Moreover, no personal identifiers were used on data collection form. The recorded data was never accessed by a third person except the principal investigator, and was kept with a firm confidentiality in a secured place.

IV. Results

a) *Descriptive Results*

From the results in Table 3.1, the average academic achievements of students measured in Mathematics, Biology, Physics, Chemistry and English subjects for non-government school students were, respectively, 2.99, 2.97, 2.50, 2.88, and 3.14 with standard deviations 0.822, 0.899, 0.942, 0.806 and 0.805, respectively, and that of government schools were 2.61, 2.73, 2.24, 2.74 and 2.77 with standard deviations 0.838, 0.866, 0.964, 0.872 and 0.802, respectively.

Table 3.1 : Cross Tabulation of School Type and Each School Vs Academic Achievement of Students in Each Selected Subject (Hawassa, 2010)

School Type	School Name		Students' Academic Achievement					Overall
			Maths	Biology	Physics	Chemistry	English	Average
	Comboni	N	35	35	35	35	35	35
		Mean	3.49	3.63	2.37	3.11	3.68	3.25
		SD.	0.743	0.598	0.877	0.758	0.471	0.689
	SOS	N	30	30	30	30	30	30
		Mean	2.83	2.87	2.23	2.83	3.07	2.77
		SD.	0.647	0.973	0.897	0.791	0.827	0.827
	Adventist	N	30	30	30	30	30	30
		Mean	2.90	2.77	2.63	2.77	2.80	2.77
		SD.	0.844	0.817	0.999	0.971	0.805	0.887
	Mount Olive	N	29	29	29	29	29	29
		Mean	2.72	2.55	3.03	2.97	3.10	2.88
		SD.	.702	.783	.778	.778	.772	0.763
Non-government	BNB	N	29	29	29	29	29	29
		Mean	2.83	2.72	2.31	2.79	2.93	2.72
		SD.	.889	.959	.967	.726	.753	0.859

	Evan	N	27	27	27	27	27	27
		Mean	3.11	3.19	2.44	2.78	3.19	2.94
		SD.	.892	.834	.974	.800	.921	0.884
	Total	N	180	180	180	180	180	180
		Mean	2.99	2.97	2.50	2.88	3.14	2.89
		SD.	.822	.899	.942	.807	.806	0.855
	% of Total N		25.0%	25.0%	25.0%	25.0%	25.0%	25%
Government	Addis Ketema	N	140	140	140	140	140	140
		Mean	2.44	2.62	2.19	2.64	2.64	2.50
		SD.	.915	.978	.853	.866	.778	0.878
	Tabor	N	144	144	144	144	144	144
		Mean	2.64	2.74	2.22	2.91	2.72	2.65
		SD.	.744	.729	1.020	.860	.848	0.840
	Alamura	N	74	74	74	74	74	74
		Mean	2.58	2.70	2.31	2.66	2.90	2.63
		SD.	0.827	0.789	0.842	.865	.847	0.834
	Tulla	N	87	87	87	87	87	87
		Mean	2.52	2.91	2.31	2.76	2.88	2.68
		SD.	.744	.923	1.015	.889	.672	0.849
	Adare	N	94	94	94	94	94	94
		Mean	2.94	2.79	2.23	2.73	2.87	2.71
		SD.	.865	.878	1.082	.869	.819	0.903
	Total	N	539	539	539	539	539	539
		Mean	2.61	2.74	2.24	2.75	2.78	2.62
		SD.	.838	.867	.964	.872	.803	0.869
	% of Total N		75.0%	75.0%	75.0%	75.0%	75.0%	75%

Table 3.2 : Descriptive Statistics Student's Achievement in Ascending Order for the Overall Sample of Students (Hawassa, 2010)

	N	Mean	SD	CV%
English	719	2.87	0.819	28.512
Biology	719	2.79	0.880	31.474
Chemistry	719	2.78	0.858	30.819
Mathematics	719	2.71	0.849	31.386
Physics	719	2.31	0.965	41.852

Table 3.2 shows the mean academic achievements and the coefficient of variations for the five subjects. In terms of coefficient of variation, the variability was the lowest for English and highest for Physics subjects. This may indicate that students' achievements were most consistent for the English subject and least consistent for Physics subject. Physics was considered as difficult subject for many students.

b) Results of Factor Analysis

Before conducting the central MVML and multivariate multiple regression analyses it is important first to establish the psychometric properties of the instrument used. Principal Component Factor Analysis was done in two steps. The first one was a general PCFA that considered the socio-economic and demographic variables with general school characteristic variables and the second was a separate PCFA relative to each achievement measures of the five subjects. This provided component factors for each of the five school subjects each based on the subject related observed items as students' responses on their personal, school and teacher characteristic variables relative to school subjects. The overall reliability was computed to be Cronbach's alpha=0.724 indicating that the questionnaire items were consistent.

Table 3.3 : KMOs and Bartlett's Tests for Factor Analyses

Responses		Separate Principal Component Factor Analysis: Maths	Biology	Physics	Chemistry	English	General PCFA
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.805	0.771	0.821	0.777	0.838	0.789
Bartlett's Test of Sphericity	Approx. Chi-Square	11170.0	8459.0	14820.0	6293.0	9703.0	8391.0
	df	153	190	231	210	231	300
	P-value	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*

**Significant (P-value <0.05)*

The KMO statistic values test if sufficient items (by partial correlation among variables) are available for each factor component in the factor analysis. KMO statistic for the separate PCFA with respect to the school subjects Mathematics, Biology, Physics, Chemistry and English were 0.81, 0.77, 0.82, 0.78 and 0.84, respectively; with the general PCFA of 0.79. These were all greater than 0.5 indicating that the sampling was adequate for factor analysis and there were significant relationships among the perceived factors of achievements in the school subjects.

The data were also checked for Bartlett's test of Sphericity to see that the original variables were sufficiently (bi-variate) correlated and these met the criteria with

$$\chi^2_{153,\text{Mathematics}} = 11170.0\ (P - value < 0.001),\ \chi^2_{190,\text{Biology}} = 8459.0\ (P - value < 0.001),$$

$$\chi^2_{231,Physics} = 14820.0\ (P - value < 0.001),\ \chi^2_{210,Chemistry} = 6293.0\ (P - value < 0.001),$$

$$\chi^2_{231,English} = 9703.0\ (P - value < 0.001) \text{ and } \chi^2_{300,General} = 8391.00\ (P - value < 0.001).$$

These indicated that the original observed variables were sufficiently correlated (the variables were not completely uncorrelated) and factor analysis was possibly appropriate in each case. The output matrixes contained the loading of each variable onto each factor. All loadings less than 0.5 were suppressed in the output and so were blank spaces for many of the loadings. Thus, the loadings were acceptable and easy for interpretation.

The results of separate factor analysis (with factor loadings greater than 0.5) are presented in Tables 4, 5, 6, 7 and 8 of Appendix-1 and Figures 1 in Appendix-2 of the Scree plots. The criteria that the required amount of explained variation accounted for being large, logical interpretability of factors and Scree plot tests were considered with Kaiser Criteria. Kaiser criteria is accurate when there are less than 30 variables with lager sample and communalities after extraction being greater than 0.6. Depending on the correlation matrix and communalities, some observed variables were rejected. Of all 140 observed items, using principal component extraction and Varimax rotation, the study found factor solution of the 28-variables for each subject. Then, six underlying common factors were obtained for each separate factor analysis of Biology, Physics, Chemistry and English related items that constituted or explained 76.67%, 78.80%, 68.64% and 73.43% of the total variability in the corresponding original observed variables, respectively. There were four common factors for Mathematics related items which constituted or explained 77.38% of the total variability in the original observed variables related to Mathematics.

Factor scores of each component factor for each of the 719 individual respondents were computed and these scores were used as data for further analysis. The common factors obtained from the general and separate PCFAs which were used as covariates,

- School facilities (SF),
- Family status (FS),
- School volume (SV),
- Safe reading (SafR),
- Interest (InterstS) to the subjects,
- Motivation (MotivS) to the subjects,
- Trouble (TroubS) to the subjects and
- Self-concept (SelfC) to the subjects.

Table 3.4 : The Generalized Principal Component Factor Analysis (Hawassa, 2010)

Accounted for 64.28%	Common Factors: Component 1	2	3	4	5	6	Communality
Eigenvalues	4.56	3.50	2.70	2.21	1.65	1.45	
Variations accounted for %	18.24	14.0	10.8	8.84	6.60	5.80	
Parent student communication	.902						0.841

Parent teacher communication	.891						0.813
Satisfaction in food at home	.853						0.767
Availability of books at home	.840						0.750
Mother education level	.824						0.685
Father education level	.795						0.702
School amenity		.763					0.606
School instructional materials		.729					0.685
School human resources		.725					0.591
School laboratory facilities		.684					0.654
School library facilities		.635					0.568
School equipment		.619					0.588
School academic and counseling services		.540					0.533
Satisfaction in school administration			.932				0.891
Student confidence			.823				0.770
School health services			.776				0.692
School rules and regulations			.533				0.554
Teacher average experience				.765			0.627
Teacher average work load				.764			0.593
School size (total number of students in the school)				.724			0.551
Class size (number of students in per class)				.603			0.643
Distance from home to school					.831		0.694
Home location for the school					.815		0.706
Comfort of studying at school						.771	0.614
Comfort of studying at home						.687	0.571

- *Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.*
- *Loadings Less than 0.5 were suppressed.*

c) *Results of Multivariate Multiple Linear Regression Analysis*

Multivariate multiple linear regression analysis was used to examine the effect of independent variables or factors on the outcome variables, i.e. academic achievement in selected subjects. Most of the explanatory variables were the common factors obtained from the general PCFA and some were the regularly appeared component factors in each separate PCFA.

Table 3.5 : Model Summary of Multivariate Multiple Linear Regression Model

	Responses				
	Mathematics	Biology	Physics	Chemistry	English
R^2	0.72	0.74	0.68	0.68	0.71
$R^2_{adj.}$	0.64	0.67	0.61	0.64	0.65

The results are shown in Table 3.6. In this analysis the overall determinants of academic achievement were assessed in terms of the five school subjects to identify the basic determinant factors for both government and non-government schools taken together. The factors sex, school type and school facilities (SF) were found to be jointly statistically significant for achievements in all the five selected school subjects. Family status (FS) was significant for achievements in the four school subjects (Biology, Physics, Chemistry and English) but statistically insignificant for achievement in Mathematics. School volume (SV) has a significant influence on achievements in the two school subjects Biology and Chemistry. Interest to the subjects (InterstS) has a significant influence on academic achievements of students in Biology and Physics. Moreover, the factors trouble to the subject (TroubS) and motivation (MotivS) to the subject in terms of Mathematics, self-concept (SelfC) in terms of Physics and students future aspiration (FutureAspira) in terms of Physics and Chemistry had significant impact on student's academic achievement as observed in overall combined data of government and non-government schools.

Moreover, the factors such as sex, interest to the subject (InterstS), motivation to the subject (MotivS), self-concept (SelfC), family status (FS), school facilities (SF) and future aspiration (Future Aspira) had positive impacts on students' academic achievements of the school subjects. However, trouble (TroubS) of the subjects and school volume (SV) showed significant negative impact on students' achievements of all the five subjects.

Table 3.6 : Parameter Estimates of Multivariate Multiple Liner Regression for Overall Samples Data

Dependent Variable	Parameter	$\hat{\beta}$	S.E.	t-value	P-value	95% CI	
						Lower	Upper
Maths	Intercept	3.042	.117	26.105	.000*	2.814	3.271
	Sex	.149	.064	2.316	.021*	.023	.275
	School Type	-.410	.080	-5.121	.000*	-.567	-.253
	Interest to the subject (InterstS)	.008	.013	.658	.511	-.017	.034
	Trouble of the subject (TroubS)	-.037	.014	-2.669	.008*	-.064	-.010
	Self-concept (SelfC)	.025	.014	1.837	.067	-.002	.052
	Motivation to the subject (MotivS)	.028	.010	2.866	.004*	.009	.048
	School facilities (SF)	.093	.032	2.917	.004*	.030	.156
	School volume (SV)	-.030	.034	-.885	.377	-.097	.037
	Family status (FS)	-.023	.032	-.720	.472	-.085	.039
	Future aspiration (FutureAspira)	-.020	.015	-1.324	.186	-.050	.010
Biology	Intercept	2.808	.119	23.621	.000*	2.575	3.041
	Sex	.291	.066	4.436	.000*	.162	.420
	School Type	-.309	.082	-3.781	.000*	-.469	-.148
	Interest to the subject (InterstS)	.027	.013	2.028	.043*	.001	.052
	Trouble of the subject (TroubS)	-.025	.014	-1.757	.079	-.053	.003
	Self-concept (SelfC)	.000	.014	.032	.974	-.027	.028
	Motivation to the subject (MotivS)	-.007	.010	-.698	.486	-.027	.013
	School facilities (SF)	.210	.033	6.447	.000*	.146	.273
	School volume (SV)	-.075	.035	-2.138	.033*	-.143	-.006
	Family status (FS)	.128	.032	3.943	.000*	.064	.191
	Future aspiration (FutureAsp)	.010	.016	.672	.502	-.020	.041
Physics	Intercept	1.941	.131	14.826	.000*	1.684	2.198
	Sex	.204	.072	2.815	.005*	.062	.345
	School Type	-.172	.090	-1.916	.056	-.349	.004
	Interest to the subject (InterstS)	.033	.014	2.293	.022*	.005	.061
	Trouble of the subject (TroubS)	.001	.016	.065	.948	-.030	.032
	Self-concept (SelfC)	.034	.015	2.242	.025*	.004	.065
	Motivation to the subject (MotivS)	.002	.011	.140	.889	-.020	.023
	School facilities (SF)	.087	.036	2.423	.016*	.016	.157
	School volume (SV)	.030	.038	.772	.440	-.046	.105
	Family status (FS)	.229	.036	6.428	.000*	.159	.299
	Future aspiration (FutureAspira)	.070	.017	4.050	.000*	.036	.103
Chemistry	Intercept	2.780	.118	23.559	.000*	2.549	3.012
	Sex	.338	.065	5.193	.000*	.210	.466
	School Type	-.274	.081	-3.375	.001*	-.433	-.114
	Interest to the subject (InterstS)	.013	.013	1.020	.308	-.012	.039
	Trouble of the subject (TroubS)	-.013	.014	-.936	.350	-.041	.014
	Self-concept (SelfC)	-.003	.014	-.231	.817	-.030	.024
	Motivation to the subject (MotivS)	-.010	.010	-1.000	.318	-.030	.010
	School facilities (SF)	.144	.032	4.471	.000*	.081	.208
	School volume (SV)	-.113	.035	-3.275	.001*	-.181	-.045
	Family status (FS)	.069	.032	2.138	.033*	.006	.132
	Future aspiration (FutureAspira)	.003	.015	.214	.831	-.027	.034
English	Intercept	3.065	.112	27.250	.000*	2.844	3.286
	Sex	.155	.062	2.497	.013*	.033	.277
	School Type	-.387	.077	-5.002	.000*	-.538	-.235
	Interest to the subject (InterstS)	.002	.012	.179	.858	-.022	.027
	Trouble(anxiety) of the subject (TroubS)	-.006	.013	-.453	.651	-.032	.020
	Self-concept (SelfC)	.004	.013	.304	.762	-.022	.030
	Motivation to the subject (MotivS)	.017	.010	1.764	.078	-.002	.035
	School facilities (SF)	.151	.031	4.920	.000*	.091	.212
	School volume (SV)	-.030	.033	-.907	.364	-.095	.035
	Family status (FS)	.061	.031	1.984	.048*	.001	.121
	Future aspiration (FutureAspira)	.002	.015	.109	.914	-.027	.031

**Significant (P-value < 0.05)*

V. Discussions and Conclusions

The PCFA technique was used as separate PCFA of items with respect to the each five responses and the general PCFA incorporated other general student with family and school with teacher characteristics variables in the data reduction. The multivariate single level multiple linear regression was applied on overall schools data. The results obtained are discussed as follows:

On an average, students, in non-government secondary schools, performed better than those in government secondary schools in almost all the achievement measures of the five school subjects. This might be because of higher availability of school and home educational supply and facilities, better study positions and higher parental involvement with teachers and students at the schools as compared to that at government schools. Moreover, on overall average, male students achieved better in almost all school subjects than female students. This implied that the school and family might treat gender differently and the variation in students' personal factors such as trouble to the subjects, self-concept, interest and motivation to the subjects showed significant impact on students' achievement ([4]; [6]; [20]).

The results obtained from the separate PCFA in each achievement measuring response indicated that about four factors related to Mathematics and six factors related to Biology, Physics, Chemistry and English were sufficient to explain the total achievement variability. Thus, factors self-concept to the subjects, motivation to the subjects, interest to the subjects, trouble (anxiety) to the subjects, teaching-learning process and absenteeism explaining most of the achievement variations in five school subjects. Moreover, the result of general PCFA indicated that the factor named as family status (FS) that encompasses parent-student communication, parent-teacher communication, availability of book at home, satisfaction in food available at home, mother educational level and father education level explained the higher variability for the overall achievement. This finding is in consistent with other studies ([13]; [18]).

The result of the multivariate multiple linear (single-level) regression analysis point to several interesting overall findings. The result indicated that the factors sex, school type, school facility (SF) which encompassed availability and satisfactoriness of school amenity, human resources, library, laboratory, equipment and academic counseling have significant impacts on achievements of the students in terms of all the selected five subjects. School volume (SV) that encompassed school size, class size, teacher workload and experience had a significant negative impact on academic achievements in terms of Biology and Chemistry. This may be due to the negative effect of school size, class size and teacher work load on academic achievement of students at school, as reported earlier ([16]; [17]).

The factor school facility (SF) that deals availability and satisfactoriness of the school instructional materials, school library, laboratories, amenities, academic counseling services and other school characteristics had significant positive impact in all five school subjects used as a measures of academic achievement. Family status (FS) which encompass parent-student communication, parent-teacher communication, availability of book at home, satisfaction in food available at home, mother educational level and father education level had a significant positive impact on academic achievements in terms of Biology, Physics, Chemistry and English subjects as observed earlier ([18]; [19]).

This study was intended to identify some factors influencing the academic achievements of students' measured by five selected subjects (Mathematics, Biology, Physics, Chemistry and English) at secondary school level based on primary and secondary data. Accordingly, factor analysis, multivariate multiple linear regression and MVML multiple linear regression techniques on the five school subjects were employed.

The factor analyses conducted in this study indicated that 4 or 6 factors (instead of twenty eight original observed variables or items) were sufficient to explain 77.4%, 76.7%, 78.8%, 68.6% and 73.4% the total variation in achievement for each separate PCFA of observed items related to Mathematics, Biology, Physics, Chemistry and English subjects, respectively. The factors self-concept, motivation, interest and trouble to the subject were the common factors explaining most of the variability of achievements in terms of each five subject, since these factors were appeared regularly in each separate PCFA. Moreover, six common factors were enough to explain about 64% of the variation using 34 originally observed variables in the generalized PCFA.

The study revealed that the factors sex, school type, family status (FS) holding parents–student communication, parent-teacher communication, satisfaction in food available at home, availability of books at home, mother educational level and father education level, and school facility (SF) enclosing school instructional materials, amenities, library and laboratory facilities had statistically significant influence on achievements of students for the selected subjects. Moreover, school volume (SV) that covers school size, class size, teacher work load and teacher experience in teaching; interests to the subject, motivation to the subject, trouble to the subject and self-concept in school subjects have been significant factors

Global Journal of Human-Social Science (G) Volume XVI Issue II Version I Year 2016

influencing students achievement on the school subjects.

VI. Acknowledgements

First, and foremost, I would like to thank my God for giving me the opportunity to allow me learning Statistical and doing researches in every fields of sciences, and for being with me in all aspects elsewhere.

I would like to gratefully and sincerely thank Dr. Ayele Taye, for his invaluable comments and suggestions that contributed to the successful realization of this study. He devoted his time to provide me insightful feedback on drafts of this manuscript and challenged me to become a better researcher and scholar. I would like to acknowledge the efforts of the School of Natural and Computional sciences and all my colleagues for their scholarly input I got through the discussions.

My special thanks also go to the staff members of the Department of Education and Capacity Building at Hawassa City Administration, school principals and home room teachers of all secondary schools in the city for their positive and genuine cooperation provided to me during data collection and those selected grade 10 students who gave me the information and data used in this study.

I owe tremendous amount of gratitude to my parents, friends and relatives; specifically, I feel greatly indebted to my brother Awol Assen, my friends Bereket Adem, Hussien Ali and Mebrat Seid who were never failed to give me support, continuous encouragement and appreciation for my studies.

At last, my deep appreciation and heart-felt thanks goes to my wife Medina Yimer Muhie who is always there to provide me with continuing motivation, encouragement and constructive advice.

References Références Referencias

1. Alyee, S. A. (1998). Ways to Evaluate the Assumption of Normality, Paper Presented at Annual Meeting of the Southwestern Psychological Association.
2. Antonia, L. (2002). Personal, Family and Academic Factors Affecting Low Achievement in Secondary School, Journal of Research in Education Psychology and Psycho Pedagogy, ISSN: 1696-2095.
3. Burdenski, T., Texas, A. and University, M. (2000). Evaluating Univariate, Bivariate, and Multivariate Normality Using Graphical and Statistical Procedures, Multiple Linear Regression View Points.
4. Christopher, L. and Sarah, T. (2006). Charter, Private, Public Schools and Academic Achievement: New Evidence from NAEP Mathematics Data, Columbia.
5. Cochran, W.G. (1997). Sampling Techniques, (3rd ed.), John Wiley and Sons, New York.
6. CSA. (2008). Summary and Statistical Report of 2007 Population and Housing Census. Addis Abeba, Ethiopia.
7. Getahun, W. (2002). School Based Factors Contributing to Differences in Students' Achievement Between High and Low Passing Rate Scorers of Government Secondary Schools of Addis Ababa at National Examination, Faculty of Education, Addis Ababa University, Ethiopia.
8. Hanna, T. (2010). Statistical Analysis of Determinants of Academic Performance of Students at Government Schools and University Levels, Thesis, Department of Statistics, Hawassa University, Ethiopia.
9. Johnson, A. (2007). Applied Multivariate Statistical Analysis (6th ed.), Pearson Prentice Hall, Trademark of Pearson Education, Upper Saddle River, New Jersey 07458.
10. Kassahun, M. and Zelalem, T. (2006). Assessment on the Impact of Plasma Television Implementation on the Teaching Learning Process of Mathematics Class: The Case on Selected High Schools, Mathematics Department, Education Faculty, Jimma University, Jimma, Ethiopia.
11. Nigussie, Y. (2010). Multilevel Analysis of Determinants of Academic Achievement of Grade Eight Students: A Case Study at Hawassa Town. M.Sc. Thesis. Hawassa University, Ethiopia.
12. Nugussie, T. (2002). The Content Validity of the Ethiopian General Secondary Education Certificate English Examination, M.Sc Thesis in Curriculum and Instruction, Addis Ababa University.
13. Tadesse, S. (2009). Understanding Female Students' Academic Performance: An Exploration of the Situation in South Nations Nationalities and Peoples Regional State, Ethiopia, International Institute of Social Studies.
14. UNESCO-WDE. (2007). World Data on Education (6 ed.), Ethiopia.
15. Zuccaro, C. (2007). Statistical Alchemy-The Use of Factor Scores in Linear Regression.
16. Adebola O. and Ademola I., (2011). School Quality Factors And Secondary School Students' Achievement In Mathematics In South-Western And North-Central Nigeria, the African Symposium: An online journal of the African Educational Research Network.
17. Anastacia, N. (20011). Factors Affecting the National Achievement Test Performance Of Selected Second Year High School Students, In Santa Maria, Bulacan.
18. Engin-Demir, C. (2009). Factors Affecting the Academic Achievement of Turkish Urban Poor,

International Journal of Educational Development, 29 (1):17 – 29.

19. Brian, H. and Russell, K. (2009). LANNA Tests and the Prediction of Year 10 English and Mathematics Results: Charles Sturt University, Educational Research, 19(1).
20. Alemayehu B., and Assaye A., (2010). An evaluation of grades 9 and 10 mathematics textbooks vis-à-vis fostering problem solving skills, Bahir Dar University, Ethiopia. Educational Expert, Bureau of Education, Amhara regional State, Ethiopia.
21. Alan J. Izenman (2010). Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning.

Appendices

Appendix-1

Table 1 : Separate Principal Component Factor Analysis of Items Related to Maths as Opinions of 719 Sample Students (Cronbach's =0.73, Hawassa, 2010)

Accounted for 77.38%	Common Factors: Components				Communality
	1	2	3	4	
Eigenvalues	4.56	3.41	3.28	2.67	
% Variations accounted for	25.42	18.93	18.19	14.84	
Need to do Maths well to get into the University	.905				0.832
The teacher prepares well for Maths daily lessons	.894				0.813
Need to do Maths well to get job	.873				0.782
Learning Maths helps me in my daily life	.845				0.717
Exam questions of Maths are standard	.839				0.726
Teaching Maths covers the whole syllabus	.828				0.692
Often study Maths in groups		.944			0. 907
Maths is difficult to learn		.912			0.861
No strength in learning Maths		.895			0.810
Need lots of hard work studying Maths to perform well		.894			0.835
Teaching method used by Maths teacher fits with the current curriculum			.933		0.880
I am satisfied with the current curriculum of Maths			.921		0.857
Maths need more time to understand			.886		0.809
Maths is Boring			.833		0.729
I usually do Maths well				.898	0.819
Enjoy learning Maths				.874	0.784
I have natural talent in Maths.				.764	0.603
Understand Maths quickly in class				.625	0.568

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

Table 2 : Separate Principal Component Factor Analysis of Items Related to Biology as Opinions of 719 Sample Students (Cronbach's α =0.72, Hawassa, 2010)

Accounted for 76.65%	Common Factors: Components						Communality
	1	2	3	4	5	6	
Eigenvalues	4.13	3.52	2.34	2.17	1.63	1.55	
%Variations accounted for	20.63	17.57	11.69	10.87	8.14	7.75	
The teacher prepares well for Biology daily lessons	.914						0.853
Need to do Biology well to get into the Preparatory or University	.907						0.847
Need to do Biology well to get job	.902						0.820
Learning Biology helps me in my daily life	.897						0.833
Teacher is efficient and skilled while teaching Biology	.864						0.754
I usually do Biology well		.949					0.922
Understand Biology quickly in class		.911					0.864
Enjoy learning Biology		.900					0.844

I have natural talent in Biology		.900					0.848
No strength in learning Biology			.893				0.845
Biology is difficult to learn			.886				0.834
Need lots of hard work studying Biology to perform well			.795				0.661
Biology need more time to understand				.880			0.786
I am satisfied with the current curriculum of Biology				.845			0.741
Biology is Boring				.798			0.697
Biology teacher often absent from class					.799		0.668
Biology teacher is often late for class					.724		0.574
Student absent from Biology class at least one per week					.626		0.644
The most preferred time of studying for Biology						.862	0.761
The most preferred study place for Biology						.855	0.740

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

Table 3 : Separate Principal Component Factor Analysis of Items Related to Physics as Opinions of 719 Sample Students (Cronbach's α =0.71, Hawassa, 2010)

Accounted for 78.80%	Common Factors: Components						Communality
	1	2	3	4	5	6	
Eigenvalues	3.90	3.75	3.64	2.86	1.60	1.58	
% Variations accounted for	17.75	17.06	16.55	13.01	7.27	7.18	
Need to do Physics well to get job	.898						0.835
Need to do Physics well to get into the Preparatory or University	.871						0.789
Teaching Physics covers the whole syllabus	.850						0.784
Learning Physics helps me in my daily life	.837						0.738
I have natural talent in Physics.	.794						0.676
I usually do Physics well		.967					0.951
Understand Physics quickly in class		.965					0.952
Enjoy learning Physics		.964					0.948
Physics need more time to understand		.956					0.923
Teacher is efficient and skilled while teaching Physics			.941				0.913
I am satisfied with the current curriculum of Physics			.940				0.933
Often study Physics in groups			.929				0.502
Need lots of hard work studying Physics to perform well			.911				0.888
Physics is Boring				.977			0.961
Physics is difficult to learn				.973			0.955
No strength in learning Physics				.947			0.899
Physics teacher is often late for class					-.730		0.609
Student get at least a onetime Physics homework /assignments/ class works per week					.697		0.565
Physics teacher often absent from class					-.688		0.596
Exam questions of Physics are standard						.778	0.641
Teaching method used by Physics teacher fits with the current curriculum						.648	0.547
The teacher prepares well for Physics daily lessons						.556	0.573

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

Global Journal of Human-Social Science (G) Volume XVI Issue II Version I Year 2016

Table 4 : Separate Principal Component Factor Analysis of Items Related to Chemistry as Opinions of 719 Sample Students (Cronbach's α =0.74, Hawassa, 2010)

Accounted for 68.64%	Common Factors: Components						Communality
	1	2	3	4	5	6	
Eigenvalues	3.03	2.70	2.64	2.32	1.88	1.85	
% Variations accounted for	14.42	12.85	12.56	11.05	8.96	8.81	
Teaching method used by Chemistry teacher fits with the current curriculum	.884						0.802
Exam questions of Chemistry are standard	.870						0.793
The teacher prepares well for Chemistry daily lessons	.837						0.742
There are enough text and reference books at school for Chemistry	.809						0.671
I usually do Chemistry well		.804					0.708
I have natural talent in Chemistry		.792					0.685
Enjoy learning Chemistry		.760					0.669
Understand Chemistry quickly in class		.744					0.577
No strength in learning Chemistry			.832				0.710
Chemistry is Boring			.824				0.742
Chemistry is difficult to learn			.804				0.723
Need lots of hard work studying Chemistry to perform well			.669				0.596
Need to do Chemistry well to get into the Preparatory or University				.867			0.781
Need to do Chemistry well to get job				.859			0.763
Learning Chemistry helps me in my daily life				.850			0.745
Chemistry teacher often absent from class					.870		0.781
Student absent from Chemistry class at least one per week					.821		0.678
Chemistry teacher is often late for class					.638		0.554
Often study Chemistry in groups						.813	0.676
I am satisfied with the current curriculum of Chemistry						.785	0.674
Chemistry need more time to understand						.708	0.552

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

Table 5 : Separate Principal Component Factor Analysis of Items Related to English as Opinions of 719 Sample Students (Cronbach's α =0.72, Hawassa, 2010)

Accounted for 73.43%	Common Factors: Components						Communality
	1	2	3	4	5	6	
Eigenvalues	3.66	3.59	3.50	2.70	1.45	1.25	
% Variations accounted for	16.65	16.31	15.93	12.28	6.59	5.67	
Need to do English well to get into the Preparatory or University	.862						0.758
Learning English helps me in my daily life	.857						0.751
The teacher prepares well for English daily lessons	.833						0.738
Need to do English well to get job	.827						0.741
Teacher is efficient and skilled while teaching English	.784						0.683
Exam questions of English are standard		.850					0.774
Need lots of hard work studying English to perform well		.845					0.732
I am satisfied with the current curriculum of English		.844					0.783
Teaching method used by English teacher fits with the current curriculum		.843					0.797
English need more time to understand		.743					0.577

I have natural talent in English			.936				0.898
Enjoy learning English			.934				0.895
Understand English quickly in class			.934				0.894
I usually do English well			.869				0.768
No strength in learning English				.868			0.789
English is difficult to learn				.842			0.723
English is Boring				.836			0.744
Student absent from English class at least one per week				.665			0.573
English teacher often absent from class					.824		0.710
English teacher is often late for class					.802		0.684
The most preferred study place for English						.782	0.629
The average time spent on studying English (hours)						.758	0.611

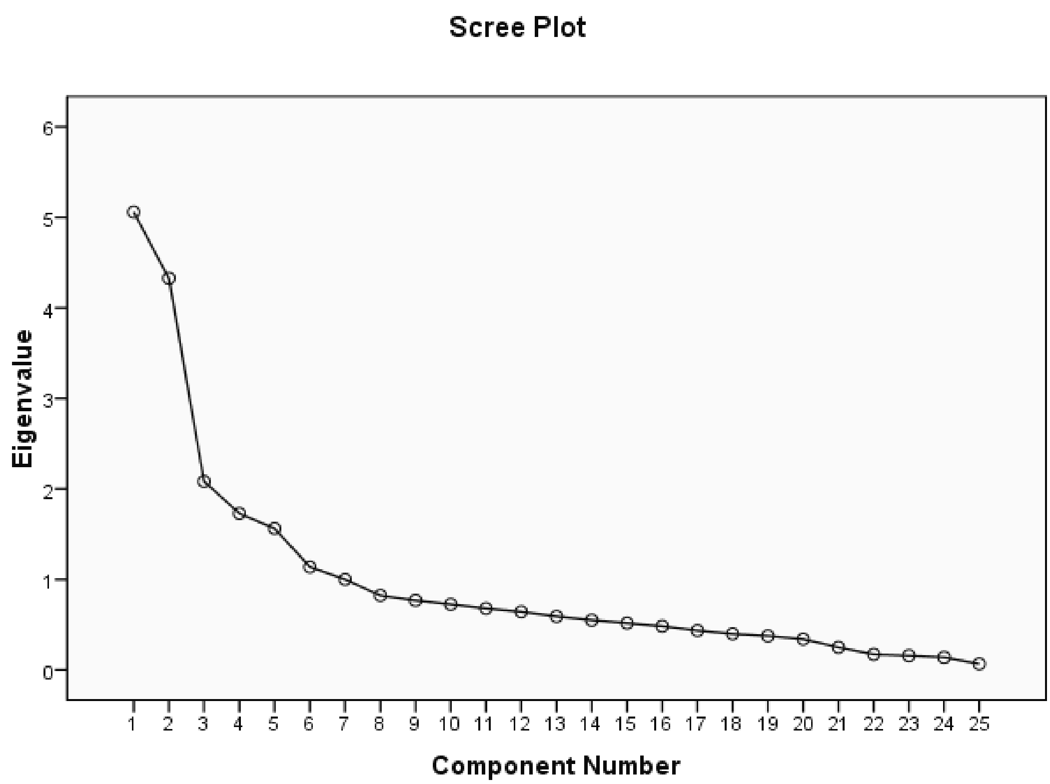
Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

Table 6 : Correlation among Component Factors That Used as Predictors (Covariates)

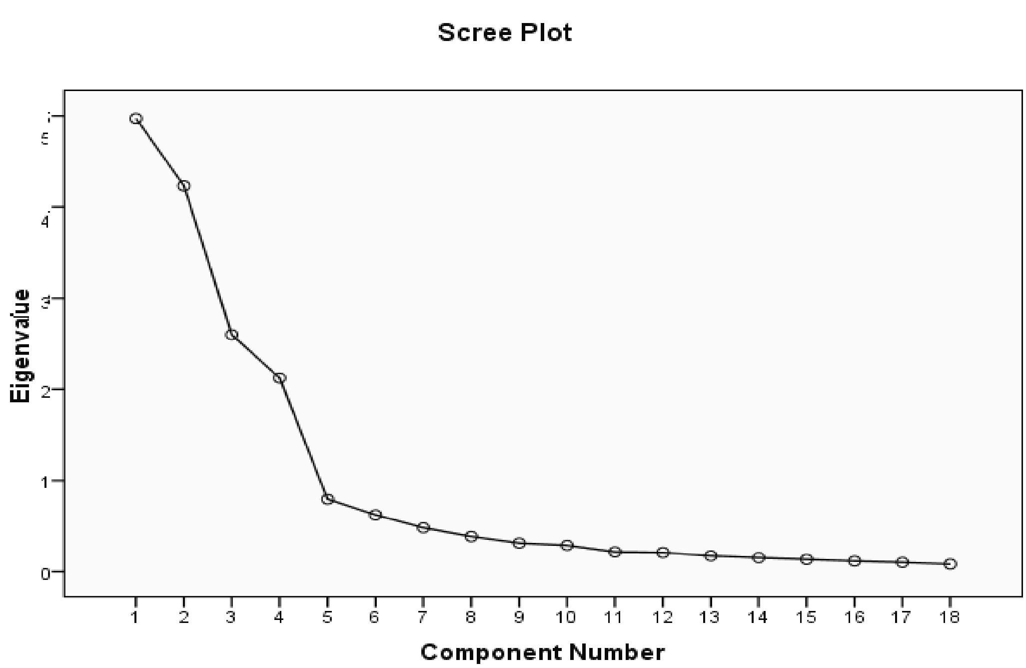
	Sex	ScType	InterstS	TroubS	SelfC	MotivS	SF	SV	FS	SafR	FutA
Sex	1										
ScType	.038	1									
InterstS	.011	-.012	1								
TroubS	-.030	.119**	.117**	1							
SelfC	-.050	.042	.261**	.007	1						
MotivS	.114**	-.114**	-.112**	-.023	-.047	1					
SF	-.059	.063	-.033	-.148**	.175**	-.120**	1				
SV	.034	-.438**	-.076*	-.019	-.064	.104**	.000	1			
FS	-.240**	-.085*	-.084*	-.092*	-.032	.021	.000	.000	1		
SafR	-.076*	-.055	.065	-.003	.005	-.061	.000	.000	.000	1	
Future Aspire	-.035	-.118**	.032	-.011	-.077*	.307**	-.142**	.083*	.045	-.063	1

***.Correlation is significant at the 0.01 level (2-tailed)*
**. Correlation is significant at the 0.05 level (2-tailed)*

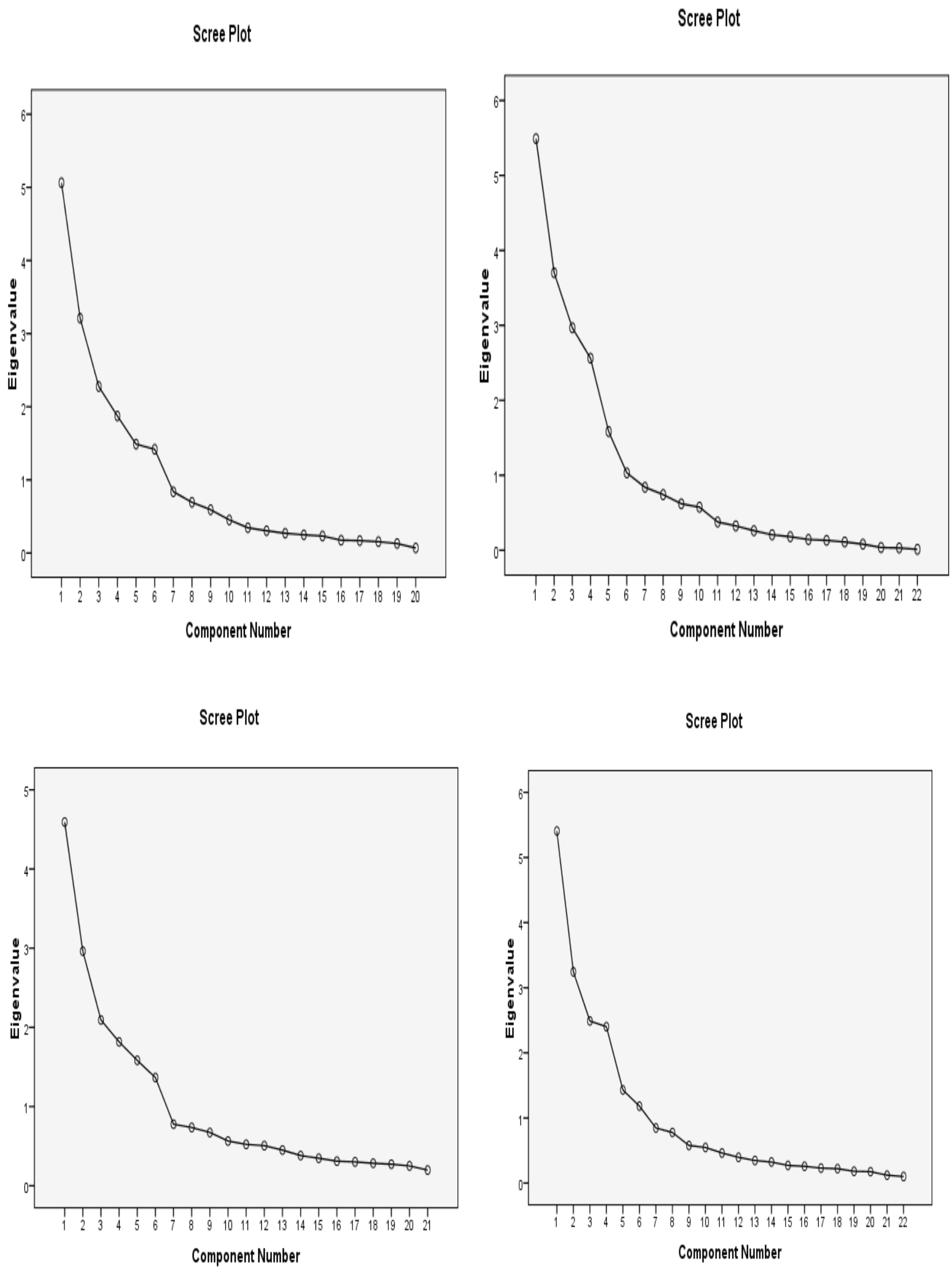
Appendix-2: Graphs/Figures for Checking Model Adequacy



The Generalized PCFA



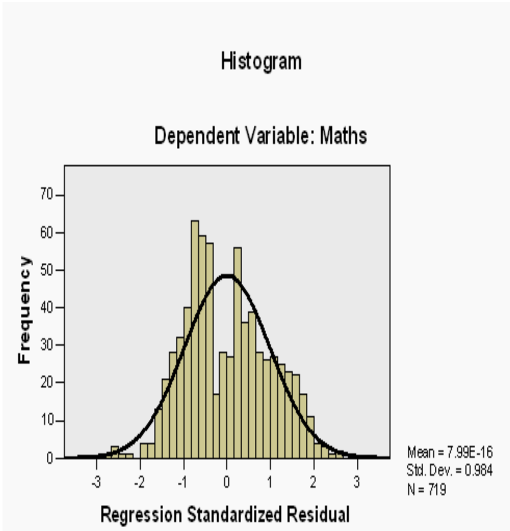
PCFA of Mathematics Related Items

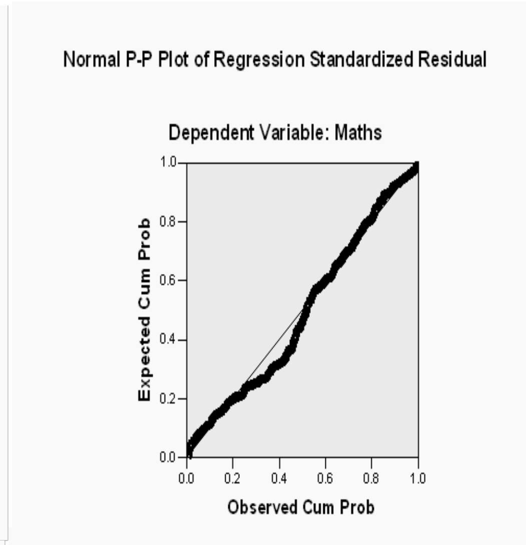


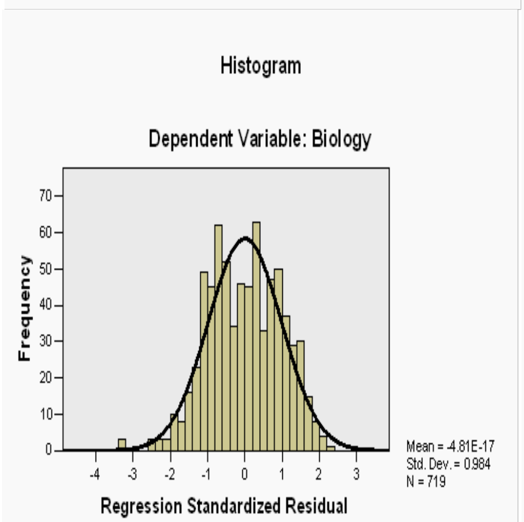
PCFA of Chemistry Related Items

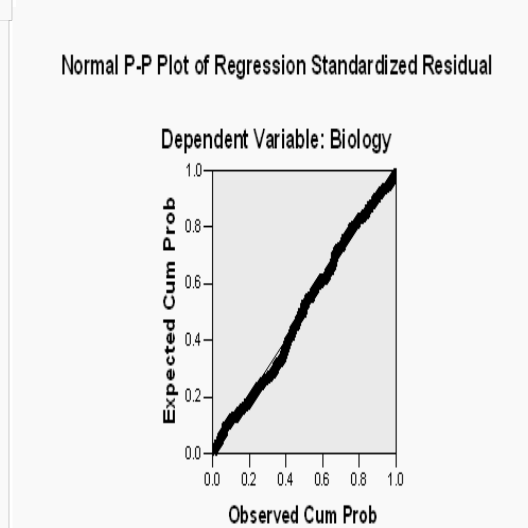
PCFA of English Related Items

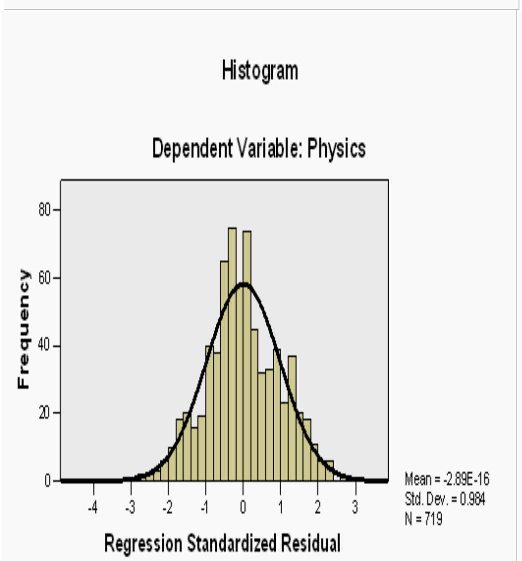
Figure 1 : The Scree Plots to Test for the Number of Factors Retained in the Generalized and Separate PCFA, Respectively

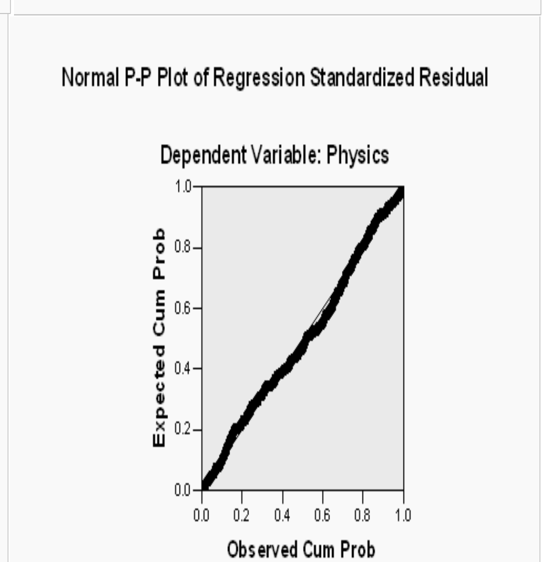

Histogram
Dependent Variable: Maths
Frequency
Regression Standardized Residual
Mean = 7.99E-16
Std. Dev. = 0.984
N = 719


Normal P-P Plot of Regression Standardized Residual
Dependent Variable: Maths
Expected Cum Prob
Observed Cum Prob


Histogram
Dependent Variable: Biology
Frequency
Regression Standardized Residual
Mean = -4.81E-17
Std. Dev. = 0.984
N = 719


Normal P-P Plot of Regression Standardized Residual
Dependent Variable: Biology
Expected Cum Prob
Observed Cum Prob


Histogram
Dependent Variable: Physics
Frequency
Regression Standardized Residual
Mean = -2.89E-16
Std. Dev. = 0.984
N = 719


Normal P-P Plot of Regression Standardized Residual
Dependent Variable: Physics
Expected Cum Prob
Observed Cum Prob

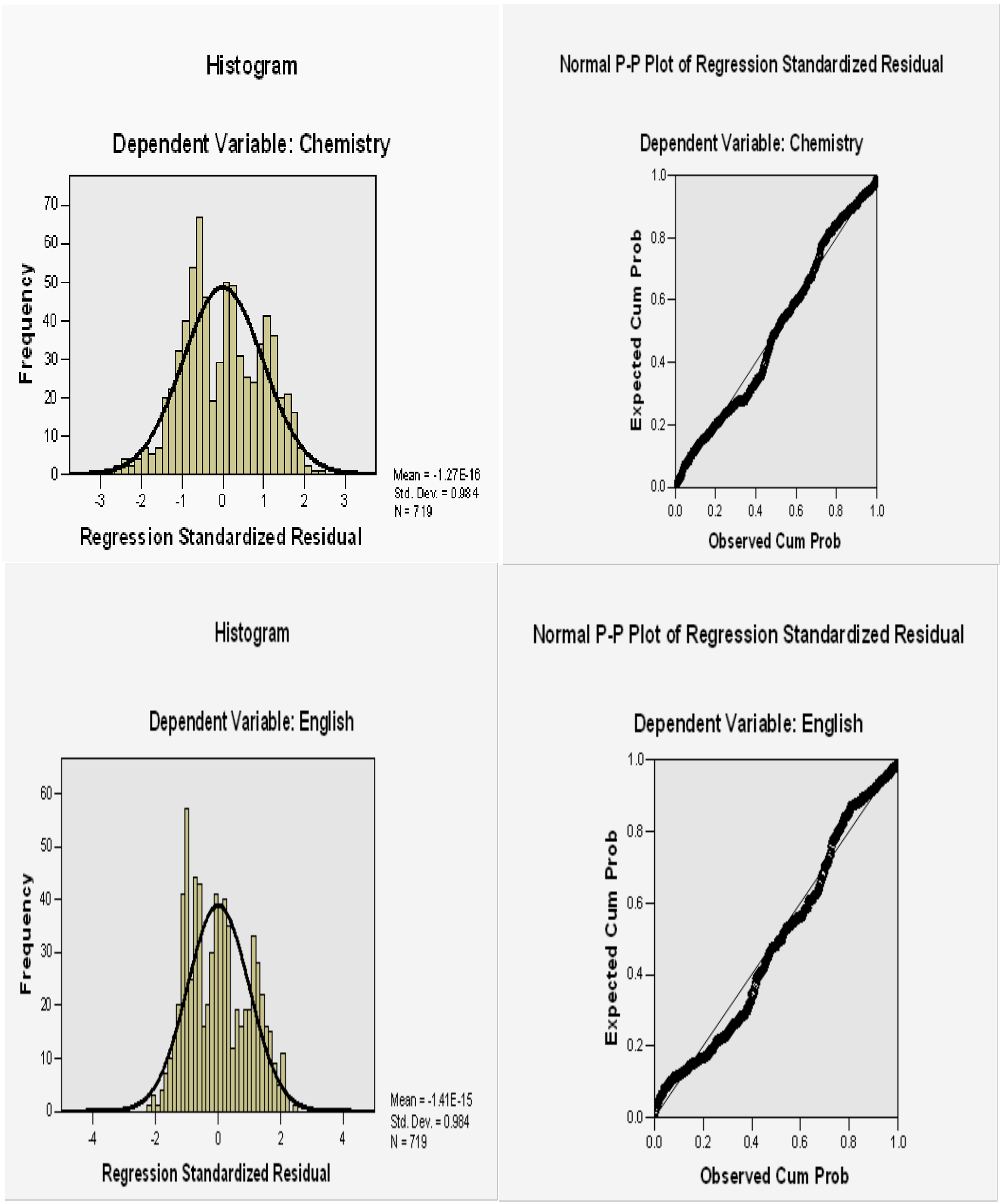


Figure 2 : Checking Model Adequacy of Multivariate (OLS) Multiple Linear Regression for Overall Sample Data