FDRE Ministry of Education

USAID | ETHIOPIA
FROM THE AMERICAN PEOPLE

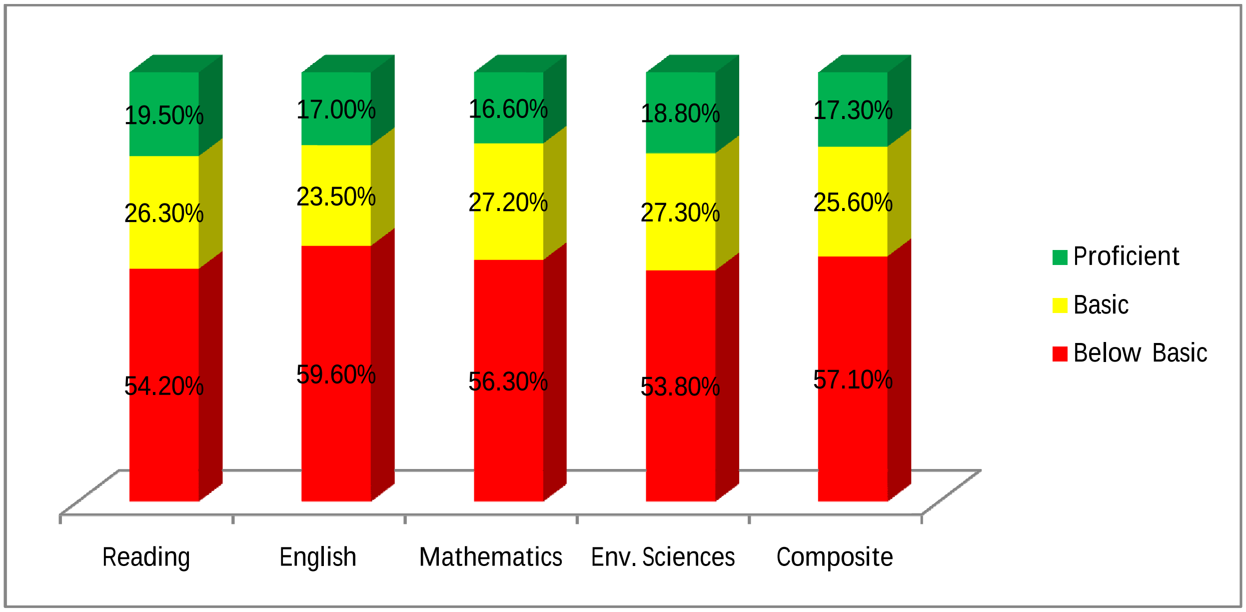
ETHIOPIAN 4th NATIONAL LEARNING ASSESSMENT OF GRADES 4 AND 8 PUPILS

Data Analytic Report

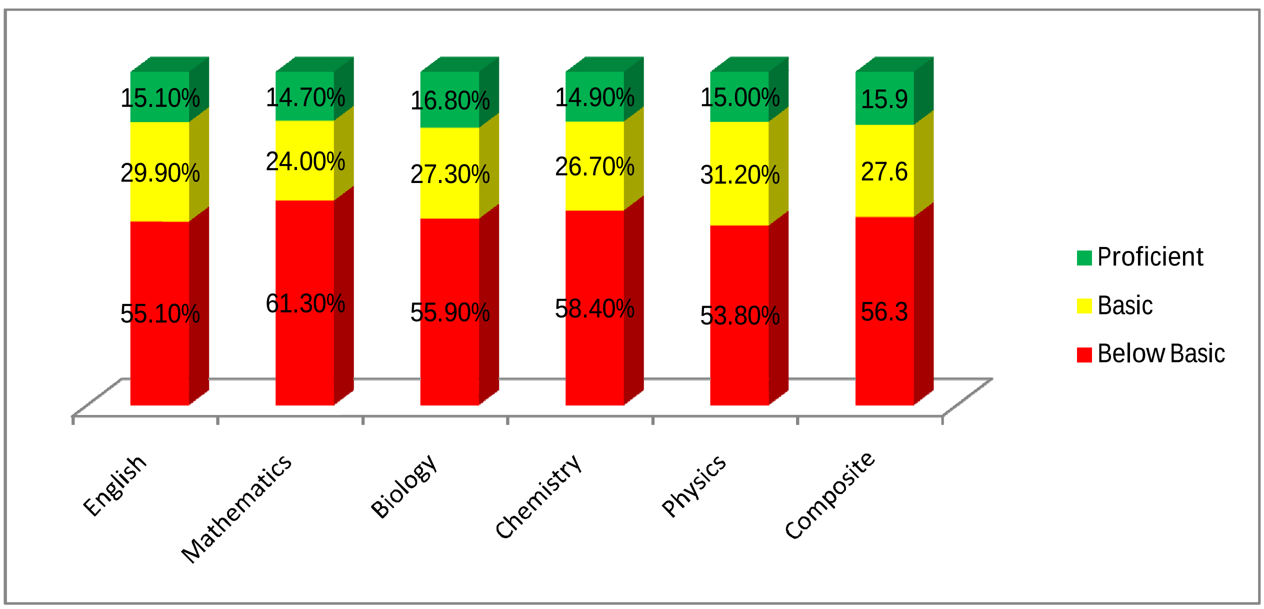
JANUARY 2013
ADDIS ABABA
ETHIOPIA

FOURTH NATIONAL LEARNING ASSESSMENT PROFICIENCY LEVELS

GRADE FOUR PUPILS



GRADE EIGHT PUPILS



The Ethiopian Fourth National Learning Assessment of Grades 4 and 8 Students is made possible by the generous support of the American People through the United States Agency for International Development (USAID).

ETHIOPIAN 4TH NATIONAL LEARNING ASSESSMENT OF GRADES 4 AND 8 PUPILS

A Study Carried out
By
Educational Assessment Directorate,
National Educational Assessment and Examinations Agency

Technical Working Group

Tamiru Zerihun, Directorate Director
Amina Menur, Assessment Expert
Arega Mamaru, Assessment Expert
Bekele Geleta, Assessment Expert
Effa Gurmu, Assessment Expert
Mengistu Admassu, Assessment Expert
Solomon Teferi, Assessment expert
Zewdu Gebrekidan, Assessment Expert
Fisseha Mikre, Consultant
Kindie Getachew, Consultant

National Educational Assessment
And Examinations Agency
January 2013
Addis Ababa, Ethiopia

Table of Contents

List of Tables Page

List of Tables Page

List of Tables Page

List of Tables Page

List of Figures Page

Acronyms

BNLA	Baseline National Learning Assessment
EGRA	Early Grade Reading Assessment
ESDP	Education Sector Development Program
FDRE	Federal Democratic Republic of Ethiopia
GEQIP	General Education Quality Improvement Program
GTP	Growth and Transformation Plan
IQPEP	Improving Quality in Primary Education Program
MLC	Minimum Learning Competencies
MLA	Monitoring Learning Achievement
MOE	Ministry of Education
NAEP	National Assessment of Educational Progress
NLA	National Learning Assessment
NOE	National Organization for Examinations
NEAEA	National Educational Assessment and Examinations Agency
OECD	Organization of Economic Co-operation for Development
PASEC	Programme d'Analyse des Systemes Educatifs de la CONFEMEN
PIRLS	Progress in International Reading Study
PISA	Programme for International Student Assessment
READ	Russia Education Aid for Development
REB	Regional Education Bureau
SACMEQ	Southern and Eastern African Consortium for Monitoring Education Quality
SIP	School Improvement Program
SNLA	Second National Learning Assessment
SNNPR	Southern Nations, Nationalities and People's Region
TIMSS	Trends in International Mathematics and Science Study
TNLA	Third National Learning Assessment
USAID	United States Agency for International Development
WEO	Woreda Education Office
ZEO	Zone Education Office

Acknowledgments

We would like to thank for the support provided by many individuals and institutions for the successful completion of the Ethiopian Fourth National Learning Assessment on Grades 4 and 8. Thanks are particularly due to USAID-IQPEP for the generous provision of the required financial and technical assistance, the test developers from the National Educational Assessment and Examinations Agency (NEAEA) and subject experts from the Curriculum Development and Implementation Directorate of MOE for reviewing the items.

We acknowledge with many thanks H.E. Ato Fuad Ibrahim, State Minister of Education, for the interest he showed in the study right from the very beginning and his continued support throughout the project. The contributions of the route coordinators of the data collection process who traveled to the regions from different departments of MoE and NEAEA, center coordinators and test administrators recruited by the regions are highly commendable. Our many thanks also go to the sample schools principals, teachers and students who enthusiastically participated in the study.

In the past few years, professionals from NEAEA have participated in a number of global conferences and workshops on National Learning Assessment and related topics funded by the Russia Education Aid for Development (READ) Trust Fund and organized by the World Bank Ethiopia Country Office. These events enabled the experts to carry out the assessment with more competence than before. Hence these institutions deserve special thanks for covering the relevant expenses and supporting the participation of these experts in those programs.

NEAEA is highly indebted to the World Bank Ethiopia Country Office and Dr. Marguerite Clarke, READ TF Program Manager, for the continued support and facilitation of the participation of NEAEA experts in global and regional conferences and workshops.

CHAPTER ONE: BACKGROUND TO THE STUDY

1.5. INTRODUCTION

Education is expanding at all levels in Ethiopia. At present, the country has embarked on providing Education for All up to 2015 based on the Jomtien 1990 agreement and the Dakar 2000 framework for action (NLA, 2004). Besides striving to realize this target, there is a great concern for the provision of quality education at all levels of the system. On the other hand, stakeholders in education would like to know what substantive knowledge and skills students gain at the completion of each sub cycle and level in the system. According to the Education and Training Policy Document (MOE, 1994), the Ethiopian education system has three levels: namely primary (that includes first cycle from grades 1-4 and second cycle from grades 5-8); secondary (including general secondary education from grades 9-10 , college preparatory, grades 11–12, and TVET); and tertiary (Colleges and Universities). Therefore, it is imperative to inquire what students can know and do while leaving each sub cycle and level of the education system. As a result, the National Agency for Examinations (NAE) has a vested upon authority to conduct learning assessments of grades 4, 8, 10, and 12.

Learning assessment can play salient roles in the provision of quality education. It could have practical implications for policy makers and practitioners in education as they help in tracing the attainment of curriculum goals, give clue to diagnose learning difficulties, motivating learners, and in improving the education system. Moreover, the sector can use educational assessments to evaluate the effectiveness of teachers, curricula, and educational systems or programs, to identify trends in educational achievement, to determine the comparative standings of school districts, states, or nations with respect to educational progress, and to aid curriculum planning and policy definition.

The Education and Training Policy of Ethiopia promulgated in 1994, states that as curriculum component, assessment should receive greater attention in order to lever the attainment of curriculum goals at all levels of education. Moreover, in the Growth and Transformation Plan (GTP) of the country and the third and forth Education Sector Development Programs (ESDP–III & IV), the ministry of education has included achievement targets among the methods of quality indicators in the system. In addition, the General Education Quality Improvement Program (GEQIP-I & II) documents has given particular importance to learning assessments as they contribute in the improvement of education quality.

Currently, learning assessment programs are becoming popular in the education systems of many countries. Some countries have come to understand the importance of measuring learning achievements for the purpose of identifying proficiency levels of students and impediments to learning in order to improve education quality. Educators recognize learning assessments as means to identify strengths and weaknesses of education systems. As a result, educators execute learning assessment either at national and/or international level. National learning assessments evaluate the learning attainment of students throughout a given country and compare the achievement among regions, locations, gender, socio-economic status, and other attributing variables to achievement. However, international learning assessments compare learning achievements across countries with selected correlating variables to achievement.

Several countries (OECD member states) have given emphasis to quantitative outputs of their education system in monitoring progress towards achieving educational goals (OECD, 2010). For instance, they use the number of students enrolled, rate of completion at a given level, the adequacy of inputs (textbooks, teachers, teacher – student ratios etc), teacher training, classroom interaction, and pedagogy as indices for the attainment of curriculum goals. Nevertheless, these become very useful indicators when supported by effective learning assessment programs.

Learning assessments conducted at national as well as international levels focus in measuring students' achievement in key curriculum areas such as reading literacy, mathematics, sciences, and social studies. Thus, the achievement of students at various proficiency scales may show the strengths and weaknesses of the learning process. Moreover, the result can show the extent to which teachers and students experienced the intended curriculum goals. Furthermore, the learning assessment practice helps to identify a host of correlating variables such as school characteristics, socio-economic backgrounds of learners, teaching-learning process, teacher preparation and quality, learning resources and the like, attributing to the variations in students achievement. Therefore, learning assessment reports become handmaiden for those intending the curriculum goals, implementing it, and experiencing it to find out areas of the curriculum process, which needs improvement.

On the other hand, the learning assessment report will help policy makers and higher officials in the sector to evaluate whether or not the education system is healthy and functions properly as per the objectives stated in the curriculum material of each key subject, supervise the education system, and plan for pre-service and in-service training of teachers along with fulfilling education resources. Moreover, if the report ranks the schools in terms of students' learning achievement, it will help parents to choose best performing schools for their children. This, in turn, creates a competitive working environment for schools to improve the implementation of the curriculum.

The Ministry of Education has recently launched the General Education Quality Improvement Program (GEQIP) in order to support the provision of quality education that prepares citizens for the competitive global market economy. The GEQIP has a comprehensive national learning assessment subcomponent for its effectiveness and efficiency. This will contribute in the implementation of the transformation agenda in the education sector. In line with this, since 2000 Ethiopia has conducted three learning assessments at the national level on grades 4 and 8. Subsequently, this study intends to report results of the fourth national learning assessment conducted in the 2010/11 academic year. The main purpose of the study is to report the attainment of curriculum goals with respect to proficiency levels of students, identify the major correlates of achievement variations, and contribute to the GEQIP and education sector reform by suggesting actions for intervention and improvement. Thus, the next subsection describes the organization of this study.

1.6. Organization of the Study Report

The study report contains three common sections and two separate ones for each grade level. Out of the three common sections, the first section discusses the introduction to learning assessments, the research questions, objectives, and the significance of conducting national learning assessments for the successful implementation of the country's third and fourth Education Sector Development

Programs. The second section focuses on the background of learning assessments and literature review to compare empirical evidences at the national and international level. The third section explains the research methodology followed to gather, analyze, and interpret data. The separate fourth sections present findings of the study in terms of achievement outcomes, attained proficiency levels, and the independent variables explaining variations in the learning achievement of students. Finally, the other separate fifth sections discuss the overall implications of the study results for appropriate measures of intervention and improvement in student learning achievements.

1.7. Research Objectives

This study intends to attain the following research objectives:

- Analyze the learning achievement of students
- Compare pupils achievement with the minimum achievement target (50%) as indicated in the Education and Training Policy
- Describe the proficiency levels of students
- Compare students' learning achievement by gender, location, region, and proficiency levels.
- Explain the variables that significantly correlate with students' learning achievement
- Discuss the implications of the findings of the fourth national learning assessment result for the improvement of learning achievement and education quality in Ethiopia.

1.8. Research Questions

This study aims to find out answers to the following research questions:

- To what extent did students achieve the curriculum goals in the key subjects (English, mathematics, biology, chemistry and physics for grade 8, and Mother tongue, English, mathematics and environmental science for grade 4)?
- How close is the learning achievement of students to the minimum achievement target (i.e. 50%), set at national level?

- To what extent do students in the different proficiency levels vary in score points?
- How do students' learning achievements differ by gender, location, region, and proficiency levels?
- What are the variables that significantly correlate with students' learning achievement?
- What are the implications of the fourth national learning assessment result to improve student learning achievement and school quality in Ethiopia?

1.6. Significance of the Study

This study is significant for providing information for policy makers and practitioners in the field of education. There are expectations from policy makers to contribute in the enabling of citizens' benefit from the globalized world economy. As a result, they concern on the improvement of curriculum goals attainment, ensure the quality of education provision, and learning opportunities and attainments of learners. In order to accomplish this, they have to rely on reliable information on how well education systems prepare students for life. Among the possibilities are conducting national learning assessments and taking part in international learning assessments.

Education is expanding at all levels in the country. There is a great endeavor to improve the education access for all citizens and make the provision equitable and efficient. Apart from this, the country is attempting to improve the quality of education provision. Thus, this fourth national learning assessment contributes in receiving feedback on the extent students achieve the curriculum goals in key subjects as compared to the national curriculum target or the minimum competency learning, which is 50% achievement in each subject as stipulated by the ministry of education.

On the other hand, national learning assessments are desirable in that educators use them to identify school factor, student factor, and teacher factor correlates of students' scholastic achievement. Therefore, this study report can provide information on the most influential correlates of students learning achievement for proper control and management.

The country expends a considerable proportion of the government budget for the provision of education to citizens. Therefore, national learning assessments are among the methods to check whether the expenditure has a return to the country's ultimate socio-economic development in terms of learners' attainment of useful knowledge and skills. Because of this, the fourth national learning assessment significantly contributes to the attainment of education sector goals.

CHAPTER TWO: REVIEW OF RELATED LITERATURE

Assessment of learning is a useful tool of information to ensure the attainment of curriculum goals in education systems. Educational assessments are useful procedures to measure the degree of curriculum goals attainment. In recent years, countries throughout the world have reached to consensus on the importance of measuring educational performance of children (Wolff, 1998). As Chinapah (2003) explains, measuring students' learning outcomes is an integral part of the educational process, and it is crucial for monitoring the implementation of educational programs and the evaluation of their impact. According to Kellaghan (2004), there are three major procedures, which can provide information on student learning. These are public (external) examinations, international assessments, and national assessments. He further stated six questions that education policy makers and educators should question concerning information about students' learning. These are:

- How well are students learning in the education system (e.g. with reference to curriculum goals, EFA goals, and in preparation for life)?
- Is there evidence of particular strengths or weaknesses in the knowledge and skills students have acquired?
- Do the achievements vary between genders, urban – rural locations?
- To what extent is achievement associated with the characteristics of the learning environment (e.g. school resources, teacher preparation, & competence, type of school, student characteristics, and so on)?
- Do the achievements of students vary across time?
- How do the achievements of students compare with the achievements of other countries? (PP, 3-4).

Thus, this section of the study briefly discusses about African experiences of international and national learning assessments including the Ethiopia's experience from the base-line, the second, and the third national learning assessment endeavors.

2.8 International experiences of learning assessments: empirical evidences from MLA studies in Africa

International learning assessments compare the learning achievement of students across countries. They are mainly conducted to evaluate the quality, equity, and efficiency of school systems. One of the commonest international learning assessments is the program for Monitoring Learning Achievement (MLA) project, which is typical for many African countries.

As discussed by Kellaghan (2004), there are four major categories of learning assessment in Africa, where three of which involve similar activities in several countries. These are: The Monitoring Learning Achievement (MLA) project, the Southern Africa Consortium for Monitoring Education Quality (SACMEQ) project, and the PASEC (Programme d'Analyse des Systems Educatifs des Pays de la COFEMEN), and the fourth is a national learning assessment conducted by individual countries. In this subsection, the study discusses MLA project as an international experience. The Monitoring Learning Achievement (MLA) project was conducted in 1999 by sampling 50,000 grade 4 pupils in response to the World Declaration of Education For All in 1990 in Jomtien to evaluate the extent to which students actually acquire useful knowledge, reasoning ability, skills and values. MLA – I was carried out for grade 4 pupils to assess reading and writing literacy, numeracy, and life skills that is relating to awareness and knowledge of health, nutrition, sanitation, and hygiene. The achievement report has shown that only four countries had met their Jomtien learning target (i.e. 80 percent of students should attain the intended learning competencies) (Kellaghan, 2004). Gender differences in achievement were small in all countries. Parents' assistance to their children's schoolwork was positively correlated to the students' achievement in many countries. Table 1 below shows grade 4 pupils' performance in literacy, numeracy and life skills from nine African countries as well as the percentage of students attaining minimum mastery learning (MML) and desirable mastery learning(DML) from each of the participating African country.

Table 1: Countries average Achievement in the Key Skills of Literacy, Numeracy, Life Skills, and the Combined

Country	Literacy			Numeracy			Life Skills			Combined	
	Average %	MML %	DML %	Average %	MML %	DML %	Average %	MML %	DML %	MML %	DML %
Mauritius	66.20	77.60	35.40	58.43	70.30	26.40	63.70	71.60	32.40	70.30	24.10
Uganda	62.00	64.30	23.30	49.63	41.90	10.20	66.23	78.80	51.10	54.40	14.40
Mali	57.70	50.40	13.10	44.10	37.90	6.20	56.63	69.80	23.70	54.40	7.30
Madagascar	57.25	56.90	20.60	41.33	34.40	5.60	76.00	97.30	60.30	66.10	11.70
Botswana	49.75	46.20	6.00	48.67	55.40	5.40	57.33	71.80	14.90	57.80	8.70
Malawi	49.67	15.30	1.40	44.00	30.70	1.40	76.00	95.40	69.40	54.90	3.00
Zambia	48.50	37.80	7.30	36.00	19.90	4.40	50.67	49.00	26.10	31.90	5.60
Senegal	52.10	45.60	6.70	35.57	22.90	3.00	45.53	36.30	7.00	31.20	2.00
Niger	51.57	39.30	3.60	41.43	15.30	5.70	47.80	44.90	7.00	25.60	2.00

As shown in Table 1 above, 9 African countries were compared with respect to their pupils' achievement in 3 key subjects, namely literacy (vocabulary, comprehension, grammar and writing); numeracy (use of numbers, measurement and geometry); and life skills (health, civic & environment, and science and technology). The result shows that grade four pupils from Mauritius (66.2 percent mean score), Uganda (62 percent mean score), Mali (57.70 percent mean score), and Madagascar (57.25 percent mean score) performed relatively better in literacy skills. Zambia, Malawi, and Botswana are relatively the least performing countries in literacy skills.

With respect to numeracy skills, which involve the use of numbers, measurement and geometry, grade 4 pupils' from Mauritius (58.43%), Uganda (49.62%), and Botswana (48.67) performed relatively better. Comparatively pupils from Senegal, Zambia, and Madagascar have shown relatively poor performances in numeracy skills.

On the other hand, pupils from the majority of the comparison countries have performed above 50% mean score in life skills sub test, which includes health skills, civic and environment, and science and technology. For instance, if we put the countries in descending order of rank, Pupils from Malawi (76%), Madagascar (76%), Mauritius (63.70%), Botswana (57.33%), Mali (56.63%), and Zambia (50.67%) have performed above a mean score of 50% in life skills. Only two countries, Senegal (45.53%) and Niger (47.80%) have performed below a mean score of 50%.

In summary, the MLA – I data depicts that grade 4 pupils of Zambia and Senegal performed poorly in two different skill types (e.g. Zambia in literacy and numeracy skills, and Senegal in numeracy and life skills). Moreover, pupils of Botswana, Malawi, Madagascar, and Niger have poor performances at least in one of the three skills (e.g. Botswana and Malawi in literacy skills, Madagascar in numeracy, and Niger in life skills).

In addition to the average scores of pupils for each country, Table 1 above shows the percentage of pupils from each country that performed at the Minimum Mastery Learning (MML) level and Desirable Mastery Learning (DML) level. The MLA – I reports pupils' achievement in three levels as below minimum mastery learning, minimum mastery learning (MML), and desirable mastery learning (DML).

Among the nine countries compared, 24.1 percent of grade 4 pupils from Mauritius have performed in the level of desirable mastery learning for the combined score of literacy, numeracy, and life skills. 14.4 percent of pupils from Uganda have also performed at this level for the combined score. Madagascar's pupils have also performed relatively better (11.7%) at the desirable mastery learning. Mauritius, Madagascar, and Botswana ranked from first to third for the combined score with respect to the minimum mastery learning (MML). If we keep on ranking the countries at least for the MML of the sub tests, Mauritius still stood first in literacy and numeracy and fifth in life skills. Uganda ranked second in literacy and third in both numeracy and life skills. Madagascar is third in literacy, fifth in numeracy, but first in life skills. While Malawi pupils performed better in life skills (second rank), they performed last in literacy skills.

In conclusion, best performing grade four pupils in MLA – I assessment are from Mauritius, where above 70 percent of them performed at MML level and above. Madagascar students are second best in which 66 percent of them have performed at MML level and above. The third best is Botswana, where 57 percent of pupils performed at MML level and above. In comparison, pupils from Niger, Senegal, and Zambia had poor performances in the MLA – I assessment (MLA, 1999).

2.9 International Experience of Learning Assessments: Empirical Evidences From PISA 2009 and MLA 2000 Studies

International learning assessments compare the learning achievement of students across countries. Organizations conduct them to evaluate the quality, equity, and efficiency of school systems in many countries. One of the commonest international learning assessments is the Program for International Students Assessment (PISA). The Organization of Economic Co-operation for Development has executed PISA (OECD) every three years since 2000. As described in the OECD (2010:8) document, PISA seeks to answer questions such as: Are students well prepared to meet the challenges of the future? Can they analyze reason and communicate their ideas effectively? Have they found the kinds of interests they can pursue throughout their lives as productive members of the economy and society?

By 2009, PISA has completed its fourth round survey by focusing on the reading skill of 15 year olds across 70 countries. Additionally, it assessed mathematics and science reasoning skills. Several countries who have participated in PISA studies were inspired to organize national efforts to help students learn better, teachers teach better, and school systems to become more effective. PISA 2009 describes the areas of assessment within a framework that includes knowledge and/or competencies in key subjects that students need to apply; contexts in which students encounter problems; and students' attitudes and dispositions towards learning (OECD, 2010: 22).

In PISA 2009 result, the GDP of a country explains 6 percent of variation in students' learning achievement. Spending in education, parents' level of education, and economic disadvantage of students (with negative correlation) explain 9 percent, 45 percent, and 46 percent of the students' reading achievement variation respectively.

2.9.1 PISA 2009 Report of Students' Performance in Reading

PISA's 2009 reports revealed that reading performance of students in six different proficiency levels with differentiation among the various kinds of student tasks and capabilities. Moreover, it classifies the proficiency levels based on a scaled score of mean 500 and standard deviation of 1.00. For instance, students who are proficient at level six can have scaled scores greater than 698 points in reading (1.98 standard deviations above the mean score). Students at this level are highly skilled readers. They are capable of analyzing reading texts with detailed comprehension and unstated implications. Only 0.8 percent of 15-year-old students perform at this level. Seven countries such as New Zealand, Australia, Japan, Canada, Finland, Singapore, and Shanghai - China have a significantly higher percentage of students performing reading at proficiency level 6. In these countries, the majority of students achieved well in reading, with less than 5 percent of the students achieving below proficiency level 1a (scores between 335 and 407). The level requires students locate pieces of explicitly stated information that are prominent in the text, recognize the theme in a text about a familiar topic, and recognize the connection between information in a text and their every day experience (OECD, 2010).

On the other hand, proficiency level 5 (scores 626 – 698) that requires students to handle and find information from unfamiliar texts was attained by 8 percent of the students. Shanghai – China has a significant share (19% of the students) for this level. Finland, Japan, Korea, Australia, Canada, Singapore, and Hong Kong – China had above 12 percent of students performing at proficiency level 5 or above. About 28 percent of students were proficient at level 4 or higher in reading skills. Level 4 has score ranges between 553 and 626 score points. Test tasks at this level require students to read difficult text and locate embedded information in the text, critically evaluate a text, and interpret the meaning of nuances of language in a section of text.

The majority (57%) of 15 year old students are proficient at level three (score ranges from 480 – 553). This level requires students capable of reading tasks of moderate complexity. Over three – fourth of the students in Shanghai – China, Korea, Hong Kong – China and Finland have successfully done the reading tasks at proficiency level 3. Nonetheless, fewer than 50 percent of the students in 30 countries including Luxembourg, Czech Republic, Austria, Turkey, Chile, and Mexico have performed at this level of proficiency.

According to OECD (2010) result, proficiency level 2 is a base-line (score ranges between 407 and 480 points). This level of reading requires students to make comparison and contrast of ideas in a text, identify the theme of a text even when the information is not prominent, and make connections between the text and personal experiences. Longitudinal studies have shown that students scoring below this level became at high risk of poor post – secondary school participation. Level 2 is the most common highest level of proficiency for many students. About 4 in 5 students (81%) are proficient at level 2 or higher. Mexico (33%), Chile (33%), and Turkey (32%) had the highest proportion of students for proficiency level 2.

Proficiency level 1a (335 – 407 score points) and proficiency level 1b (262 – 335 score points) were below the base-line. About 6 percent of students were in proficiency level 1a or below. In countries such as Indonesia, Azerbaijan, Kazakhstan, Panama, Peru, Brazil, Albania, and Qatar, level 1a is the commonest proficiency level for their students (OECD, 2010). This data has particularly relevant for the above-mentioned countries to critically evaluate their education system in reference to the development of reading competence for students. Furthermore, more than one in hundred students is below the last proficiency, which is level 1b, with score points below 262 on the PISA scale. In countries such as Qatar, Panama, Peru, and Azerbaijan, more than 34 percent of the students had performed at or below level 1b. As a result, these countries were expected to investigate their curriculum and the implementation in order to compete with other countries.

2.9.2 Gender Difference in Reading Achievement

Analysis of the reading achievement data from PISA 2009 indicated that on the average girls outperformed boys in all participating countries by an average of 39 score points. The gender gap in achievement was much wider in some countries, for example 55 score points difference in Finland. When we compare girls and boys with respect to the different proficiency levels, 1.2 percent of girls and 0.5 percent of boys were in the highest proficiency level 6. At proficiency level 5, we can find 8.8 percent of girls and only 4.8 percent of boys. To the contrary, in the lowest two proficiency levels 1a and 1b, and below the last (below 1b), we find significant proportion of boys than girls.

For instance, at proficiency level 1a, we find 16.6 percent of boys and only 9.5 percent of girls. Moreover, at proficiency level 1b, we find 6.6 boys and only 2.6 percent of girls. Furthermore, below the last proficiency level 1b, we find 1.8 percent of boys and only 0.5 percent of girls. In general, 25 percent of boys and 12.6 percent of girls are below the base-line proficiency level, which is level 2. In this case, the proportion of boys below the base-line proficiency doubles that of the girls.

2.9.3 PISA 2009 Report of Student Performance in Mathematics

To what extent do students reason mathematically? To what extent do they utilize mathematical concepts, procedures, facts, and tools to describe, explain and predict phenomena? PISA studies were concerned in finding answers to questions mentioned above. The proficiency levels for mathematics achievement are similar to that of the reading. Across all participating countries, 3.1 percent of students had performed at the highest level 6(with scores higher than 669 points). Students in Shanghai – China were the top performers at this level, which was more than 25 percent of the students. To mention example performances in mathematics at this level, Singapore = 15.6 percent, Chinese – Taipei = 11.3 percent, Hong Kong – China = 10.8 percent, Korea and Switzerland = 8 percent, Japan, Belgium, and New Zealand =5 percent in each country (OECD, 2010). On the other hand, in Mexico, Chile, Greece, and Ireland less than 1 percent of the students reached at proficiency level 6. Alarmingly, the percentage of students at this level was close to zero for countries such as Indonesia, Colombia, Jordan, Albania, Tunisia, and Panama.

When students can develop and work in complex situations using their mathematical skills, they are assigned to proficiency level 5 (scores higher than 607 and lower or equal to 669). An average of 12.7 percent of students reached to this level or higher across all countries. Korea had the highest percentage of students (25.6 %) at this level. Chile and Mexico had less than 5 percent of their students at this level. Level 4 was attained by 31.6 percent of students across all countries. Students at this level of attainment can work with explicit models for complex and concrete situations that may involve constraint or call for making assumptions. The majority of students from Korea, Shanghai – China, Singapore, Hong Kong – Chine, and Chinese – Taipei performed at this level. Around 40 percent of the students from Finland, Switzerland, Japan, the Netherlands, Canada, Belgium, New Zealand, Liechtenstein, and Macao – China also performed at this level of proficiency.

In the extreme case, an average of 14 percent of students performed at level 1 (358 – 420 score points). This level of proficiency requires students answer questions involving familiar contexts where all relevant information is present and clearly defined. Nonetheless, about 8 percent of students across all countries performed below level 1. These students might have difficulties using mathematical concepts to benefit from further education and learning opportunities in the future.

2.9.4 Gender Difference in Mathematics Achievement

In contrast to the reading achievement, boys had performed better than girls with a difference of 12 score points in mathematics. Belgium, Chile, the United Kingdom, and the United States had shown the largest gender differences, with an advantage of 20 score points for boys. This had its own implication for countries to work towards improving girls' performance in mathematics.

2.9.5 PISA 2009 Report of Students' Performance in Science

School systems should prepare students with understanding of science and technology for life in modern society. PISA defines scientific literacy as an individual's scientific knowledge and use of the knowledge to identify questions, acquire new knowledge, and explain scientific phenomena (OECD, 2010. As in the reading and mathematics scores, PISA reports science scores in six proficiency levels. At the highest level, which is level 6 (scores higher than 708 points), students can identify, explain and apply scientific knowledge in a variety of complex situations. On the average, 1.1 percent of students performed at this level, where 3.6 percent in New Zealand, 3.3 percent in Finland, 3.1 percent in Australia, 2.6 percent in Japan, 4.6 percent in Singapore, 3.9 percent in Shanghai – China performed at proficiency level 6. In many countries, including Mexico, Chile, and Turkey, no one student reached this level.

From all countries 5 percent of students performed below level 1, (less than 335 score points). Level 1 requires students to apply their scientific knowledge to a few, familiar situations. Students scoring below level 1 have difficulties in using science to benefit from learning opportunities and further education. About 47.7 percent of students from Mexico scored at or below proficiency level 1 in science. This shows that Mexico has to look ways of improvement in the teaching of science for students.

2.9.6 Gender Difference in Science Achievement

There was a very small gender difference in science achievement across countries. In some countries such as the United States, Denmark, Colombia, and Liechtenstein boys outperformed with an average difference of 12 to 21 points. On the other hand, girls' outperformed boys with 10 to 15 score points difference in Finland, Slovenia, Turkey, and Greece.

2.10 African Experiences in International Learning Assessments

As discussed previously, African states have experienced international learning assessments very recently and have been involved in four major categories of learning assessments where three of which involve similar activities in several countries. This subsection discusses MLA project as an international experience specifically in relation to grade 8 students' achievements.

MLA – II assessed the achievement of grade 8 pupils in 2000 compared students of six selected African countries in mathematics, physics, chemistry, and other sciences. Table 2 below shows the performance of pupils in the six countries on the mentioned key subjects.

Table 2: Performance of Students in Key Subjects

Country	Mathematics%	Physics%	Chemistry%	Other sciences%
Burkina Faso	31.91	40.15	41.92	57.00
Cameroon	29.20	39.06	44.79	54.85
Mali	24.52	34.13	32.54	38.17
Mauritania	34.19	27.35	32.69	45.69
Niger	27.34	33.13	35.34	45.81
Senegal	36.29	36.76	39.15	50.56

As shown in the Table above, the mean percentage scores are very low for many of the countries, which is below a mean score of 50%. In fact, pupils from Senegal, Cameroon, and Burkina Faso are exceptions for the other sciences since they have average scores of 50.56%, 54.85%, and 57% respectively. Countries with relatively highest averages in mathematics were Senegal and Mauritania with average scores of 36.29% and 34.19%, and countries with the lowest averages were Mali and Niger with average scores of 24.52% and 27.34%.

In physics, the relatively highest averages go to Burkina Faso (40.15%), Cameroon (39.06%), and Senegal (36.76%). The lowest averages in physics were for pupils from Mauritania (27.35%), Niger (33.13%), and Mali (34.13%). In chemistry, relatively highest average scores were for Cameroon (44.79%), Burkina Faso (41.92%), and Senegal (39.15%). On the other hand, lowest averages in chemistry were for pupils from Niger (32.34%), Mali (32.54%), and Mauritania (32.69%). Finally, in the other sciences, relatively better performing students were from Burkina Faso (57%), Cameroon (54.85%), and Senegal (50.56%). The relatively least performing pupils in other sciences were from Mali with a mean score of 38.17%.

In summary, grade 8 pupils from the mentioned countries had relatively poor average scores in mathematics, physics, chemistry, and other sciences. Particularly average scores in the key subjects were relatively low for Mali, Niger, and Mauritania.

On the other hand, the achievement data for the two genders showed that boys outperformed girls in mathematics for all countries except Burkina Faso and Niger. In physics, still boys outperformed girls in all the six countries compared. In chemistry, boy students in Cameroon, Mali, and Mauritania had significantly outperformed girl students. In the other sciences, except in Mauritania and Niger, in the remaining comparison countries boy students significantly outperformed girl students.

Furthermore, there were urban – rural differences in learner performance in almost all African countries compared. Results from the MLA surveys clearly showed that urban schoolchildren scored better than rural schoolchildren overall and in learning areas (Chinapah, 2003).

2.10.1 National Learning Assessments: Empirical Data from TIMMS and African Experience

National learning assessments are procedures designed to describe the level of achievement of the education system in a particular country. Many countries around the world do have national learning assessments. They attempt to measure their students' achievement in key subjects while leaving a certain level of the education structure (e.g. grades 4 or 8, or 10). National assessments commonly used to check the attainment of curriculum goals by the whole education system. In many of the national assessment programs, data gathered to determine students' learning achievement and the correlates or impacts of in school factors such as teacher preparation, class size, and textbook availability and so on, as well as non-school factors such as parent education, socioeconomic status, rural-urban location, distance to school, home language and so on. Data for national learning assessments either administered by a particular country as in the National Assessment of Educational Progress (NAEP) example in the US America or extracted from international assessment programs. This section briefly discusses national learning assessment experiences of some countries that use the Trends in International Measurement of Mathematics and Sciences (TIMMS) and Monitoring of Learning Achievement (MLA).

2.11 Empirical Data from the US America TIMMS 2007 Study

Results from national assessment programs are important resource for education policy makers and practitioners in the field. In 2007, America has taken part in the TIMMS study to compare the mathematics and science achievement of pupils with other countries pupils. Gonzales et.al (2009: 3-4), reported the key findings as summarized below:

- The average mathematics scores of U.S grade 8 pupils (508) were higher than the TIMMS scaled average (i.e. 500 score points).
- The average mathematics score for grade 8 students of US America was higher than those of students in 37 of the 47 participating countries.
- The average mathematics score for grade 8 in 2007 was 508, which is 16 points higher than the average of 492 in 1995.

- The average science score of grade 8 pupils (520) was higher than the TIMMS scaled average (i.e. 500 score points).
- The average U.S fourth grade science score was higher than those of students from 25 of the 35 other countries.
- At eighth grade, the average U.S science score was higher than the average scores of students in 35 of the 47 countries.
- The average science scores for grade 8 were not measurably different between 1997 and 2007. It was 513 in 1997 and 520 in 2007.

2.12 National Learning Assessment in Ethiopia

In Ethiopia, since 2000, the then National Organization for examinations (NOE) has conducted three National Learning Assessments to appraise both grade 4 and 8 pupils' academic performances in core subjects (Mathematics, English, Biology, Chemistry, and Physics for grade 8, and Mathematics, English, Mother Tongue Reading and Environmental Science for grade 4). The National Learning Assessments compared pupils' performance across regions, sex and location, and explored factors that can have significant influence on academic achievements of learners.

The first National Learning Assessment was taken place in the year 2000 for both grade 4 and 8 students to serve as a baseline assessment. The second National Learning Assessment was conducted in 2004 by sampling 8,059 pupils from 213 schools of 10 regions and 1 city administration. The third National Learning Assessment that was conducted in 2007 sampled 10,806 pupils from 280 schools. Currently, the fourth National Learning Assessment has been carried out in 2010/2011 academic year. 10,335 pupils were participated in the assessment process. They took tests on the core subjects which the NOE prepared and have filled out a questionnaire that consisted of items on general information, background variables, and schools and out of school variables that might relate to their learning and academic achievement. Chapter 4 of this report discusses the findings of the fourth National Learning Assessment. Following are brief discussions on the major findings of the two National Learning Assessments. The discussion aims to give a highlight for the reader to see the trends on students' test performances on the Second and Third National Learning Assessments.

2.12.1 Trends in Students' Academic Performances in Ethiopian NLAs

2.12.1.1 Grade 4 Students Academic Performance from 2000-2008

In Ethiopia, beginning from 2000 the NOE has carried out three National Learning Assessments to appraise grade 4 pupils' academic performances in Reading, English, Mathematics, and Environmental Sciences. The National Learning Assessments in grade 4 compared pupils' performance across regions, sex and location, and explored factors that influence academic achievements of learners.

The NOE conducted the first National Learning Assessment in the year 2000 by taking a sample of about 10,357 grade 4 pupils from 305 schools, where its intent was to serve as a baseline assessment. The second National Learning Assessment was conducted in 2004 by sampling 13,346 pupils from about 407 schools of 9 regions and two city administrations. The third in 2008 on 11,493 pupils sampled from 305 schools of the 10 regional states and one city administration. Currently, the NOE has conducted the Fourth National Learning Assessment by taking a sample of 10,357 Pupils from 305 schools. They have taken tests on Reading, English, Mathematics and Environmental Science, which were prepared by the National Agency for Examinations. Moreover, the pupils filled out questionnaires consisting items on general information, background variables, and school and out of school variables that may explain variations in the pupils' test performance. Hence, this research paper discusses the findings of the fourth National Learning Assessment in chapter 4.

Following is a brief discussion on the major findings of the two previous National Learning Assessments of grade 4 pupils. The discussion intends to give a highlight so that the readers of this report capture the trends in pupils' academic performance from the second, third and fourth National Learning Assessments.

2.12.1.2 Pupils Test Performance in 2004 and 2008

When comparing the SNLA (2004) with TNLA (2008), there was a marked decline of performance in the key subjects. Declines in test performance are common for reading with a mean difference of 20.59% though the two reading tests were not equivalent, English, and Environmental Sciences with mean differences of 2.18 % and 9.14%

respectively. Consequently, the overall test performance has fallen in 2008 as compared to the 2004 with a composite mean score decline of 7.58%. Nevertheless, as an exceptional case, test performance of grade 4 pupils in Mathematics has shown a slight improvement with a percentage increase of 0.6%.

2.12.1.3 Pupils' Test Performance by Gender in 2004 and 2008

As the SNLA data reveals, boys (with a composite mean score of 50.14%) outperformed girls (with a composite mean score of 46.12%) in all key subjects assessed. In the same manner, the TNLA result showed the same trend. For instance, in the TNLA, boys achieved a mean score that is greater than the girls by 1.12% for the composite score are. Nevertheless, the gap was smaller as compared to the 2004 result that was 4.02%.

Data from the third national learning assessment depicts statistically significant differential performance between boys and girls only for English (a mean score difference of 2.36%) and environmental science (a mean score difference of 1.58%). However, there were no statistically significant performance differences between boys and girls for Mathematics and reading. In general, the gender gap in test performance became narrower in 2008 than in 2004. A case in point is that the gaps in achievement for 2004 were all statistically significant between boys and girls and range from the lowest 2.7% in reading to 5.02% in mathematics. The range of achievement difference between boys and girls in 2008 was from 0.23% in Mathematics to 1.58% in environmental science.

In summary, the gaps in academic achievement between the two sexes have narrowed significantly in almost all subjects for the 2008 national learning assessment as compared to that of the 2004.

2.12.1.4 Pupils Test Performance by Location in 2004 and 2008

In the 2004 national learning assessment, the overall test achievement was in favor of the urban pupils by a mean score point of 1.19%. Pupils from urban outperformed in all key subjects except the environmental science achievement that favors the rural by a mean score point of 0.38% though their mean difference was not statistically significant.

Moreover, mean difference for English test score (0.28%) between urban and rural pupils was not statistically significant. Nonetheless, mean differences for reading (4.43%) and for Mathematics (1.02%) were statistically significant and favoring pupils from the urban location.

On the other hand, in the 2008 national learning assessment, pupils from rural areas have outperformed pupils from urban areas in the composite mean score, with a mean difference of 1.59% that was statistically significant. Moreover, the rural pupils mean scores were higher in all of the key subjects from a lowest of 0.04 in Mathematics to a highest of 2.48 in reading. Furthermore, all mean differences were statistically significant in all key subjects except mathematics.

In summary, test performances in the 2004 national learning assessment favored pupil from urban schools and test performances in the 2008 national learning assessment favored pupils from rural schools. Moreover, the mean differences were statistically significant for reading and Mathematics in the 2004 national assessment result. On the other hand, for the 2008 national learning assessment the mean differences were statistically significant for reading, English, and environmental sciences.

2.12.1.5 Pupils' Performance at Varying Levels of Performance Standards in 2004 and 2008

Mean scores of the composite and mean scores mainly reported pupils' achievement for the second national learning assessment for each key subject. Based on this, the mean score was 48.5% for the composite, ranging from the lowest 39.7% in Mathematics to the highest 64.5% in reading. On the other hand, the mean composite score in the TNLA was 40.9%, which was lower from the SNLA by 7.6% score points. Moreover, the mean scores for key subjects range from 36.5% in English to 43.9% in reading.

In addition to reporting test performances using mean scores for composite and key subjects, test performance in the third national learning assessment for grade 4 pupils was reported using three different proficiency levels as below basic, for pupils who have performed below the mean of the scaled score; basic for pupils who have performed between the mean and one standard deviation above the mean, and proficient that is for pupils who have performed greater than one standard deviation from the mean of the scaled score and beyond. Based on these levels of proficiencies, out of 11493 pupils who sat for the 2008 national learning assessment tests, only 14.7% of them achieved at the proficient level, 37.8% achieved at the basic level, and 47.4% of pupils achieved below the basic level, which involves scores lower than the mean of the scaled score.

In summary, we use mean scores were used for comparing the results of the two national assessments. To this end, there was a 7.6% score point decline of the composite mean score in the 2008 national learning assessment from the 2004. Unfortunately, it is not possible to compare results from the two learning assessments using proficiency levels, because results of the 2004 national learning assessment were not reported in terms of proficiency levels.

2.12.1.6 Summary of Pupils' performance by region, in the SNLA and TNLA

Table 3: Pupils' performance by region in the SNLA and TNLA

Region	SNLA				TNLA			
	Reading comprehension (%)	English (%)	Maths (%)	Env'tal science (%)	Reading comprehension (%)	English (%)	Maths (%)	Env'tal science (%)
Tigray	63.37	32.95	34.50	45.87	45.40	35.20	41.40	45.20
Afar	71.92	35.30	40.29	51.60	41.30	35.00	37.60	42.90
Amahara	70.92	37.87	43.27	60.05	56.70	40.50	34.90	48.10
Oromia	63.73	39.49	38.30	50.73	44.70	37.50	44.60	42.80
Somali	50.53	50.36	43.09	45.83	33.70	34.30	35.20	31.10
Benshangul-Gumuz	60.34	32.44	34.34	47.84	43.20	34.10	38.40	41.80
SNNPR	60.31	38.93	39.70	49.75	43.80	39.60	41.90	44.20
Harari	68.60	42.46	40.71	53.64	38.00	34.80	40.70	42.20
Addis Ababa	76.90	39.94	44.92	54.64	50.40	36.90	41.40	47.30
Dire Dawa	62.24	37.19	37.83	47.86	41.30	35.80	42.40	43.90
Gambella	-	-	-	-	32.50	32.50	40.00	31.80
National Average	64.49	38.68	39.70	51.74	43.90	36.50	40.30	42.60

As shown in Table 3, Reading Comprehension achievement was the highest for each of the regions in the SNLA when compared to the other key subjects (mean percentage score of 64.49). The top average was for pupils from Addis Ababa (mean percentage score of 76.90), and the least average was for pupils from Somali (mean percentage score of 50.53). In the TNLA, average score for Reading Comprehension was only 43.90, which has a difference of 20.59 mean score points from the SNLA result. In

Reading Comprehension of the TNLA, pupils from Amhara region had achieved the highest average score (mean percentage score of 56.70). The least performing were those pupils from Gambella region (mean percentage score of 32.50).

Pupils from Somali region had achieved the highest average in English (mean percentage score of 50.36) in the SNLA. The least performing pupils were from Benshangul-Gumuz (mean percentage score of 32.44). When it comes to the TNLA, pupils from Amhara region had achieved the highest average in English (mean percentage score of 40.50). The least performing pupils were from Gambella region (mean percentage score of 32.50).

In the SNLA, average Mathematics achievement was the highest for pupils from Addis Ababa region (mean percentage score of 44.92), and the least average was for pupils from Benshangul-Gumuz region with mean percentage score of 34.34. Whereas in the TNLA, pupils from Oromia region had the highest average in Mathematics (mean percentage score of 44.60), and the least average was for pupils from Amhara region (mean percentage score of 34.90).

With respect to Environmental Science result of the SNLA, pupils from Amhara region had achieved the highest average (mean percentage score of 60.05), and the least average goes for pupils from Somali region (mean percentage score of 45.83). In the TNLA of environmental pupils from Amhara had achieved the highest average was for pupils from Somali region (mean percentage score of 31.10).

2.13 Factors Influencing the Academic Achievement of Pupils

Many educators do believe that academic achievement of pupils is influenced by several factors happening within and outside of the school. This subsection of the study briefly discusses the major factors both within and outside of the school, and influencing the academic achievement of pupils. Several factors operate to produce pupils' academic achievement. The pupil's home environment, learning facilities, the instructional language, time of instruction, and frequency of homework can be good examples. Moreover, what is going on in the school, such as pupils' motivation, teachers' perceptions, their education level, relationship behavior and teaching experiences, all potentially influence pupils' achievement. In line with this, there is a great concern on the part of teachers, education planners, government officials and researchers in education to know about and identify the factors that significantly influence pupils' academic performance.

2.13.1 Within School Factors Influencing Academic Achievement

School based researches have been reporting the significant influence of within school factors on pupils' academic performance. Scholars in the field of education (e.g. Stockard and Mayberry, 1992) classify within school factors as school environments and school resources. On the other hand, Bruce (1992) as cited in Getahun (2002) categorizes the within school factors that influence achievement into four. These are (1) teacher quality – that includes the level of qualification, experience, verbal proficiency, in-service training, and low absenteeism. (2) Classroom organization – that involves homework frequency, length of instruction, student participation in learning, and teacher preparations for lessons; material inputs such as class size, availability of instructional materials, library size, instructional media, and laboratories. And, (3) school management, such as the leadership quality of the principal.

Within school factors are highly influential to academic performance. For instance, Heyneman and Loxley (1982) reported that in some Indian schools, the overwhelming proportion i.e. 90% of the variance in pupils' science achievement was explained by school variables.

Teacher quality and characteristics such as years of schooling, preparation, in-service training, and verbal proficiency have great influence on pupils' achievement in developing countries (Fuller, 1986). If we consider years of schooling as an instance, pupils having teachers with more years of post – secondary education perform better than those having teachers with fewer years of post secondary education (Aslam and Kingdon, 2008).

In many studies, teachers level of education become among the dominant factors affecting the pupils' academic gains and performance. As Kingdon (1999) states it, pupils learn more from teachers who hold higher degrees in subjects they are teaching, because the level of teacher's qualification to a lesser or greater degree affects classroom interaction. Teachers with advanced qualifications and experience are more likely to communicate easily and better, thereby enhancing the performance of their pupils (Bishop, 1996). Nevertheless, there are also research findings contrary to the mentioned claims (Getahun, 2002).

On the other hand, effective teaching – learning process or the classroom organization as a whole is likely to influence academic achievement. The way the teacher presents the lesson, motivation of learners, teacher–pupil relationship, expectation of standards, and pupils' level of understanding are highly influential to the pupils' academic achievement. Moreover, classroom organization refers to conditions such as home work frequency, length of instruction, active learning, and teacher preparation for lessons. According to Aslam and Kingdon (2008), from all within school factors influencing academic performance, teachers classroom practices and the teaching process such as student participation, teacher preparation, frequency of home work and checking student work matters a lot. Similarly, Anderson (1999) reported that teachers who regularly monitor and supervise pupils learning by checking their work and helping individual pupil to overcome errors and learning difficulties are likely to have pupils who exhibit higher level of achievement.

The other important factor within the school to influence pupils' test performance is material input. Material input includes resources that facilitate the learning process. Availability, relevance, and adequacy of educational resources such as textbooks and reading materials used by pupils and teachers contribute to academic achievement (Hallack, 1990). Fuller (1985) reported that students who had used two or more books in their learning were almost three times better than those who had no textbooks. However, the availability of textbooks and reading materials in the school's library and store does not guarantee the quality of schooling, unless they are given to learners on time during a given academic year (Getahun, 2002). Furthermore, Hallack listed resources that contribute to poor academic performance as unattractive school building, crowded classrooms, non-availability of playgrounds, and school environment that has no aesthetic beauty.

Still another crucial factor to influence academic performance of pupils is the overall managerial potential and instructional leadership skill of the school principal. The principal is the main actor in the school improvement process. He/she has to coordinate instructional innovations and facilitations that will contribute in the enhancement of student learning and the quality of education. Visionary planning and closer supervision of the teaching – learning process by the school head also contributes to the academic achievement of pupils. Moreover, the principal's leadership in generating internal revenue of the school is an important factor for the school's effectiveness. In general, as Getahun (2002: 57) citing Rencher (1992) discussed it, the principal can create positive school environment to influence pupils and teachers by undertaking the following activities:

- Stress goal setting and self – regulation,
- Offer students choices in instructional settings,
- Reward students for attaining best goals,
- Foster team work through group learning and problem solving experiences,
- Teach time management skills,
- Show students that success is important,
- Involve parents in the issue of motivation and give them guidance in fostering in their children,
- Analyze the ways that motivation operates on own life and develop a clear way of communicating it to teachers and students.

In summary, Getahun (2002) concluded that the quality of school management has strong and positive relationship with pupils' performance. Thus, the school principal should be effective and competent in realizing the school goals objectively.

2.13.2 Out of school factors Influencing Academic Achievement

Other than within school factors, a number of out of school factors influence the academic achievement of pupils. The main ones that influence performance include pupil's personal characteristics and home background.

2.13.2.1 Pupil Personal Characteristics

Regarding the pupil characteristics there is a common recognition that gender, intrinsic motivation, level of understanding, reading habits and the like influence academic performance. For instance, the BNLA (2010) on grades 10 and 12 students reported that boys significantly outperformed girls in all the subjects tested. To the contrary, Erkyhun et.al (2004) citing Fagerlind and Saha (1989) reported that the academic performance of girls to be equal with the boys at primary school level, but at secondary level girls begin to do more poorly. Furthermore, several studies in the area of learning achievement report that girls do poorly in Mathematics and the sciences, and do better than boys in key subjects requiring language skills and comprehension. In addition, the pupils reading habits, intrinsic motivation, level of understanding, and a host of other factors do have significant influence on academic performance.

2.13.2.2 Home Variables

The pupils' home background can have significant influence on academic performance. Home background refers to family characteristics including parent's level of education, occupation, socio-economic status, and support to the child, family size, and number of siblings and so on. Pupils come to school from different home backgrounds. The family background, which they come from, should be an environment in which children have the opportunity to grow, succeed, and be happy. As Adeyamo (2010) stated it, a facilitative home influence manifests itself further in the school environment and academic performance of pupils. There is also growing evidence on the considerable influence of the social, cultural, and learning experiences, attitudes, and aspirations of pupils' home background on their academic performance (Coleman, 1990; Symeou, 2007).

According to Adeyemo (2010) the interplay of family factors such as parental educational level, income, occupation, support to the child, and parental relationship with each other greatly determine the child's readiness to learn and performance at school. For instance, broken homes may cause unhappiness that may in turn affect the child's academic achievement. In short, home backgrounds of pupils exert significant influence on their academic achievement.

2.13.3 Grade 8 Students Academic Performance in 2004 and 2008

Comparing 2008 academic performance of pupils with that of 2004, students' academic performance in Biology, Chemistry, English, Mathematics, and Physics showed a marked decline. Particularly the decline was significant in Chemistry and Mathematics. As a result, the overall academic achievement of students had shown a marked drop in 2008 learning assessment as compared to that of 2004.

Table 4: Trends in Academic Performances in Various Subjects and Overall Academic Performances of Grade 8 Students in 2004 and 2008.

Year		
	2004	**2008**
SUBJECTS	**Mean (%)**	**Mean (%)**
Biology	41.34	38.3
Chemistry	40.10	34.7
English	41.07	38.4
Mathematics	40.93	34.1
Physics	35.32	32.2
Composite	39.74	35.6

2.13.4 Trends in Academic Performance by Gender

In 2004 and 2008 national learning assessments, boys outperformed girls. However, comparing the 2008 and 2004 national Learning assessments, the gaps in academic achievement between the two sexes had narrowed slightly with the exception of English. In English, the difference between the two sexes had slightly increased in 2008. Because of narrowing the gap in academic achievement between the two sexes in 2008, the overall academic achievement difference between the two sexes had dropped to a score of 4.3% in 2008 as compared to 5.43% in 2004.

2.13.5 Trends in Academic Performance by Location

When we examine grade 8 pupils test performance by location, those from rural areas outperformed those from urban areas in all subjects except English, in 2004 and 2007 national learning assessments. Urban pupils performed better in English than the rural ones in 2004 and 2007 assessments even though the mean difference for 2004 was not statistically significant. The gap between urban and rural area pupils had narrowed in 2008 as compared to 2004 except for the English test. Consequently, the overall academic achievement gap of grade 8 pupils had dropped to 1.0 % in 2008 as compared to 2.05 % in 2004.

2.13.6 Trends of Students' Performance at Varying Levels of Performance Standards

In 2008, above 62% of pupils performed in the "below basic" category, and scored below the mean of the scaled score for the composite. Moreover, from the lowest percentage of 54.2 in physics and the highest percentage of 59.7 in English, mathematics, and biology performed in the below basic proficiency level. On the other hand, 25.1% of the pupils were in the basic proficiency level for English, mathematics, and biology, where the basic proficiency level lies between the mean and one standard deviation to the right on the scaled score. Furthermore, the proportion of pupils attaining at the proficient level, which is two standard deviations from the mean and beyond was only 13.9% for the composite and from 14.2% to 17% for the key subjects tested. In general, 62.1 %, 24.0 % and 13.9 % of the pupils performed in the "below basic," "basic" and "proficient" categories respectively for the composite test performance in the 2008 national learning assessment.

2.13.7 Trends of Students' Academic Performance in Biology and Chemistry by Region

When we examine grade 8 pupils performance in Biology in 2004, those from Tigray (mean percentage score of 49.08%) performed better followed by students from Oromia (mean percentage score of 48.43%) and Amhara (mean percentage score of 48.33%). On the other hand, pupils from Afar (mean percentage score of 31.96%) and Addis Ababa (mean percentage score of 33.67%) had performed relatively the least.

In the 2008 national learning assessment of biology, pupils from Tigray (mean percentage score of 48%) performed relatively better followed by pupils from Oromia (mean percentage of 45.5%). Pupils from Benshangul Gumuz (mean percentage score of 32.9%) and Dire Dawa (mean percentage score of 33.7%) were relatively the least performing in biology.

In 2008, pupils from Amhara region have significantly dropped in their performance in Biology and students from Harari, Dire Dawa, Oromia, and Benshangul Gumuz had shown marked drops in their performance in Biology in comparison to the 2004. Students from Afar and Somali regions had shown improvements in their biology performance in the 2008 learning assessment. However, students' overall test performance in biology had shown a drop of 7.24% in comparison to the 2004 national learning assessment.

When we examine grade 8 students' performance in Chemistry, students from Amhara regions performed better followed by those from Oromia, Tigray, and Harari regions while students from Afar and Benshangul Gumuz had performed relatively the least in Chemistry test of 2004.

In 2007, students from Tigray relatively performed better followed by students from Amhara, Harari and Oromia regions while students from Afar, Somalia, Benshangul Gumuz, Addis Ababa and Dire Dawa were the least performing in Chemistry.

When we examine the changes in academic performance, students in all regions except those from Tigray region had shown marked declines in their Chemistry test performance in 2007/8 when compared to their performance in 2004.In general the overall academic performance of grade 8 students in Chemistry declined by 5.4% in comparison with the 2004 performance.

2.13.8 Trends of Students' Academic Performance in Mathematics And Physics by Region

When we examine grade 8 students performance in Mathematics, students from Tigray performed better followed by students from Harari, Oromia and Somali regions while students from Benshangul Gumuz and Afar regions performed relatively the least in Mathematics test of the 2004 national learning assessment..

In 2008, students from Tigray performed better followed by students from Afar and Amhara regions while students from Gambela, Benshangul Gumuz and Somali were the least performing regions in Mathematics.

When we examine the changes in academic performance in Mathematics, students in all regions except those from Afar region have showed a marked drop in their performance in 2008 as compared to their performance in 2004 national learning assessment result. As a result, students over all academic performance in Mathematics in 2007 dropped by 6.83 % in comparison with the 2004 national learning assessment.

When we examine grade 8 students performance in Physics, those from Amhara performed better followed by students from Tigray and Oromia regions, while students from Benshangul Gumuz, Afar , Addis Ababa and SNNPR regions performed relatively least in Physics in 2004.

In 2007, students from Tigray performed better followed by students from Amhara and Oromia regions while students from Gambela, Benshangul Gumuz, and Somali were the least performing regions in Physics.

When we examine the changes in academic performance in Physics, students in all regions have showed a marked drop in their performance in Physics in 2007 as compared to their performance in 2004 except students from Benshangul Gumuz and SNNPR regions. As a result, students overall academic performance in Physics in 2008 dropped by 3.12 % in comparison with 2004 national learning assessment.

2.13.9 Trends of Students' Academic Performance in English by Region

Pupils from Harari performed better in English followed by those from Somali, Dire Dawa and Addis Ababa regions while those from Tigray, Amhara, Afar and Benshangul Gumuz regions performed least in English in 2004.

In 2008, students from Amhara, SNNPR, and Harari regions performed better while students from Benshangul Gumuz, Oromia, and Dire Dawa were the least performing regions in English.

When we examine the changes in academic performance in English, students in all regions showed a marked decline in their performance in English in 2008 as compared to their performance in 2004 in English except students from Amhara region. As a result, students over all academic performance in English in 2008 have dropped by 2.67 % in comparison with 2004 national learning assessment in English.

2.13.10 Trends in Students' Overall Academic Performances by Region

When we examine grade 8 students overall academic performances in 2004 and 2008 across regions, students from Oromia, Tigray and Amhara regions performed better as compared to students from Benshangul Gumuz, and Afar regions, which performed least in 2004. In 2008, students from Tigray regions performed better followed by Oromia and Amahra while students from Benshangul Gumuz, Dire Dawa, and Addis Ababa were the least performing in overall academic performances.

When we examine the changes in overall academic performance, students in all regions except students from Afar regions have showed a marked drop in the overall academic performance as compared to the overall performance in 2004 national learning assessment. The overall academic performance of grade 8 students in 2008 declined by 4.14% when it is compared with 2004 national learning assessment. Students from Oromia regions have showed the highest drop followed by students from Amhara, Harari, and Dire Dawa regions in overall academic performances.

2.13.11 Trends in Students' Overall Academic Performances by Location and Region

In 2004, the overall academic performance of rural students from Tigray, Benshangiul Gumuz and Harari, Somali , Dire Dawa , SNNPR and Amhara regions excel the urban ones while urban students from Afar and Oromia regions excel the rural ones by the 2004 national learning assessment.

In 2007, rural students from Tigray, Afar, SNNPR, Oromia, and Amhara regions excel the urban ones in overall academic achievement, while urban students from Gambela, Somali, Harari, and Benshangul Gumuz regions excel the rural ones in overall academic achievement.

When we investigate the changes in 2007 in comparison with the 2004 overall academic performances between urban and rural students across regions, the gap observed in 2004 in students of Tigray and Dire Dawa has narrowed slightly while the gap in students of SNNPR and Amhara regions has widened slightly in 2007 .On the other hand, the observed differences between urban and rural students in 2004 was reversed in students of Harari, Afar, Somali, Bensahngul-Gumuz and Oromia in 2007.

2.13.12 Trends in Overall Academic Performances of Pupils by Sex and Region

When we see to grade 8 students' overall academic performances in 2004 across regions and by sex, boys in all regions excelled the girls. There were highest differences in academic performances between the two sexes for students of Harari region followed by Dire Dawa, Tigray, Amhara, Oromia, SNNPR, and Addis Ababa.

In 2008, boys in all regions excelled the girls in overall academic achievement. When we investigate the changes in 2008 in comparison with the 2004 overall academic performances between the two sexes across regions, the gap between boys and girls in Harari, Diredawa ,Addis Ababa , Tigray ,Afar, and Oromia regions is decreased . There was a highest drop in Harari followed by Dire Dawa and Addis Ababa. On the other hand, the gap between boys and girls was wide in SNNPR, Somali, Bensahngul-Gumuz, Afar, Amhara, and Bensahngul- Gumuz regions. The highest gap in overall academic achievement between the two sexes was observed in students of SNNPR and Somali regions.

2.14 Factors Influencing the Academic Achievement of Pupils

Many educators do believe that several factors happening within and outside of the school influence academic achievement. This subsection of the study briefly discusses the major factors both within and outside of the school, and influencing the academic achievement of pupils. Several factors operate to produce pupil achievement. The pupil's home environment, learning facilities, the instructional language, time of instruction, and frequency of homework can be good examples. Moreover, what is going on in the school such as pupils' motivation, teachers' perceptions, their education level, relationship behavior and teaching experiences influence pupils' achievement. In line with this, there is a great concern on the part of teachers, education planners, government officials, and researchers in education to know about and identify the factors that influence pupils' academic performance.

2.14.1 Within the School Factors

School based researches have been reporting the significant influence of within school factors on pupils' academic performance. Scholars in the field of education (e.g. Stockyard and Mayberry, 1992) classify within school factors as school environments and school resources. On the other hand, Bruce (1992) as cited in Getahun (2002) categorizes the within school factors that influence achievement into four. These are teacher quality – that includes qualification, experience, verbal proficiency, in service training, and low absenteeism. Classroom organization – that involves homework frequency, length of instruction, student participation in learning, and teacher preparations for lessons. In addition, material inputs such as class size, availability of instructional materials, library size, instructional media, and laboratories; and school management, such as the leadership quality of the principal.

Within school, factors are highly influential to academic performance. For instance, Heyneman and Loxley (1982) reported that in some Indian schools, the overwhelming proportion i.e. 90% of the variance in pupils' science achievement was explained by school variables.

Teacher quality and characteristics such as years of schooling, preparation, in-service training, and verbal proficiency have great influence on pupils' achievement in developing countries (Fuller, 1986). If we consider years of schooling as an instance, pupils having teachers with more years of post – secondary education perform better than those having teachers with fewer years of post secondary education (Aslam and Kingdon, 2008).

In many studies, teachers level of education become among the dominant factors affecting students' academic gains and performance. As Kingdon (2008) states it, pupils learn more from teachers who hold higher degrees in subjects they are teaching because the level of teacher's qualification to a lesser or greater degree affects classroom interaction. Teachers with advanced qualifications and experience are more likely to communicate easily and better thereby enhancing the performance of their pupils (Bishop, 1996). Nevertheless, there are also research findings contrary to the mentioned claims (Getahun, 2002).

On the other hand, educators recognize that effective teaching – learning process or the classroom organization as a whole influences achievement. The way the teacher presents the lesson, motivation of learners, teacher – pupil relationship, expectation of standards, and pupils' level of understanding are highly influential to the pupils' academic achievement. Moreover, classroom organization refers to conditions such as homework frequency, length of instruction, active learning, and teacher preparation for lessons. According to Aslam and Kingdon(2008), from all within school factors influencing academic performance, teachers classroom practices and the teaching process such as student participation, teacher preparation, frequency of home work and checking student work matters a lot. Similarly, Anderson (1999) reported that teachers who regularly monitor and supervise pupils learning by checking their work and helping individual pupil to overcome errors and learning difficulties are likely to have pupils who exhibit higher level of achievement.

The other important within the school factor to influence pupils' test performance is material input. Material inputs include resources that facilitate the learning process. Availability, relevance, and adequacy of educational resources such as textbooks and reading materials used by pupils and teachers contribute to academic achievement (Hallack, 1990). Fuller (1985) reported that students who had used two or more books in their learning were almost three times better than those who had no textbooks. However, the availability of textbooks and reading materials in the school's library and store does not guarantee the quality of schooling, unless they are given to learners on time during a given academic year (Getahun, 2002). Furthermore, Hallack listed resources that contribute to poor academic performance as unattractive school building, crowded classrooms, non-availability of playgrounds, and school environment that has no aesthetic beauty.

Still another crucial factor to influence academic performance of pupils is the overall managerial potential and instructional leadership skill of the school principal. The principal is the main actor in the school improvement process. He/she has to coordinate instructional innovations and facilitations that will contribute in the enhancement of student learning and the quality of education. Visionary planning and closer supervision of the teaching – learning process by the school head also contributes to the academic achievement of pupils. Moreover, the principal's leadership in generating internal revenue of the school is an important factor for the school's effectiveness. In general, as Getahun (2002: 57) citing Rencher (1992) discussed it, the principal can create positive school environment to influence students and teachers by undertaking the following activities:

- Stress goal setting and self – regulation,
- Offer students choices in instructional settings,
- Reward students for attaining best goals,
- Foster team work through group learning and problem solving experiences,
- Teach time management skills,
- Show students that success is important,
- Involve parents in the issue of motivation and give them guidance in fostering in their children,
- Analyze the ways that motivation operates on own life and develop a clear way of communicating it to teachers and students.

In summary, Getahun (2002) concluded that the quality of school management has strong and positive relationship with pupils' performance. Thus, the school principal should be effective and competent in realizing the school goals objectively.

2.14.2 Out of School Factors

Other than within school factors, a number of out of school factors influence the academic achievement of pupils. The main ones that influence performance include pupil's personal characteristics and home background.

2.14.2.1 Student Personal Characteristics

Regarding the pupil characteristics there is a common recognition that gender, intrinsic motivation, level of understanding, reading habits and the like influence academic performance. For instance, For instance, the BNLA (2010) on grades 10 and 12 students reported that boys significantly outperformed girls in all the subjects tested. To the contrary, Erkyhun et.al (2004) citing Fagerlind and Saha(1989) reported that the academic performance of girls to be equal with the boys at primary school level, but at secondary level girls begin to do more poorly. Furthermore, several studies in the area of learning achievement report that girls do poorly in mathematics and the sciences, in fact they do better than boys in key subjects requiring language skills and comprehension. In addition, the pupils reading habits, intrinsic motivation, level of understanding, and a host of other factors do have significant influence on academic performance.

2.14.2.2 Home Variables

The pupils' home background can have significant influence on academic performance. Home background refers to family characteristics including parent's level of education, occupation, socio-economic status, and support to the child, family size, and number of siblings and so on. Pupils come to school from different home backgrounds. The family background, which they come from, should be an environment in which children have the opportunity to grow, succeed, and be happy. As Adeyemo (2010) stated it, a facilitative home influence manifests itself further in the school environment and academic performance of pupils. There is also growing evidence on the considerable influence of the social, cultural, and learning experiences, attitudes, and aspirations of pupils' home background on their academic performance (Coleman, 1990; Symeou, 2007).

According to Adeyemo (2010) the interplay of family factors such as parental educational level, income, occupation, support to the child, and parental relationship with each other greatly determine the child's readiness to learn and performance at school. For instance, broken homes may cause unhappiness that may in turn affect the child's academic achievement. In short, home backgrounds of pupils exert significant influence on their academic achievement.

CHAPTER THREE: RESEARCH METHODOLOGY

3.5 Sampling Technique

The target populations of this study were Grades 4 and 8 students at the end of the 2010/2011 academic year (2003 E.C.). In order to identify the students involved in the study, a two-stage stratified cluster sampling procedure was followed. At the first stage, a fixed number of schools proportional to the number of schools in the regions were randomly selected for each grade. At the second stage 40 students were taken from each sample school. The students involved in the study sat for an achievement tests in the four and five core subjects at grades 4 and 8 respectively. Moreover, students of both grades filled out questionnaires pertinent to their personal characteristics, home background, and school situations. In addition, those teachers teaching the sampled students were also asked to fill out questionnaires.

The Tables below show the distribution of sample schools and students across regions. Looking at planned and achieved sample sizes, the response rates at the first and the second stages were 95.22% and 85.88% for grade 4 and 96.68% and 92.87% for grade 8 respectively.

Table 5: Distribution of grade 4 Sample Schools and Students by Region

Region	Schools		Students	
	Planned	Achieved	Planned	Achieved
Addis Ababa	25	26	1000	914
Afar	25	24	1000	877
Amhara	30	29	1200	1055
Benshangul-Gumuz	25	24	1000	819
Dire Dawa	25	22	1000	757
Gambella	25	25	1000	912
Harari	25	23	1000	776
Oromia	42	40	1680	1379
SNNPR	42	42	1680	1570
Somali	25	19	1000	752
Tigray	25	25	1000	976
Grand Total	**314**	**299**	**12,560**	**10,787**

Table 6: Distribution of grade 8 Sample Schools and Students by Region

Region	Schools		Students	
	Planned	Achieved	Planned	Achieved
Addis Ababa	25	24	1000	915
Afar	25	24	1000	921
Amhara	32	32	1280	1240
Benshangul-Gumuz	25	23	1000	921
Dire Dawa	21	21	840	833
Gambella	25	25	1000	921
Harari	24	24	960	919
Oromia	42	42	1680	1626
SNNPR	32	32	1280	1254
Somali	25	19	1000	635
Tigray	25	25	1000	996
Grand Total	**301**	**291**	**12,040**	**11,181**

3.6 Data Collecting Instruments

Two kinds of instruments were used for the collection of data, namely achievement tests in English, Mathematics, Biology, Chemistry and Physics for grade 8, English, Mathematics, Mother Tongue Reading and Environmental Science for grade 4 and background questionnaires for students, teachers, and directors of both grades. The achievement tests were developed based on the National Curriculum using the Minimum Level of Competency (MLC). These instruments were pilot tested in November, 2011 in selected schools. Item and Test analyses were carried out. Based on the results of the analyses, items were further improved and final versions produced. The questionnaires contained questions in order to identify the association of various personal, home, school, and teaching-learning variables with student achievement.

3.7 Data Collection and Organization

Experts form NOE and MOE served as route coordinators. Data collectors were recruited from the regions. They attended a training program on how to administer the tests and questionnaires. In each school, a data collector stayed for three days to administer the tests and questionnaires.

Data from each grade tests and questionnaires were captured using MS Access. For the purposes of data cleaning and checking consistency, MS Excel 2007 and SPSS v 15 were used.

3.8 Data Analysis

Descriptive summary statistics to summarize central tendencies and dispersion were computed to each grade and subject and to the average score. Correlation and statistical tests of significance were also computed to detect relationships and differences. One-way analysis of variance followed by Post Hoc test was computed to identify homogenous subset groups. Statistical analyses were carried out using SPSS v 15. All tests of statistical significance in this study considered significant at a level of alpha = 0.05.

CHAPTER FOUR: FINDINGS OF THE STUDY

This chapter presents findings of the fourth national learning assessment for both grades 4 and 8. It reports overall academic performance and typical performance of pupils in the key subjects tested. Moreover, the report includes presentation of achievements in terms of the pupils' attainment of proficiency levels. In addition to this, it describes the relationship of test performance with factors such as home background, pupil characteristics, school variables that potentially influence pupil performance. As a result, since grades 4 and 8 are the terminations to the first cycle of primary education and the second cycle of the same respectively, having the aims of providing basic education, they can both be considered as an indication to the state of learning at the lower and higher levels of providing general education to citizens in Ethiopia.

4.3 Grade 4 Students' Academic Performance across Key Subjects and Overall Academic Performance

4.3.1 Summary of the Descriptive Statistics

Table 7: Performance of Pupils in the Key subjects Tested

Subjects	N	Mean (%)	Median (%)	Std. Dev.	Std. Error	T	df	Sig.	MD
Reading	11229	42.96	40	19.93	.19	-37.44	11228	.000	-7.04
English	10787	38.87	35	18.67	.18	-61.93	10786	.000	-11.13
Mathematics	11228	37.06	32.50	16.82	.16	-81.54	11227	.000	-12.94
Environmental Sciences	10787	41.21	40	16.27	.16	-56.12	10786	.000	-8.79
Composite	**10365**	**40.06**	**38.96**	**15.01**	**.15**	**-67.38**	**10364**	**.000**	**-9.94**

Similar to results of the second and third national learning assessments, the mean composite score of the 4 key subjects was below the minimum expected average (50%). The one sample t test also indicated that there was statistically significant mean difference between the composite score of each subject and the minimum competency level (50%).

The composite median score (38.96%) is less than the composite mean score (40.06%) by score points of 1.10%. This shows that 50% of the 10365 grade 4 pupils who sat for the 4th national learning assessment tests have achieved 38.96% or below (see Table 7).

As seen in Table 7 above, the mean and median scores were relatively the least for Mathematics test. Moreover, in the key subjects tested the median scores were less than the mean scores in score points ranging from 1.21% in environmental science to 4.56% in mathematics. Compared to the previous national learning assessments of grade 4, particularly to the second and the third, the average scores in the 4th national learning assessment are decreasing for the composite as well as the key subjects except for English. The English mean score (38.87%) was slightly greater than the second and the third assessment results, which were 38.68% and 36.50% respectively.

4.3.2 Performance Standard of Pupils in Each Subject

Table 8: Proportion of Pupils in the Different Performance Standard

Subject	Below Basic	Basic	Proficient
Reading	54.2%	26.3%	19.5%
English	59.6 %	23.5%	17.0 %
Mathematics	56.3 %	27.2 %	16.6 %
Environmental Sciences	53.8 %	27.3%	18.8 %
Composite	**57.1%**	**25.6%**	**17.3%**

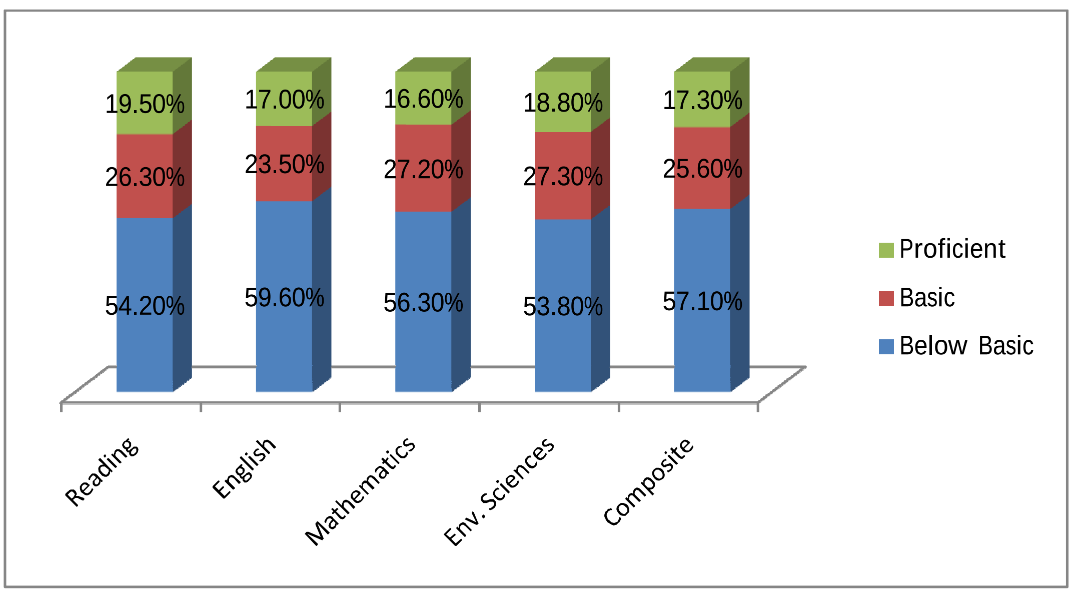


Figure 1: Performance Standard of Pupils in Each Subject

The performances of grade 4 pupils were categorized into three proficiency levels based on their test achievements. The categorization was made in reference to standard z – distributions, namely below basic, basic and proficient. The below basic category includes those students who fall at or below a z standard score of zero; the basic category is within a z standard score of zero and one standard deviation above the mean; and proficient category includes those students who fall above a z standard score of one standard deviation above the mean. Based on these categories, the proportion of pupils attaining at each level was reported for the composite score as well as for each key subject at the national level

For the composite average, 17.3% of the pupils' achievement is at the proficient level, which is a better performance when compared to the result of 3rd national learning assessment, which was only 14.7%.With respect to the key subjects, the proportion of grade 4 pupils attaining at the proficient level was improved for environmental science (with a proportion of 2.5% of pupils), English (with a proportion of 0.1% pupils), and Reading (with a proportion of 4.9% pupils) when compared to the 3rd national learning assessment result. On the other hand, in Mathematics there was a decrease of 0.50% in the proportion of pupils performing at the proficient level between the 3rd and the 4th national learning assessments.

At the basic performance level, the proportion of pupils for the composite score decreased by 12.2% in comparison with the third national assessment. Moreover, considerable decline was observed in the proportion of pupils who have performed at the basic level for the key subjects. For instance, the decline in proportion of pupils became 1.8% in mathematics, 7.4% in reading, 8% in English, and 11.2% in Environmental Science. These decreases in the proportion of pupils performing at the basic level may be an indication of poor performances of schools.

Finally at the below basic level, which is lower than the mean of the standard z score, the proportion of pupils performing at the level accounts for 57.1% with an increase of 9.7% from that of 3rd national learning assessment. For the key subjects the proportion of pupil who have performed at below basic level raised by 2.4% in mathematics, 2.5% in reading, 8% in English, and 8.7% in environmental science.

In summary, the trend of pupils performance at the different performance standards follows an increase for the proficient level, a decrease at the basic level, and an increase for the below basic level. Even though the increasing pattern in the proportion of pupil for the proficient level of achievement is regarded as encouraging, the decrease on the proportion of pupil at the basic level and an increase at the below basic level are indicators of poor performance of grade 4 pupils in the 4th national learning assessment.

4.3.3 Range of Achievement Scores of Grade 4 Students at five Different marker points

Table 9: Range of Achievement Scores at five Percentile Ranks

%ile Rank	Environmental Sciences(%)	English(%)	Reading(%)	Mathematics(%)	Average(%)
10th	22.5	20.0	20.0	20.0	23.75
25th	27.5	25.0	26.67	25.0	28.55
50th	40.0	35.0	40.0	32.5	37.08
75th	52.5	50.0	56.67	47.5	49.8
90th	62.5	67.5	73.33	62.5	61.67

Table 9 shows the range of achievements in five different percentile ranks for the respective key subjects tested. The data in the table depicts the cut-off scores for the 4 key subjects at the 10th, 25th, 50th, 75th, and 90th percentile. For instance, if we consider the 90th percentile for the average score only 10% the pupils could have achieved an average score of 61.67 and above. On the other hand, the 10th percentile shows that about 10% of grade 4 pupils scored an average of 23.75 and below. The highest percentile score was attained in Reading (i.e. about 10% of the students have achieved an average score of 73.33 in the Reading test). Whereas about 10% of the students has an average score of 20 and below in reading test. Moreover, these percentile ranks indicate the presence of highest variation among pupils in reading test achievement. With respect to Environmental Science test achievement, about 10% of the pupils achieved a score of 22.5 and below. And another 10% have achieved a score point of 62.5 and above. In English, about 10% of the pupils achieved a score point of 20 and below. On the other hand 10% of the top performing pupils achieved a score point of 67.5 and above. Finally, the percentile ranking for mathematics test achievement showed that about 10% of the pupils to have a maximum score of 20 and below that. On the other hand, about 10% of the top performing pupils achieved a minimum of 62.5 and above in mathematics test achievement. For further clarity, look at figures 2 and 3 below.

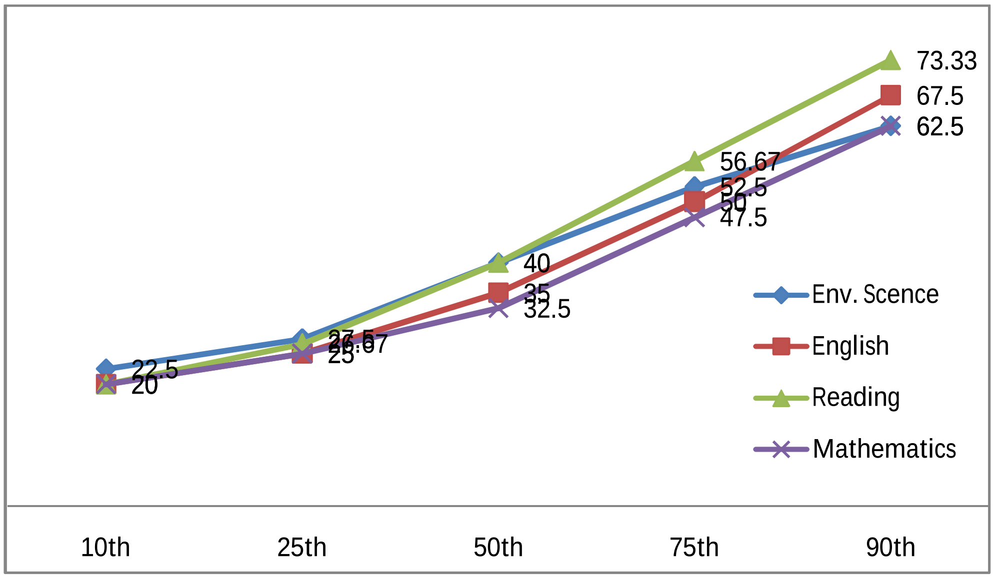


Figure 2: Range of achievement scores in key subjects at five key marker points

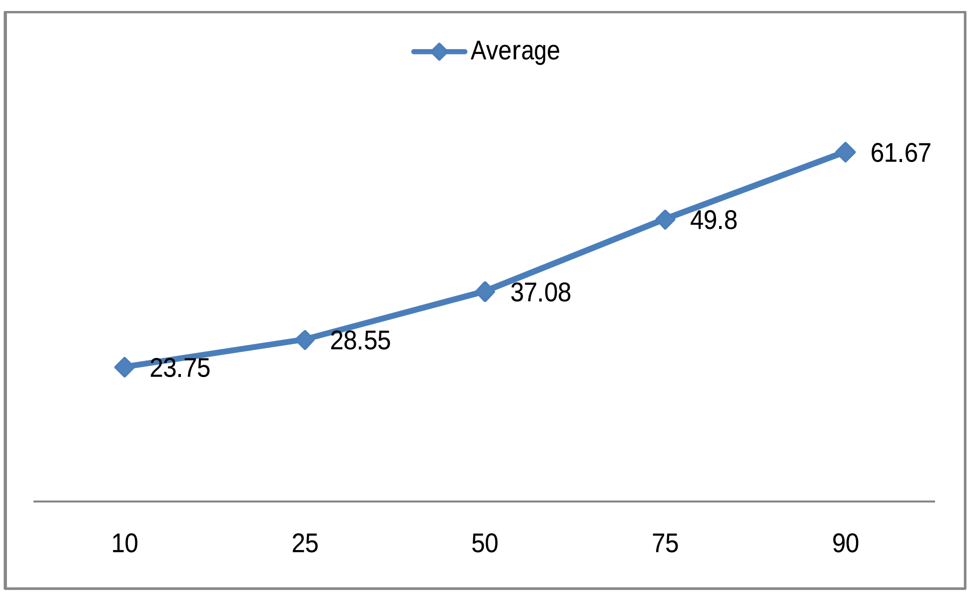


Figure 3: Range of overall achievement scores at five key markers

4.3.4 Group Differences in Academic Performances across Subjects and Overal Academic Achievement of Grade 4 Students

4.3.4.1 Test Performance of Boys and Girls

Table 10: Performance of Boys and Girls in Key Subjects T Value and Sig. Level

Subject	Gender	N	Mean (%)	Std. Error	Std. Deviation	T	Df	Sig.	MD
Mathematics	Boys	5873	37.88	.22	17.09	5.606	10,539	.000	1.84
	Girls	4668	36.04	.23	16.24				
Reading	Boys	5873	43.23	.26	19.92	-.093	10,540	.926	-.04
	Girls	4669	43.26	.29	20.08				
English	Boys	5632	39.30	.24	18.68	3.154	10,114	.002	1.17
	Girls	4484	38.13	.28	18.46				
Environmental Sciences	Boys	5632	41.52	.21	16.17	1.610	10,114	.107	.52
	Girls	4484	41.00	.24	16.20				
Composite	**Boys**	**5529**	**40.45**	**.20**	**15.13**	**2.939**	**9,937**	**.003**	**.89**
	Girls	**4410**	**39.56**	**.22**	**14.81**				

Table 10 above shows performance variation between boys and girls. For the composite score, boys perform better than girls do. Even though the mean difference is smaller in size (0.89), it is found to be statistically significant ($p<.003$). When it comes to the four key subjects tested, boys' performance became relatively higher for mathematics, English and environmental sciences. Moreover, the mean differences are statistically significant between the two sexes in Mathematics and English ($p <.000$ and $p <.002$ respectively) but not for environmental sciences ($p >.01$).On the other hand, girls outperformed boys in reading skill. Nevertheless, mean differences for reading are not statistically significant ($p >.01$). In comparison to the 3rd national learning assessment results that boys excelled females in all key subjects tested, there is a change in the fourth national learning assessment in which girls also outperformed boys in test of reading skills though the difference are not statistically significant. In the third national learning assessment, statistically significant mean differences were observed for English and environmental science in favor of boys. Whilst in the 4th national learning assessment, statistically significant mean differences are observed for Mathematics and English, still favoring boys.

4.3.4.2 Test Performance of Urban and Rural Pupils

Subject	Location	N	Mean	Std. Error	Std Deviation	T	Df	Sig.	MD
Mathematics	Rural	5251	36.32	.23	16.47	-3.61	9607	.447	-1.22
	Urban	4358	37.54	.25	16.59				
Reading	Rural	5251	40.25	.26	18.69	-14.57	9608	.000	-5.89
	Urban	4359	46.15	.32	20.96				
English	Rural	5480	37.53	.24	18.02	-7.77	9951	.000	-2.91
	Urban	4473	40.44	.29	19.30				
Environmental Sciences	Rural	5480	39.65	.21	15.63	-11.18	9951	.000	-3.66
	Urban	4473	43.32	.25	17.00				
Composite	Rural	5254	38.42	.20	14.47	-11.29	9597	.000	-3.46
	Urban	4345	41.87	.23	15.46				

Table 11: Performance of Urban and Rural Pupils in Key Subjects, T Value and Significance Level

When we examine pupils' test performance by location, urban pupils excelled rural students in the composite as well as in the key subjects tested. Large differences were observed for environmental science (t= -11.8, DF=9951 and p=.000) and reading (t=-14.57, DF= 9608, and p =.000). Except for mathematics, observed mean differences were statistically significant. For instance, mean difference for the composite is 3.46 score points (t=-11.29, DF=9597=, and p=.000). In contrast to the 3rd national learning assessment where rural pupils outperformed the urban for the composite and the key subjects, results in the 4th national learning assessment reversed in favor of the urban pupils.

4.3.4.3 Pupils Performance for the Composite Score across Regions

Region	N	Mean (%)	SD	Std. Error
Tigray	968	38.20	13.53	.43
Afar	872	36.31	12.96	.44
Amhara	919	43.51	15.88	.52
Oromia	1286	38.50	14.55	.41
Somali	659	41.21	13.35	.52
Benshangul-Gumuz	812	36.32	14.70	.52
SNNPR	1541	42.21	14.47	.37
Gambella	897	32.05	9.97	.33
Harari	748	40.23	14.55	.53
Addis Ababa	905	52.26	16.45	.55
Dire Dawa	756	38.60	14.58	.53
Composite	**10363**	**40.06**	**15.02**	**.15**

Table 12: Composite Scores by Region

The data in Table 12 depicts the average performance of pupils from the different regions. Average performances were below 50% for all regions except Addis Ababa. The average performance of pupil from Addis Ababa (i.e. 52.26%) is the highest of all other regions, followed by pupils' from Amhara (i.e. 43.51%), and pupils' from SNNPR (i.e. 42.21%). On the other hand, average scores became the least for pupils from Afar (i.e. 36.31%), Benshangul-Gumuz (36.32%), and Gambella (32.05%) regions.

Table 13: ANOVA Summary for Overall Academic Achievement by Regions

	Sum of squares	DF	Mean square	F ratio	Sig.
Sum square between	242830.819	10	24283.082	120.004	.000
Sum square within	2094756.418	10352	202.353		
Sum square total	2337587.236	10362	24283.082		

Furthermore, the one-way analysis of variance (ANOVA) that assumes homogeneity of variances among the scores of pupils from the different regions revealed the presence of a statistically significant differences ($F (10, 10352) = 120.004$, $p= .000$) for the overall academic performance of pupils across regions.

Table 14: Homogenous Subset Grouping for Composite Score across Regions

Region	Number of pupils	Subset for alpha = 0.005						
		1	2	3	4	5	6	7
Gambella	897	32.05						
Afar	872		36.31					
Benshangul-Gumuz	812		36.31					
Afar	968		38.20	38.20				
Oromia	1286		38.50	38.50	38.50			
Dire Dawa	756		38.60	38.60	38.60			
Harari	748			40.23	40.23	40.23		
Somali	659				41.21	41.21	41.21	
SNNPR	1541					42.21	42.21	
Amhara	919						43.51	
Addis Ababa	905							52.26
Significance		**1.000**	**.314**	**.520**	**.094**	**.566**	**.312**	**1.000**

The Scheffe test of homogenous mean subset grouping (see Table 14 above) showed statistically significant mean differences of pupils' overall performances between Gambella and the other regions; and between pupils of Addis Ababa and the other regions. Pupils from Addis Ababa are relatively top performers, while those from Gambella are relatively least performers. For further statistically significant mean differences of pupils' overall performances, please see Table 10 above.

4.3.4.4 Pupils Performance for Key Subjects across Regions

Reading Academic Achievement

Table 15: **Pupils' Achievement in Reading by Region**

Region	N	Mean(%)	SD	Std. error
Tigray	969	40.87	17.39	.56
Afar	872	41.27	19.64	.67
Amhara	917	48.60	20.25	.67
Oromia	1284	39.57	17.73	.50
Somali	660	38.04	15.34	.60
Benshangul-Gumuz	812	38.99	20.53	.72
SNNPR	1541	42.56	18.73	.48
Gambella	896	35.33	15.62	.52
Harari	764	43.98	19.61	.70
Addis Ababa	905	62.62	20.52	.68
Dire Dawa	756	41.49	19.26	.70
Composite	**10376**	**43.04**	**19.88**	**.19**

Table 15 above shows grade 4 pupils average performances for reading across the regions. The top 3 regions where grade 4 pupils performed better in reading skills are Addis Ababa(with average score of 62.62%), Amhara (with average score of 48.60%), and Harari (with average score of 43.98%). On the other hand, the average performance in reading skills is relatively the least for pupils from Gambella (with average score of 35.33%), Somali (with average score of 38.04%), and Oromia (with average score of 39.57%). Only pupils from Addis Ababa scored an average above 50%.

Table 16: ANOVA Summery for Reading Test Scores

	Sum of squares	DF	Mean square	F ratio	Sig.
Sum square between	483935.758	10	48393.576	138.658	.000
Sum square within	3617539.372	10365	349.015		
Sum square total	4101475.130	10375			

The test of one way ANOVA conducted (see Table 16) for reading skills test performances across regions revealed the presence of statistically significant mean differences ($F(10, 10365) = 138.658$ $p= .000$).

Table 17: Scheffe Test for Reading Score

Region	Number of pupils	Subset for alpha = 0.005					
		1	2	3	4	5	6
Gambella	896	35.33					
Somali	660	38.05	38.05				
Benishangul- Gumuz	812	38.99	38.99	38.99			
Oromia	1284		39.57	39.57			
Tigray	969		40.87	40.87	40.87		
Afar	872		41.27	41.27	41.27		
Dire Adwa	756		41.49	41.49	41.49		
SNNPR	1541			42.56	42.56		
Harare	764				43.98		
Amhara	917					48.56	
Addis Ababa	905						62.62
Significance		**.072**	**.125**	**.090**	**.260**	**1.000**	**1.000**

In addition, based on the Scheffe test of subset grouping, the average reading skills performance of pupils from Addis Ababa have shown statistically significant differences from the rest of the regions; and the average reading skills performance of pupils from Amhara regions have shown statistically significant differences from the rest of the regions. Pupils from Gambella have shown a statistically significant variation in reading performance from the rest of the regions except those pupils from Benshangul Gumuz and Somali regions. On the other hand, pupils from the six regions, namely Benshangul Gumuz, Oromia , Tigray, Afar, Dire Dawa, and SNNPR became homogenous in their average score of reading skills test (for further information see Table 17 above).

English Academic Achievement

Table 18: Average Scores of Regions for English

Region	N	Mean (%)	SD
Tigray	976	34.46	15.99
Afar	877	31.91	14.66
Amhara	1055	37.91	18.50
Oomiya	1379	38.37	18.22
Somali	752	47.75	18.55
Benishangul-Gumuz	819	33.45	17.08
SNNPR	1570	43.97	19.36
Gambella	912	34.13	16.07
Harari	776	38.58	18.42
Addis Ababa	914	45.69	20.67
Dire Dawa	757	39.05	19.29
Composite	10787	38.87	18.67

As shown in Table 18 above, the pupils' average test performance in English is below 50% in all of the regions. Average performances are relatively better for pupils from Somali (47.75%), Addis Ababa (45.69%), and SNNPR (43.97%) regions. The lowest average scores in English test are reported for pupil from Afar (31.91%), Benshangul-Gumuz (33.45%), and Gambella (34.14%) regions.

Table 19: ANOVA Summery for English Test Scores across Regions

	Sum of squares	Df	Mean square	F ratio	Sig.
Sum square between	203559.898	10	22617.766	71.478	.000
Sum square within	3120933.498	9863	316.428		
Sum square total	3324493.396	9872			

The ANOVA summary shown above in Table 16 indicates the existence of statistically significant variation ($F_{(10, 9863)} = 71.48$, $p < .001$) in the English test performance of pupils across regions.

Table 20: Scheffe Test for English Score

Region	Number of pupils	Subset for alpha = 0.005				
		1	2	3	4	5
Afar	877	31.91				
Benshangul-Gumuz	819	33.45				
Gambella	912	34.13				
Tigray	976	34.46	34.46			
Amhara	1055		37.91	37.91		
Oromia	1379			38.37		
Harari	776			38.58		
Dire Dawa	757			39.05		
SNNPR	1570				43.97	
Addis Ababa	914				45.69	45.69
Somali	752					47.75
Significance		**.510**	**.075**	**.997**	**.935**	**.810**

Moreover, the Scheffe test of homogenous subset grouping revealed statistically significant mean differences in English between pupils' of Somali region and the rest of regions except Addis Ababa and the test also revealed statistically significant mean differences in English between pupils' of Addis Ababa and SNNPR regions and the rest of regions . However, there were not statistically significant mean differences in English among pupils of Afar, Benshangul Gumuz, Gambella, and Tigray regions. In addition, the data reveals the presence of a wide achievement gap (average score points of 15.84%) of the English test between pupils of Somali and Afar regions (for details See Table 20 above).

Mathematics Academic Achievement

Table 21: Average Scores of Regions for Mathematics

Region	N	Mean (%)	SD
Tigray	969	36.84	15.02
Afar	871	32.56	13.61
Amhara	917	42.67	19.16
Oromia	1284	36.26	17.25
Somali	660	40.55	18.80
Benshangul-Gumuz	812	34.83	16.27
SNNPR	1541	39.59	16.20
Gambella	896	27.83	9.82
Harari	764	37.47	15.95
Addis Ababa	905	42.67	18.18
Dire Dawa	756	34.83	15.06
Composite	**10375**	**37.04**	**16.65**

As shown in Table 21 above, the pupils' average test performance in Mathematics is below 50% in each of the regions. Average performances in Mathematics are relatively better for pupils from Addis Ababa (42.67%), Amhara (42.67%), and Somali (40.55%) regions. The lowest average scores in Mathematics test are for pupil from Gambella (27.83%), Afar (32.56%), Benshangul-Gumuz (34.83%), and Dire Dawa (34.83%) regions.

Table 22: ANOVA Summery for Average Score in Mathematics across Different Regions

	Sum of squares	DF	Mean square	F ratio	Sig.
Sum square between	146615.466	10	17797.622	68.329	.000
Sum square within	2400681.865	10364	260.468		
Sum square total	2547297.331	10374			

The ANOVA summary in Table 19 indicates the presence of statistically significant variation ($F_{(10,\ 9863)} = 68.329$, $p=.001$) in the test performance of pupils in mathematics.

Table 23: Scheffe Test for Mathematics Score

Region	Number of pupils	Subset for alpha = 0.005					
		1	**2**	**3**	**4**	**5**	**6**
Gambella	896	27.83					
Afar	871		32.56				
Benshangul -Gumuz	812		34.83	34.83			
Dire Dawa	756		34.83	34.83			
Oromia	1284			36.26			
Tigray	969			36.84	36.84		
Harari	764			37.47	37.47	37.47	
SNNPR	1541				39.59	39.59	39.59
Somali	660					40.55	40.55
Addis Ababa	905						42.67
Amhara	917						42.67
Significance		1.000	.546	.287	.223	.092	.093

Moreover, the Scheffe test of homogenous subset grouping revealed statistically significant mean differences in Mathematics between pupils' of SNNPR , Somali, Addis Ababa, and Amhara regions and the rest of regions ; and the test also revealed pupils' of Gambella region to have statistically significant mean differences in Mathematics from the rest of regions . However, there were not statistically significant mean differences in Mathematics among pupils of Benshangul Gumuz, Dire Dawa, Oromia, Tigray and Harari regions. In addition, the data reveals the presence of a wide achievement gap (average score points of 14.84%) of the Mathematics test between pupils of Amhara or Addis Ababa and Gambella regions. For further detail on statistical significant differences, please see Table 23 above.

Environmental Sciences

Table 24: Average Scores of Regions for Environmental Science

Region	N	Mean (%)	SD	Std. error
Tigray	976	40.58	14.87	.476
Afar	877	39.32	15.97	.539
Amhara	1055	45.32	17.06	.525
Oromia	1379	39.11	14.76	.397
Somali	752	38.29	14.41	.525
Benshangul -Gumuz	819	37.83	16.05	.560
SNNPR	1570	42.40	14.70	.371
Gambella	912	30.92	10.23	.339
Harari	776	40.75	15.27	.548
Addis Ababa	914	58.02	16.32	.540
Dire Dawa	757	38.96	15.428	.560
Composite	**10787**	**41.21**	**16.27**	**.157**

As shown in Table 24 above, the pupils' average test performance in environmental science is below 50% for each of the regions participated except Addis Ababa, in which the average performance is 58.02%. Average performances in environmental science are relatively better for pupils from Addis Ababa (58.02 %.), Amhara (45.32%), and SNNPR (42.40%) regions. The lowest average scores in environmental science test are that of pupils from Gambella (30.92%), Benshangul-Gumuz (37.83%) and Somali (38.29%) regions.

Table 25: ANOVA Summery for Environmental Science Scores

	Sum of squares	DF	Mean square	F ratio	Sig.
Sum square between	404223.133	10	40422.313	177.697	.000
Sum square within	2451313.300	10776	227.479		
Sum square total	2855536.433	10786			

The one-way ANOVA test in Table 25 above revealed a statistically significant mean differences in Environmental Science test among pupils in the different regions (F(10, 10786) = 177.70, and p=000) .

Table 26: Scheffe Test for Environmental Science Score

Region	Number of pupils	Subset for alpha = 0.005				
		1	2	3	4	5
Gambella	912	30.92				
Benshangul-Gumuz	819		37.83			
Somali	752		38.28			
Dire Dawa	757		38.96			
Oromia	1379		39.11			
Afar	877		39.32			
Tigray	976		40.58	40.58		
Harari	776		40.75	40.75		
SNNPR	1570			42.39	42.39	
Amhara	1055				45.32	
Addis Ababa	914					58.02
Significance		**1.000**	**.065**	**.751**	**.064**	**1.000**

The Scheffe test below also revealed that mean performance of pupils in environmental science from Addis Ababa were statistically significantly different from pupils of other regions and those pupils from Gambella region were the least performing and statistically significantly different from pupils of other regions. See Table 26 above for further information.

Reading and English Academic Achievements in Relation to Performance Standard across Regions

Table 27: Pupils' Performance Standard in Reading and English by Regions

REGION	Reading (%)			English (%)		
	Below Basic	Basic	Proficient	Below Basic	Basic	Proficient
Tigray	57.1	28.9	14.0	70.8	19.1	10.1
Afar	57.1	24.7	18.2	75.0	18.8	6.2
Amhara	42.1	30.0	27.9	62.8	21.9	15.3
Oromia	58.8	28.3	12.9	60.2	24.4	15.4
Somali	59.2	33.6	7.1	35.0	31.6	33.4
Benishangul-Gumuz	63.7	19.2	17.1	71.7	17.9	10.4
SNNPR	54.3	27.2	18.6	46.9	28.7	24.4
Gambella	71.2	21.9	6.9	70.6	19.8	9.5
Harari	51.8	27.5	20.7	61.1	22.7	16.2
Addis Ababa	19.0	26.2	54.8	45.5	26.9	27.6
Dire Dawa	58.3	24.6	17.1	61.2	23.0	15.9
Total (%)	**53.8**	**26.6**	**19.6**	**59.6**	**23.5**	**17.0**

The data in Table 27 above shows pupils' performance in reading and English test at the different attainment levels by region. With respect to the reading test pupil from Addis Ababa (54.8%) and from Amhara (27.9%) have performed at the proficient level, which are top performances relative to the others. In all of the regions, except Addis Ababa and Amhara, highest proportion of grade 4 pupils (i.e. from 51.8% in Harari to 71.2% in Gambella) have performed at the below basic level in reading.

In English, relatively highest proportions of pupils from Somali (33.4%), Addis Ababa (27.6%), and SNNPR (24.4%) have performed at the proficient level of attainment. Similar to the reading test, highest proportions of grade 4 pupils from many regions have performed at the below basic level of attainment.

In summary, the regional level data analysis of pupils' performance in reading and English at the different attainment levels reveals that majority of them have performed at the below basic level.

Mathematics and Environmental Science Academic Achievements in Relation to Performance Standard across Regions

Table 28: Pupils' Performance Standard in Mathematics and Environmental Sciences by Regions

REGION	Mathematics (%)			Environmental Sciences (%)		
	Below Basic	Basic	Proficient	Below Basic	Basic	Proficient
Tigray	54.5	31.5	14.0	54.6	29.9	15.5
Afar	68.4	23.4	8.2	58.6	25.3	16.0
Amhara	45.7	26.9	27.4	43.6	29.8	26.6
Oromia	56.9	27.4	15.7	57.0	29.0	14.0
Somali	45.2	27.4	27.4	60.2	26.5	13.3
Benshangul-Gumuz	62.1	24.3	13.7	60.8	24.9	14.3
SNNPR	47.8	32.9	19.3	49.1	31.8	19.0
Gambella	82.3	17.0	.8	84.9	14.0	1.1
Harari	54.7	30.8	14.5	54.6	29.0	16.4
Addis Ababa	42.9	29.8	27.3	16.2	29.2	54.6
Dire Dawa	62.7	24.1	13.2	58.8	26.2	15.1
Total (%)	**56.2**	**27.3**	**16.5**	**53.8**	**27.3**	**18.8**

The data in Table 28 above shows, pupils' performance in Mathematics and environmental sciences tests at the different attainment levels by region. With respect to the Mathematics test more than a quarter proportion of pupil from Addis Ababa (27.3%) and from Amhara (27.4%), and Somali (27.4%) have performed at the proficient level, which are top performances relative to the others. In some of the regions such as Gambella , Afar, Dire Dawa, and Benshangul Gumuz significantly highest proportions of pupils have performed at the below basic level in mathematics, while those from Gambella take the lions share(82.3%) at this lowest level of performance.

In the environmental science test, grade 4 pupils from Addis Ababa (54.6%) followed by those from Amhara (26.6%) have performed at the proficient level of attainment. Significantly high proportion of pupils from some of the regions such as Gambella , Benshangul Gumuz, Somali, Dire Dawa, and Afar have performed at the below basic level of attainment in environmental science. Similar to the Mathematics test, very high proportion of pupils from Gambella (84.9%) have performed at this lowest level of attainment.

In summary, the regional level data analysis of pupils' performance in Mathematics and environmental science at the different attainment levels reveals that majority of them have performed at the below basic level.

4.3.5 Factors that Predict Overall Academic performance of Grade 4 Students

4.3.5.1 Personal Factors Predicting Overall Test Performance

Table 29: A Zero Order Correlation between Overall Academic Achievements and Personal Factors

	1	2	3	4	5
1. Academic achievement	1.00				
2. Gender	-.029**	1.00			
3. Age	.072**	-.028**	1.00		
4. Language	-.103**	-.022*	.043**	1.00	
5. Additional reading	-.031**	-.003	-.040**	-.018	1.00

** Significant at .01 level

*Significant at .05 level

As indicated in Table 29 above, pupils' personal characteristics such as gender, age, language spoken at home, and additional reading have shown statistically significant relationships with academic achievements though correlation coefficients are very low. The correlation coefficients for gender (where boys are coded 0 and girls coded 1 = -0.029), language (-0.103) and having additional reading at home (-0.031) have weak negative relationships while age (.072) have a weak positive relationship.

Table 30: Personal Factors as Predictors of Academic Achievement

Variables	R^2	B	SE	T	Sig.
Constant		38.572	.641	60.191	.000
Language	.011	-3.239	.304	-10.645	.000
Age	.016	.363	.050	7.327	.000
Gender	.017	-.939	.305	-3.077	.002
Additional reading	.018	-.404	.136	-2.978	.003

The stepwise regression analysis indicated the mentioned personal factors predicted significantly the overall academic achievement of grade 4 pupils. As depicted in Table 30 above only 1.8 % of the variance in grade 4 pupils' academic achievement was explained for by the linear combination of language, age, gender, and additional reading at home other than textbooks. Even though the coefficient of determination is considerably low, the personal variables explanation of pupils academic achievement is statistically significant (R squared = 0.018, p< 0.003).

4.3.5.2 Home Factors Predicting Overall Test Performance

Table 31: A Zero Order Correlation between Academic Achievement and Home Variable

	1	2	3	4	5	6	7
1. Academic achievement	1.00						
2. Family size	.041**	1.00					
3. Education attended	-.039**	.220**	1.00				
4. Father education	.026*	-.044**	.056**	1.00			
5. Tutorial given	-.052**	-.023*	.076**	.135**	1.00		
6. Listening radio	-.037**	.000	.070*	.070**	.203**	1.00	
7. Meal per day	.073**	.021*	.039**	.129**	.085**	.000	1.00

** Significant at .01 level
*Significant at .05 level

As indicated in Table 31 above , home variables such as family size, father education, number of meals per day, number of people attending education at home, number of times a tutorial is given per week and number of times listening to a radio had a statistically significant relationship with overall academic achievement. On the other hand, mother education did not show statistically significant relationships with test performance. Family size, father education and number of meals per day have shown a very weak positive relationship while number of people attending education at home excluding the child , number of times a tutorial is given per week , and number of times listening to radio in a week have a very weak relationship with overall students' academic achievement.

Table 32: Home Background as Predictors of Overall Academic Achievement

Variables	**B**	**SE**	**T**	**R^2 change**	**Sig.**
How often do you eat	1.560	.261	5.980	.004	.000
How often if tutored	-.912	.204	-4.477	.008	.000
Family size	.720	.141	5.096	.010	.000
Education attended excluding you	-.617	.138	-4.486	.013	.000
How often do you listen radio	-.459	.118	-3.874	.015	.000
Father education	.262	.108	2.434	.016	.015
Constant	**39.484**	**. 599**	**65.881**		**000**

The step–wise regression analysis made to see the extent that home variables explain variations in test performance have revealed the presence of considerably weak explanation (0.4%) of the variables to test performance. As shown in Table 32 above, 0.4% of grade 4 pupils test performance variation was explained by the linear combination of number of meals per day (B = 1.56, t = 5.90, p< .05), number of tutorials given per week (B= - .912, t= - 4.48, p < .05), number of times listening to the radio (B= - .46, t= - 3.87, p< .05), and father education (B=. 11, t= - 3.87, p< .05).

4.3.5.3 School Factors Predicting Overall Test Performance

Table 33: Relationship between School Variable and Pupils Overall Academic Achievement

	1	2	3
1. Overall academic achievement	1.00		
2. Absent in a semester	-.165*	1.00	
3. Time taking from school to home	-.113**	.075**	1.00

** Significant at .01 level

*Significant at .05 level

As the zero order correlation matrixes shows, the numbers of days pupils become absent from school and the time it takes to travel from home to school have statistically significant negative relationships with their test performance. However, the extent of relationship is found to be very weak. As the number of days the pupil become absent from school increases, the test performance decreases. On the other hand, as the distance of the school from the pupil's home and the time it takes to reach from home to school increases, the test performance decreases

Table 34: School Variable as Predictors of Academic Achievement

Variables	B	SE	T	R^2 change	Sig.
Absent in a semester	-2.297	.149	-15.460	.026	.000
Time taking from school to home	-1.233	.121	-10.195	.037	.000
Constant	**43.791**	**.240**	**182.510**		**.000**

The step-wise regression analysis in Table 34 above indicates that only 3.7 % of the pupils' variation in the overall test performance at grade 4 was explained for by the linear combination of variations on number of absents from school in a semester and the time spent on going from home to school. Number of absents in a semester (B=2.297, t= 15.460, $p < .05$) and time spent on going from home to school (B =-1.233, t= -10.195, $p < .05$) significantly predicted the overall test performance of pupils at grade 4.

4.3.6 Factors that predict academic achievement in key subjects

Table 35: Relationship between English Inputs and Pupils' Test Performance in English

	1	2	3	4
1. Academic achievement in English	1.00			
2. Home work given in English	.124*	1.00		
3. Pupils' understanding English	.166**	.063**	1.00	
4. Number of sharing text book in English	-.111**	-.059**	-.381**	1.00

** Significant at .01 level
*Significant at .05 level

As indicated in Table 35 above home work given in English, availability of text book in English, and pupils' understanding of English have significant positive relationships with the pupils' test performance in English. However, the degrees of relationships are considerably weak. As the data in the zero order correlation matrix below depicts, as the number of home works given for the English subject, and the pupils' understanding to English increase, so does their test performance in English. On the other hand, as the number of students who shared English text books increases, their test performance in English seems to show a decrease.

Table 36: English Inputs as Predictors of Academic Achievement

Variables	B	SE	T	R^2 change	Sig.
Understanding English	2.492	.205	12.176	.028	.000
Number of home work given in English	1.655	.166	9.988	.040	.000
Availability of text book in English	-.537	.120	-4.471	.042	.000
Constant	30.462	.972	31.333		.000

The step-wise regression analysis below indicates that only 4.2 % of the pupils variation on English test performance at grade 4 was explained by the linear combination of the variation on pupils' understanding English, number of home works given in a week ,and availability of textbook in English. Understanding English (B=2.492, t= 12.176, p < .000), number of home works given (B =1.655, t= 9.988, p < .000) and text book in English (B =-.537, t = -4.471 and p < .000) have significantly predicted the pupils achievement in the English test.

Table 37: Relationship between Mathematics Input and Pupils Test Performance Mathematics

	1	2	3	4
1. Academic achievement in Mathematics	1.00			
2. Home work given in Mathematics	.133**	1.00		
3. Students understanding of Mathematics	.169**	.050**	1.00	
4. Number of sharing text book in Mathematics	-.159*	-.015	-.439**	1.00

** Significant at .01 level

*Significant at .05 level

As indicated above in Table 37 the frequency of home work given in mathematics, and understanding Mathematics have shown significant positive relationships with the pupils test performance. Nevertheless, the relationship was found to be considerably weak. As the number of home works given and the level of pupils understanding in Mathematics increase, so does their test performance. On the other hand, sharing text book In Mathematics related negatively to academic achievement in mathematics. The increase in the number of pupils sharing Mathematics textbook shows a decreasing pattern in their test performance for mathematics.

Table 38: Mathematics Input as Predictors of Academic Achievement

Variables	B	SE	T	R^2 change	Sig.
Understanding Mathematics	1.886	.184	10.261	.030	.000
Home work given in Mathematics	1.624	.152	10.717	.043	.000
Availability of Mathematics text book	-.959	.111	-8.614	.052	.000
Constant	31.579	.896	35.244		.000

The step-wise regression analysis below indicates that only 5.2 % of the variation of grade 4 pupils' test performance in Mathematics was explained by the linear combination of the variations on understanding mathematics, number of home works given, and availability of textbook in Mathematics. Understanding Mathematics (B=1.886, t= 10.261, p < .000), number of home works given (B =1.624, t= 10.717, p < .000), availability of textbook (B=-.959, t= -8.614, p < .000) were able to significantly predict the test performance of pupils in Mathematics.

Table 39: Relationship between Environmental Sciences Input and Test Performance in Environmental Sciences

	1	2	3	4
1. Academic achievement in Environmental Sc.	1.00			
2. Home works given in Environmental Sc.	.084**	1.00		
3. Students understanding in Environmental Sc.	.207**	.077**	1.00	
4. Sharing text book in Environmental Sc.	-.224**	-.048**	-.466**	1.00

** Significant at .01 level

*Significant at .05 level

As indicated in Table 39 above home works given in Environmental Sciences and understanding Environmental Science had significant relationships with the pupils test performance. However, similar to the other key subjects, the relationship between the mentioned variables and test performance in Environmental Science was found to be considerably weak. As the number of home works given and the pupils' level of understanding Environmental Science increases, so does their test performance. On the other hand, sharing textbook in Environmental Science had negative relationship with academic achievement in Environmental Sciences. As the number of pupils who shared Environmental Science textbook increases, the pupils test achievement in Environmental Science decreases.

Table 40: Environmental Science Input as Predictors of Environmental Science Test Performance

Variables	B	SE	T	R^2 change	Sig.
Understanding environmental science	1.729	.169.	10.240	.055	.000
Number of home work given in environmental science	.615	.126	4.889	.068	.000
Availability of text book in environmental science	-1.570	.109	-14.3891	.070	.000
Constant	**42.156**	**.839**	**50.231**		**.000**

The stepwise regression analysis in Table 37 above indicates that 7.0% of the variation in pupils' test performance in environmental science was explained by the linear combination of the variation on understanding environmental science, number of home works given, and availability of textbooks in the subject. Understanding Environmental science (B=1.729, t= 10.240, $p < .000$), number of homework given (B =.615, t= 4.889, $p < .000$), availability of textbook (B=-1.570 t= -14.3891, $p < .000$) predicted test performance significantly.

4.3.7 Personal, Home and School Variables as a Function of Overall Academic Achievement

Different factors could influence pupils' academic achievement. In this study these factors were categorized as personal factors that are pertinent to the pupil, home factors and school factors. The following section describes the influence of these factors on academic achievement.

4.3.7.1 Personal Variable as a Function of Academic Achievement

In this study personal variables include the pupils' gender, language used at home, and reading additional text books.

Table 41: Personal Variables as a Function of Academic Achievement

Variable		Mean	Std. Dev.	F test/t test	Sig
Gender	boys	40.45	15.13178	2.94	.003
	girls	39.57	14.80759		
Language at home is different from instructional language	No	41.71	15.68784	10.183	.000
	Yes	38.60	13.93334		
Reading additional material	Never	36.1340	14.45200	153.499	.000
	Sometimes	42.7997	15.60326		
	Frequently	38.8211	13.77591		

As the data in Table 38 above shows, on average boys' test performance becomes better than that of the girls. The independent sample t test also indicate that there were a significant mean difference between boys and girls composite scores (t=2.94, p =.003). Pupil who responded that their language at home is different from instructional language performed below those who responded that language at home is similar to the instructional language. The independent sample t test also indicates that there was significant mean difference between pupils who responded that language at home is similar to instructional language and language at home is different from instructional language (t=10.18, p=.000). The average composite score of pupil who read additional materials sometimes becomes the highest (mean = 42.80), followed by those who read frequently (mean = 38.82), and those who do not read at all (mean = 36.13). Moreover, the test of one way ANOVA that assumes equality of variances has shown the presence of statistically significant mean differences of test performance among pupils who read additional material other than textbook frequently, sometime and never at all. In addition, the Scheffe test of homogeneity confirms results of the ANOVA test.

4.3.7.2 Home Variable as a Function of Academic Achievement

Table 42: Home Variables 1 as a Function of Pupil's Academic Achievement

Variable		Mean	Std. Dev.	F test	Sig
With whom you are living?	With my mother and my father	40.46	14.50	32.36	.000
	With my father only	35.54	14.51		
	With my mother only	39.87	15.93		
	With relatives	44.09	17.35		
	With others	35.32	14.29		
Family size	2	38.37	14.86	6.20	.000
	3	39.75	15.27		
	4	40.60	14.80		
	5	39.70	14.76		
	More than 5	40.93	15.06		
Father occupation	Farmer	40.08	14.77	14.03	.000
	Government employee	39.50	14.73		
	Merchant	39.42	14.66		
	Unemployed	36.36	13.79		
	I don't know	42.12	16.34		
	Other	42.98	16.03		
Mother occupation	Farmer	38.34	14.79	21.97	.000
	Government employee	38.19	14.54		
	Merchant	39.71	14.00		
	House wife	41.82	15.13		
	I don't know	38.73	14.46		
	Other	41.71	16.79		

For the items that probe "with whom you are living", pupils who are living with their relatives performed better than those who are living with their mother and father or with their mother or father only, followed by those who live with their mother and father. The one way ANOVA test also indicated statistically significant mean differences in test performance among those who are living within different family structures. The Scheffe test result depicted that there were statistically significant mean differences between those who are living with their relatives and the other groups , between students who are living with their mother and with their father only and the others, and between those who are living with their mother and father together, and father only and others .

For the items that ask family size, pupils who are living with a family size of 4 performed better than the other group. The ANOVA test indicated statistically significant mean differences of academic performance among those who are living within a different family size. The Scheffe test depicted that there were statistically significant mean differences between those who are living with a family size of 4 and above 5 and those who are living with family size of 2. The rest were not statistically significant.

Regarding the relationship between father's occupations test performance, pupils whose fathers are farmers performed better than the other group. The ANOVA test also indicated statistically significant mean differences in academic achievement among pupils whose fathers have different occupations. The Scheffe test depicted that there were a statistical significant mean differences between pupils whose father were farmers, government employee, merchant and between those whose father were unemployed . The rest were not statistically significant.

Regarding the relationship between mother's occupational status and academic achievement, those pupils who have housewife mothers performed better than the other groups. The ANOVA test also indicated statistically significant mean differences in academic achievement among pupils whose mother have different occupations. The Scheffe test depicted that there were statistically significant mean differences among pupils having the mother as a housewife.

Table 43: Home Variables 2 as Related to Pupil's Academic Achievement

Variable		Mean	Std. Dev.	F test	Sig.
When do you help your family with chores?	Sometimes absent from school	35.98	13.95	86.08	.000
	Always after school	43.09	15.73		
	Sometimes after school	40.28	15.78		
	Always Sunday and Saturday	40.32	13.71		
	I don't help	33.41	11.93		
How many family member attended school other than you?	One	40.17	15.08	7.75	.000
	Two	41.33	15.55		
	Three	40.95	14.81		
	More than three	39.67	14.62		
	No one	38.96	14.51		
Father's educational level	Illiterate	39.80	15.18	12.79	.000
	Write and read	41.08	15.16		
	1-8	39.43	14.11		
	9-12	38.27	13.93		
	Above 12 grade	42.25	16.48		
Mother educational level	Illiterate	40.74	15.09	12.477	.000
	Write and read	40.13	15.48		
	1-8	39.56	14.51		
	9-12	37.29	13.71		
	Above 12 grade	41.19	15.72		

Table 40 above shows data for some of the home variables that may have potential influences on pupils' test performances. When we examine test performance by family support, pupil who responded "I help my family always after school" performed better than those who help their mother sometimes after school, pupils who help always on Sunday and Saturday, pupils who help sometimes being absent from school and those who responded I don't help. The ANOVA test also indicated statistically significant mean differences on academic achievement on family support. The Scheffe depicted that there were a statistical significant mean differences between those who helped their family always after school and other groups, between those who helped sometimes after school and Sunday and Saturday, and between students who helped their family being absent and not at all.

For the relationship between other family members attending school and academic performance of pupils, those who responded two family members scored higher in overall academic achievement followed by those who responded three and one. The ANOVA test also indicated statistically significant mean differences on academic achievement on number of family members attending school other than the respondents. The Scheffe test depicted that there were statistically significant mean differences between those who had one family member attending school and those who had more than three and none, and there were statistically significant mean differences between those who had three family members attending school and those who had none.

Pupils whom their fathers had above 12 grade education scored more than the other groups followed by those who had fathers who could only read and write. The ANOVA teat also indicated statistically significant mean differences on academic achievement on father educational level. The Scheffe test depicted that there were statistically significant mean differences between pupils whose fathers are above 12 grade and the rest of the groups.

Pupils who had mothers with above 12 grade education scored more than the other group followed by pupils who had illiterate mothers and who could only read and write. The ANOVA test also indicated statistically significant mean differences on academic achievement for the mothers' educational level. The Schaffer test depicted that there were statistically significant mean differences between pupils whose mothers were above 12 grade education, illiterate, could read and write and those whose mothers are from grade 1-8, and from grade 9 and 12.

Table 44: Home Variable 3 as Related to Pupil's Academic Achievement

Variable		Mean	Std. Dev.	F test / t test	Sig.
Reading additional material other than text books	Never	36.13	14.45	113.85	.000
	Sometimes	42.80	15.60		
	Frequently	38.82	13.78		
Home tutor	No	38.00	15.26	-5.85	.000
	Yes	40.90	14.78		

Variable		Mean	Std. Dev.	F test / t test	Sig.
How many times tutored per week?	1-3 days	41.40	15.15	31.25	.000
	4-6 days	38.27	14.36		
	7 days	39.7	14.35		
Radio available	Yes	38.45	14.46	-8.79	.000
	No	41.29	15.14		
How many times you listen radio in a week?	Don't listen	40.31	14.81	16.09	.000
	1 -2 days	41.09	14.94		
	3 to 4 days	41.22	15.57		
	5 to 6 days	37.06	14.37		
	All days	39.79	14.73		
TV available	Yes	40.20	14.67	-3.00	.003
	No	41.17	15.52		
How many times watch TV per week?	Don't watch	41.30	15.03	36.35	.000
	1 -2 days	37.78	13.74		
	3 to 4 days	37.69	14.51		
	5 to 6 days	37.07	14.80		
	All days	42.18	15.57		
How many times you eat per day?	2 times	36.93	14.33	86.38	.000
	3 times	41.76	15.07		
	More than 3 times	39.95	14.89		

As the data in Table 41 shows, pupils who have the experience of reading additional reading materials some of the times relatively attained higher test scores than those who read additional reading material frequently, and those who do not read at all. The ANOVA test also indicated statistically significant mean differences on academic achievement because of reading additional material. The Scheffe test depicted that there were statistically significant mean differences between pupils who read additional reading material for some time (occasionally) and frequently, and there were also statistically significant mean differences between those who read additional reading material frequently and none at all.

Pupils who have the chance of tutorial in their study have achieved better scores than those who do not have the chance of tutorial while studying. Moreover, the independent sample mean difference test (t test) shows a statistically significant mean difference of pupils' test scores between those who had tutors while studying and those without.

On the other hand, the frequency of tutorial explains variation on the pupils test performance. For instance, pupil who have been tutored between 1-3 days a week showed a better test performance than those who have been tutored for 4 to 6 days, and 7 days a week. The ANOVA test also indicated statistically significant mean differences of test performance because of the number of days the pupils tutored in a week. In addition, the Scheffe test of significance depicted the presence of statistically significant mean differences between those who were tutored for 1 to 3 days, 4 to 6 days, and 7 days in a week.

Students who do not have a radio performed better than students who had a radio. The independent sample t test also depicted that there was a statistically significant mean difference on overall test performance between those who had a radio and who do not have a radio.

Pupils who have the habit of listening to the radio 3 to 4 days a week and 1 to 2 days a week showed a relatively better test performance than those who do not have the habit of listening to the radio, and those who listen 5 to 6 days a week. The ANOVA test also indicated statistically significant mean differences of academic achievement on number of days pupils listen to the radio in a week. The Scheffe test also depicted the presence of statistically significant mean differences among pupils who listen to the radio for 5 to 6 days and the other groupings who listen for all days in a week, 3 to 4 days a week, 1 to 2 days, and not at all.

Pupil whom there is a TV set at their home generally have better test scores than those who did not have. The independent sample mean test (t test) also indicated the presence of statistically significant mean difference on overall test achievement between those who have a TV set at their home and those who have not.

Pupil who watch TV all days and those who do not watch at all showed a better test performance than those who watch TV for 1 to 2 days, 3 to 4 days, and 5 to 6 days a week. The ANOVA test also indicated statistically significant mean differences of academic achievement for the number of days the pupil were watching TV in a week time. Similarly, the Scheffe test of significance for homogenous subset grouping showed statistically significant mean differences between those who watch TV for all days and those who do not watch at all, and the other groups who watch TV for 1 to 2 days, 3 to 4 days, and 5 to 6 days in a week.

With respect to the relationship of number of meals a day and performance on test, those who eat 3 meals a day have got better scores than those who eat 2 meals a day and, more than three meals a day. The ANOVA test also indicated statistically significant mean differences of academic achievement for the number of meals students eat in a day. The Scheffe test of significance also reveals statistically significant mean differences between students who have 3 meals a day and those groups who have 2 meals a day, and more than three meals a day.

4.3.7.3 School Variables as a Function of Overall Academic Achievement

Table 45: School Variables as a Function of Overall Academic Achievement

Variable		Mean	Std. Dev.	F test	Sig
How long does it take from home to school?	15 minutes at most	41.31	14.96	36.14	.000
	20 to 30 minute	40.67	15.25		
	30 to 60 minute	40.05	15.41		
	1 hour to 1:30 minute	36.81	13.58		
	More than 1:30 minute	35.11	12.90		
Absent in a semester	None	42.38	15.45		
	1 to 3 days	40.79	14.76	99.16	.000
	4 to 6 days	35.96	13.38		
	7 to 10 days	35.85	13.66		
School attended before joining first grade	Church school	40.39	15.40	43.79	.000
	Quran school	37.07	13.46		
	Kindergarten	41.71	14.76		
	Attended none	40.94	15.51		

As shown in Table 42 above, the distance pupil travel from their home to school has relationship with their test performance. Pupils who travel only for 15 minutes from their home to school have shown a better test performance than those who travel for 20 to 30 minutes, 30 to 60 minutes , 1 hour to 1:30 minutes. The ANOVA test also indicated statistically significant mean differences of academic achievement for the length of time the pupil travel to reach to school. The Scheffe test of subset grouping also depicted that there were statistically significant mean differences between those groups who travel for 1 hour to 1:30, and 30 to 60 minutes and those groups who travel for 15 minutes at most, 20 to 30 minutes, 30 to 60 minutes respectively.

With respect to the relationship of absentees from school and test performance, those pupil who were not absent from school even for a day showed a better performance than those who were absent from 1 to 3 days, 4 to 6 days and 7 to 10 days in a semester. The ANOVA test also indicated statistically significant mean differences of academic achievement based on the number of days pupils are absent from school in a semester. The Scheffe test of subset grouping also depicted that there were statistically significant mean differences between those who attended school regularly and those who were absent from 1 to 3 days, 4 to 6 days and 7 to 10 days. There were also statistically significant mean differences on the academic achievement of pupils who were absent for 1 to 3 days and those pupils who were absent for 4 to 6 days and 7 to 10 days in a semester. In general, the number of days pupils became absent from school showed a negative relationship to their test performance.

The type of pre-school attended by the pupil has relationship to academic performance. Pupils who attended Kindergarten as their pre- school experience have shown a relatively better performance than those who have attended other forms of preschool education. The ANOVA test also indicated statistically significant mean differences of test performance for the type school attended before joining the first grade. The Scheffe test of homogenous subset grouping also disclosed the presence of statistically significant mean differences between those who attended kindergarten and those who attended other forms of preschool education.

4.3.8 Other School Variables as a Function of Test Performance in Key Subjects

This sub section intends describing the relationship of school variables and test performance of pupil at grade 4 in key subjects. The main variables described are the frequency of home works for each key subject, availability of text book, and the pupils' perception whether they can understand lessons in the key subjects tested in the 4th national learning assessment.

English Test Score

Table 46: School Variables as a Function of English Test Score

Variable		Mean	Std. Dev.	F test	Sig
Home work given in English in a semester	None	31.47	15.69	44.7	.000
	Only once	34.92	17.72		
	Twice	37.67	18.48		
	Three times	40.43	18.33		
	Above three times	40.17	18.82		
Text book in English	Don't have	36.70	17.77	34.15	.000
	One for me	41.77	19.54		
	Shared for two	38.38	18.15		
	Shared for three	39.19	18.61		
	Shared for four	41.14	19.21		
Understanding English	I don't understand	37.10	18.50	82.71	.000
	Slight understanding	36.20	17.31		
	Average	39.71	18.26		
	Very well	45.50	20.15		

English test performance of pupils has considerably related to the frequency of home works given. As the data reports, pupil whom their teachers given them home works three times and above have shown better performances than those whom their teachers given them home works less than three times and none at all.

Furthermore, the one way ANOVA test confirmed the presence of statistically significant mean differences of test performance for the number of times home works given by teachers. The Scheffe test of homogenous subset grouping also showed the presence of statistically significant mean differences between pupils who were given home works three times and above, and three times, and those pupils who were given home works only twice, once, or not at all. Furthermore, the Scheffe test of subset grouping revealed statistically significant difference of test performance between those pupils whom their teachers given them home works at least twice a week and those who were given home works only once a week, and none of the time.

Regarding the relationship of text book possession in English and pupils test performance, the data in Table 43 above shows the presence of significant relationships between the two. For instance, pupils who shared their English text book with no one else have achieved better marks in English test than those who shared their English text book with other pupils, and to those who did not possess English text book at all either for their own or in a form of sharing with others. The one way ANOVA test also indicated statistically significant mean differences of English test performance for possession of English text book. The Scheffe test of subset grouping confirmed statistically significant mean differences between those who have a text book for their own and those pupils who shared for two and possess no text book at all.

One of the other school variable related to the pupils test performance is whether they understand English lessons well. Pupils who reported that they understood English well scored higher in the English test than those who reported average understanding, slight understanding and not at all. The one way ANOVA test also indicated statistically significant mean differences of English test performance for understanding English. The Scheffe test of subset grouping also confirmed that there were statistically significant mean differences between those who have a very good understanding of English lessons and the rest of the groups.

Among the school variables in the study, only portion coverage was found to have a significantly positive relationship with the pupils test performance in English ($r= .102$, $p=.05$).

Math Test Score

Table 47: School Variables as a Function of Mathematics Test Score

Variable		Mean	Std. Dev.	F test	Sig.
Home work given in Mathematics in a semester	None	29.67	13.55	53.20	.000
	Only once	32.87	14.10		
	Twice	36.64	17.14		
	Three times	38.20	16.44		
	Above three times	38.29	16.68		
Text book in Mathematics	Don't have	34.58	15.58	64.54	.000
	One for me	40.72	17.63		
	Shared for two	36.79	15.88		
	Shared for three	34.26	13.83		
	Shared for four	36.42	16.67		
Understanding Mathematics	I don't understand	34.67	16.11	73.01	.000
	Slight understanding	34.67	15.36		
	Average	37.69	16.41		
	Very well	41.90	17.58		

Pupils whom their Mathematics teacher give them home works for three times and above in a week achieved better than those whom their teacher given them home works only twice, once, or not at all. The one way ANOVA test also indicated statistically significant mean differences of test performance in Mathematics for the number of home works given in a week. The Scheffe test of subset grouping also depicted that there were statistically significant mean differences between those who were given home works three times and more and twice and those pupils who were given only once and none at all. The Scheffe test also depicted that pupils who were given home works once significantly differ in their Mathematics achievement from those who were not given home works at all**.**

Regarding the relationship of text book possession in Mathematics and pupils test performance, the data in Table 44 above shows the presence of significant relationships between the two. For instance, pupils who shared their Mathematics textbook with no one else have achieved better marks in Mathematics test than those who shared their textbook with other pupils, and to those who did not possess Mathematics textbook at all either for their own or in a form of sharing with others. The one-way ANOVA test also indicated statistically significant mean differences of Mathematics test performance for possession of textbook. The Scheffe test of subset grouping confirmed statistically significant mean differences between those who have a textbook without sharing and those pupils who shared for two, three, and four and possess no textbook at all.

Similar to the English test result, one of the school variables related to the pupils test performance in Mathematics is whether they understand the lessons very well. Pupils who reported that they understood Mathematics well scored higher than those who reported average understanding, slight understanding, and no understanding at all. The one-way ANOVA test also indicated statistically significant mean differences of Mathematics test performance for understanding lessons. The Scheffe test of subset grouping also confirmed that there were statistically significant mean differences between those who have a very good understanding of Mathematics lessons and the rest of the groups.

Environmental Science Test Score

Table 48: School Variables as a Function of Environmental Science Test Score

Variable		Mean	Std. Dev.	F test	Sig
Home work given in environmental sciences in a semester	None	37.69	15.51	23.23	.000
	Only once	41.18	16.18		
	Twice	41.46	16.00		
	Three times	41.00	15.70		
	Above three times	42.93	16.38		
Text book in environmental sciences	Don't have	38.18	14.72	115.96	.000
	One for me	46.27	17.17		
	Shared for two	43.50	15.68		
	Shared for three	39.92	16.02		
	Shared for four	39.61	14.41		
Understanding environmental sciences	I don't understand	38.30	14.48	109.12	.000
	Slight understanding	38.10	15.58		
	Average	41.65	16.17		
	Very well	46.30	16.65		

The same to the English and Mathematics test results, pupils whom their environmental science teacher give them home works for three times and above in a week achieved better than those whom their teacher given them home works only twice, once, or not at all. The one-way ANOVA test also indicated statistically significant mean differences of test performance in environmental science for the number of home works given in a week. The Scheffe test of subset grouping also depicted the presence of statistically significant mean differences between those who were given home works three times and more, twice, once and three times significantly differ from those pupils who were given home works not at all.

Regarding the relationship of text book possession in environmental science and pupils test performance, the data in Table 45 above shows the presence of significant relationships between the two. For instance, pupils who shared their environmental science textbook with no one else have achieved better marks in the test than those who shared their textbook with other pupils, and to those who did not possess environmental science textbook at all either for their own or in a form of sharing with others. The one-way ANOVA test also indicated statistically significant mean differences of environmental science test performance for possession of textbook. The Scheffe test of subset grouping confirmed statistically significant mean differences between those who have textbook without sharing and those pupils who shared for two and possess no textbook at all.

One of the school variables related to the pupils test performance in environmental science is whether they understand the lessons very well. Pupils who reported that they understood environmental science well scored higher in the test than those who reported average understanding, slight understanding, and no understanding at all. The one-way ANOVA test also indicated statistically significant mean differences of environmental science test performance for understanding lessons. The Scheffe test of subset grouping also confirmed that there were statistically significant mean differences between those who have a very good understanding of environmental science lessons and the rest of the groups.

4.3.9 Teacher Variables

4.3.9.1 Background Variables of Grade 4 Teachers

The data from the 783 grade 4 teachers who were involved in the 4th national learning assessment reveals that 487 (62.7%) of teachers at grade 4 were males and the remaining 288 (37.1%) were females. Regarding their age, 67.9 % are from 18-30 years old, 18.8 % are from 31 to 45 years old, and 5.6 % are 46 years old and above. Concerning grade 4 teachers qualifications 47.1% were certificate holders, 41.8% were diploma holders and 8.6 % were first-degree holders. When it comes to the years of teaching experience, 47.0% had 1 to 5 years of teaching experiences, 28.1 % had 6 to 10 years of teaching experiences, and 7.5 % had 11 to 15 years of teaching experiences.

4.3.9.2 Reasons for Joining and Professional Satisfaction of Grade 4 Teachers

Table 49: Reasons for Teacher Dissatisfaction

Reason for dissatisfaction	Yes	No
Low emphasis and prestige for the profession	125 (36.3%)	219(63.7 %)
No reinforcement/reward	195(59.5%)	133(40.5%)
No professional advancement and training	154(48.7%)	162(51.3%)
Low school service	103(33.4%)	205(66.6%)
Students misbehavior	124(39.5%)	190(60.5%)
Other	28(82.4%)	6(17.6%)

For a questionnaire items that aimed to elicit the main reasons the teachers joined the profession, majority of them 552 (70.5%) responded that they joined the profession because they were interested in the profession. Others, 106 (13.5 %) responded that they joined the profession as means of living as they failed to get other opportunities. The remaining, 113 (14.4 %) of them reported joining the profession by coincidence.

With regard to an item that probe professional/teaching satisfaction, 75 (9.6%) said that they did not get any satisfaction from the profession, 287 (36.7%) said that they were somehow satisfied being in the teaching profession. On the other hand, above half of them 415 (53 %) reported as they had a high satisfaction with the teaching profession.

For those items that probe reasons for teachers dissatisfaction towards the profession, about a quarter of them 195 (59.5 %) said that “there are no reinforcements/rewards”, followed by no professional advancement and training 154 (48.7%) , low emphasis and prestige for the profession 125 (36.3 %) , students’ misbehavior at school 124 (39.5 %) , and low school services 103 (33.4 %). Furthermore, 221 (29.7 %) of the respondents replied that they want to leave the profession while 523 (70.3 %) replied that they did not want to leave the profession. See Table 45 for details.

4.3.9.3 Teacher Related Variable that Influence Teaching Learning Process

With respect to the distance teachers’ travel from their home to school, about 323 (41.0%) of grade 4 teachers from the sampled schools reported to travel fewer than 15 minutes. Above a quarter, 230 (29.5%) of them testified traveling between 15 to 30 minutes to reach to school. Others, 108 (13.9 %) travel between 30 to 60 minutes to reach to school. Significant proportion of them i.e. 120 (15.3 %) evidenced traveling more than an hour to reach to school.

The teachers’ teaching workload from the sampled schools ranges from 16 to 20 periods a week for about 254 (32.7%) of them; and 26 to 30 periods a week for 240 (30.9 %) of them. About 120 (15.5%) of the teachers had workloads ranging between 21 to 25 periods; and the rest 10% reported to teach more than 30 periods a week. On the other hand, only 10% of the teachers said that they had fewer than 15 periods in a week time.

Table 50: Rating of Teachers Perceived Difficult Task

Task	Yes	No
Preparing a lesson plan	119 (17.8 %)	551 (82.2%)
Presenting the lesson	87 (13.2%)	571 (86.8 %)
Assessing the student and preparing exam	86 (13.1 %)	572 (86.9 %)

When teachers asked to identify their perception of a difficult task in relation to teaching learning process, about 119 (17.8 %) of them rated the preparation of lesson plans as the most difficult one. About equivalent proportions of teachers (i.e., 13.2% of them) reported presentation of lessons, and assessing student work and preparation of exams as most difficult tasks in relation to teaching learning.

Classroom observation and supervision of teachers' activities has contribution for the improvement of instruction. Based on this, teachers from the sampled schools reported different frequency of supervisions by school directors or supervisors. For instance, about 147 (19%) of them were observed only once in a semester. Others, i.e. 254 (51.8 %) of them evidenced at least twice classroom supervision in a semester. A considerable number of teachers i.e. 316 (40.8%) reported to have been observed and supervised for three and more than three times a semester. To the contrary, 57(7.4%) of grade 4 teachers replied that they were not observed and supervised while they are teaching in classes.

There is an acknowledgement on the contribution of teacher – parent discussions on matters of pupils learning and behavior development. The teacher data of the 4th national learning assessment at grade 4 level revealed that about 409 (53.1 %) of the teachers had discussions with pupils' parents more than three times in a semester. Whereas 155 (20.1%) of them had discussed with parents twice a semester. 73 (9.5 %) of them had discussion with pupils' parents only once in a semester. The remaining 32 (4.2%) of the teachers reported no discussion with parents in a semester time.

Participation of teachers in skills development programs revealed that 312 (51.7 %) of them had attended training on methods of teaching, about 150 (24.8 %) on curriculum issues. Whereas, about 90 (14.9 %), 31(5.1%), and 21 (3.5%) of them had attended training on contents such as student assessment, classroom management, and students' behavior respectively.

With respect to the subjects they teach at their school, 327 (48.5 %) of the teachers reported that they teach all types of subjects. About 117 (17.4 %) of them were teaching English, 102 (15.1 %) were teaching math, 103 (15.3 %) were teaching environmental sciences.

In addition to this, the teachers reported the number of years they thought the subjects mentioned. Thus, 284 (38.3 %) thought more than five years, 107 (14.4 %) of them for four years and 81 (10.9 %) taught the subject for about three years. 133 (17.9 %) responded that they taught the course for about a year and two years.

4.3.10 Availability of Teaching Material in School as Perceived by Teachers

Table 51: Teaching Material for Mother Tongue

	Yes	No
Syllabus	243 (46.4 %)	281 (53.6 %)
Teachers guide	266 (50 %)	266 (50 %)
Students text book	368 (70.1 %)	157 (29.9 %)

Regarding teaching material for mother tongue as indicated in Table 48 above , 243 (46.4 %) teachers reported that they get syllabus from school, while 281 (53.6 %) did not get a syllabus. With respect to teacher guide, 266 (50 %) reported that they have got it from school, while the remaining 266 (50 %) did not get it. Similarly, about 368 (70.1%) of the teachers reported that they have got text book for the subjects they teach from school, while 157 (29.9 %) of them said that they did not get text book in mother tongue.

Table 52: English Teaching Materials

	Yes	No
Syllabus	184 (41.3 %)	262 (58.7 %)
Teachers guide	244 (53 %)	216 (47 %)
Students text book	315 (69.2 %)	140(30.8 %)

The proportions of grade 4 teachers who have accessed English syllabus were only 184 (41.3%). Above half of the teachers, i.e. 262 (58.7 %) did not access English syllabus to support their teaching activity. About 244 (53%) teachers confirmed that they have got teacher guide from school, while the remaining 216 (47%) did not get a teacher guide in English. For access to English textbook, 315 (69.2%) of them reported that they have got text book from school, while about 140 (30.8%) of them did not get text book in English as revealed in Table 48.

Table 53: Mathematics Teaching Materials

	Yes	No
Syllabus	180 (41.0 %)	259 (59.0%)
Teachers guide	214 (46.0 %)	243 (53.2%)
Students text book	214 (46.8 %)	243 (53.2%)

As the data in Table 50 above shows, the proportions of grade 4 teachers who have accessed Mathematics syllabus were only 180 (41.0 %). Above half of the teachers, i.e. 259 (59.0 %) did not access Mathematics syllabus to support their teaching activity. About 214 (46 %) teachers confirmed that they have got Mathematics teacher guide from school, while the remaining 243 (53.2 %) did not get Mathematics teacher guide. In accessing Mathematics textbook, 214 (46.8 %) of them reported that they have got text book from school, while about 243 (53.2%) of them have not got Mathematics text book.

Table 54: Environmental Sciences Teaching Material

	Yes	No
Syllabus	182 (41.6 %)	255 (58.4%)
Teachers guide	220 (47.9 %)	239 (52.1%)
Students text book	354 (75.3%)	116 (24.7%)

As the data in Table 51 above shows, the proportions of grade 4 teachers who have accessed environmental science syllabus were only 182 (41.6 %). Above half of them, i.e. 255 (58.4%) did not access environmental science syllabus to support their teaching activity. About 220 (47.9 %) teachers confirmed that they have got teacher guide for the subject from their school, while above half of them, 239 (52.1%) did not get teacher guide for environmental science subject. However above a quarter of the teachers, 354 (75.3%) were able to access textbook in environmental science from their school, while the remaining 116 (24.7%) of the teachers were not able to access text book.

Table 55: Sharing of Books among Pupils

Subject	One to one	One for two	One for three	One for four and more than four
Mother tongue	298 (58.8%)	81 (10.3%)	31(4.0%)	97 (12.4%)
English	236(45.2%)	103(19.7%)	53(10.2%)	130(16.6%)
Mathematics	284(53.3%)	92(17.3%)	51(9.6%)	106(13.5%)
Environmental Sciences	313 (55.9%)	79(14.1%)	50(8.9%)	118(21.1%)

Furthermore, pupils' access to textbooks in the key subjects tested as reported by the teachers was presented in Table 52 above. As seen in the data, from a minimum of 236(45.2%) of teachers in English to a maximum of 298 (58.8%) of teachers in mother tongue witnessed that pupils possess text books for their own. On the other hand, on average 10% to 20% of teachers witnessed that pupils share text books in the key subjects for two, three, and four persons accordingly.

Table 56: Portion Coverage by Teachers

Subject	<60%	60% - 75%	76% - 90%	>90 %
Mother tongue	63 (13.0%)	52 (10.7%)	161 (33.3 %)	208 (43.0%)
English	74 (14.8%)	65 (13.0%)	166 (33.3 %)	194(38.9 %)
Math	56(11.1%)	66(13.1%)	198(39.3%)	184(36.5 %)
Environmental sciences	52 (10.3%)	43(8.5 %)	187(36.9 %)	225(55.6%)

With regard to portion coverage, 55.6 % of teachers responded that above 90 % of the portion in Environmental sciences was covered while 43.0 %, 38.9 % and 36.5 % of teachers covered the portion above 90 % in Mother tongue, English and math respectively. On the other hand, more than 10 % of grade 4 teachers responded that the portion covered in all subjects were below 60 %. Please consult Table 53 above.

4.3.11 Correlation between Teacher Variables and Pupils Test Performance

Table 57: The Multiple Regression Analysis

Item	Pearson r	Sig.
1. Background Variable		
Qualification	.124	.005
Experience in teaching	-.035	>.05
2. Reason for dissatisfaction		
Satisfaction	.074	>.05
Low emphasis and prestige for the profession	-.049	>.05
No reinforcement/reward	.049	>.05
No professional advancement and training	-.074	>.05
Low school service	-.104	>.05
Students misbehavior	.014	>.05
3. Leaving the profession	.011	>.05
4. Teaching variable		
Teaching load in a semester	.073	>.05
Time taking to reach school	.053	>.05
Presenting the lesson	.073	>.05
Lesson plan preparation	.017	>.05
Assessment	.033	>.05
5. Discussion with parents and Supervisory Support		
Supervisors observation	-.035	>.05
Discussion with parents	.125	.004

Out of several teacher variables included in the study only two had statistically significant association with the pupils test performance. The qualification of teachers and their discussion with parents had significant positive correlation with the pupils test performance even though the correlation coefficients are considerably low. The higher the qualification of teachers, the better was the pupils test performance. Moreover, the more frequent teachers discuss with parents, the better the pupils test performance. However, several of the teacher variables included in the study did not have statistically significant correlation with the pupils' performance

Table 58: Teacher Variable as Predictors of Overall Academic Achievement

Variables	B	SE	T	Sig.	R^2 change
Experience in teaching	-2.026	.631	-3.214	.000	.042
School service	-3.286	1.434	-2.292	.000	.068
Constant	47.957	2.574	18.635		

As shown in Table 55 above, the step wise regression analysis showed an approximately 7% of pupils performance differences explained for by the liner combination of experience in teaching and school service.

4.3.12 Correlation Between School Input as Perceived by Teachers and Pupils Test Performance in Key Subjects

Table 59: Correlation between School Input in English and Achievement in English

School Input	Pearson r	Sig.
Syllabus	-.120	.029
Teachers guide	.011	.843
Students text book	.019	.730
Sharing of books among students	-.064	.203
Portion coverage	.102	.053

From the five school variables, the only variables that showed significant positive relationship with the pupils test performance in English was the availability of syllabus in English.

Table 60: School Variables as Predictors of English Achievement

Variables	B	SE	T	Sig.	R2 change
Portion coverage	2.135	.829	2.574	.011	.022

The step wise regression analysis shows 2.2% variation of the English test performance was accounted for by the variation in portion coverage (See Table 57).

Table 61: Correlation between School Input in Mathematics and Achievement in Mathematics

School Input	Pearson r	Sig.
Syllabus	-.011	.843
Teachers guide	.158	.003
Students text book	.059	.273
Sharing of books among students	-.028	.573
Portion coverage	-.026	.624

From the five school variables, the only variables that showed significant positive relationship with the pupils test performance in Mathematics was the availability of teachers guide in Mathematics (See Table 58).

Table 62: School Variables as Predictors of Mathematics Achievement

Variables	B	SE	t	Sig.	R^2 change
Teachers' guide	3.234	1.306	2.477	.014	.014
Sharing of books among students	-1.043	.527	-1.979	.049	.024
Constant	35.227	2.086	16.416	.000	

The step-wise regression analysis indicated that 2.4% of the variation in Mathematics test performance among pupils was explained for by linear combination of teacher's guide and sharing of books among students (see Table 59 above).

Table 63: Correlation between School Input in Environmental Science and Achievement in Environmental Science

School Input	Pearson r	Sig.
Syllabus	-.141	.011
Teachers guide	.071	.188
Students text book	.103	.054
sharing of books among students	-.196	.000
portion coverage	.019	718

Sharing of books among pupils, availability of teacher guide and portion coverage found to have a statistically significant relationship with the pupils academic achievement in environmental science (See Table 60 above).

Table 64: School Variables as Predictors of Environmental Science Achievement

Variables	B	SE	t	Sig.	R2 change
Sharing of books among students	-1.643	.524	-3.136	.002	.024
Teachers guide	3.639	1.323	2.750	.006	.039
Syllabus	-3.074	1.340	-2.294	.023	.053
Constant	46.321	2.735	16.937	.000	

The step-wise regression analysis conducted explained that approximately 5.3 % of the variation in pupils' test performance in environmental science attributed for the linear combination of sharing of books among students, availability of teacher's guide and syllabus (See Table 61 above).

4.4 Grade 8 Students' Academic Performance across Key Subjects and Overall Academic Performance

4.4.1 Summary of the Descriptive Statistics

Table 65: Summery of Statistics in Key subjects

Subjects	N	Mean (%)	Median (%)	Std. Dev.	Std. Error	T	df	Sig.	MD
English	11322	36.87	35	15.82	.15	-88.34	11321	.000	-13.13
Mathematics	11184	25.53	22.50	11.35	.11	-228.09	11183	.000	-24.47
Biology	11322	42.10	40	16.73	.16	-50.26	11321	.000	-7.90
Chemistry	11224	36.44	32.5	14.53	.14	-98.885	11223	.000	-13.56
Physics	11224	34.45	32.5	12.11	.11	-136.10	11223	.000	-15.55
Composite	9435	35.32	34.00	8.59	.09	-165.90	9434	.000	-14.67

Similar to results of the second and third national learning assessments, the mean composite score of the five key subjects was below the minimum expected average of (50%). The one sample t test also indicated that there were statistically significant mean differences between each subject and the minimum competency level (50%). The median composite score (34%) is less than the composite mean score (35.32%) by score points of 1.32%. This shows that 50% of grade 8 students who sat for the fourth national learning assessment tests have achieved 34% and below in the overall achievement (see Table 4 above).

As seen in the data, mean and median scores were relatively the least for Mathematics test, 25.53%, and 22.50% respectively. Compared to the previous national learning assessments of grade 8, particularly to the second and the third, average scores in the fourth national learning assessment are decreasing for the composite as well as the key subjects except for Physics. Physics mean score (34.45%) was slightly greater than the third national learning assessment result, which was 32.20%, and less than the result in the second national learning assessment result of 35.32%.

4.4.2 Performance Standard of Grade 8 Pupils in Each Subject

Subject	Below Basic	Basic	Proficient
English	55.1%	29.9%	15.1%
Mathematics	61.3%	24.0 %	14.7%
Biology	55.9%	27.3%	16.8%
Chemistry	58.4%	26.7%	14.9 %
Physics	53.8%	31.2%	15.0%
Composite	56.3%	27.6%	15.9%

Table 66: Performance Standard of Pupils in Each Subject

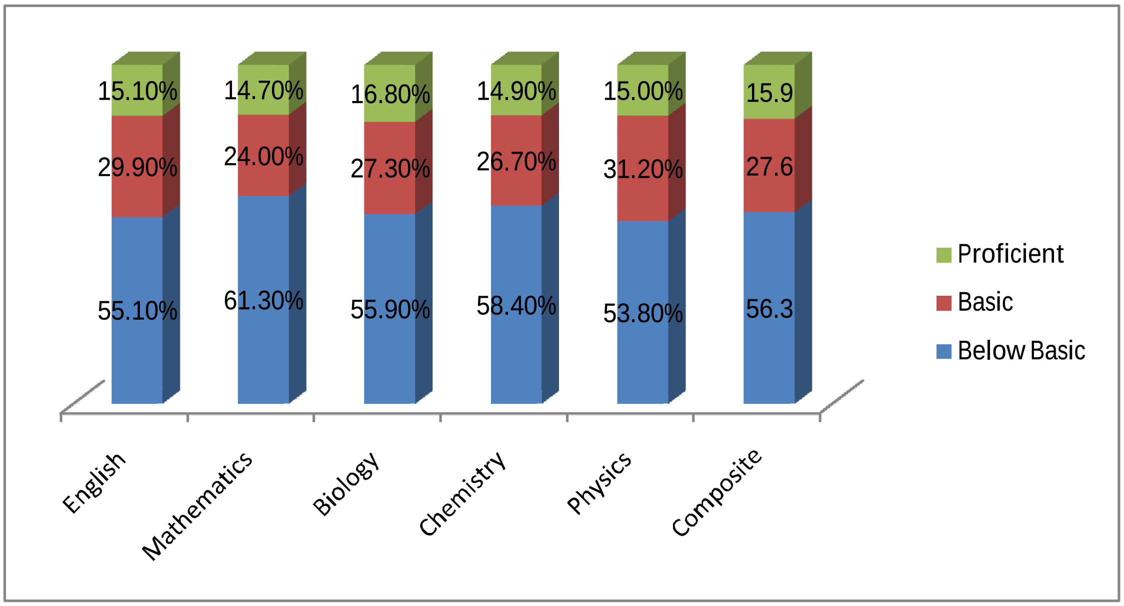


Figure 4: Performance Standard of Pupils in Each Subject

The performances of grade 8 pupils were categorized into three proficiency levels based on their test achievements. The categorization was made in reference to standard z – distributions, namely below basic, basic and proficient. The below basic category includes those students who fall at or below a z standard score of zero; the basic category is within a z standard score of zero and one standard deviation above the mean; and proficient category includes those students who fall above a z standard score of one standard deviation above the mean. Based on these categories, the proportion of pupils attaining at each level was reported for the composite score as well as for each key subject at the national level. For the composite average, 15.9% of the students' achievement is at the proficient level, which is a better performance when compared to the result of third national learning assessment, which was only 13.9%.

When compared to the third national learning assessment result, the proportion of grade 8 students attaining at the proficient level has shown a slight improvement for Chemistry (0.1%), Mathematics (0.2%), and Physics (0.8%), and slight decreases for English and Biology (with a proportion of 0.1% and 0.2%). The highest increase in the proportion of students at the proficient level was for the Physics test.

At the basic performance level, the proportion of students for the composite score increased by 3.6%. Moreover, slight declines observed in the proportion of pupils who have performed at the basic level for some of the key subjects. For instance, the decline in proportion of pupils became 0.4% in Physics, 2% in Mathematics, and 2.5% in Chemistry. On the other hand, for the English and Biology test, the proportion of students attaining at the basic level increased by 4.8% and 2.2% in comparison to the third national learning assessment results.

Finally at the below basic level, which is lower than the mean of the standard z score, the proportion of pupils performing at the level accounts for 56.3% with a decrease of 5.8% from that of 3rd national learning assessment. For the key subjects tested, the proportion of students who have performed at below basic level raised by 1.8% in Mathematics and 2.3% in Chemistry. In the rest of the subjects, the proportions of students who have performed at the below basic level has shown a decrease when compared to the third national learning assessment result.

In summary, the trend of students' performance at the different performance standards follows both an increase and a decrease for the different subjects and the various proficiency levels. For instance, performance standards improved for English and Biology test and a decrease for Mathematics, Chemistry, and Physics.

4.2.3. Range of Achievement Scores of grade 8 students at five Different marker points

Percentiles	Physics (%)	Chemistry (%)	Biology (%)	English (%)	Mathematics (%)	Average score(%)
10th	20	20	22.5	20	12.5	25.5
25th	25	25	30	25	17.5	29
50th	32.5	32.5	40	35	22.5	34
75th	40	45	52.5	47.5	30	40
90th	50	57.5	67.5	60	40	47

Table 67: Range of Achievement Scores at Five Percentile Ranks

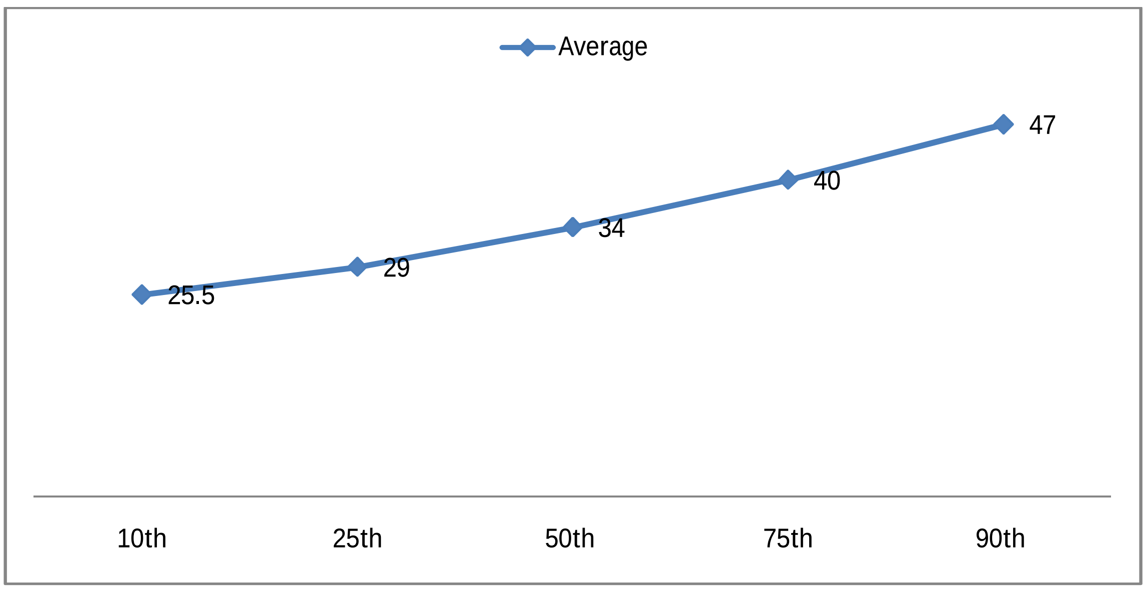


Figure 5: Range of overall achievement scores at five key markers

Table 6 shows the range of achievements in five percentile ranks for the respective key subjects tested. The data in the table depicts the cut-off scores for the 5 key subjects at the 10th, 25th, 50th, 75th, and 90th percentile. For instance, if we consider the 90th percentile for the average score, only 10% of the pupils could have achieved an average score of 47 and above. On the other hand, the 10th percentile shows that about 10% of the pupils scored an average of 25.5 and below. For further clarity, look at figure 2 above.

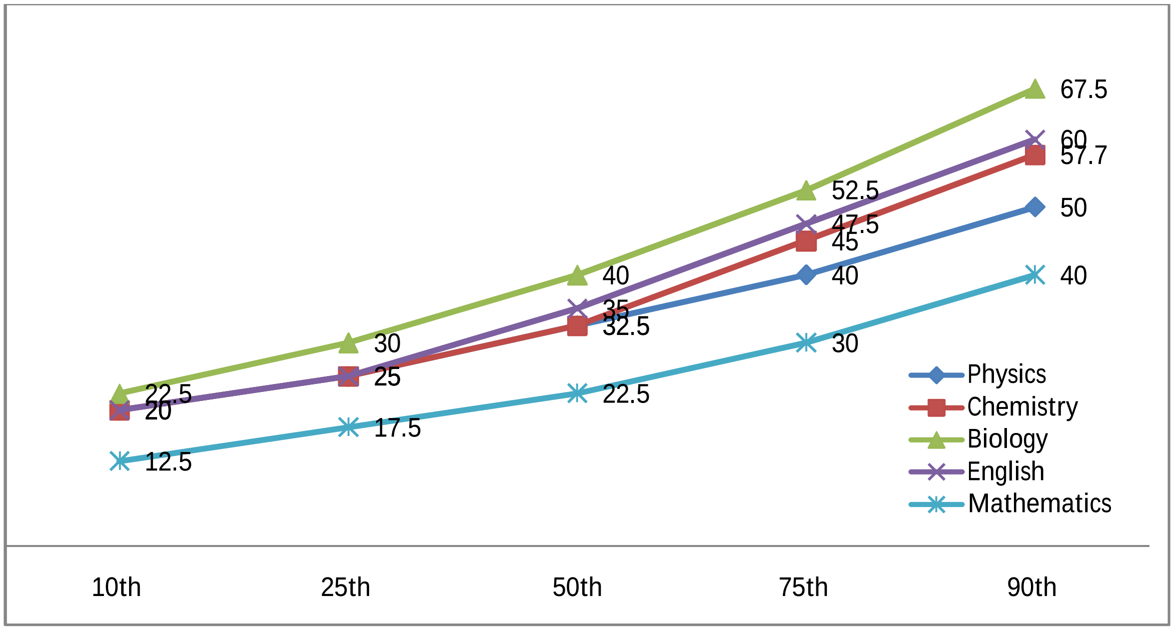


Figure 6: Range of achievement scores in the key subjects at five key marker points

With respect to key subjects tested, the highest percentile score was attained in Biology (i.e. about 10% of the students have achieved an average score of 67.5 in the Biology test). Whereas about 10% of the students have an average score of 22.5 and below in Biology test. Moreover, these percentile ranks indicate the presence of highest variation among pupils in Biology test achievement. With respect to English test achievement, about 10% of the pupils achieved a score of 20 and below. And another 10% have achieved a score point of 60 and above. In Chemistry, about 10% of the pupils achieved a score point of 20 and below. On the other hand 10% of the top performing pupils achieved a score point of 57.5 and above. The percentile ranking for Physics test achievement showed that about 10% of the pupils to have a maximum score of 20 and below that. On the other hand, about 10% of the top performing pupils achieved a minimum of 50 and above in Physics test achievement. Finally, when we examine the data in Table 6 above, about 10% of high performing pupils could have achieved a score of 40 and above in Mathematics, which is the least compared to the other key subjects. On the other hand 10% of the lowest performing pupils have achieved a maximum of 12.5 and below that, in mathematics test. For further clarity, look at figure 3 above.

4.2.4 Group Differences In Academic Performances across Subject and Overall Academic Achievement

4.2.4.1 Test Performance of Boys and Girls

Table 68: Performance of Boys and Girls in Each Subject, T value and Sig. Level

Subject	Gender	N	Mean (%)	Std. Error	Std Deviation	T	Df	Sig.	MD
English	boys	5198	38.69	.23	16.31	9.063	9506	.000	2.96
	girls	4310	35.73	.23	15.30				
Mathematics	boys	4653	25.88	.17	11.69	2.003	8523	.045	.50
	girls	3872	25.38	.18	11.30				
Biology	boys	5198	44.16	.25	17.69	11.163	9506	.000	3.85
	girls	4310	40.32	.24	15.46				
Chemistry	boys	5207	36.37	.20	14.54	-1.637	9518	.102	-.50
	girls	4313	36.87	.23	14.93				
Physics	boys	5207	36.19	.18	12.96	12.108	9518	.000	3.03
	girls	4313	33.16	.17	11.10				
Composite	boys	4461	36.36	.14	9.00	9.82	8184	.000	1.86
	girls	3725	34.50	.13	7.95				

Table 67 above shows performance variation between boys and girls. For the composite score, boys outperformed girls (with a difference of 1.86% score points). The mean difference found to be statistically significant. When it comes to the five key subjects tested, boys' performance became relatively higher for Mathematics, English, Biology, and Physics. Moreover, the mean differences are statistically significant. On the other hand, girls outperformed boys in Chemistry even though the mean difference was not statistically significant. In the third national learning assessment, statistically significant mean differences observed for all of the key subjects in favor of boys. Whilst in the fourth national learning assessment, statistically significant mean differences observed for all subjects except that of Chemistry. This indicates that the gender gap in academic achievement is still recurrent.

4.2.4.2 Test Performance of Urban and Rural Pupils

Table 69: Performance of Urban and Rural Pupils in Each Subject, T value, and Significance Level by Each Subject

Subject	Location	N	Mean (%)	Std. Error	Std Dev.	T	df	Sig.	MD (%)
Mathematics	Rural	4164	25.63	.18	11.34	-.003	8497	.998	-.0006
	Urban	4335	25.63	.18	11.72				
English	Rural	4758	35.35	.22	15.08	-10.988	9398	.000	-3.600
	Urban	4642	38.95	.24	16.66				
Biology	Rural	4758	41.68	.24	16.83	-3.663	9398	.000	-1.273
	Urban	4642	42.95	.25	16.88				
Chemistry	Rural	4643	37.12	.23	15.49	4.120	9419	.000	1.248
	Urban	4778	35.87	.20	13.89				
Physics	Rural	4643	34.26	.18	11.98	-3.179	9419	.001	-.8048
	Urban	4778	35.06	.19	12.58				
Composite	**Rural**	**4065**	**34.95**	**.13**	**8.52**	**-4.72**	**8144**	**.000**	**-.902**
	Urban	**4081**	**35.85**	**.14**	**8.74**				

When we examine test performance by location, urban students excelled rural students in the composite as well as in the key subjects tested. With respect to the subjects, urban students excelled rural students in all subjects except Chemistry. A large difference was observed for English (MD= 3.600 %, t =10.99, df= 9398 and p = .000). Except for Mathematics test result, mean differences between urban and rural students became all statistically significant. In contrast to the third national learning assessment where rural pupils outperformed the urban for the composite and the key subjects except English, results in the fourth national learning assessment reversed in favor of the urban pupils except for the subject of Chemistry.

4.2.4.3 Pupils Performance for the Composite Score across Regions

Table 70: Summery Statistics of Composite Scores by Region

Region	N	Mean(%)	SD	Std.Error
Tigray	954	38.67	8.98	.29
Afar	777	31.60	7.53	.27
Amhara	1114	35.47	8.67	.26
Oromiya	1067	36.54	8.36	.26
Somali	556	33.74	7.68	.33
Benshangul-Gumuz	804	34.16	7.96	.28
SNNPR	1047	36.93	8.67	.27
Gambela	802	31.15	6.70	.24
Harari	751	34.46	8.23	.30
Addis Ababa	747	38.36	9.03	.33
Dire Dawa	815	35.51	8.39	.29
Composite	9434	35.33	8.59	.09

The data in Table 69 above depicts the average performance of pupils from the different regions. Average performances were below 50% for students of all regions. The average performance of pupil from Tigray (i.e. 38.67%) is the highest of all other regions, followed by pupils' from Addis Ababa (i.e. 38.36%), and students' from SNNPR (i.e. 36.93%). On the other hand, average scores became the least for student from Gambella (i.e. 31.15%), Afar (31.60%), and Somali (33.74%) regions.

Table 71: ANOVA Summery for Overall Academic Achievement by Region

	Sum of squares	DF	Mean square	F ratio	Sig.
Sum square between	49723.057	10	4972.306	72.479	.000
Sum square within	646450.836	9423	68.604		
Sum square total	696173.893	9433			

Furthermore, the one-way analysis of variance (ANOVA) that assumes homogeneity of variances among the scores of students from the different regions revealed the presence of statistically significant mean differences ($F (10, 9423) = 72.479$, $p = .000$) for the overall academic performance of students across regions.

Table 72: Homogenous Subset Grouping for the Composite Score across Regions

Region	Number of pupils	Subset for alpha = 0.005					
		1	2	3	4	5	6
Gambela	802	31.15					
Afar	777	31.60					
Somali	556		33.74				
Benshangul Gumuz	804		34.16	34.16			
Harari	751		34.46	34.46			
Amhara	1114		35.47	35.47	35.47		
Dire Dawa	815			35.51	35.51		
Oromiya	1067				36.54		
SNNPR	1047				36.93	36.93	
Addis Ababa	747					38.36	38.36
Tigray	954						38.67
Significance		**1.000**	**.055**	**.359**	**.241**	**.258**	**1.000**

The Scheffe test of homogenous mean subset grouping used to identify which regions are statistically significantly different from one another. Test performance students from Addis Ababa and Tigray regions were found to be different from the other regions. It was also found out that students from Gambella and Afar regions were found to be least performer and statistically significantly different from other regions. For students' performance differences in other regions, see table 71 above.

4.2.4.4 Students' Performance for Key Subjects across Regions

English Academic Achievement

Table 73: **Students' Achievement in English by Region**

Region	N	Mean (%)	SD	Std. Error
Tigray	1040	38.60	17.23	.53
Afar	912	34.45	14.90	.49
Amhara	1233	36.62	14.93	.43
Oromiya	1627	33.45	14.08	.35
Somali	664	36.10	15.92	.62
Benshangul-Gumuz	890	35.35	14.60	.49
SNNPR	1308	39.54	15.90	.44
Gambela	944	31.13	12.71	.41
Harari	920	35.97	14.99	.49
Addis Ababa	945	44.55	17.71	.58
Dire Dawa	835	41.16	16.88	.58
Composite	**11318**	**36.86**	**15.82**	**.15**

Table 72 above shows grade 8 students' average performances for English test across the regions. The top three regions where grade 8 students performed better in English were Addis Ababa (with average score of 44.55%), Dire Dawa (with average score of 41.16%), and SNNPR (with average score of 39.54%). On the other hand, the average performance in English was relatively the least for pupils from Gambela (with average score of 31.13%), Oromia (with average score of 33.45%), and Afar (with average score of 34.45%). The average performance of all regions is below the 50% minimum achievement level average standard as indicated in the Education and Training Policy.

Table 74: ANOVA Summery for English Test Scores by Regions

	Sum of squares	DF	Mean square	F ratio	Sig.
Sum square between	142356.593	10	14235.659	59.865	.000
Sum square within	2688738.093	11307	237.794		
Sum square total	2831094.685	11317			

The test of one way ANOVA (see Table 73 above) for English test performances across regions revealed the presence of statistically significant mean differences (F (10, 11307) = 59.865, p= .000).

Table 75: Scheffe Test for English across Regions

Region	Number of pupils	Subset for alpha = 0.005						
		1	**2**	**3**	**4**	**5**	**6**	**7**
Gambela	944	31.13						
Oromia	1627	33.45	33.45					
Afar	912		34.45	34.45				
Benshangul Gumuz	890		35.35	35.35				
Harari	920		35.97	35.97	35.97			
Somali	664		36.10	36.10	36.10			
Amhara	1233			36.62	36.62	36.62		
Tigray	1040				38.60	38.60	38.60	
SNNPR	1308					39.54	39.54	
Dire Dawa	835						41.16	
Addis Ababa	945							44.55
Significance		**.353**	**.155**	**.476**	**.168**	**.063**	**.199**	**1.000**

In addition, based on the Scheffe test of subset grouping, the average English test performance of students from Addis Ababa has shown statistically significant difference from the rest of the regions. Students from Gambella have shown a statistically significant variation in English test performance from the rest of the regions except those students from Oromia. On the other hand, students from the five regions, namely Afar, Benshangul Gumuz, Harari, Somali, and Amhara became homogenous in their average score of English test (for further information see Table 74 above).

Mathematics Academic Achievement

Table 76: Average Scores of Regions for Mathematics

Region	N	Mean	SD	Std. Error
Tigray	995	25.97	10.53	.33
Afar	803	23.94	11.26	.40
Amhara	1160	25.55	11.40	.33
Oromia	1150	27.75	13.08	.39
Somali	551	22.77	9.55	.41
Benshangul-Gumuz	845	25.48	10.83	.37
SNNPR	1098	27.63	13.62	.41
Gambla	873	25.22	10.11	.34
Harari	773	23.39	10.85	.39
Addis Ababa	793	26.03	10.62	.38
Dire Dawa	835	25.86	10.64	.37
Composite	**9876**	**25.65**	**11.46**	**.12**

As shown in Table 75 above, the students average test performance in Mathematics is below 50% in all the regions, alarmingly half way to the set minimum achievement standard as indicated in the Education and Training Policy.

Table 77: ANOVA Summery for Mathematics Test Scores by Regions

	Sum of squares	Df	Mean square	F ratio	Sig.
Sum square between	20684.934	10	2068.493	15.999	.000
Sum square within	1275460.942	9865	129.292		
Sum square total	1296145.876	9875			

The ANOVA summary Table shown below indicates the existence of statistically significant variation (F (10, 9865) = 15.999, p < .001) in the Mathematics test performance of students across regions.

Table 78: Scheffe Test for Mathematics score

Region	Number of pupils	Subset for alpha = 0.005			
		1	2	3	4
Somali	551	22.77			
Harari	773	23.39	23.39		
Afar	803	23.94	23.94	23.94	
Gambela	873		25.22	25.22	
Benshangul-Gumuz	845		25.48	25.48	25.48
Amhara	1160		25.55	25.55	25.55
Dire Dawa	835			25.86	25.86
Tigray	995			25.97	25.97
Addis Ababa	793			26.03	26.03
SNNPR	1098				27.63
Oromia	1150				27.75
Sig.		**.919**	**.114**	**.152**	**.072**

Moreover, the Scheffe test of homogenous subset grouping revealed that students from Oromia and SNNPR have statistically significant mean differences on Mathematics achievement from students of Addis Ababa , Tigray , Dire Dawa, Amhara , and Bensahngul Gumuz regions. The Schefee test also revealed that students from Somali region have statistically significant mean differences on Mathematics achievement from students of other regions except students of Harari and Somali regions. For further students' performance differences in Mathematics subject in other regions, see table 77 above.

Biology Academic Achievement

Table 79: Average Scores of Regions for Biology

Region	N	Mean (%)	SD	Std. Error
Tigray	1040	55.24	16.51	.51
Afar	912	35.24	13.35	.44
Amhara	1233	39.16	16.22	.46
Oromia	1627	50.35	16.46	.41
Somali	664	39.43	11.68	.45
Benshangul–Gumuz	890	37.59	14.41	.48
SNNPR	1308	43.17	16.44	.45
Gambela	944	33.02	13.02	.42
Harari	920	37.33	15.18	.50
Addis Ababa	945	43.92	17.49	.57
Dire Dawa	835	40.15	15.37	.53
Composite	**11318**	**42.10**	**16.73**	**.16**

As shown in Table 78, the students' average test performance in Biology is below 50% in many of the regions except Tigray (55.24%) and Oromia (50.35%). The lowest average scores in Biology test are for students from Gambela (33.02%), Afar (35.24%), and Harari (37.33%) regions.

Table 80: ANOVA Summery Average Score in Biology across Different Regions

	Sum of squares	DF	Mean square	F ratio	Sig.
Sum square between	473258.643	10	47325.864	198.497	.000
Sum square within	2695829.344	11307	238.421		
Sum square total	3169087.987	11317			

The ANOVA summary Table for Biology score (Table 79) indicates the presence of statistically significant variation ($F(10, 11307) = 198.497$, $p < .001$) in the Biology test performance of students from the different regions.

Table 81: Scheffe Test for Biology Score

Region	Number of pupils	Subset for alpha = 0.005					
		1	2	3	4	5	6
Gambela	944	33.02					
Afar	912	35.24	35.24				
Harari	920		37.33	37.33			
Benshangul Gumuz	890		37.59	37.59			
Amhara	1233			39.16			
Somali	664			39.43			
Dire Dawa	835			40.15			
SNNP	1308				43.17		
Addis Ababa	945				43.92		
Oromia	1627					50.35	
Tigray	1040						55.24
Significance		**.430**	**.336**	**.092**	**1.000**	**1.000**	**1.000**

The Scheffe test of homogenous subset grouping revealed trends of significant variation of Biology test performance for students from Tigray and Oromia regions. Students' performance from these two regions significantly varied with each other and with students of the remaining regions. On the other hand, Biology test performance of students from Gambella became homogenous only to those students from Afar. Students from Addis Ababa and SNNPR regions have shown homogeneity in their Biology test performance. In addition, the data reveals the presence of a wide achievement gap (average score points of 22.22%) of the Biology test between students of Tigray and Gambela regions.

Chemistry Academic Achievement

Table 82: Average Scores of Regions for Chemistry

Region	N	Mean (%)	SD	Std. error
Tigray	998	34.23	14.35	.45
Afar	891	32.88	11.54	.39
Amhara	1228	41.34	17.34	.49
Oromiya	1633	34.84	13.37	.33
Somali	671	33.44	10.52	.41
Bensahngul-Gumuz	884	39.48	14.80	.50
SNNPR	1266	37.00	15.05	.42
Gambela	946	35.78	14.38	.47
Harari	892	38.02	15.29	.51
Addis Ababa	935	37.34	14.65	.48
Dire Dawa	837	34.95	13.39	.46
Composite	**11181**	**36.42**	**14.54**	**.14**

As shown in Table 81, the students average test performance in Chemistry is below 50% across all the regions. Average performances in Chemistry are relatively better for students of Amhara (41.34%), Benshangul Gumuz (39.48%), and Harari (38.02%) regions. The relatively lowest average scores in Chemistry test belong to students from Afar (32.88%), Somali (33.44%), and Tigray (34.23%) regions.

Table 83: ANOVA Summery for Chemistry Scores by Region

	Sum of squares	DF	Mean square	F ratio	Sig.
Sum square between	69784.709	10	6978.471	33.997	.000
Sum square within	2292848.107	11170	205.268		
Sum square total	2362632.816	11180			

The one-way ANOVA test in Table 82 above revealed a statistically significant mean differences in Chemistry test among students in the different regions (F (10, 11170) = 33.997, and p=000) .

Table 84: **Scheffe Test for Chemistry Score**

Region	Number of pupils	Subset for alpha = 0.005						
		1	2	3	4	5	6	7
Afar	891	32.88						
Somali	671	33.44	33.44					
Tigray	998	34.23	34.23	34.23				
Oromia	1633	34.83	34.83	34.83	34.83			
Dire Dawa	837	34.95	34.95	34.95	34.95			
Gambela	946		35.78	35.78	35.78	35.78		
SNNPR	1266			37.00	37.00	37.00	37.00	
Addis Ababa	935				37.34	37.34	37.34	
Harari	892					38.02	38.02	
Benshangul-Gumuz	884						39.48	39.48
Amhara	1228							41.34
Significance		**.439**	**.231**	**.055**	**.138**	**.296**	**.153**	**.617**

The Scheffe test of subset grouping above also revealed that Chemistry test mean performance of students from Amhara were statistically significantly different from students of other regions except Benshangul Gumuz while students from Afar region also have showed a statistically significant mean differences in Chemistry achievement from students of all regions except students from Somali , Tigray, Oromia and Dire Dawa regions . Similar to the Mathematics test performance, students test result in Chemistry far from the attainment of curriculum goals in the subject. For students' performance differences in other regions, see table 83 above.

Physics Academic Achievement

Table 85: Average Scores of Regions for Physics

Region	N	Mean (%)	SD	Std. error
Tigray	998	38.30	13.91	.44
Afar	891	32.25	11.58	.39
Amhara	1228	35.64	12.57	.36
Oromia	1633	36.09	12.11	.30
Somali	671	32.21	9.48	.37
Benshangul-Gumuz	886	32.03	11.23	.38
SNNPR	1266	35.23	11.75	.33
Gambela	946	28.75	8.67	.28
Harari	892	34.66	11.29	.38
Addis Ababa	935	36.37	13.18	.43
Dire Dawa	837	34.82	12.60	.44
Composite	**11183**	**34.47**	**12.11**	**.11**

As shown in Table 84, the students' average test performance in Physics is far below 50% for all the regions. Average performances in Physics are relatively better for students from Tigray (38.30%), Addis Ababa (36.37%), and Oromia (36.09%) regions. The relatively lowest average scores in Physics test observed for students from Gambela (28.75%), Benshangul-Gumez (32.03%) and Somali (32.21%) regions.

Table 86: ANOVA Summery Table Physics Scores

	Sum of squares	DF	Mean square	F ratio	Sig.
Sum square between	68984.968	10	6898.497	49.037	.000
Sum square within	1571657.545	11172	140.678		
Sum square total	1640642.513	11182			

The one-way ANOVA test in Table 85 revealed a statistically significant mean differences in Physics test among pupils in the different regions ($F_{(10, 10786)}$ = 177.70, and p=000).

Table 87: **Scheffe Test for Physics Score**

Region	Number of pupils	Subset for alpha = 0.005			
		1	2	3	4
Gambela	946	28.75			
Benshangul Gumuz	886		32.03		
Somali	671		32.21		
Afar	891		32.25		
Harari	892			34.66	
Dire Dawa	837			34.82	
SNNPR	1266			35.22	
Amhara	1228			35.64	
Oromiya	1633			36.09	36.09
Addis Ababa	935			36.37	36.37
Tigray	998				38.30
Sig.		**1.000**	**1.000**	**.432**	**.080**

The Scheffe test of subset grouping below also revealed that mean performance of students in Physics from Gambella is significantly different from students of other regions and the lowest of all. Meanwhile students from Tigray have showed statistically significant mean differences from students of other regions except student of Oromia and Addis Ababa regions. For students' performance differences in other regions, see table 86 above.

English and Mathematics Academic Achievements in Relation to Performance Standards across Regions

Table 88: Students performance standard in English and Mathematics by Region

Region	English			Mathematics		
	Below Basic(%)	Basic (%)	Proficient (%)	Below Basic(%)	Basic (%)	Proficient (%)
Tigray	51.5	28.0	20.5	57.2	27.0	15.8
Afar	59.4	27.6	12.9	69.6	19.8	10.6
Amhara	54.0	33.3	12.7	61.6	24.6	13.8
Oromia	64.7	26.2	9.0	54.7	25.1	20.2
Somali	57.4	27.9	14.8	68.6	22.3	9.1
Benshangul Gumuz	59.3	28.5	12.1	61.2	24.3	14.6
SNNPR	47.1	36.2	16.7	57.2	22.2	20.6
Gambela	72.8	21.1	6.1	61.3	26.0	12.7
Harari	56.7	29.6	13.7	71.2	18.1	10.7
Addis Ababa	35.7	34.8	29.5	56.9	28.2	14.9
Dire Dawa	44.0	34.1	21.9	58.6	25.4	16.0
Total	**55.1**	**29.8**	**15.1**	**61.0**	**24.1**	**15.0**

The data in Table 87 shows students' performance in English and Mathematics test at the different attainment levels across region. With respect to the English test, students from Addis Ababa (29.5%) and from Dire Dawa (21.9%) have performed at the proficient level, which are top performances relative to the others. In all of the regions, except Addis Ababa, Dire Dawa, and SNNPR highest proportion of grade 8 students (i.e. from 51.5% in Tigray to 64.7% in Oromia) have performed at the below basic level in the English test.

In Mathematics, relatively highest proportions of students from Oromia (20.20%) and SNNPR (20.6%) have performed at the proficient level of attainment. Similar to the English test, significantly highest proportions of grade 8 students from many regions have performed at the below basic level of attainment.

In summary, the regional level data analysis of students' performance in English and Mathematics at the different attainment levels reveals that majority of them have performed at the below basic level.

Biology and Chemistry Academic Achievements in Relation to Performance Standards across Regions

Table 89: Pupils' Performance at the Various Attainment Levels in Biology and Chemistry

REGION	Biology			Chemistry		
	Below Basic (%)	Basic (%)	Proficient (%)	Below Basic (%)	Basic (%)	Proficient (%)
Tigray	21.5	37.9	40.6	66.7	21.8	11.4
Afar	73.6	19.8	6.6	68.4	24.5	7.2
Amhara	65.4	21.5	13.1	47.6	26.1	26.4
Oromia	32.2	38.7	29.1	63.6	23.9	12.6
Somali	58.3	36.0	5.7	63.5	30.7	5.8
Benshangul Gumuz	71.3	18.9	9.8	46.5	34.2	19.3
SNNPR	53.6	28.8	17.6	56.3	27.4	16.3
Gambela	80.3	14.9	4.8	61.0	24.3	14.7
Harari	68.7	21.0	10.3	52.4	29.6	18.0
Addis Ababa	51.7	28.1	20.1	56.1	27.7	16.1
Dire Dawa	59.6	28.3	12.1	62.4	26.3	11.4
Composite	**55.9**	**27.3**	**16.8**	**58.5**	**26.6**	**14.9**

The data in Table 88 shows, pupils' performance in Biology and Chemistry test at the different attainment levels by region. With respect to the Biology test above significant proportion of students from Tigray (40.6%) and from Oromia (29.1%) have performed at the proficient level, which are top performances relative to the others. In the rest of the regions significantly highest proportions of students have performed at the below basic level in Biology, whilst those from Gambella take the lions share (80.3%) at this lowest level of performance.

In the Chemistry test, grade 8 students from Amhara (26.4%) followed by those from Benshangul Gumuz (19.3%) have performed at the proficient level of attainment. Significantly high proportion of students from some of the regions such as Gambela, Tigray, Oromiya, Somali, Dire Dawa, and Afar have performed at the below basic level of attainment in Chemistry. A relatively high proportion of students from Afar (68.4%) have performed at this lowest level of attainment.

In summary, the regional level data analysis of students performance in Biology and Chemistry at the different attainment levels reveals that majority of them have performed at the below basic level.

Physics Academic Achievement in Relation to Performance Standards across Regions

Table 90: Pupils' Performance at the Various Attainment Levels in Physics

REGION	Physics		
	Below Basic (%)	**Basic (%)**	**Proficient (%)**
Tigray	43.1	29.9	27.1
Afar	60.4	29.1	10.5
Amhara	50.7	31.2	18.1
Oromia	47.3	35.1	17.6
Somali	59.2	32.6	8.2
Benshangul-Gumuz	62.0	28.4	9.6
SNNPR	50.7	33.6	15.6
Gambela	75.3	22.2	2.5
Harari	51.2	34.5	14.2
Addis Ababa	47.4	33.8	18.8
Dire Dawa	53.3	29.9	16.8
Composite	**53.7**	**31.2**	**15.0%**

Table 89 depicts the Physics test of grade 8 students from Tigray (27.1%) followed by those from Addis Ababa (18.8%) have performed at the proficient level of attainment. Significantly high proportion of students from some of the regions such as Gambella, Benshangul-Gumuz, and Afar have performed at the below basic level of attainment in Physics. Similar to the Mathematics performance, a relatively high proportion of students from Gambella (75.3%) have performed at this lowest level of attainment.

4.2.5 Factors that Predict Overall Academic Performance of Students

- Is there a relationship between personal factor and overall academic performance of grade 8 students?

4.2.5.1 Personal Factors Predicting Overall Test Performance

Table 91: A Zero Order Correlation Between Overall Academic Achievement and Personal Factors

	1	2	3	4	5
1. Academic achievement	1.00				
2. Gender	-.018 **	1.00			
3. Age	-.051**	-.086**	1.00		
4. Language	-.005	.042*	-.020	1.00	
5. Additional reading	-.028**	.022	-.028**	.043**	1.00

** correlation is significant at.01 * correlation is significant at .05

As indicated in Table 90 personal characteristics such as gender, age, and additional reading have shown a statistically significant relationships with academic achievements though correlation coefficients are considerably low. The correlation coefficients for gender (where boys are coded 0 and girls coded 1 = - 0.018), age (-.051) and having additional reading at home (-0.028) have weak negative relationships. Language spoken at home did not show a significant relationship with overall academic achievement

Table 92: Personal Factors as Predictors of Academic Achievement

Variables	R^2	B	SE	T	Sig.
Gender	.012	-1.996	.192	-10.376	.000
Age	.017	-.165	.027	-6.203	.000
Constant		**38.986**	**.432**	**90.179**	**.000**

The stepwise regression analysis found out the mentioned personal factors to significantly predict the overall academic achievement of grade 8 students. As depicted in the Table 81 only 1.7 % of the variance in grade 8 students' overall academic achievement was explained by the linear combination of age and gender. Even though the coefficient of determination is considerably low, the personal variables explanation of pupils academic achievement is statistically significant(R squared = 0.017, $p< 0.000$).

4.2.5.2 Home Factors Predicting Overall Test Performance

Table 93: A Zero Order Correlation between Academic Achievement and Home Variables

	1	2	3	4
1. Academic achievement	1.00			
2. Family size	.039**	1.00		
3. Father education	.034*	-.060**	1.00	
4. Meal per day	.032**	.000	.199**	1.00

** Correlation is significant at.01 * correlation is significant at .05

As indicated in Table 82, home variables such as family size, father education, and number of meals per day had a statistically significant relationship with overall academic achievement. On the other hand, education attended other than you at home, tutorial given, listening radio, mother education and number of times listening to the radio did not show statistically significant relationships with test performance. Family size, father education and number of meals per day have shown a very weak positive relationship.

Table 94: Home Background as Predictors of Overall Academic Achievement

	B	SE	T	Sig.	R^2 change
Family size	.314	.079	3.962	.000	.002
Father education	.173	.063	2.756	.006	.003
How often do you eat per day	.360	.154	2.330	.020	.004
Constant	33.802	.311	108.706	.000	

The step–wise regression analysis made to see the extent that home variables explain variations in test performance have revealed the presence of considerably weak explanation (0.5%) of the variables to test performance. As shown in Table 33 , 0.4% of grade 8 students test performance variation in overall academic achievement was explained by the linear combination of number of meals per day (B = .360, t = 2.33, p< .020), family size (B=.314, t=3.962, p<.002) and father education (B=. 173, t= 2.756, p< .003).

4.2.5.3 School Factors Predicting Overall Test Performance

Table 95: Relationship between School Variable and Grade 8 Students Overall Academic Achievement

	1	2	3
4. Overall academic achievement	1.00		
5. Absent in a semester	-.077**	1.00	
6. Time taking from school to home	-.022**	.111**	1.00

** Correlation is significant at.01 * correlation is significant at .05

As the zero order correlation matrixes below shows, the numbers of days pupils become absent from school and the time taking to travel from home to school have statistically significant negative relationships with their test performance. However, the extent of relationship found to be very weak. As the number of days the pupil become absent from school increases, the test performance decreases. On the other hand, time taking to reach from home to school increases, the test performance decreases.

Table 96: School Variable as Predictors of Academic Achievement

Variables	B	SE	T	R^2 change	Sig.
Absent in a semester	-.729	.101	-7.192	.006	.000
Constant	35.977	.122	294.727		.000

The step wise regression analysis in Table 35 indicates that only .6 % of the pupils' variation in the overall test performance at grade 8 was explained for by the variation on number of absents from school in a semester (B=-.729, t= -7.192, p < .000) and time spent on going from home to school (B =-1.233, t= -10.195, p < .05) significantly predicted the overall test performance of pupils at grade 4.

4.2.6 Factors that Predict Academic Achievement in Key Subjects

Table 97: Relationship between English Inputs and Pupils' Test Performance in English

	1	2	3	4
1. Academic achievement in English	1.00			
2. Home work given in English	.092**	1.00		
3. Pupils' understanding English	-.069**	-.081**	1.00	
4. Number of sharing textbook in English	.067**	-.045**	.061**	1.00

As indicated below home works given in English and availability of textbook in English have shown significant positive relationships with the students' test performance in English. Whereas students report of understanding English has significant negative correlation with test performance. However, the degrees of relationships in either case are considerably weak. As the data in the zero order correlation matrix below depicts, as the number of home works given for the English subject and availability of English textbook increase, so does their test performance in English.

Table 98: English Inputs as Predictors of Academic Achievement

Variables	B	SE	T	R^2 change	Sig.
Understanding English	-1.095	.104	-10.534	.010	.000
Home work given in English	.934	.130	7.176	.015	.000
Availability of Textbook in English	.314	.099	3.162	.016	.002
Constant	35.856	.541	66.267		.000

The step-wise regression analysis below indicates that only 1.6% of the variations on English test performance at grade 8 explained by the linear combination of the variation on understanding English, number of home works given in a week, and availability of textbook in English. Understanding English (B=-1.095, t=-10.534, p<000), number of home works given (B =0.934, t= 7.176, p < .000) and availability of textbook in English (B=0.314, t = 3.162 and p < .000) have significantly predicted the students' achievement in the English test.

Table 99: Relationship between Mathematics Input and Pupils Test Performance

	1	2	3	4
1. Academic achievement in mathematics	1.00			
2. Home work given in mathematics	.004	1.00		
3. Students understanding in mathematics	-.024**	-.095**	1.00	
4. Number of sharing Textbook in math	-.022*	-.038**	.086**	1.00

As indicated in table 88, the frequency of home works given in Mathematics has shown a very weak and insignificant relationship with test performance. As the number of home works given in Mathematics increase, so does their test performance. The students' report of understanding Mathematics and increase in the number of pupils sharing Mathematics textbook have shown significant negative relationships with test performance in Mathematics.

Table 100: Mathematics Input as Predictors of Academic Achievement

Variables	B	SE	T	R^2 change	Sig.
Understanding mathematics	-.164	.073	-2.254	.001	.024
Constant	26.072	.222	117.413		000

The stepwise regression analysis below indicates that only 0.1% of the variation of grade 8 students test performance in Mathematics explained by the linear combination of the variations on understanding Mathematics, number of home works given, and availability of textbook in Mathematics.

Table 101: Relationship between Biology Input and Test Performance

	1	2	3	4
1. Academic achievement in Biology	1.00			
2. Home work given in Biology	-.019	1.00		
3. Students understanding in Biology	-.025*	-.088**	1.00	
4. Number of sharing Textbook in Biology	-.051**	-.005	.127**	1.00

As indicated in Table 90, home works given in Biology, understanding biology, and sharing of textbook negatively correlated with test performance in biology. However, similar to the other key subjects, the relationship between the mentioned variables and test performance in Biology found to be considerably weak.

Table 102: Biology Input as Predictors of Biology Score

Variables	B	SE	T	R^2 change	Sig.
Availability of Textbook in Biology	-.838	.169	-4.952	.003	.000
Constant	44.371	.398	111.388		.000

The step-wise regression analysis in Table 91 indicates that about 0.3% of the variation of test performance in Biology was explained by the variation on availability of text books (B = -0.838, t=4.95, p<000).

Table 103: Relationship between Chemistry Input and Test Performance in Chemistry

	1	2	3	4
1. Academic achievement in Chemistry	1.00			
2. Home work given Chemistry	.020	1.00		
3. Students understanding in Chemistry	-.020	-.132**	1.00	
4. Number of sharing Textbook in Chemistry	.014	.016	.125	1.00

As indicated in table 92 the frequency of home works given in Chemistry, students' report of understanding chemistry,, and availability of textbook have very weak and insignificant correlation with test performance in chemistry.

Table 104: Chemistry Input as Predictors of Chemistry Score

Variables	B	SE	T	R^2 change	Sig.
Home work given Chemistry	.249	.113	2.209	.001	.027
Availability of Textbook in Chemistry	.349	.162	2.155	.001	.031
Constant	**35.48**	**.441**	**80.519**		**.000**

The step-wise regression analysis indicates that only .1% of the variation of grade 8 students test performance in Chemistry explained by the linear combination of the variations on the number of home works given, and availability of textbook in Chemistry.

Table 105: Relationship between Physics Input and Test Performance in Physics

	1	2	3	4
1 Academic achievement in Physics	1.00			
2. Home work given Physics	.022*	1.00		
3. Students understanding in Physics	-.070**	-.130**	1.00	
4. Number of sharing Textbook in Physics.	-.005	-.005	.126**	1.00

As indicated in Table 94 the frequency of home works given in Physics, students' report of understanding Physics, and availability of textbook have weak, but significant correlation with test performance in Physics.

Table 106: Physics Input as Predictors of Physics Score

Variables	**B**	**SE**	**T**	**R^2 change**	**Sig.**
Understanding Physics	-.948	.142	-6.692	.005	.000
Constant	37.073	.329	112.796		.000

The step-wise regression analysis in Table 95 indicates that about .5% of the variation of grade 8 students test performance in Physics was explained by the variations of students' understanding Physics.

4.2.7 Personal, Home and School Variables as a Function of Overall Academic Achievement

Different factors could influence academic achievement of students. In this study, the factors were categorized as personal factors that are pertinent to the pupil, home factors and school factors. The following section describes the influence of these factors on the academic achievement of students.

4.2.7.1 Personal Variable as a Function of Academic Achievement

In this study, personal variables include the pupils' gender, language used at home, and reading additional textbooks.

Table 107: Personal Variables as Related to Academic Achievement

Variable	Gender	Mean	Std. Dev.	t test/F test	Sig.
Gender	boys	36.36	9.00	9.82	.000
	girls	34.50	7.95		
Language at home is different from instructional language	No	35.44	8.79	.455	.642
	Yes	35.36	8.40		

As the data in Table 96 above shows, on average boys' test performance becomes better than that of the girls. The independent sample t test also indicate that there was a significant mean difference between boys' and girls' composite scores (t=9.82, p =.000).The language spoken at home has no statistically significant relationship with the students test performance.

4.2.7.2 Home Variable as a Function of Academic Achievement

Table 108: Home Variables 1 as a Function of Pupils Performance

Variable		Mean	Std. Dev.	F test	Sig.
With whom you are living	With my mother and my father	35.40	8.59	1.759	.134
	With my father only	34.94	8.62		
	With my mother only	35.52	8.57		
	With relatives	35.62	8.48		
	With others	34.36	8.64		
Family size	2	34.04	8.47	5.040	.000
	3	35.03	8.44		
	4	34.94	8.23		
	5	35.75	8.65		
	More than 5	35.53	8.68		
Father occupation	Farmer	35.18	8.56	1.613	.153
	Government employee	35.58	8.69		
	Merchant	35.40	8.82		
	Unemployed	35.10	8.27		
	I don't know	35.39	8.38		
	Other	35.97	8.51		
Mother occupation	Farmer	35.37	8.84	.658	.656
	Government employee	35.50	9.39		
	Merchant	35.00	8.29		
	House wife	35.44	8.42		
	I don't know	35.15	8.50		
	Other	35.63	8.65		

For the items that probe "with whom you are living," the data showed no difference in the students' test performance. The one-way ANOVA test also indicated that there was not statistically significant mean differences in test performance among those who are living within different family structures. For the items that ask family size, students who are living with a family size of five performed better than the other group. The ANOVA test indicated statistically significant mean differences of academic performance among those who are living within a different family size ($F=5.04$, $p< 000$). The Scheffe test also depicted that students who are living with a family of five and more than five have showed a statistically significant mean differences in overall academic achievement from students of two family sizes. Regarding the relationship between father's occupations and test performance, the data showed that there were no differences in test performance of students. Similar result was obtained with mother occupation too. That means, variation in mother occupation did not show variation in students test performance (for detail, see Table 97).

Table 109: Home Variables 2 as a Function of Pupils Performance

Variable		Mean	Std. Dev.	F test	Sig.
When do you help your family with chores?	Sometimes absent from school	35.20	8.38	19.078	.000
	Always after school	36.23	8.94		
	Sometimes after school	36.56	8.89		
	Always Sunday and Saturday	34.16	8.20		
	I don't help	34.81	8.62		
How many family member attended school other than you?	One	35.38	8.07	.841	.499
	Two	35.45	8.73		
	Three	35.56	8.68		
	More than three	35.11	8.82		
	No one	35.52	8.28		
Father's educational level	Illiterate	35.18	8.56	1.613	.153
	Write and read	35.58	8.69		
	1-8	35.40	8.82		
	9-12	35.10	8.27		
	Above 12 grade	35.39	8.38		
Mother educational level	Illiterate	35.37	8.84	.658	.656
	Write and read	35.49	9.39		
	1-8	35.00	8.29		
	9-12	35.44	8.41		
	Above 12 grade	35.15	8.50		

Table 98 shows data for some of the home variables that may have potential influence on test performance. When we examine test performance by family support, pupils who responded “I help my family sometimes after school” performed better than the rest of the group. The ANOVA test also indicated statistically significant mean differences on test performance in relation to family support (F= 19.08, p< 000). The Schefee test depicted that there were statistically significant mean differences on overall academic achievement between those who helped their family sometimes after school and the rest of the group except student who helped their family always after school.

For the relationship between other family members attending school and academic performance of pupils, the data showed that there was no statistically significant difference among the group.

The same to the number of family members attending school, variation in the level of father education and mother education did not show variation in students’ test performance (see Table 98 for detail).

Table 110: Home Variables 3 as a Function of Pupils Performance

Variable		Mean	Std. Dev.	F test / t test	Sig.
Home tutor	No	35.62	8.76	1.771	.077
	Yes	35.29	8.49		
How many times tutored per week?	1-3 days	35.04	8.27	2.543	.079
	4-6 days	35.29	8.75		
	7 days	34.62	8.33		
Radio available	Yes	34.79	8.48	-3.867	.000
	No	35.62	8.63		
How many times you listen radio in a week?	Don’t listen	35.04	8.54	5.958	.000
	1 -2 days	35.48	8.72		
	3 to 4 days	36.00	8.67		
	5 to 6 days	36.13	8.81		
	All days	34.79	8.16		
TV available	Yes	35.32	8.52	-1.027	.305
	No	35.51	8.70		
How many times watch TV per week?	Don’t watch	35.45	8.54	.731	.571
	1 -2 days	35.21	8.58		
	3 to 4 days	35.66	8.37		
	5 to 6 days	34.99	8.72		
	All days	35.24	8.71		
How many times you eat per day?	2 times	34.58	8.17	10.234	.000
	3 times	35.63	8.62		
	More than 3 times	35.39	8.86		

Among other home variables that predict students test performance, those significantly predicted test performance of students were availability of radio at home (t= -3.867, p< 000), the number of times a student listen to the radio in a week (F=5.958, p< 000) and number of meals a day (F=10.234, p<000). The other variables such as home tutor, number of times the student tutored in a week, availability of television set, and number of times watching TV in a week did not explain variation in students test performance.

Students who do not have a radio performed better than students who had a radio. The independent sample t test also depicted presence of statistically significant mean difference on overall test performance between those who had a radio and those who do not have a radio.

Students who have the habit of listening to the radio 5 to 6 days a week showed a relatively better test performance than those who do not have the habit of listening to the radio, and the other groups. The ANOVA test revealed the presence of statistically significant mean difference in overall academic achievement for radio listening. The Scheffe test also indicated that mean performance of students who listened radio for five to six days had a statistically significant mean differences on overall academic achievement from students who listened none at all and listened radio all days in a week. Moreover, the Scheffe test revealed that mean performance of students who listened radio all days had a statistically significant mean differences on overall academic achievement from students who listened three to four days per week.

With respect to the relationship of number of meals a day and performance on test, those who eat more than 3 meals a day have better scores than those who eat 3 meals a day and, 2 meals a day. The ANOVA test also indicated statistically significant mean differences of academic achievement for the number of meals students eat in a day. The Scheffe test revealed that mean performance of students who eat twice a day was statistically significantly different from students who eat three times, and above three times.

4.2.7.3 School Variables as a Function of Overall Academic Achievement

Table 111: School Variables as a Function of Overall Academic Achievement

Variable		Mean	Std. Dev.	F test	Sig.
How long does it take from home to school?	15 minutes at most	35.33	8.67	4.350	.002
	20 to 30 minuet	35.84	8.82		
	30 to 60 minute	35.42	8.40		
	1 hour to 1:30 minute	35.21	8.29		
	More than 1:30 minute	34.17	7.91		
Absent in a semester	None	36.10	9.00	21.224	.000
	1 to 3 days	35.07	8.21		
	4 to 6 days	33.87	7.86		
	7 to 10 days	34.56	8.29		
School attended before joining first grade	Church school	35.50	8.51	18.729	.000
	Quran school	34.10	8.00		
	Kindergarten	35.80	8.59		
	Attended none	35.80	8.87		

As shown in Table 100, the distance pupil travel from their home to school has relationship with their test performance. Pupils who travel for 20 to 30 minutes from their home to school have shown a better test performance than those who travel for 15 minutes at most, 30 to 60 minutes, 1 hour to 1:30 minutes, and more than 1: 30 minutes. The ANOVA test also indicated statistically significant mean differences of overall academic achievement for the length of time the pupil travel to reach to school. The Scheffe test of subset grouping also depicted that there were statistically significant mean differences between those who travel more than 1 hour and 30 minutes and the rest of students.

With respect to the relationship of absentees from school and test performance, those students who were not absent from school even for a day showed a better performance than those who were absent from 1 to 3 days, 4 to 6 days and 7 to 10 days in a semester. The ANOVA test also indicated statistically significant mean differences of academic achievement based on the number of days pupils are absent from school in a semester. The Scheffe test of sub set grouping showed that those students who had no any absentee significantly vary in their test performance from the rest of the group.

Moreover, the Scheffe test showed significant variation in overall academic achievement among students who were absent from school for about one to three days and for about four to six days in a semester.

The type of pre-school attended by students has relationship to academic performance. Students who attended Kindergarten and students who attended none as their pre-school experience have shown a relatively better performance than those who have attended other forms of preschool education. The ANOVA test also indicated statistically significant mean differences of test performance for the type of school attended before joining the first grade. The Scheffe test revealed that students who attended Kindergarten, attended none at all, and attended Church school had a statistically significant mean variation on overall academic achievement from students who attended Quran School.

4.2.8 Other School Variables as a Function of Test Performance in Key Subjects

This sub section intends describing the relationship of school variables and test performance of grade 8 students in key subjects. The main variables described are the frequency of home works for each key subject, availability of textbook, and the pupils' perception whether they can understand lessons in the key subjects tested in the 4th national learning assessment.

English Academic Achievement

Table 112: School Variables Relationship with English Test Score

Variable		Mean	Std. Dev.	F test	Sig
Home work given in English in a semester?	None	32.96	14.43	22.029	.000
	Only once	33.78	14.73		
	Twice	36.04	15.91		
	Three times	36.58	15.55		
	Above three times	38.15	16.10		
Textbook in English?	Don't have	33.77	14.33	56.420	.000
	One for me	39.33	16.52		
	Shared for two	36.53	15.48		
	Shared for three	37.09	16.38		
	Shared for four	38.50	15.99		
Understanding English?	Very well	39.60	17.40	31.125	.000
	Average	36.75	15.13		
	Slight understanding	35.37	15.05		
	I don't understand	38.21	16.73		

English test performance of students has considerably related to the frequency of home works given. As the data reports, pupil whom their teachers given them home works three times and above have shown better test performances than those whom their teachers given them home works less than three times and none at all.

Furthermore, the one way ANOVA test confirmed the presence of statistically significant mean differences of test performance for the number of times home works given by teachers ($F=22.03$, $p<000$). The Scheffe test of homogenous subset grouping also showed that students who were given no homework at all in Mathematics had statistically significant mean differences in Mathematics from the rest of students except those students who were given only once. The Schefee test also indicated that students who were given homework only once had a statistically significant mean difference in Mathematics from students who were given homework in Mathematics twice and three times .

Regarding the relationship of textbook possession in English and pupils test performance, the data in Table 101 shows the presence of significant relationships between the two. For instance, students who shared their English textbook with no one else have achieved better marks in English test than those who shared their English textbook with other pupils, and to those who did not possess English textbook at all either for their own or in a form of sharing with others. The one way ANOVA test also indicated statistically significant mean differences of English test performance for possession of English textbook ($F=56.4$, $p<000$). The Schefee test of subset grouping revealed that students who had English textbook for their own showed a significantly different test performance from students who had no textbook at all and students who shared English text book for two. The Scheffe test also indicated that students who had no English text book at all had statistically significant mean differences from the students who have text book for their own and shared textbooks.

One of the other school variable related to test performance is whether students understand English lessons well. Those who reported that they understood English well scored higher than those who reported average understanding, slight understanding and not at all. The one way ANOVA test also indicated statistically significant mean differences of English test performance for understanding English (F=31.13, P<000). The Schefee test of subset grouping confirmed that students who reported understanding English very well showed a significantly different test performance than the others except students who said I don't understand Mathematics.

Mathematics Academic Achievement

Table 113: School Variables Relationship with Mathematics Test Score

Variable		Mean	Std. Dev.	F test	Sig.
Home work given in Mathematics in a semester?	None	25.33	10.48	1.092	.358
	Only once	26.40	11.41		
	Twice	25.50	11.34		
	Three times	25.35	11.14		
	Above three times	25.78	11.72		
Textbook in Mathematics	Don't have	26.98	12.33	8.783	.000
	One for me	25.24	11.03		
	Shared for two	25.00	11.45		
	Shared for three	25.93	11.90		
	Shared for four	26.22	12.77		
Understanding Mathematics	Very well	26.12	11.88	1.686	.168
	Average	25.50	11.54		
	Slight understanding	25.81	11.60		
	I don't understand	25.24	10.48		

Mathematics test performance has no significant variation for the number of home works given to students. The one way ANOVA test indicated there was no statistically significant mean differences of test performance in Mathematics for the number of home works given in a week (F=1.09, p>0.05).

Regarding the relationship of textbook possession in Mathematics and test performance, the data in Table 102 shows the presence of significant relationships between the two. The one-way ANOVA test indicated statistically significant mean differences of Mathematics test performance for possession of textbook (8.78, p<000). Surprisingly, those without Mathematics textbook have performed better than the other groups. The Schefee test of sub set grouping revealed that students who did not possess Mathematics textbook had statistically significant mean differences from students who shared Mathematics text book for two.

Similar to the English test result, one of the school variables related to test performance in Mathematics is whether students understand the lessons in Mathematics very well. The one-way ANOVA test indicated that there are no statistically significant mean differences of Mathematics test performance for understanding lessons.

Biology Academic Achievement

Table 114: School Variables Relationship with Biology Test Score

Variable		Mean	Std. Dev.	F test	Sig
Home work given in Biology in a semester	None	40.21	16.04	22.553	.000
	Only once	43.94	17.52		
	Twice	44.09	17.68		
	Three times	42.91	16.73		
	Above three times	40.82	16.01		
Textbook in Biology	Don't have	43.41	17.02	10.890	.000
	One for me	42.79	17.19		
	Shared for two	39.97	15.33		
	Shared for three	39.98	17.45		
	Shared for four	41.75	15.26		
Understanding Biology	Very well	45.20	18.03	104.159	.000
	Average	40.01	15.29		
	Slight understanding	39.49	15.49		
	I don't understand	47.18	17.96		

Grade 8 students whom their teachers given them Biology home works twice a week have the highest mean score than the rest of the group. The one-way ANOVA test also indicated statistically significant mean differences of test performance in Biology for the number of home works given in a week ($F=22.5$, $p<000$).

Furthermore, the Scheffe test revealed those students whom their teachers given home works above three times and none at all had statistically significant mean differences on Biology score from students who were given homework only once, twice and three times.

Regarding the relationship of textbook possession in Biology and test performance, the data in Table 103 shows the presence of significant relationships between the two. For instance, students who have no Biology textbook at all have the highest average score. The one-way ANOVA test indicated statistically significant mean differences of Biology test mean score for possession of textbook (F=10.890,p<000). The Schefee test revealed that students who had Biology textbook for their own and did not possess any textbook at all showed statistically significant mean differences on Biology from students who shared a Biology textbook for two and three.

One of the school variables related to the pupils test performance in Biology is whether they understand the lessons very well. Students who reported that they understood Biology very well scored higher in the test than those who reported average understanding, slight understanding, and no understanding at all. The one way ANOVA test also indicated statistically significant mean differences of Biology test performance for understanding lessons (F=104.6, p<000). The Scheffe test also indicated students who understood Biology very well showed statistically significant mean differences in their score from students who had average or slight understanding in Biology except student who replied as don't understand biology.

Chemistry Academic Achievement

Table 115: School Variables as a Function Chemistry Test Score

Variable		Mean	Std. Dev.	F test	Sig
Home work given in Chemistry in a semester	None	36.15	14.63	2.416	.047
	Only once	36.06	14.38		
	Twice	37.34	15.24		
	Three times	36.80	15.36		
	Above three times	36.95	14.59		
Textbook in Chemistry	Don't have	35.40	13.90	8.684	.000
	One for me	37.22	15.10		
	Shared for two	37.53	15.20		
	Shared for three	36.24	13.81		
	Shared for four	35.33	14.35		
Understanding Chemistry	Very well	36.67	15.03	2.109	.097
	Average	36.87	14.75		
	Slight understanding	36.18	14.15		
	I don't understand	35.81	14.52		

Grade 8 students whom their teachers given them Chemistry home works twice a week have the highest mean score than the rest of the group. The one-way ANOVA test also indicated statistically significant mean differences of test performance in Biology for the number of home works given in a week ($F=2.416$, $p<05$).

Regarding the relationship of textbook possession in Chemistry and test performance, the data in Table 104 shows the presence of significant relationships between the two. For instance, students who have Chemistry textbook shared for two students have the highest average score, followed by those who have Chemistry textbook for their own. Those students who have no textbook in Chemistry performed the least. The one-way ANOVA test indicated statistically significant mean differences of Chemistry test mean score for possession of textbook ($F=8.684, p<000$). The Scheffe test of sub set grouping showed a statistically significant variation of students test performance among those having Chemistry textbook for their own and those did not possess Chemistry textbook at all and also between those who did not possess Chemistry text book at all and those students who shared for two.

One of the school variables related to the students test performance in Chemistry is whether they understand the lessons very well. As the data in Table 104 shows, there are no statistically significant test performance variations among students who have different levels of understanding in Chemistry**.**

Physics Academic Achievement

Table 116: School Variables as a Function of Physics Test Score

Variable		Mean	Std. Dev.	F test	Sig
Home work given in Physics in a semester	None	33.00	11.88	8.445	.000
	Only once	34.86	12.48		
	Twice	35.50	12.86		
	Three times	35.36	12.48		
	Above three times	34.62	11.89		
Textbook in Physics	Don't have	33.86	11.73	10.532	.000
	One for me	35.51	12.54		
	Shared for two	34.13	12.14		
	Shared for three	33.80	12.41		
	Shared for four	33.78	11.15		
Understanding Physics	Very well	36.81	13.31	25.535	.000
	Average	34.05	11.46		
	Slight understanding	33.77	11.61		
	I don't understand	34.72	12.95		

Grade 8 students whom their teachers given them Physics home works twice a week have the highest mean score than the rest of the group. The one-way ANOVA test also indicated statistically significant mean differences of test performance in Physics for the number of home works given in a week (F=8.445, p<000).

Furthermore, the Scheffe test has shown that students whom their teachers given homework none at all in Physics had a statistical significant mean differences in Physics from students whom their teachers given homework once, twice, three times and above three times in Physics.

Regarding the relationship of textbook possession in Physics and test performance in Physics, the data in Table 105 shows the presence of significant relationships between the two. For instance, students who have Physics textbook for their own have the highest average score in physics. The one-way ANOVA test indicated statistically significant mean differences of Physics test mean score for possession of textbook (F=10.532,p<000). The Scheffe test of subset grouping revealed significant variation in students test performance in Physics among students who possessed textbook for their own and those who had no textbook at all, and between students who possessed text book for their own and those students who shared Physics textbook for two.

One of the school variables related to the students test performance in Physics is whether they understand the lessons in Physics very well. Students who reported that they understood Physics very well scored higher in the test than those who reported average understanding, slight understanding, and no understanding at all. The one way ANOVA test also indicated statistically significant mean differences of Physics test performance for understanding lessons (F=25.535, p<000). The Schefee test sub set grouping revealed that there was a statistically significant mean difference on students' test performance in Physics among students who understood Physics very well, and those students who reported that they understand lessons in Physics on average, slightly, and none at all.

4.3 Teacher Variables

4.3.7 Background Variables of Grade 8 Teachers

The data from the 1,685 grade 8 teachers who were involved in the fourth national learning assessment reveals that 1247 (80.5%) of teachers were males and the remaining 302 (19.5 %) were females. The average age score was 29.47 with a standard deviation of 11.340.

Concerning grade 8 teachers' levels of qualification, 8.3% of the teachers involved in the study were certificate holders, 14.4% were attending diploma program, 39.6 % were graduates in diploma, 34.1 % were attending first-degree in service program, and 3.4 % replied other. When it comes to the years of teaching experience, 7.9% had 1 and 2 years of teaching experiences, 8.3% had 3 years of teaching experiences, and 15% had 4 years of teaching experiences and 60.9 % had five years and above teaching experiences. From this we can conclude that majority of grade 8 teachers who were involved in the fourth national learning assessment had teaching experiences above five years.

4.3.8 Teacher Related Variables that Link to Teaching Learning Process

The time it takes teachers to travel from home to their work place, (school) ranges from the minimum of less than 15 minutes to the maximum 1 hour and 30 minutes. Above half (906, 54.2%) of teachers involved in the study reported that they travel for fewer than 15 minutes from their home to school. About 385 (23%) of them travel from home to school between 15 to 30 minutes. Others, 253 (15.1 %) travel between 30 to 60 minutes to reach from home to school. Few of them, i.e. 127 (7.5 %) reported to travel more than an hour from their home to school.

Concerning the weekly workload of teachers from the participating schools, about 547 (32.6%) reported to shoulder a weekly teaching load of 16 to 20 periods. Other 305 (18.2%) had weekly teaching loads of 26 to 30. About 505 (30.1 %) of the teachers had workloads ranging between 21 to 25 periods a week. The highest workload, which was more than 30 periods a week, goes to 122 (7.3 %) teachers. On the other hand, only 199 (11.8%) participating teachers had reported to teach fewer than 15 periods a week.

Regarding additional commitment in the community beside teaching, 453 (27.4%) of the participating teachers reported a lot, 605 (36.6%) reported to some extent, 239 (14.5 %) reported very little and the rest 355 (21.5 %) had reported to have no commitment at all.

Classroom observation and supervision of teachers' activities has contribution for the improvement of instruction. Based on this, teachers from the sampled schools reported different frequencies of supervision by school directors or supervisors. For instance, about 217 (13.1%) of them were observed only once in a semester. About 751 (64.9%) of the teachers reported experiencing classroom supervision at least twice in a semester. A considerable number of teachers i.e. 583 (40.8%) reported to have been observed and supervised for three and more than three times a semester. To the contrary, 109(6.6%) of grade 8 teachers replied that they were not observed and supervised while they are teaching in classes.

When it comes to missed periods of the class contact in a semester, on the average the teachers missed 4.78 periods.

There is a general recognition on the contribution of teacher – parent discussions on matters of students' learning and behavior development. The teacher data of the fourth national learning assessment at grade 8 revealed that about 1547 (92.0 %) of the teachers had discussions with pupils' parents whereas the remaining 32 (4.2%) of the teachers reported that they did not discuss with students' parents in a semester time.

With respect to the participation of teachers in skill development programs, the data revealed that 1115 (67.7%) had attended training on methods of teaching, whereas 532 (32.3%) had not attended, about 724 (44.7%) had attended training on curriculum issues and 895 (55.3%) had not attended. About 1017 (62.2 %) attended training on assessment technique whereas 618 (37.8%) did not attend training on assessment technique. 934 (57.4 %) responded that they had training on student learning behavior whereas 691 (42.6 %) did not attend any training on students learning behavior.

With respect to the subjects they teach at their school in the 2003 academic calendar, 703 (62.6%) reported that they taught English, 692 (62.3 %) reported that they taught math, 696 (62.6%) reported that they taught biology, 682 (60.4%) teachers reported that they taught chemistry, and 639 (69.0%) reported that they taught physics.

4.3.9 Availability of Teaching Material in School as Perceived by Teachers

Table 117: Sharing of Books among Students by Subject

Subject	One to one	One for two	One for three	One for four and more than four
English	177 (42.1 %)	53 (12.6%)	24(5.7%)	152(36.2 %)
Mathematics	205(48.9%)	56(13.4%)	21(5.0%)	126(30.1%)
Biology	190(45.8%)	60(14.5%)	25(6.0%)	124(29.9%)
Chemistry	220(49.2%)	55(12.3%)	19(4.3%)	137(30, 6%)
Physics	139(48.45%)	39(13.6%)	17(5.9%)	83(28.9%)

Student access to textbooks of key subjects at grade 8 may have contributions in learning and academic achievement. As participating teachers in this study reported, highest proportions of their students accessed textbooks for their own. A minimum of 42.1% English teachers and a maximum of 49.2% of Chemistry teachers reported that their students had textbooks of the subjects for their own. On the other hand, from a minimum of 28.9% teachers of Physics to a maximum of 36.2% teachers in English reported that their students access textbooks by sharing for four students and above.

4.3.10 Portion Coverage in the Key Subjects

Table 118: Content Area Coverage by Subject

Subject	Above 90 %	80%-90%	60%-70%	50%-60%
English	201 (47.9)	152 (36.2%)	59(14.0%)	4(1.0%)
Mathematics	224 (53.5%)	131(31.3%)	50(11.9%)	12(2.9%)
Biology	196(47.2%)	148(35.7%)	52(12.5%)	12(2.9%)
Chemistry	219(49.0%)	153(34.2%)	65(14.5%)	6(1.3%)
Physics	175 (61.0%)	85(29.6%)	18 (6.3%)	2(.7%)

Portion coverage by teachers may facilitate for better academic achievement of students. As the data in Table 107 shows from a minimum of 47.9% of English teachers to a maximum of 61% teachers in Physics reported that, they had covered above 90% of the curriculum content in the subject they teach. Considerably less number of teachers from each subject reported to cover the curriculum content between 50%-60%.

4.3.11 Student Active Involvement

Table 119: **Frequency of Students' Participation in a Class as Perceived by Teachers**

Item	Rarely	Sometime	Most of the time
Ask and answering questions	174 (10.4%)	649(38.9%)	845(50.7%)
Complete home work regularly	210 (12.6%)	840(50.2%)	623(37.2%)
Complete class work on time	286 (17.2%)	803(48.3%)	572(34.4%)
Speaks to the class with confidence	391 (23.2%)	624(37.0%)	655(38.9%)
Attend the class regularly	176 (10.5%)	517(31.0%)	976(58.5%)
Give much attention to learning	231 (13.8%)	624(37.3%)	816(48.8%)
Keep class discipline	154 (9.2%)	457(27.4%)	1057(63.4%)
Participate in group works	183 (11.0%)	728(43.6%)	758(45.4%)
Respect their teachers	174 (10.5%)	433(26.0%)	1058(63.5%)
Solve their problems by discussion	308 (18.4%)	721(43.0%)	647(38.6%)

Student active involvement in the learning process improves learning for better academic achievement. Grade 8 teachers who have participated in the fourth national learning assessment reported their student class participation in different frequencies. For instance, majority of teachers identified respect to teachers (63.5), class discipline (63.4%), regular attendance (58.5%) and asking and answering questions (50.7%) to become the most common behaviors that characterize students. On the other hand, considerable proportion of teachers rated that their students show behaviors such as expressing their ideas with confidence (23.2 %), tendency of solving their problems by discussion (18.4%) and on time completion of class works (17.2%) rarely.

4.3.12 Availability of Curriculum Materials

Regarding the availability of curriculum materials for the subjects they teach, 1059 (70 %), teachers responded that they had not provided with the subject syllabus. Another 454 (30%) teachers responded that they had syllabus of the subject to support their teaching. When it comes to teacher guide, 776(50.6%) teachers responded that they did not access teacher guide for the subject they teach, while 758 (49.4%) responded that they were provided with teachers' guide. Majority of teachers 1197(71%) reported to access textbook in the subject they teach whereas about 370 (23.6%) of them responded that they were not provided with student textbook.

4.3.13 Correlation between Teacher Variables and Students' Test Performance

Table 120: Correlation between Teacher Background Variable and Overall Academic Achievement

Variable	Pearson r	Sig.
1. Background Variable		
Qualification	.012	>.05
Experience in teaching	.007	>.05
2. Teaching variable		
Teaching load in a semester	-.005	>.05
Time taking	-.032	>.05
Additional commitment	.017	>.05
Missed period	-.007	>.05
Content coverage	-.008	>.05
Student Textbook ratio	-.048	>.05
3. Discussion with parents and Supervisory Support		
Supervisors observation by the school principal or supervisors	.091*	<.05
Discussion with parents	.007	>.05

As shown in the Table 109, out of the teacher variables included in the study only one of them had statistically significant association with the students' test performance. This is the teachers' report about the frequency of classroom supervision by the school principal or education supervisors. The correlation coefficient obtained (r =0.085) depicts the presence of significant and positive correlation between teachers' report and experience of classroom supervision and the students' overall test performance. However, the relationship is considerably weak in magnitude. The more frequent supervision the teachers experienced, the better the students test performance for the composite score. Other teacher variables such as teacher qualification, experience in teaching, teaching load of teachers, discussion with parents and the like, did not show statistically significant correlation with the students' test performance.

Table 121: Teacher Variable as Predictors of Overall Academic Achievement

Variables	**B**	**SE**	**t**	**Sig.**	**R^2 change**
Supervision	.537	.197	2.723	.007	.007
Constant	**34.143**	**.440**	**77.523**	**.000**	

The stepwise regression analysis as indicated in Table 110 showed that approximately 0.7% of students' variation in test performance is attributed for by the frequency of teacher supervision.

CHAPTER FIVE: CONCLUSIONS AND RECOMMANDATIONS

5.1 Conclusions

The objectives of this study were to assess grade 8 pupils' academic achievement with respect to curriculum goals in key subjects and find out the factors that could explain variation in academic achievement. To achieve these objectives, data were collected from grade 8 pupils and teachers. Data were gathered through a questionnaire and administration of various tests in the key subjects. Though data gathered through questionnaire and comparing assessments of SNLA, TNLA, and FNLA might have some limitations on the grounds that the tests were not equated and there is absence of information on the psychometric qualities of the instruments used, the following major findings were obtained.

Specifically, the data analyses and interpretations of the study lead to drawing the following conclusions:

- The overall test performance of grade 8 students in the fourth national learning assessment was far below the minimum standard, where about 50 percent of the students achieved an average score of 34% and below for the 5 key subjects.
- Compared to the second and third national learning assessment results, the composite average is the least in the current national learning assessment. However, the decline in achievement gap is narrower between TNLA and FNLA as compared to that of between the SNLA and TNLA. This resistance to decline may become attributable to the current school reform program in practice by the school system. Nevertheless, this may need further research investigation in the future.
- The achievement of students became poor in all key subjects; particularly similar to the result of grade 4 the students' achievement in Mathematics was very poor (the composite average became 25.65%), where 50 percent of the students scored at or below a score of 22.5%.
- The composite average performance for Physics (34.47%) was far below 50%, where 50 percent of the students scored at or below a score of 32.5%.

- The composite average performance for Chemistry (36.42%) was far below 50%, where 50 percent of the students scored at or below a score of 32.5%.
- The composite average performance for English (36.86%) was far below 50%, where 50 percent of the students scored at or below a score of 35%.
- The composite average performance for Biology (42.10%) was far below 50%, where 50 percent of the students scored at or below a score of 40%.
- Regarding the students' performance at the various proficiency levels, the proportion of grade 8 students who have performed at the highest or proficient level showed a 2 percent increase (from 13.9 to 15.9) compared to the TNLA result. Moreover, the proportion of students who have performed at the basic level of performance increased by 3.6 percent.
- Boy students significantly outperformed girl students in the composite average as well in the four key subjects such as English, Mathematics, Biology, and Physics. Even though girls performed better than boys in Chemistry did, the mean difference was not statistically significant.
- Students of urban location found to excel rural students by the mean composite score as well as by four key subjects out of the five. The wider gap in mean scores observed for English (3.6% score points). Except for Mathematics test performance, which was the poorest performance for the majority, mean differences of scores between urban and rural students became statistically significant. This becomes quite opposite to the result of the TNLA result.
- Statistically significant achievement variations were observed across regions. Composite score averages were found to become far below 50 for students of all regions. The Scheffe test of homogenous sub test grouping for the composite score grouped students from Tigray region only with students of Addis Ababa, as they were relatively highest performing, and Gambella and Afar as relatively least performing.
- Grade 8 students from Addis Ababa region significantly vary from the rest in their English test performance (44.55%), which was the highest compared to other regions.

- In Mathematics test, mean scores were relatively homogenous across regions, where performance in all regions was poor, almost half-way from 50 and even below that for some of the regions.
- Students from Tigray and Oromia regions have mean scores above 50 in Biology test. Moreover, students from Tigray region significantly vary from the rest in their Biology test performance (52.24%), which was the highest, compared to the other regions.
- In Chemistry test performance, mean scores were relatively homogenous for the majority of the regions except that of Amahara region, which identified itself only with Benshangul Gumuz region.
- Based on their mean score in physics, the Scheffe test grouped students of the different regions into four, where Physics test performance of students from Gambella (with mean score of 28.75%) identifies them as significantly least performing.
- Personal variables such as the student gender, language used at home, and additional reading materials had significantly predicted the overall test performance. However, the variation in overall test performance as explained by the personal variables was extraordinarily low (1.1%).
- Home variables such as family size, father education, number of meals a day, and home tutorial had significantly predicted the overall test performance of students although the explained variation found to become considerably low (0.5%).
- From the school variables, only student absentees in a semester and the time it takes by the student to reach from home to school significantly predicted academic achievement. Nevertheless, the variation of test performance explained by those variables became considerably low (0.6%).
- With respect to the relationship of test performance in the key subjects and input factors such as understanding the subject, frequency of home works given in a week, and availability of textbook for the subject, only English test result significantly predicted by the linear combination of the three in puts mentioned. Chemistry test performance significantly predicted from the frequency of home

works given and availability of textbook in chemistry. Whereas test performances in Mathematics and in Physics were significantly predicted only by understanding the subjects. On the other hand, test performance in Biology significantly predicted from the availability of textbook in biology.

5.2 Recommendations

Based on the findings, the study forwards the following recommendations:

1. There has to become effort to strengthen the school reform program focusing on the six education quality indicators. Because, even though it needs further investigation in the future, the resistance to decline achievement gap between the TNLA and FNLA may become indicative of the school reform impacts. Thus, the school reform program initiated has to become further strengthened for it will contribute to pupils' learning achievement and quality education.

- The fourth national learning assessment result has shown that students' test performance became far below the national minimum achievement standard. Because of this, there have to be coordinated efforts of all stakeholders to improve the performance of the school system. Particularly school personnel should refocus on the education quality indicators and implement them towards the attainment of the curriculum goals of general education at the level. Moreover, there have to be efforts to make parental and home situations, student personal behaviors, and school variables to be facilitative of enhanced learning and improved academic achievement of students.
- As observed from the findings of the FNLA result, students' test performance in Mathematics and the sciences became very poor. Thus, the education of Mathematics and the sciences should receive special emphasis in all deliberations and reforms of the school system.

- Efforts of the school personnel reforms should emphasize on continuous professional development of teachers about innovative pedagogies, learning, and assessment techniques.

- Teacher – parent discussions on matters of student learning should be encouraged and practiced, and there is a need to emphasize on the contribution of parental involvement in the improvement of student learning and performance in tests.

- As the government already initiated and committed to reform the school system in meeting EFA goals and providing quality education, schools should become resourced with sufficient learning materials such as syllabus in mother tongue, teacher guide, textbooks in key subjects, and library facilities for these may bring in the improvement of student learning and test performance.

- Teacher reforms may focus on training them to the highest level of qualification for the cycle, on- job- trainings, close supervision and technical support, mentorship, reinforcement and recognition for best performances, and providing facility of housing to teachers near to schools become important.

- There need to introduce student and school reward scheme as part of the national learning assessment program. This may encourage students and schools to perform at the highest level possible, and creates an achievement oriented and competitive school system.

2. Gender disparity in test performance still needs the attention of the concerned stakeholders.

   - The previous learning assessment performance of grade 8 students showed gender gap in favor of boys. The fourth national learning assessment also follows a similar pattern particularly for the key subjects of English, Mathematics, and the sciences except chemistry. When all concerned stakeholders in the education system and parents give attention and make a special support to the improvement and quality of girls' education, their learning, and test performance may improve and the achievement gap will minimize in the future. The recommended actions in this respect are provisions of assertiveness training, study skills training, and tutorial classes for girls.

- Students in rural schools may need more support than what they received at present. In the fourth national learning assessment test, achievement scores are in favor of the urban students, which is the opposite to that of the second and third national learning assessment results. Thus, reform attempts at the student, family, teacher, and school level should address learning and test performance improvement of students in the rural schools. In addition, the governments' commitment on the school reform program, if strengthened and targets rural schools, the significant gap in the test achievement of urban and rural pupils will diminish.

3. There is a need to progress in academic achievement levels based on the recommendations given by the successive national learning assessments. There is a need to have a simultaneous look at the results of NLAs and school reform efforts in order to enhance and assure the of quality education.

   - The proportion of grade 8 students attaining at the proficient level of performance increased when compared to result of third national learning assessment. However, schoolteachers, students, parents, and all others concerned should strive towards decreasing the proportion of pupils at the below basic level of performance, which is considerably large in proportion compared to the proficient level of performance.

   - Parents should be encouraged and consulted by schoolteachers and directors to avail home materials and provide assistances that could support in the improvement of student learning and achievement.

   - The school – community relationship scheme should focus in the encouragement of parents to avail materials for their child that may support in the improvement of student learning and test performance when appropriately used. Such materials may include additional reading materials other than textbooks, radio, TV set and so on.

4. Housing for teachers and school directors should built around schools. Local governments and the community should consider facilitating conditions of housing for teachers and school directors near and around schools. Traveling long distance from home to school may take the time and energy of teachers and directors, which in turn negatively influence teaching effectiveness and student learning achievement.
5. There should be emphasis to improve students' performance of the least performing regions. The disparity of students' test performance was a threat to equitable provision of quality education. Thus, school systems in the least performing regions should pay special attention to the improvement of student learning and academic achievement to minimize the huge gap. Moreover, the implementation of the school reform program needs to give particular attention to the regions which have relatively large proportion of pupils performing poorly in the FNLA

6. Priorities for comprehensive school reform have to be holistic. Attempts of school reforms in the provision of quality general education at the level and students' learning improvement and achievement necessitates consideration of the variables that have positive influence on performance. As discussed in this study the variables of interest that could influence student learning and achievement are student variables, school variables, teacher variables, and home variables.

7. Achievement gaps within (government) schools and regions should be addressed and minimized in order to maintain equity in student learning and achievement.

Unexpectedly, the current findings have shown that the contribution of inter-school differences to differences in academic performance has been high. This study has taken the samples from government schools. The implication is that there are wide variations among government schools in terms of students test performance. Therefore, the Ministry of Education and other stakeholders in the sector have to give priority to creating equal access for all schools in all regions to conditions that facilitate the improvement of quality education in general and students' academic performance in particular.

References

- Aslam, M. and Kingdon, G. (2008). What can teachers do to raise pupil achievement? Research Consortium on Educational Outcomes and Poverty: University of Oxford.
- Bishop, J.H. (1996). The impact of curriculum based external examinations on school priorities and student learning. International Journal of Education Research.
- Cinapah, V. et.al. (2000). With Africa for Africa: Towards quality education for all. Pretoria: Human Sciences research Council.
- Chinapah, V. (2003). Monitoring Learning Achievement (MLA) project in Africa. Mauritius: Conference paper presented at ADEA.
- Coleman, J.S. (1990). Foundations of Social Theory. Cambridge, MA: Belknap.
- Erkyhun, D and collegues. (2004). ETHIOPIAN SECOND NATIONAL LEARNING ASSESSMENT OF GRADE FOUR STUDENTS , Addis Ababa
- Erkyhun, D and collegues. (2008). ETHIOPIAN THIRD NATIONAL LEARNING ASSESSMENT OF GRADE FOUR STUDENTS , Addis Ababa
- FDRE. (1994). Education and Training Policy (first edition). Addis Ababa.
- Fuller, B. (1985). Raising school quality in developing countries: What investments boost learning? Washington, D.C: The World Bank.
- Gonzales, P. et.al. (2009). Highlights from TIMMS 2007: Mathematics and Science Achievement of U.S fourth and Eighth- Grade Students in an international context.
- Hallack, J. (1990). The analysis of educational cost expenditure. Paris: UNESCO, IIEP.
- Heynman, S. and Loxley, W. (1982). Influences on high academic achievement across high and low income countries: A re-analysis of IEA data. Sociology of education 55(1), 13 – 21.
- Kellaghan, T. (2004). Public Examinations, national and international assessments and educational policy. Dublin: St Patrick's College.

- Kingdon, G. (1999). How much do Schools Matter to Pupil Achievement in India. University of Oxford. Download from http://www.JEPA.school pupil achievement
- MOE. (2005). Education Sector Development Program III (ESDP – III), 2005/06 – 2010/11, Program action Plan (PAP). Addis ababa
- OECD (2010), PISA 2009 Results: What students know and can do – Student performance in Reading, Mathematics, and Science (Volume I). http://dx.doi.org
- Raymond S. Nickerson (1989). New Directions in Educational Assessment. Educational Researcher, Vol. 18, No. 9, pp. 3-7
- Symeou, L. (2007). Cultural capital and family involvement in children's education: Tales from two primary schools in Cyprus. British Journal of Sociology and Education, 28 (4), 473 - 487
- Wiliam, D. (2000). The meanings and consequences of educational assessments.

  Critical Quarterly, 42 (1), 105 - 127
- Wolff, L. (1998). Educational Assessment in Latin America: Current Progress and Future Challenges, Newton: Massachusetts.
- Wright, et.al. (1997). Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation. University of Tennesse: value – Added Research and Assessment center.