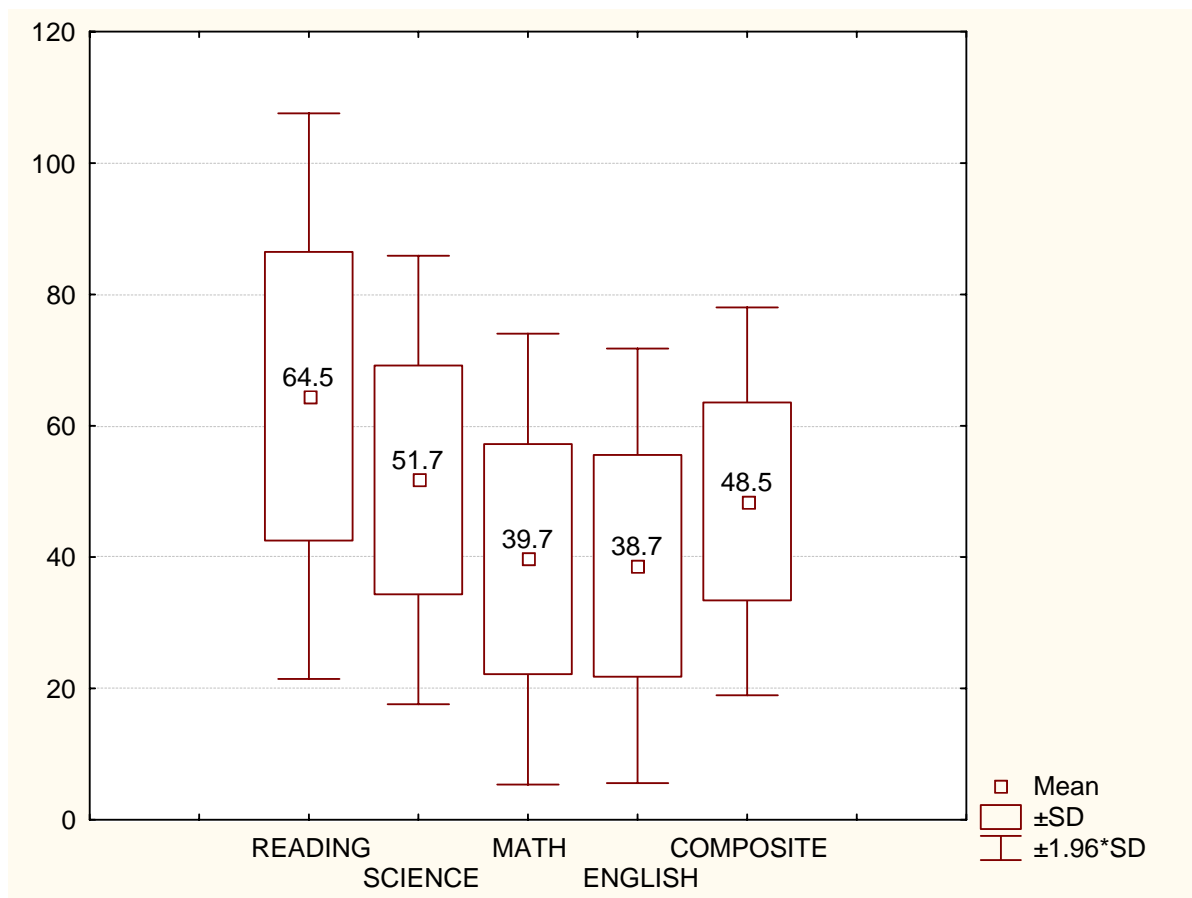


ETHIOPIAN SECOND NATIONAL LEARNING ASSESSMENT OF GRADE 4 STUDENTS



National Organization for Examinations

Addis Ababa

December 2004



Funded under USAID Contract No. 663-C-00-02-00349-00

BASIC EDUCATION STRATEGIC OBJECTIVE



USAID | ETHIOPIA
FROM THE AMERICAN PEOPLE

Technical Working Group

<i>Erkihun Desta</i>	<i>Chairman</i>	<i>NOE</i>
<i>Endale Shewangizaw</i>	<i>Secretary</i>	<i>>></i>
<i>Getahun Worgineh</i>	<i>Member</i>	<i>>></i>
<i>Hailu Chane</i>	<i>>></i>	<i>>></i>
<i>Kasu Lewdie</i>	<i>>></i>	<i>>></i>
<i>Kefeleghn Tsegie</i>	<i>>></i>	<i>>></i>
<i>Yhannes Afework</i>	<i>>></i>	<i>>></i>
<i>Lewdu Gebrekidan</i>	<i>>></i>	<i>>></i>
<i>Teshome Nekatibebe</i>	<i>Technical Advisor</i>	<i>AED/BESO II</i>

A national assessment may be defined as an exercise designed to describe the level of achievement, not of individual students, but of a whole education system or a clearly defined part of one (such as grade 4 pupils or 11 – year- olds). (Kellaghan and Greany, 2004)

Table of Contents

List of Tables	iv
Acronyms	vi
Acknowledgments	vii
CHAPTER ONE: INTRODUCTION	1
1.1 <i>Ethiopia</i>	1
1.2 <i>Purpose of the Study</i>	3
1.3 <i>Key Research Questions</i>	4
1.4 <i>Specific Objectives of the Study</i>	4
1.5 <i>Significance of the Study</i>	5
1.6 <i>Limitations of the Study</i>	6
1.7 <i>Definition of Terms</i>	7
CHAPTER 2: A REVIEW OF RELATED LITERATURE	9
2.1 <i>Introduction</i>	9
2.2 <i>The Concept of National Learning Assessment</i>	9
2.3 <i>Purposes and Functions of National Learning Assessment</i>	10
2.4 <i>Methodological and Technical Considerations in National Learning Assessment</i>	11
2.5 <i>Experiences of Other Countries Regarding National Learning Assessment</i>	12
2.6 <i>Background and Community Factors that Influence Student Achievement</i>	15
2.7 <i>School Based Factors that Influence Student Achievement</i>	16
2.8 <i>Historical Overview of Learning Assessments in Ethiopia</i>	18
CHAPTER THREE: FRAME OF ANALYSIS AND METHODOLOGY	20
3.1 <i>Frame of Analysis</i>	20
3.2 <i>Design</i>	21
3.3 <i>Sampling Procedures</i>	21
3.4 <i>Instruments of Data Collection</i>	25
3.5 <i>Validity and Reliability of Instruments</i>	26
3.6 <i>Data Collection and Administration of Instruments</i>	27
3.7 <i>Methods of Data Analysis</i>	28

CHAPTER 4: FINDINGS OF THE STUDY	29
4.1 <i>Findings on Grade 4 Achievement Tests</i>	29
4.2 <i>Overall Achievement of the Students by Subject</i>	31
4.3 <i>Grade 4 Pupils' Background and Academic Achievement</i>	46
4.4 <i>Grade 4 Pupils' Attitudes in Socially Relevant Issues</i>	51
4.5 <i>Relationships between Grade 4 Pupils' Achievement and Factors</i>	66
4.6 <i>Comparative Analysis between EBNLA and ESNLA</i>	72
4.7 <i>Analysis of Qualitative Data</i>	77
CHAPTER 5: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	89
5.1 <i>Summary</i>	89
5.2 <i>Recommendations</i>	95
REFERENCES	97

List of Tables

Table 1. Description of variables	20
Table 2. Sample schools planned and visited by region	23
Table 3. Sample students, teachers and directors by region	24
Table 4. Participants of focus group discussions by region.....	24
Table 5. Number of participants involved in field work by route	28
Table 6. Summary of the descriptive statistics for all subjects at national level	29
Table 7. T-test for equality of means by gender at national level.....	30
Table 8. T-test for equality of means by location at national level	30
Table 9. Summary of the descriptive statistics for reading comprehension by region	31
Table 10. One- way analysis of variance on reading comprehension by region	31
Table 11. Homogenous subset groups for reading comprehension	32
Table 12. T-test for equality of means for reading comprehension by gender within regions	32
Table 13. T - test for equality of means for reading comprehension by location within regions.....	33
Table 14. Summary of the descriptive statistics for English by region.....	33
Table 15. One way analysis of variance of English across regions	34
Table 16. Homogeneous subset groups of English by region	34
Table 17. T-test for equality of means for English by gender within regions	35
Table 18. T-test for equality of means in English by location within regions	36
Table 19. Summary of the descriptive statistics for mathematics by region.....	36
Table 20. One way analysis of variance on mathematics by region	37
Table 21. Homogenous subset groups of mathematics by region	37
Table 22. T-test of mean difference for mathematics by gender within regions	38
Table 23. T-test of mean difference for mathematics by location within regions.....	39
Table 24. Summary of the descriptive statistics for environmental science by region	39
Table 25. One way analysis of variance on environmental science by region.....	40
Table 26. Homogenous subset groups of environmental science by region.....	40
Table 27. T -test of mean difference for environmental science by gender within regions	41
Table 28. T-test for mean difference in environmental science by location within regions	42
Table 29. Summary of the descriptive statistics for composite mean by region	42
Table 30. One way analysis of variance on composite mean by region	43
Table 31. Homogenous subset groups of mean composite score by region	43
Table 32. T -test of mean difference for composite mean by gender within regions.....	44
Table 33. T-test of mean difference for composite mean by location within regions	45
Table 34. Time needed to reach school and academic achievement.....	46
Table 35. ANOVA on time required to reach school	46
Table 36. Watching non-educational television and student achievement	47
Table 37. Listening to the radio (other than educational radio) per week	47
Table 38. Absenteeism and achievement	48
Table 39. Frequency at which attendance is taken.....	48
Table 40. Daily meals and achievement.....	48
Table 41. Interest in learning English.....	49
Table 42. Interest in learning mathematics	49
Table 43. Interest in learning environmental science	49
Table 44. English textbook distribution.....	50
Table 45. Mathematics textbook distribution	50
Table 46. Environmental science textbook distribution	50
Table 47. Attitude towards health care.....	52
Table 48. Attitude towards environmental protection.....	52
Table 49. Attitude towards civics and ethics.....	53
Table 50. Attitudes towards harmful cultural practices.....	53
Table 51. Attitude towards education.....	53
Table 52. We can prevent the spread of diseases through personal and environmental hygiene	54
Table 53. I would like my parents to give birth over and over again so that I can have many brothers and sisters.	54
Table 54. As far as we are aware of the modes of transmission of HIV/AIDS and take care, we will not be infected by the virus.	55
Table 55. Blood tests cannot help identify HIV/AIDS infection	55
Table 56. Helping HIV/AIDS infected people is a social obligation	55
Table 57. Learning with those with HIV/AIDS causes no problem	56

Table 58. Eating meat regularly is a proper dietary practice.....	56
Table 59. Drinking river water can expose one to water borne diseases.....	56
Table 60. Visiting historical places is wasting both money and time.....	57
Table 61. Gardening gives me pleasure.....	57
Table 62. One should plant trees to prevent erosion.....	58
Table 63. Keeping the environment clean is everyone's duty.....	58
Table 64. It is possible to prevent the consequences of drought.....	58
Table 65. It is necessary to burn forests to drive away wild animals that destroy crops.....	59
Table 66. Women are equal to men in knowledge and thinking.....	59
Table 67. It is appropriate to return property found lost either to its owner or to the school administration.....	60
Table 68. I believe that group work is important for academic success.....	60
Table 69. Gathering information about our surrounding is necessary to have knowledge about it.....	61
Table 70. Solving differences through dialogue is good culture.....	61
Table 71. Boys and girls should equally participate in household chores after school.....	62
Table 72. Early marriage is a practice that should be abolished.....	62
Table 73. Female circumcision is a harmful traditional practice that should be abolished.....	63
Table 74. Absenteeism and coming late to school are signs of laziness.....	63
Table 75. Student participation in teachers' evaluation is appropriate.....	64
Table 76. Going to school and staying at home are the same to me.....	64
Table 77. Learning is a means to earn a higher status.....	65
Table 78. Because I am a student, I am more knowledgeable and mature than those not in school.....	65
Table 79 Spearman Rho Correlation with Grade 4 Composite Score.....	66
Table 80. School-level multiple regression results on the Grade 4 composite achievement score details on reduced models.....	68
Table 81 School level multiple regression results on the composite score model summary.....	69
Table 82 Contributions of each block to R-square of completely reduced model.....	70
Table 83. G4 EBNLA national baseline mean score (by the end of 1992 E.C.).....	72
Table 84. G4 ESNLA National mean score (by the end of 1996 E.C.).....	72
Table 85. Learning Progress: Comparison between the two national learning assessments.....	73
Table 86 Reading: average change in percent correct scores from Baseline to SNLA.....	74
Table 87. English: average change in percent correct scores from Baseline to ESNLA.....	75
Table 88. Mathematics: average change in percent correct scores from Baseline to SNLA.....	75
Table 89. Environmental science: average change in percent correct scores from Baseline to SNLA.....	76
Table 90. Average difference in percent correct scores across Grade 4 achievement tests.....	76
Table 91. Student learning as viewed by teachers, students and parents by region.....	78
Table 92. School supplies and school facilities.....	80
Table 93. Teacher behavior, availability and competence.....	81
Table 94. Student behaviors influencing academic achievement.....	83
Table 95 The influence of the curriculum on students achievement.....	85
Table 96. The influence of parental support on school achievement in Ethiopia.....	87
Table 97. Summary of Grade 4 achievement scores by region.....	90

Acronyms

BESO I:	Basic Education System Overhaul
BESO II:	Basic Education Strategic Objective
ENBA:	Ethiopian National Baseline Assessment
ERGESE:	Evaluative Research of the General Education System in Ethiopia
ESDP:	Education Sector Development Program
ESR:	Education Sector Review
ETP:	Education and Training Policy
FDRE:	Federal Democratic Republic of Ethiopia
ICDR:	Institute of Curriculum Development and Research
EMIS	Educational Management Information Service
IEA:	International Association for the Evaluation of Educational Achievement
IIEP:	International Institute for Education Planning
MLA	Monitoring Learning Achievement
MOE:	Ministry of Education
NAC:	National Advisory Council
NCES	National Center for Education Statistics
NAEP:	National Assessment of Educational Progress
NLA:	National Learning Assessment
ESNLA:	Ethiopia Second National Learning Assessment
NOE:	National Organization for Examinations
PISA:	Program for the International Student Assessment
EFA	Education for All
REB:	Regional Education Bureaus
SACMEQ:	South African Consortium for Monitoring Educational Quality
SNNPR:	Southern Nations, Nationalities and People's Region
SPC:	School Pedagogical Center
TWG:	Technical Working Group
USAID:	United States Agency for International Development

Acknowledgments

A number of people and institutions have contributed to the successful completion of the Second Ethiopian National Learning Assessment. AED/BESO II provided the required financial and technical assistance for carrying out the study. Dr. Elizabeth Randolph and Mr. Benjamin Piper have provided an immense technical support as external consultants to the project. Dr. Randolph has displayed a great deal of dedication in the design of the study, analysis of data, and training of personnel and provision of resource materials required for the assessment. Mr. Piper from Harvard University has provided a technical support in the analysis of statistical data. Ato Mulu Nega from the Institute of Educational Research (Addis Ababa University) has contributed to the study on several occasions as a local consultant. Ato Gizaw Tasissa of the Department of Curriculum and Instruction (Addis Ababa University) has assisted in the analysis of qualitative data upon request. Due acknowledgment also goes to Ato Bedru Kedir and Ato Fiseha W/Michael who contributed to writing the report and preparing guidelines for field work respectively.

Members of the National Advisory Council of the Ethiopian Second National Learning Assessment devoted their time to advise and support the Technical Working Group. In addition to his duties as a chairman of the National Advisory Council of the Second Ethiopian National Learning Assessment, His Excellency Ato Dereje Terefe, Vice Minister of Education, has constructively responded on different occasions to our requests to provide comments on draft reports. There is no doubt that this study would not have been completed without the full support of the data collectors from NOE and regions, in particular Zone and Woreda Education department heads, school directors, teachers, parents and students. Finally, USAID/Ethiopia deserves special thanks for making the whole study possible in cooperation with the Ministry of Education of Ethiopia.

CHAPTER ONE: INTRODUCTION

1.1. Ethiopia

Ethiopia, covering an area of 1,133,380 sq. km (Keller, 2004), is located in the North Eastern part of Africa. The country is bounded in the North by Eritrea and Djibouti, in the East and South East by Somalia, in the South West by Kenya, and in the West and North by Sudan.

Located in the sub tropical zone, Ethiopia possesses varying topography mainly plateaus, highland chains at the center, and low lands at the coastal areas. The climate varies from very hot to very cold temperature and from arid to high rainfall areas following the topography. The principal rainy season of the country is between June and September, whereas dry season is from October to May with short lived and weak rain in February and March.

Wide variations in land structure, soil type and climate are the main determinants for the diversified agricultural products, vegetation, wild life and mineral resources in Ethiopia. The nature of the plateaus and other factors like deforestation, over grazing, poor land management and the cultivation system in the high lands accelerate the degree of erosion of the top soil causing low productivity.

1.1.1 Population and Culture

With an estimated population of 68.6 million (World Bank, 2003), Ethiopia is the third populous country in Africa next to Nigeria and Egypt. Different nationalities and ethnic groups having diversified cultures and more than 80 spoken languages inhabit the country. Of these languages, about 22 are used as media of instruction in schools. Amharic is the Federal Government's working language. There are two dominant religions, Christianity and Islam, which have a long history and many followers in the country.

Since August 1995, Ethiopia has become a Federal Democratic Republic constitutionally. It comprises 9 regional states and three city administrations. 4 of the regions are not only large in size but also have better infrastructure. Two of the regions are city administrations while the rest are considered as emerging regions.

1.1.2 Economy

Traditional agricultural and pastoralist practices employ about 80% of the population. The most important food crops grown primarily for local consumption are cereal grains. Coffee, cotton, oil seeds, pulses, fruits, vegetables as well as leather and skin are some of the commercial products used for internal processing as well.

In addition to the backwardness and the low productivity of the agricultural sector, the mining and industrial sectors are also at a lower level of development. But some promising potentials and developments are being observed around the production of hydro-electricity, airlines and other communication sectors. The main financial sources of the government are some local revenues, loans and donor funds. The Ethiopian currency is called Birr; about 8.60 Birr is equivalent to one US dollar.

1.1.3 Education System

Ethiopia is one of the least developed countries in the world. Social indicators (health and education) have remained at a lower level compared to other Sub-Saharan countries. The country is hard hit by recurrent drought. About 44% of the total population is living below poverty line (MOFED, 2002). Infant mortality rate is relatively high. Life expectancy is about 42.1 years (World Bank, 2002) with a declining trend due to HIV/AIDS.

The government of Ethiopia (FDRE) has set up various policies and strategies that could alleviate poverty and guide the overall development of the society. Special focus in the development policies has been given to rural and agricultural development. In general, there are 4 specific areas of emphasis in the development strategy of the government: agriculture led industrialization, justice and civil service reform, government decentralization and empowerment, and capacity building in public and private sectors.

The strategy and program framework of capacity building, in particular, focuses on education in order to increase the production of trained his manpower and upGrade the competence of the existing labor force in Ethiopia. Special attention in the development of the education and training system has been given to ensure access to primary education for all citizens, and simultaneously improve the quality and relevance of the sector for the expected socio-economic development.

To this end, the government of Ethiopia issued a new Education and Training Policy (ETP) in 1994. Overall, ETP was a response to the observed crises in the Ethiopian education sector, and it aimed at ensuring equity, accessibility, quality and relevance of the education. The following strategies were formulated to realize the goals of ETP:

- Restructuring of the education system (new educational structure)
- Administrative change through decentralized management
- Curriculum change – development of new curriculum
- Use of local languages as media of instruction
- Changes in teacher training and professional development opportunities (new teaching career structure)
- New approaches to classroom and national examinations
- Change in educational finance (introducing cost sharing in upper secondary and tertiary education)

Particular priority areas of the ETP are curriculum change, decentralization and teacher education.

The Ministry of Education has formulated a 20-year development strategy to implement the ETP. The Education Sector Development Program I (ESDPI), implemented from 1997-2002, was the first phase of the 20-year program followed by the ESDP II (2002-2005).

The overall objectives of the ESDP II regarding primary education focus on the inculcation of good citizenship and actualization of the goal of universal primary education by the year 2015.

The program has also set directions regarding assessment and quality improvement. These include:

- conducting assessment studies on selected languages of nations and nationalities that are used as media of instruction,
- developing learning competency guidelines to assess the profile of education standards at each level for each subject,
- producing manuals for promoting continuous assessment, and
- carrying out national educational assessment on Grades 4 and 8 to assess the skill development level of students in key subject areas.

The implementation of the ETP has been underway for the last 12 years. One major measure taken is the decentralization of educational management following the federal structure of administration. According to this arrangement, all educational institutions, except the universities, are under the jurisdiction of the regional governments.

The previous twelve years of general education structure (6+2+4) has been changed into ten years. The new structure comprises eight years of primary education subdivided into two cycles of basic education (1-4) and general education (5-8); a general secondary education (9-10) with a terminal (school leaving) examination and a preparatory education (11-12). Those who could not join the preparatory program go to technical and vocational training that takes one to three years to complete. Those who successfully complete the preparatory program go to higher education of undergraduate studies (3-6 years).

The implementation of ETP has brought about encouraging results in terms of access to primary education and schooling in the country. Some of the milestones include the following.

- The Gross Enrolment Rate for primary education (1-8) increased from 45.8% to 64.4% between 1998/99 and 2002/03. The increase has been 18.7% for boys, 18.5% for girls and 18.6% for the total.
- The Net Enrolment Rate for primary education has increased from 39.6% in 1998/99 to 54% in 2002/03. The gender gap in net enrolment rate has decreased from 15.1% in 1998/99 to 13.4% in 2002/03.
- The proportion of qualified teachers has reached 30.9% in 2002/03 for upper primary (5-8).
- The pupil teacher ratio for primary education has risen from 51:1 to 64:1 between 1998/99 and 2002/03, which is a 25.5% increase. This indicates a rapid growth of enrolment rate at this level.
- The education share of government budget expenditure rose from 15.6% in 1997/98 to 18.8% in 2002/03.

1.2 Purpose of the Study

Education is moving from being a privilege for the few to becoming the right for all. However, this quantitative expansion has brought about serious challenges to its quality. Quality does not mean only what goes into schools, but also what goes in the mental and physical changes of children. It is important to develop the knowledge, skills, attitudes and habits of pupils in addition to giving emphasis to input factors.

Some developing countries have tried to assess and measure student achievement and improve their educational systems. Nevertheless, most countries still apply public

examinations for certification, selection and promotion. Improving student learning has remained one of the most desired goals of educational processes.

In Ethiopia, quality assurance has been an important part of the reform process. To this end, the Ethiopian National Baseline Learning Assessment for primary education was carried out in 1999/2000. Currently, the Second National Learning Assessment is taking place in the country.

The main purpose of conducting the Second National Learning Assessment is to provide information about learning attainments by students and the factors that determine those attainments in the Ethiopian primary education so that attention is paid to the improvement of the system as a whole.

1.3 Key Research Questions

1. To what extent did Grade 4 students achieve the stated curriculum in key subjects and to what degree does their performance vary across regions, gender, location and medium of instruction?
2. What do the students' background information and interest look like in relation to their overall achievement?
3. What do Grade 4 students' general attitudes, beliefs and preferences look like in relation to pro-social behavior and socially relevant issues at national and regional levels?
4. What are the factors that influence students' achievement in the primary schools of Ethiopia?
5. Is there any progress from the baseline regarding pupils' learning achievement?
6. What are the qualitative assessments and judgments of different groups (directors, teachers, students and the community) on the efficiency, problems and solutions concerning student learning in schools?
7. What are the possible implications of the findings of the study for improving student performance and school quality in Ethiopia?

1.4 Specific Objectives of the Study

The Ethiopian Second National Learning Assessment of students in Grade 4 has the following specific objectives:

1. Analyze the national student learning achievement results in Grade 4 mathematics, English, environmental science and basic reading comprehension and student attitudes in socially relevant issues
2. Analyze Grade 4 students' achievement in mathematics, English, environmental science and basic reading comprehension results by gender, location, and region
3. Find out Grade 4 students' general attitude towards socially relevant issues at national and regional level
4. Determine the relationship between Grade 4 students' background variables and their overall achievement in the given subjects
5. Determine the level of Grade 4 students' learning progress from baseline by comparing scores obtained in the first and second national learning assessment

6. Provide baseline data for Grade 4 student attitudes on socially relevant issues
7. Explain the factors that influence Grade 4 students' achievement
8. Assess the opinions and judgments of different groups: directors, teachers, students and the community on the efficiency and problems of learning in schools
9. Summarize the implications of the findings of the Second National Learning Assessment for the improvement of school quality and effectiveness in Ethiopia

1.5 Significance of the Study

Student learning assessment involves a systematic process of collecting relevant, valid and timely information about the outcomes of schooling so that decisions are made about the learning and development of students, curriculum, educational programs and educational policy. Student learning assessment provides the necessary feedback and objective evidence required to maximize the outcomes of educational efforts. Such assessments summarize what learners know, understand, and can do in relation to some or all of the learning goals determined in the curricula.

Over the last decade, substantial attempts have been made to expand primary education, and improve access, equity and efficiency in Ethiopia. Now the emphasis has shifted towards improving quality in all areas and in particular towards student learning achievement. This national learning assessment, therefore, provides an indication or feedback of where students' achievement in the country stands in relation to the stated profiles of the curriculum.

A student learning assessment can provide baseline information from which progress can be measured during and at the end of a key stage in education. Since it focuses on actual learning, it enables one to find out the extent to which an educational system is effective as a whole. If it is properly integrated into the system of education, student learning assessment can help actors and stakeholders to focus their collective attention, examine their assumptions, and create a shared academic culture dedicated to assuring and improving the quality of education. The Ethiopian Second National Learning Assessment is a nationwide program and a continuation of the Ethiopian National Baseline Learning Assessment. The first national assessment has provided a benchmark from which improvement can be measured. In this respect, the Second National Learning Assessment will serve as a key tool for monitoring changes or improvements since the time the Ethiopian National Baseline Learning Assessment was conducted. Since it includes new subjects, it also creates baseline data.

The Education and Training Policy of the Federal Government decentralizes education in the sense that regions plan and administer primary education under the guidelines and standards set by the Ministry of Education. Moreover, the policy states that primary school children should learn in their mother tongues. Regions implement the Education and Training Policy by taking into consideration their own specific conditions. This implies that some of the features of these regions affect the practice of primary education in relation to curriculum development, material preparation, teacher education, school management, teacher practices, school-community relations, language of instruction etc., and the extent to which students learn from their schooling. The Ethiopian Second National Learning Assessment contributes to monitor how expected national standards have been implemented and if each of the regions has developed realistic mechanisms to convert national guidelines into local tools for school development.

Information on the relationship between student learning outcomes and school inputs provides an immense potential to policy makers to identify, allocate and manage the resources of education to improve quality. The Ethiopian Second National Learning

Assessment provides such information along side the achievement results so that the most influential determinants of learning are properly known and managed.

Ethiopia expends a considerable amount of its public finance on education. In order for the education sector to justify this expenditure and retain support, both the government and the public require that the money expended produce the required skills. The Ethiopian Second National Learning Assessment provides access to such relevant information and this, hopefully, promotes accountability in the system.

1.6. Limitations of the Study

Tests were more of an objective type and this did not allow the measurement of high order thinking and skills students acquired in schools. The curriculum intends several proficiencies other than the ability to recognize and know the proper answers from the given alternatives. Future studies may consider the measurement of various domains of learning as much as possible.

Sample schools were not visited before the field work started. The consequence had been that current changes in the status of the schools demanded the replacement of few of the sampled schools. Data obtained from EMIS showed that some schools were full primary schools having Grades 1-8, but it was found out during the field study that these schools had only Grades 1-4.

Gambella region was not included in the study due the problem that prevailed in the area. The inclusion of such an emerging region would have provided further insight into how schools function in the relatively less developed regions of the nation.

1.7. Definition of Terms

National learning assessment: Sometimes is known as system assessments, assessments of learning outcomes or learning assessments. This activity may be defined as an exercise designed to describe the level of achievements, not of individual students, but a whole of the education system or a clearly defined part of it (Kellaghan and Greany, 2004). In the Ethiopian Second National Learning Assessment, the main exercise was to measure the level of student achievement of the whole primary education system by taking Grades 4 and 8 as particular parts for investigation. The assumption was that Grade 4 was a terminal point for investigating the outcomes of the first cycle primary education and Grade 8 for investigating the outcomes of the whole primary sector of education.

Student home background: There is the recognition that student home background variables are important determinants of student achievement. Chen (1996) reviews a lot of literature to identify three sets of variables related to home background that influence student achievement. These are home socio-economic status (SES), family configuration and parental support. Home socio-economic status is measured using such indicators as parental occupation, level of parental education, parental income and the prestige of parental occupation. Family configuration measures include family size, sibling sex and spacing, and birth position in a family. Measures of family support are both psychological and practical and include such factors as academic guidance and support, stimulation to explore and discuss ideas and events, language environment, academic aspirations and expectations, and work habits of a family. In the Ethiopian Second National Learning Assessment, student home background variables included the agreement between home and instructional language, student chore time, distance to school, number of family members in school, availability of television at home, availability of radio at home, and daily meals.

Student behavior: Activities, thoughts and feelings students demonstrate during their learning are included in this conceptual construct in the Ethiopian Second National Learning Assessment. Notable variables in this category included students' sense of ownership in school property, time spent listening to the radio, time spent watching television, leisure reading (reading non-textbook materials), interest in English, interest in mathematics, interest in chemistry, interest in biology, interest in physics and the use of school library.

Government and non-government primary schools in Ethiopia: Primary schools in Ethiopia are divided into government and non-government schools in terms of ownership. Government primary schools are administered and controlled by the government. Non-Government schools are all schools controlled by institutions other than the government, and these may include religious institutions, NGOs, the private sector, communities, etc. In the Ethiopian Second National Learning Assessment government schools are those schools that are administered and controlled by the government of Ethiopia.

Levels/cycles of primary schooling in Ethiopia: The Ministry of Education (2004) classifies the levels or cycles of primary school in Ethiopia into two structures. These are the First Cycle Primary School (Grades 1-4) structure and the Second Cycle Primary School (Grades 5-8) structure. The Ethiopian Second National Learning Assessment used this definition as a reference for its operations. But, schools can also be First Cycle to comprise Grades 1-4, and Full Primary comprising Grades 1-8.

School structure: Chen (1996) includes three main variables in school structure. These are school size, class size and cycles of schooling. School structure in the Ethiopian Second National Learning Assessment was used to include the following variables: time to reach the main road from the school, time to reach the Woreda education office from the school,

availability and condition of school pedagogical center, availability of football and volleyball fields, availability and conditions of school library and school pedagogical center, income generated from the sale of hay and vegetables, and school location.

Curriculum materials: These are materials that support the implementation of the curricula. Curriculum materials are those resources that serve as daily guides for students and teachers in directing activities related to instruction and learning. In the current study, curriculum materials include the syllabus, teacher's guide, student textbooks and reference books.

Teacher variables: Fuller (1986) classifies teacher variables into teacher quality and teacher behavior in the classroom. Teacher quality variables include schooling, social background, verbal proficiency, and motivation of teachers. Teacher behaviors in the classroom include the efficient use of instructional time, the level of performance standards and expectations set for students, the extent to which teachers evaluate students' performance, and teachers' ability to motivate students to learn. In the Ethiopian Second National Learning Assessment, teacher variables include both quality and behavior indicators. Quality variables include teacher qualifications, teachers' total experience, experience at the school, distance to school, teacher training on new syllabus, teacher training on new teaching methods, teacher training on new assessment techniques, teacher training on student discipline and teacher training on textbooks. From teacher behaviors, teacher perceptions or judgment of textbook difficulty and student learning attitudes are included.

School management: is to mean the responsibility of the director to plan, organize, lead and control schools to work well (Sergiovanni, 1995) as well as his own personal quality as a leader (Fuller, 1986). In the Ethiopian Second National Learning Assessment, this component includes director's qualification, teaching experience, social obligations outside of directing, director's supervising of teachers, the distance the director has to walk home, director's meeting with teachers, and director's ability in generating funds from different sources.

School quality: The term "school quality" has no agreed upon definition. However, in relation to the Ethiopian Second National Learning Assessment, school quality refers to the outcomes of education as measured by student's level of academic performance or achievement.

Instruction: This refers to teachers' teaching practice and classroom organization (Chen, 1996, Fuller, 1996). Such variables as efficient use of instructional time, teachers' expectations of pupils' performances, time spent for preparing lessons, frequency of homework, marking students' homework, keeping order in the classroom, organizing feedback in a systematic way, use of correctives in helping children to learn what they failed to learn are studied in this variable. In the Ethiopian Second National Learning Assessment, instruction included teacher's instructional quality, the sum of homework assigned in subjects by teachers, student understanding of subjects, parental academic support, and the percentage of contents of curriculum taught so far.

Language of instruction: A medium in which instruction takes place in the classroom.

CHAPTER 2: A REVIEW OF RELATED LITERATURE

2.1 Introduction

The main purpose of education, especially of primary education, is to enhance economic and social development of a country by creating learning opportunities at individual, community, and national levels, and to expand literacy and give basis for further training and self-education. To attain such major aims, various countries have been designing and employing different strategies for expanding access and improving quality of schooling.

In developing nations, the actual reality of schooling is getting worse from time to time due to the decreasing educational expenditure, on one hand, and the rapid increase of enrolment rates on the other. This in turn contributes to the poor quality of schooling in general and to the low level of students' achievement in particular. Consequently, there is a growing awareness and shift of attention towards improving students' achievement and the quality of education. To this end, educational research has become the best tool to identify factors (variables) that can affect students' achievement and seek ways to design, implement and then measure improvements.

2.2 The Concept of National Learning Assessment

Educational systems are known for public examinations, but national learning assessments are relatively new (Kellaghan and Greany, 2004). The importance of monitoring learning achievement grew rapidly after the 1990 World Declaration of Education for All (EFA) in Jomtien. This declaration necessitated the introduction of a system or national assessment to determine if children were acquiring useful knowledge, reasoning ability, skills and values that schools promised to deliver.

The term assessment is used to refer to the process of gathering, interpreting, and applying outcomes data on programs or entire curricula to improve program effectiveness, particularly as measured by student learning outcomes. It is an ongoing process aimed at understanding and improving student learning. It involves making expectations explicit; setting appropriate criteria and high standards for learning; systematically gathering, analyzing, and interpreting evidence to determine how well performance matches those expectations and standards, and using the resulting information to document, explain, and improve performance (Angelo, 1999).

Learning outcomes are changes that occur within the student, and not what the instructor does. That is, learning outcomes are a level of knowledge, skills, abilities that a student has attained. They are essential and enduring knowledge, abilities (skills), and attitudes (values, dispositions) that constitute the integrated learning needed by a graduate of schools or programs.

National learning assessment may be described as a systematic process of collecting relevant, valid and timely information about the outcomes of schooling used for making decisions about the development and learning of students. According to Kellaghan and Greaney (2001), it is an exercise designed to describe the level of achievement, not of individual students, but of a whole education system, or a clearly defined part of it. That is, national learning assessment is meant to discover how well an educational system is progressing in general and

students are acquiring the knowledge and skills delivered by the educational system. Greaney and Kellaghan (1996) indicate that all national assessments seek answers to one or more of the following questions:

1. Do particular groups in the population perform differently? Are there disparities between students' achievements of different regions, of boys and girls, of rural and urban locations, and of different language groups?
2. How well are students learning in the education system? (aims of curriculum)
3. Is there evidence of particular strengths and weaknesses in students' learning?

National learning assessment emphasizes the measuring of students' acquisition of knowledge, level of understanding, attitudes, values, skills, satisfactions, actions and intellectual growth.

2.3 Purposes and Functions of National Learning Assessment

National Learning Assessment seeks to determine how well students are learning, and it is an integral part of the quest for improved education. National assessments are systematic, regular measures of learning achievements in a country that are designed to assist policy making. Indeed, national learning assessment provides feedback not only to policy makers, but also to educators, parents, students and the public at large about the effectiveness of educational services and quality of students' learning.

Information obtained from national learning assessment may be used for a variety of purposes. According to Greaney and Kellaghan (1996), these may include the following:

- *Monitoring standards:* Information about students' learning and achievement in school subjects collected on a regular basis that help to monitor changes of achievement over time.
- *Informing policy makers:* Results of the analysis of the learning outcomes can have implications for curriculum design, teacher training, planning and allocation of resource.
- *Introducing realistic standards:* The results of national learning assessment can foster a sense of realism in the discussion about the appropriate achievement levels we should aim for.
- *Identifying correlates:* Information on relationships between student learning outcomes and system input and characteristic variables may help policy makers to identify factors over which they can exercise some control to direct the scarce resources and materials. Therefore; data, on some of those variables which can potentially be manipulated, have to be collected along with achievement data at the time of the national assessment.
- *Promoting accountability:* Governments spend a substantial part of their public resources on education. To retain the public support for education, the government and the people need information which substantiates that the money spent on education is producing skills that are desirable in the students. So, governments need access to relevant information on the operation of their education system.
- *Increasing public awareness:* Results of national learning assessment can be used to bring an aid of reality, a level of integrity to discussions about the education system and increase public support for quality.

In general, national learning assessment is used to manage expectations, direct the curricula, and create a culture of continuous improvement, culture of accountability and improve planning and adjust policy.

2.4 Methodological and Technical Considerations in National Learning Assessment

National learning assessment is a highly complex process that involves a number of interrelated dimensions, each guided by theory and practice. Its design and implementation process requires a coordinated effort and consensus of a wide range of stakeholders and decision makers.

According to Greaney and Kellaghan (1996), the design and implementation of a successful national learning assessment requires the involvement and consensus of major stakeholders. It includes those organizations and individuals who are responsible for administering the national assessment, those who will consider the results for policy making, those responsible for funding the exercise, and those who involve in policy making and reform activities.

All these and other concerned bodies have to have a room or chance for participating at the different levels of the assessment process through various mechanisms. For example, by establishing national steering committee, organizing a highly technical working group, selecting item-writing groups and data collectors, etc; each would have several functions right from the stage of designing the assessment conceptual framework to the levels of data analysis and report production.

Any national learning assessment program requires collecting and systematizing a host of empirical data that can be analyzed and interpreted as indicative of learning in specific curriculum areas at clearly identified levels of students' performances.

The scope of the data needed to conduct a national learning assessment can determine the nature and techniques of the sampling process. Most of the national learning assessment programs so far conducted around the globe are based on a representative sample of both students and items.

In general, the sampling design has three stages: 1) selection of primary sampling units (eg. regions); 2) selection of secondary sampling units (Schools), and 3) selection of tertiary sampling units (students). This type of sampling design is known as a multi (three) stage stratified sampling procedure, which is too complex to analyze data and interpret findings in relation with different variables.

The size of the sample on which data are collected depends on the purpose of the national assessment and the number and homogeneity of students, teachers, schools and regions that the national assessment plan addresses. For example, the United States NAEP's sample consists of approximately 0.4 percent of the total student population for each Grade (NCES, 2003).

The type of information to be collected for conducting a national learning assessment should cover the issues and subjects based on policy needs. Available literature indicates that most of the national assessment programs focus on knowledge, conceptual understanding, investigation and practical reasoning, attitudes, outcomes, and background variables of both school and non-school factors. For example, the National Assessment of Educational Progress conducted in USA assessed reading and mathematics every two years in Grades 4 and 8 (NAEP, 2003). In general, language, mathematic and science subjects are the most

widely considered areas of curricula by various NLAs. Language, mathematics and science subjects are emphasized in most of the national assessment programs because:

- the language skills are essential tools not only because they serve as the necessary bases for further learning and career development but also because they enable the human spirit to be enriched, foster responsible citizenship, and preserve the collective memory of a nation,
- mathematics is critical for all students in that it trains the mind to be analytic-providing the foundation for intelligent and precise thinking, and
- Science is an organized body of knowledge and a method of proceeding to an extension of this knowledge through hypothesis and experiment.

The instruments used in national assessments should undergo extensive scrutiny, review, and pre-testing before operational use. National assessment analysis methods combine results across different test books within an assessment year and produce results that are comparable across assessment years despite changes in the item pool. The results provide information about students' proficiency in subject area and content sub-domains in terms of average scores and percentages of students at or above the achievement levels for the subject in addition to providing information about trends.

Data analyses may be designed to compare group performances by location, gender and type of school or it may be required to provide information about relationships between student achievement and characteristics of students, schools, and teachers over time (Greaney and Kellaghan, 1996).

In the United States, for example, National Assessment of Educational Progress (NAEP) results were reported in terms of predetermined achievement levels because each assessment reflects current standards of performance. The NAEP reports on student performance with comprehensive information about what students at Grades 4, 8, and 12 know and can do in various subject areas. It also describes relationships between achievement and certain background variables.

In 1999, the Uganda National Examinations Board (UNEB) conducted a sample-based assessment on Grade six students to measure the mastery levels of reading and writing in English. The findings of this assessment indicated that 98% of the Grade 6 pupils in Uganda failed to achieve the advanced grade, and that 35% of the pupils failed to achieve the minimum acceptable level of competency in English. And, pupils in urban schools have a distinct advantage over their peers in rural schools (Kellaghan and Greaney, 2001). Regarding gender disparity, at both Grades 4 and 8, male students scored higher on average than female students.

2.5 Experiences of Other Countries Regarding National Learning Assessment

Experiences in conducting national learning assessments vary among countries. The longest – running and best-known national learning assessments are in the United Kingdom, which existed since 1948; the United States which conducted its first national learning assessment in 1969; and France which introduced the system in 1979. From developing countries, the longest-running national learning assessment is that of Chile, and it dates back to 1978. Most economically less developed countries including Colombia, Vietnam, Thailand, and many African nations started national learning assessments in the 1990s.

There are two basic models for the implementation of national assessments. One is sample based (analytical view of achievement) derived from USA and the other is a census type

(holistic performances) derived from the United Kingdom assessment (Greaney and Kellaghan, 1996).

In the United States, the National Assessment for Educational Progress (NAEP), which is mandated by the National Congress, has become a standard feature of the education system since 1969. The objective of the program is to measure students' achievements at specified ages and Grades (4, 8, and 12) on 11 instructional areas. Since 1969, assessments have been conducted periodically in reading, mathematics, science, writing, and other subjects. By making objective information of student performance available to policymakers at the national, state, and local levels, NAEP is an important part of the nation's evaluation of the condition and progress of education.

England and Wales first applied a large scale survey or national assessment in 1948 at the age levels of 9, 11, and 15 years. In 1978, an improved assessment system was made on three main areas; language, mathematics and science at ages of 11, 13, and 15 years. It was given much weight by politicians in the 1980s and exhibited the various functions of assessment such as formative, diagnosis, summative, and evaluative (Greaney and Kellaghan, 1996).

National learning assessment in France has been introduced since 1979 using both sample and census models of USA and United Kingdom respectively. On the sample based, students are assessed about every five years at the end of Grades 7, 9 and 10 to provide information on achievements at the system level in relation to the curriculum. In the other method, the total population of students in Grades 3, 6, and 10 are provided with diagnostic assessment designed to provide information on individual schools and feedback is sent to schools, students, and parents with assisting teachers to adapt their pedagogical skills to the needs of their students.

The assessments conducted in six countries of Latin America -Chile, Argentina, Brazil, Costa Rica, Mexico, and Colombia- showed similar results in that students scored far below expectations, and students from urban and private schools scored better than their counterparts. Chile is the most experienced country in conducting national learning assessment among these countries. The first national assessment, School Performance Assessment Program (PER), was carried out from 1978 to 1984 at Grades 4 and 8 with the general goal of providing information on the extent to which students achieve learning standards or targets. Achievement tests on Spanish (reading, writing), mathematics, natural and social sciences were used for 90% of the population, and data on contextual variables including home, teacher and classroom, principal and school and instructional practices were collected (Himmel, 1984). After 4 years of interruption, national learning assessment was reintroduced in 1988 under the National Program for Quality Assessment of Basic Chilean Education (SIMCE). Findings showed that the performances of students were poor on curriculum objectives in general, but better in urban schools than in rural, in large schools than in small, and in private schools than in government schools. As a result of the findings, low scoring schools were given special fund, curriculum reform was attempted and percentile rank system was changed to percentage.

Colombia is another country in Latin America that conducted an assessment in 1991 on Grades 3 and 5. The purpose was to find out to what extent students attained the minimum standards set in mathematics and language. Findings showed negative correlation for grade repetition, absenteeism, time spent getting to school and family size. Number of courses teachers took was not a significant predictor of achievement. The final result was released through mass media for discussion at national level, at local workshops, newsletters, brochures and friendly documents.

In Asia, learning assessment was made in a number of countries including Sri Lanka, Bangladesh, Nepal, Pakistan, Cambodia, Thailand and India. Notable among these

assessments was the one conducted in Thailand. After two years of abolishing the public certification examination, Thailand introduced national assessment in 1983 at Grade 12. The main purpose of the assessment was to determine the national education standards schools reach and encourage schools to broaden their objectives and instructional practices. Students were assessed on mathematics, the sciences and career education in both cognitive and affective areas. The task was given to the Office of Educational Assessment and Testing Services in the Department of Curriculum and Instruction Development. The assessments continued in subsequent years. Every second year, samples of Grades 6 and 9 were assessed and expanded to include measures of school process. The result was provided to each school, region and province and for the public too. As Prawalpruk (1996) indicated, the findings of national learning assessment in Thailand had been used for school and provincial planning, monitoring of achievement levels of students over time and increasing the interests of teachers in learning outcomes.

Greany and Kellaghan (1996) identified 4 major categories or approaches to the implementation of national learning assessments in Africa: Monitoring Learning Achievement (MLA), South Africa Consortium for Monitoring Educational Quality (SACMEQ), Francophone countries (PASEC), and others.

The Monitoring Learning Achievement (MLA) project was initiated in 1992 by a joint UNESCO/UNICEF support. It was an attempt to monitor the extent to which students achieved in literacy, numeracy and life skills for Grade 4, and the knowledge, ability and skills in mathematics and the sciences for Grade 8 students. In both Grades, data were collected through tests and background questionnaires for students on school characteristics and home issues. MLA was conducted twice: MLA I focused on Grade 4 and MLA II on Grade 8. About 47 African countries were involved and by 2003 reports were published on MLA I assessments of 18 Sub-Saharan countries, and MLA II of 2 countries. Separate MLA I reports were prepared in 11 countries (Botswana, Madagascar, Malawi, Mali, Mauritius, Morocco, Niger, Senegal, Tunisia, Uganda and Zambia). Results indicate that only 4 countries had met the learning target set in Jomtien (i.e., that 80 percent of learners should attain the defined learning competencies) for Grade 4 pupils in life skills, only two countries met the target in literacy, and one in numeracy. In all the countries, gender difference was not found to be insignificant. The ability of parents to assist their children with their schoolwork positively correlated with student achievement.

The other approach to the national learning assessment is that of SACMEQ. The Southern African Consortium for Monitoring Educational Quality (SACMEQ) is a collaborative effort of 15 Ministries of Education and the International Institute for Educational Planning (IIEP). The first SACMEQ assessment was conducted between 1995 and 1998, and only 8 Ministries (Kenya, Malawi, Namibia, Tanzania/Zanzibar, Zambia, Mauritius and Zimbabwe) collected information on baseline indicators for educational inputs, general conditions of schooling, equity assessments for human and material resource allocations, and literacy levels among Grade 6 students. The second SACMEQ study was conducted between 1999 and 2002, and 15 countries in the region participated. One of the major aims was also to promote capacity building by equipping educational planners in member countries with the technical ability to monitor and evaluate schooling and the equity of education.

Like in the MLA project, findings from SACMEQ indicated that education systems were failing to meet performance standards of ministries. For instance, in Namibia and Zambia, fewer than 30 percent of Grade 6 pupils met the specified minimum literacy standards. According to these findings, some problems such as lack of equipment and facilities (Zambia, Zimbabwe), lack of text books (Zimbabwe), unqualified teachers (Zambia and Zanzibar), lack of books and writing materials (Zimbabwe), high repeaters, un-repaired buildings and high inequity (Zanzibar), high pupil-teacher ratio and lack of seats

(Zimbabwe), low frequency of tests/ homework (Zambia, Zimbabwe), district variation (Mauritius), and students' absenteeism were observed as the major determinants of students' performance.

The third type of national learning assessment in Africa is PASEC (*Programme d'Analyse des Systemes Educatives des Pays de la CONFEMEN*). PASEC was established as a network for sharing information on educational evaluation, instruments and results among 12 Francophone countries (Burkina Faso, Cameroon, The Republic of Congo, Cote d'Ivoire, Djibouti, Guinea, Niger, Central Africa Republic, Mali, Senegal and Togo). It encourages the involvement of senior decision makers and stakeholders to identify policy issues, base decisions on data and follow up decisions for actions. At first, the assessment began on Grades 2 and 5 for literacy (French) and mathematics. Starting from 1994, the study included all Grades from 2 to 6 and same tests were used in all countries. Other data on school and background factors were also collected from schools. This network differs from the others in that the assessment is made twice in an academic year, at the beginning of November and at the end of May. As a result, this approach reflected a strong research orientation of the program and it attempted to identify causal relationships on which to base policy that seeks to improve the efficiency and effectiveness of education. Efforts have also been made to determine the impacts of in-school factors (teacher training, class size, textbooks availability) and out-of-school factors (parents education, home language, distance from school).

The 4th group includes some countries which attempted national assessments of their own: Burundi, Eritrea, Mali, Senegal, Uganda and Zambia. For example, in Namibia, National Learner Baseline Assessment measured students' achievements in English and mathematics at Grades 4 and 7. In Eritrea, students were assessed in six regions to determine if they had mastered basic skills in the official curriculum, basic skills in their mother tongue (Grade 1) and English and mathematics (Grade 4). The assessment identified particular areas of the curriculum causing problems. It found that boys outperformed girls, and identified implications for teacher education and teaching. The experiences of some counties can be found in Greaney and Kllanghan (1996).

2.6 Background and Community Factors that Influence Student Achievement

Various studies emphasize a host of factors, both external and internal to schools, to have influences on student achievement. Among the external factors, the influence of student home background, personal and community characteristics are widely acknowledged.

With reference to student home background, well known influences to student achievement relate to parental socio-economic status (SES), family configuration and parental support. Among the indices of SES are parental occupation, parental level of education, parental income and occupational prestige. Husen (1967) reported that in the first IEA mathematics study, the total mathematics score at 13-year-old level correlated 0.16 with father's education and 0.12 with mother's education over 12 countries. The correlations between father's occupational status and the mathematics scores were 0.25 at the same age level of the same populations. In the IEA six-subject studies, it was found that achievements in science, reading comprehension, literature and the cognitive side of civic education were linked to the socio-economic background of the home. In family configuration, family size was found to have a negative relationship with student achievement. Increased spacing between children's birth reduced the decrease in scholastic performance, although this did not hold true for developing countries. With regard to parental support, studies indicated that parental attitudes towards education, their interests in their children's education, and their beliefs in the values of schooling are related to students' scholastic performance.

When it comes to students' personal characteristics, it has been observed that the relationship between gender and achievement varies based on the subject and the level of students under investigation. Parelius and Parelius (1987) noted that girls' grades in both elementary and high schools tend to be higher than that of boys. However, boys often scored higher on standardized tests. Fagerlind and Saha (1989) observed that in most developed countries, girls performed equally with boys at primary school level, but at secondary school, girls began to do more poorly. Subjects also make a difference in achievement of boys and girls. For instance, a comparison of the reading performance of boys and girls in many countries has frequently shown differences in favor of girls (Thorndike, 1973). However, girls fall behind boys in mathematics and science. In the second IEA science study in 17 countries, it was found that boys scored higher than girls at all levels, and the gap increased from the 10-year-olds to the 14-year-olds (Chen, 1996). TV watching, as the other student characteristics, is inversely associated with student achievement (Comstock, 1994). However, leisure or voluntary reading as well as the average amount of time spent on homework of all countries have shown positive relations with student achievement.

Just as students may be influenced by their classmates, their teachers and the norms of their schools, they are influenced by the communities in which they live. Some of this influence can be transmitted through families, but some may happen through direct influence of communities. Variations among communities can take various forms and this has a great deal of influence on student achievement. Traditionally, communities are divided in geographical neighborhoods taking urban, rural or suburban forms. Several studies indicate that such divisions have impact on student achievement, attainment, aspirations, and school climate (Stockard and Mayberry, 1992). Just as the geographical space of schools influences learning, it also affects the resources that a school can offer its students. The resources available to schools depend to a large extent on the economic basis of its inhabitants. Parental and community involvement in schools, community participation in school control, and community values all have their influences on student achievement. In developing countries, demand for child's labor and opportunity costs of attending schools also exert a significant impact on academic achievement.

2.7 School Based Factors that Influence Student Achievement

The analysis of contextual or external factors provides only a partial view of the influences on student learning. To fully understand how higher achievement is facilitated, it is important to examine internal factors at school and classroom levels. Stockard and Mayberry (1992) discuss these factors under two generic divisions: school and classroom environment, and school resources. In the school climate/environment, they include academic expectations and excellence, school leadership, orderly environment, and school coherence, teacher and student morale, and effective teaching. School resources comprise school facilities and per-pupil expenditures, teacher qualifications, classroom size, and school size. Fuller (1986) analyzes these factors under the generic concept of school quality. This construct comprises material inputs, teacher quality, teaching practices and classroom organization, school management and structure. Chen (1996) discusses internal factors influencing student achievement as school background factors. In this category, he included school structure, instructional facilities, teachers' instructional time, and the curriculum.

With reference to school environment, Stockard and Mayberry (1992) refer to various materials to indicate the significance of 4 main dimensions. These are the *ecology* (the physical environment), *milieu* (the social characteristics of individuals and groups participating in organizations), *social system* (patterned relationships of persons and groups), and *culture* (the collectively accepted beliefs, values, and meanings of the group). School environments are commonly measured using indices containing a variety of items based upon

the above factors. Various studies in the area demonstrate that school environment is linked to school achievement.

According to McDill *et.al*, (1967), McDill and Rigsby (1973), schools with teachers and students who see or expect higher achievement and academic excellence as a real and attainable goal actually do have higher achievement. Among the most important attitudinal factors related to increasing school achievement in this respect is staff commitment to improved academic performance, high or increasing expectations of teachers about students, high opinions of students' abilities and peer norms emphasizing academics.

Strong school leadership and management is another internal characteristic of schools which determines student achievement. Several investigations indicate that principals, who are involved in instruction, communicate high expectations, promote good feelings and collegiality between faculty and administrators and among faculty members. They also encourage teacher participation in the school's decision-making processes which is effective in promoting success. Effective administrators promote cohesive relations within schools performing a balancing act. They promote higher achievement and develop ethos or culture that enhances morale, mutual trust and respect, and shared norms and values (Ellet and Walberg, 1979; Fullan, 1990).

Orderly environments and teacher and student morale are important ingredients for success in school. Edmonds (1979) indicates that student achievement is enhanced in schools that maintain a consistent set of rules and values that clearly map out school goals and policies while also promoting purposefulness and pleasure in learning. Teacher's morale in terms of job satisfaction and students' recognition of teachers' morale in their work exhibit high level of attendance and achievement (Brookover and Lezotte, 1979). Schools that can nurture high morale among students and staff maximize the chances of developing attitudes about individual abilities and learning environment that promotes higher level of achievement.

Specific material inputs enhance student learning. According to Fuller (1986), material inputs that consistently relate to student achievement in developing countries include instructional materials (texts, desks and reading materials), instructional media, and school building quality, library size and quality and nutrition and feeding programs. His report indicates that class size, school size and laboratories do not consistently relate to student achievement.

Teachers' characteristics and quality have great effects on students' achievement. As Fuller (1986) mentions, teacher quality elements that consistently impact student achievement in developing countries include total years of teacher's schooling, years of tertiary and teacher training, in-service teacher training, teacher's verbal proficiency, and teacher's social background. Among the elements of teacher quality that show no consistent relationship with student learning in developing countries are individual teacher's salary level, total years of experience and punctuality. Studies in developed countries show mixed results for the impact of teacher qualifications on student learning. According to Stockard and Mayberry (1992), studies which focus on the impact of teacher qualification on student achievement show only 50 % confirmation rate, while those studies that investigated the impact of teacher's attending prestigious schools show very little relation with student achievement. A little more than 50% confirmation rate was obtained from studies which examined the impact of teacher's experience. However, a significant positive relation was observed between teacher verbal proficiency and longevity on the one hand, and student achievement on the other. Again, teacher's salary has not shown a consistent or significant impact on learning.

Effective teaching practice makes a difference in student achievement. For developing countries, the most important influences come from the length of instructional time, frequency of homework given, teacher's expectation of pupil performance, and time spent by teachers on classroom preparation (see Fullan, 1986). In developed nations, studies suggest

that the quantity and pacing of instruction (i.e. curriculum covered and time spent on instruction), the way in which teachers give information (i.e. well-organized, well structured presentations), the way in which teachers question students and wait for responses, and the way they handle seat work and homework all influence student achievement (see Puff, 1978; Klitgard and Hall, 1973; Rosenshine, 1983).

The length of instructional time is an important indicator/school factor that boosts achievement. Instructional time varies from the number of school days in the school year to the hours the subject is studied during a week. This instructional time is bounded by the variety of material resources and management practices. Teacher's knowledge of the subject is strongly and consistently related to student performance. Although there is no specific teaching practice that is universally effective, teachers without a limited repertoire of teaching skills appear to be more effective than those with a limited repertoire (Chen, 1996).

Active teaching - learning interaction in the classroom and the frequency of evaluating students produced better effects on student achievement (Brook, et al., 1979). Clear explanation of the materials by the teacher and more time spent on concrete learning also affected student learning positively (Heyneman and Loxily, 1983). A focus on students' perceptions rather than on teachers' is likely to be more productive; it attempts to improve and understand classroom learning. Students' perceptions vary with such factors as teacher personality, class size, Grades level, student gender, subject matter, the school level, environment, and the type of school.

2.8 Historical Overview of Learning Assessments in Ethiopia

2.8.1 ESR and ERGESE

In Ethiopia, very limited attempts have been made to review the education sector since the introduction of modern education in 1908. The Education Sector Review (ESR), the Evaluative Research of the General Education System in Ethiopia (ERGESE) in 1983 to 1986, and the Ethiopian National Baseline Assessment (ENBA) on Grades 4 and eight students' achievement launched in 1999/2000 are the major attempts made to review the education sector in the history of the Ethiopian education system.

The Education Sector Review (ESR) did not investigate the quality of education system, rather it devoted its attention to strategies for a rapid expansion of primary education with the view of achieving universal literacy before the year 2000 (Tekeste Negash, 1990). The ERGESE project was initiated by the Ministry of Education based on the resolution passed by the government in 1983 to review the education sector. The purpose of the project was to investigate the quality of primary and secondary schooling in Ethiopia and to make recommendations for improvement (MOE, 1986). The findings of this study indicated that while educational opportunities had widened, there had not been a corresponding improvement in quality.

These two projects ESR and ERGESE, therefore, could not be considered as national learning assessment programs of the Ethiopian education system. This was because these two studies did not collect and analyze data on the actual learning of students systematically and thoroughly. Rather such attempts could be considered as survey studies of the overall status of general education system of the country at that particular time.

2.8.2 EBNLA (1999 – 2000)

The Ethiopian National Baseline Learning Assessment was launched in 1999/2000 (1992 E.C.) initiated by the Ministry of Education (MOE) with the need to evaluate and improve the results of the ongoing reform in the education system over certain periods of time. In the process, the National Organization for Examinations (NOE), and the Basic Education System Overhaul (BESO I) project played major roles.

The main objective of this National Baseline Learning Assessment was to determine the various levels of students' performances at both Grades 4 and eight in 4 key academic subject areas. Grade 4 students were tested in English, basic reading, mathematics and environmental science subjects, all prepared in the different instructional languages; and Grade 8 students were assessed in English, mathematics, chemistry, and biology. Moreover, teachers and head teachers and the overall conditions of school compounds, in addition to students, were considered as major sources of the data collected for the study (NOE, 2000).

A three stage stratified random sampling design was used to select sample regions, schools and students at both grade levels (4 and 8). Accordingly, 256 schools for Grade 4, and 136 schools for Grade 8 studies were sampled. About 10,506 students for Grade 4, and 5099 samples for Grade 8 studies were tested across ten regions of the nation (NOE, 2000).

Information on the background characteristics of students and teachers were also collected and analyzed. The implications of the findings for participating regions were indicated. According to the findings of the study, all participating regions have room for improvement in all key subject areas since no one region scored above the acceptable minimum level of 50% achievement (NOE, 2000).

The results also indicated that schools with high achievement at Grade 4 level also tended to obtain high achievement at Grade 8. The results of the study were also discussed in a workshop and constructive recommendations were provided for the improvement of the quality of students' learning in the nation. The reports were disseminated to regions and other stakeholders.

CHAPTER THREE: FRAME OF ANALYSIS AND METHODOLOGY

3.1 Frame of Analysis

The main purpose of the Second National Learning Assessment was to find out the extent to which learning takes place in the Ethiopian primary education, and determine the main conditions that influence the learning outcomes of students. It was also to generate information on school improvement from the first national learning assessment conducted in 1999/2000.

The dependent variable, student learning, was measured by achievement tests. Another dependent variable, student's attitude towards socially relevant issues, was measured by questionnaires. The independent variables that refer to factors which affect the achievement of student learning in this study included school condition/environment, teachers behavior, school management, school structure and supply, instructional support, language of instruction, students' background, and community opinions. The following table shows their relations and descriptions.

Table 1. Description of variables

Variables		Description
Dependent Variables	Students achievement	Total achievement mean score for Grades 4 and 8 Mean score of basic reading comprehension, English, mathematics and environmental science for Grade 4 Mean score of English, mathematics, biology, physics, and chemistry for Grade 8
	Students' attitude	Attitude of Grade 4 and 8 students towards socially relevant issues included health, environment, civics and ethics, and the school.
Independent Variables	Students' home background Student personal characteristics	Family size, parents' education and occupation, language at home, distance from school, attendance, learning support. Students' gender, students' sense of ownership of school property, time spent listening to the radio, time spent watching television, leisure (reading non-textbook materials), interest in English, interest in mathematics, interest in chemistry, interest in biology, interest in physics and the use of school library
	School structure and curriculum materials	Location (urban-rural), school program, level, instructional language, class size, program, classroom condition, supplies, facilities and equipment, provision of instructional materials, period allotment, school construction, class size
	Instructional inputs	Textbook-pupil ratio, availability of basic instructional materials (other than textbooks), availability and use of pedagogical center, lab., library
	Teacher variables	Teacher's qualifications, years of experience, knowledge of subject matter, distance from school, teaching load, attendance or absenteeism
	School management	School directors' qualifications, organization, evaluation of teachers, meetings conducted with staff and community
	Parent/community views	Attitude towards students' behavior, learning and schools, collaboration with the school to solve problems, parents' involvement in school affairs

3.2 Design

In order to obtain the required information for the proposed research questions, both quantitative and qualitative research approaches were used. In the quantitative approach, a cross-sectional data using achievement tests were collected to determine the extent to which learning takes place in primary schools. Moreover, a baseline study design which compared student achievement results of the Ethiopian Baseline National Learning Assessment with the Second National Learning Assessment was used in order to monitor school improvement. Since the tests were given in different years to two independent samples of the same Grade, threats to internal validity due to history or maturation effects were controlled. A qualitative study design was used to supplement the quantitative study. It mainly involved a collection of cross-sectional data on similar issues addressed by the quantitative study. Wherever necessary, a historical approach has been used to understand the change of events over time in the qualitative design.

The Second National Learning Assessment has been carried out in three stages. The first stage involved institutional arrangement for carrying out the study. In addition to mapping out the mission of the entire study, this task led to establishing the National Advisory Council, the Technical Working Group and the Secretariat of the Second National Learning Assessment as responsible key structures for its implementation. Terms of Reference for each of these structures were prepared. Accordingly, the National Advisory Council (NAC) was responsible for providing overall leadership to the study. The chairman of the NAC was the Vice Minister for General Education and the Manager of NOE was the secretary. The members of NAC included leaders of central institutions of the Ministry of Education (MOE) and the Heads of Regional Education Bureaus, AED/BESO II and USAID. The Technical Working Group (TWG) was established to provide leadership and direction on everyday basis to the development and implementation of The Ethiopia Second National Learning Assessment. The secretariat coordinated the required logistic and other support to the study.

The second stage involved planning, development and field testing of the Second National Learning Assessment activities, materials and instruments respectively. Planning the second national learning assessment activities included the identification of preparatory activities, making decisions on the design of the main study, fieldwork as well as dissemination. Development activities included reviewing the literature and the previous assessment documents, review of curricular profiles, initial revision of instruments, pilot testing of instruments (in three languages), translation, printing and packaging of instruments.

The third stage involved field work, data analysis and interpretation. The field work, which began by selecting and training of data collectors, focused on the collection of both quantitative and qualitative data. This was followed by data organization, cleaning, analysis and reporting.

3.3 Sampling Procedures

In order to provide national and regional estimates of student achievement results and attitudes on selected curriculum based topics with group comparisons across gender, location of school (i.e., urban/rural), and language of instruction, all students of Grades 4 and 8 in the Ethiopian primary schools were taken as the target population. To be able to obtain data on independent variables using quantitative methods, teachers and head teachers were targeted. The target population for the qualitative study included the purposefully sampled schools focusing on students, teachers, head teachers and community representatives.

3.3.1 School Sampling

The sample size used in the Ethiopian National Baseline Learning Assessment provided the minimum number of schools to be sampled as a starting point. In that study, the total number of schools included was 272. By taking into consideration the expansion of primary schools since this assessment in 1999, an assumption to consider 300 sample schools was initially accepted. After using a simulation procedure to find out the acceptable minimum number of samples to make strong estimation of achievement results from the national sampling frame, and the representation of regions considered, the number of sample schools was determined at 407. The selection of samples was done for all regions except for Gambella.

To meet the goal of sound statistical estimates of performance for the nation, a multi-stage stratified random sampling technique was used. The number of sample schools in each region was determined based on the relative proportion of its school population. Prior to selection, the sampling frame comprising the number of schools by regions, levels and location was collected from EMIS (2002/3) statistical document. Following this, schools were stratified based on region. The decisions to use a stratified sampling procedure were due to the following reasons.

1. To accurately represent individual regions and the geographical locations (urban/rural)
2. To have a good representation of various linguistic groups
3. To ensure adequate representation of specific groups of the target population in the sample
4. To obtain reliable estimates for each stratum, if required

As language, in general, is tied to geographic region or zone, the stratification on zone was assumed. Emphasis was given to the proportion of schools per region (and zone to SNNPR only).

A minimum of 25 primary schools were randomly selected from each region. This sample size was determined from confidence intervals estimated using population statistics from baseline data. Using simulation procedures, a 90% confidence interval of +/- 5% from the mean was calculated for a minimum regional sample of 25 schools. A 90% confidence interval of +/- 8% of the mean was calculated for a minimum number of 15 schools in a region, the minimum number that would be tolerated for selecting the set of full primary schools where Grade 8 tests were conducted.

To make school based paired comparison or to investigate what changes have been made from the Ethiopian Baseline National Learning Assessment study to the Ethiopian Second National Learning Assessment, it was decided to consider 10 of the sample schools from each region. This sampling procedure was not aimed at making generalizations, but to determine school improvement using the schools under investigation as units of analysis.

When stratifying the minimum number of schools, regions with less than 25 schools were “topped up” in order to meet the minimum representation of regions by a sample of 25 schools. In order to ensure representation in the context of extreme diversity of nationalities in the SNNPR, the sample size for this region was “topped up” to 80 schools.

The randomization of sample schools for each region was conducted by assigning a random number to each of the schools in the region, and then by sorting the schools according to the assigned random number. The target number of schools was counted from the top of the sorted schools in the spreadsheet.

Based upon the above sampling procedure, a total of 407 sample schools from 12471 national totals was selected. Oromia contributed the highest proportion of the sample schools with 38.15% followed by Amhara 24.14% and SNNPR 19.45%.

Table 2. Sample schools planned and visited by region

Region	Total Primary Schools*	Actual Proportion	Topped up Sampled Schools	Sample schools by Grade			
				Data Planned for collection		Actual Data Collected	
				Grade-4	Grade- 8	Grade- 4	Grade- 8
Tigray	967	7.75%	25	25	18	25	12
Afar	156	1.21%	25	25	14	25	11
Amhara	3,011	24.14%	50	50	30	47	26
Oromia	4,758	38.15%	77	77	53	76	52
Somali	301	2.41%	25	25	10	24	9
Ben-Gumuz	275	2.21%	25	25	14	25	12
SNNPR	2,426	19.45%	80	80	61	80	43
Gambella	149	1.19%	25	25	14	X	X
Harari	48	0.38%	25	25	16	24	11
Addis Ababa	313	2.50%	25	25	21	25	21
Dire Dawa	67	0.54%	25	25	16	25	16
Total	12,471	100%	407	407	267	376	213

*Source: EMIS 2002/3

The above table shows that it was planned to cover 407 primary schools from all regions. Of these schools, it was planned to include 267 schools having Grade 8 and 407 or all primary schools having Grade 4. However, the actual number of schools from which data were collected in the field was 376 for Grade 4, and 213 for Grade 8 (excluding Gambella Region). Furthermore, the number of sample schools planned by location was 136 for urban and 271 for rural but data were actually collected from 126 and 256 respectively.

The selection of schools for the qualitative study was accomplished on the field. Prior to this, however, a decision was made to include up to 50 schools overall. It was managed to nationally cover 44 schools for this purpose with varying degree of coverage from one region to the other. In Tigray, Harari, Addis Ababa and Dire Dawa the number of schools included was 2 each. In Afar and Beneshangul Gumuz 4 schools each were included while in Oromia and Amhara the number was 6 each. In Somali, only 1 school was used for collecting qualitative data while the number was 15 for quantitative data. The details are given in Table 2.

3.3.2. Sampling of School Directors, Teachers and Students

The sampling of schools was followed by another decision to select students, teachers and directors. It was decided to randomly select only one section from each grade. Within each section the maximum number of students randomly selected was 40 and less in the cases where there were fewer students to meet the maximum number. The decision to include a maximum of 40 students was made by the study team assuming this to be an average number that can be managed during the test administration.

The school director and all teachers of the sampled students in the respective grades (4 and 8) were also selected to fill in the questionnaires. The following table provides a summary of sample teachers, students and school directors by region.

Table 3. Sample students, teachers and directors by region

Region	Teachers			Grade4 Students		Grade 8 Students	
	School directors	Grade4	Grade8	Collected	Cleaned	Collected	Cleaned
Tigray	25	74	60	953	951	480	479
Afar	25	42	52	658	648	385	384
Amhara	47	48	73	1721	1721	1028	1027
Oromia	76	102	222	3014	2967	1964	1957
Somali	24	77	42	694	687	310	307
Ben-Gumuz	25	42	27	776	776	448	446
SNNPR	80	21	215	3025	3004	1787	1736
Harari	24	76	96	727	723	412	412
Addis Ababa	25	28	105	987	983	825	823
Dire Dawa	25	47	94	791	791	488	488
Total	376	751	988	13346	13248	8127	8059

The above table indicates that the national sample of students who took part in the Ethiopian Second National Learning Assessment was 13346 in Grade 4, and 8,127 in Grade 8. Similarly, 751 Grade 4 teachers, 988 Grade 8 teachers and 375 school directors were included in the provision of the required information for the Second National Learning Assessment.

For the qualitative study, it was decided to include school directors, teachers, Grade 8 students, parents and community representatives in focus group discussions and/or interviews. For focus groups, it was decided to form different groups comprising 5-10 individuals at least in two school areas from each data collection route. The following table shows the number of respondents who participated in these discussions by region.

Table 4. Participants of focus group discussions by region

Region	Participants			Number of Schools
	Students	Teachers	Parents	
Tigray	22	15	6	2
Afar	15	22	14	4
Amhara	33	44	38	6
Oromia	70	59	53	6
Somali	5	9	3	1
Ben-Gumuz	26	34	22	4
SNNPR	111	109	76	15
Harari	16	10	16	2
Addis Ababa	13	15	10	2
Dire Dawa	10	9	8	2
Total	321	326	246	44

Table 4 above shows that the number of Grade 8 student participants in focus group discussions was 321 while the number of teachers and directors was 326 and 246

respectively. The largest number of schools and respondents was from the SNNPR due to the larger size included from this region.

3.4 Instruments of Data Collection

In this study, all the previous instruments developed during the National Baseline Learning Assessment were adopted with some modification and inclusion of other three new instruments (Physics test and attitude survey for Grade 8, and background questionnaire for Grade 4 students). In the achievement test, adequate number of items (60-80) was prepared and organized in two parallel forms for each subject to be piloted. The main rationale for pre-testing parallel tests was to replace items that require replacement from the first study. The various types of instruments used for data collection are described below.

Achievement tests: The achievement tests for both Grades contain 40 items each (except for reading comprehension which is 20) and cover different content areas of the respective subject. The tests include reading comprehension, English, mathematics, and environmental science for Grade 4 and English, mathematics, physics, biology and chemistry for Grade 8 students.

Attitude survey: This instrument was prepared for students of both Grades and it contains two parts meant to find out the background information and attitude of students towards some socially relevant issues. The background information deals with students' personal characteristics, family size and education, provision of textbooks, learning support provided, learning and assignments at school, distance walked to school and interests in subject areas.

School checklist form: This instrument was used for recording the general conditions of the school level, location, size, number of classrooms, school facilities, equipment, services and the classroom conditions of grades 4 and 8.

Questionnaire for teachers: This instrument focused on general background information of Grade 4 and 8 teachers; sex, qualification, experiences, family education, their opinions towards the teaching profession, school management, curriculum materials and students, teachers' load and provision of instructional materials, and meetings made with parents and others.

Questionnaire for school directors: This questionnaire is similar to that of the teachers and seeks data about background information of the directors, manpower in the school, evaluation of teachers' performances, conducting meetings with the staff and parents, provision of curricular materials, and improvements made on the school.

School semester result: This is a form used to collect the first semester result of sample students from each sample school on the selected subject areas for comparing school results and results from the Second National Learning Assessment.

Group discussion guide: This instrument was used to conduct discussion with school teachers about the strength and weakness of the school on various issues: characteristics of students and teachers, availability of facilities and instructional materials, the school's program and period allocation, organization and administration of the school, and involvement of different groups in supporting the school and other environmental constraints that affect the teaching learning process.

3.5 Validity and Reliability of Instruments

Taking into account the decentralization of the curricula, one of the methods used for the validation of instruments involved a national workshop which brought together regional curriculum experts, specialists from ICDR and NOE to review the extent to which the instruments measure student learning in each region. Prior to this, all instruments mentioned above were critically reviewed by the TWG members and test developers, and this was followed by comments from an international consultant. In order to ensure the representativeness of the contents of the instruments, the workshop weighted items that were prepared on the basis of the table of specification following the syllabi of primary education. In other words, the national workshop ensured the content validity and relevance of the test items prepared by subject experts or the item writers by relating the items to the curriculum objectives.

Another measure used to check the predictive validity of the Second National Learning Assessment was collection of teacher evaluation of students from rosters. A comparison between the composite achievement tests from the Second National Learning Assessment and school teachers' evaluation had shown strong correlation. Since samples were randomly selected and they were representatives of the Grades 4 and 8 Ethiopian student population, the internal and external validity of the assessment was ensured.

The Second National Learning Assessment was conducted in 14 nationality languages in Ethiopia. The instruments were first developed in the English language after which they were translated into other languages. In fact, the process of translation was made two times, first for pre-testing and second for the final instruments. Based on the agreement made with the Region Education Bureaus during the NAC meeting held at MOE, the pre-testing of instruments was made in Addis Ababa, Adama, Debre Birhan and Mekele in three major language areas (Afan Oromo, Amharic and Tigrigna). The translation of instruments was made by subject teachers and curriculum experts who have experience in dealing with the languages in the respective places. On translation, two persons were assigned for different tasks, one to translate direct from the original version to the respective language and the other to do the backward translation.

Pre-testing of instruments was made in sixteen schools of 4 regions (Amhara, Addis Ababa, Oromia and Tigray) to identify item clarity, difficulty and the problem areas of each item. The 4 regions were first identified by the study team at the centre. Next, the regions were contacted for selecting 4 schools (primary and secondary) to carry out the tests on Grades 5 and 9 students (for grades 4 and 8 instruments respectively). Questionnaires or instruments were pre-tested only in primary schools. In each school, the piloting was a two-day task by two persons (data collectors). After the data were collected, item analyses were carried out in order to make additional improvements on items. Before decisions were made to modify, accept or reject items, the difficulty level, discrimination index and point bi-serial correlation were thoroughly investigated.

Finalized instruments, in particular, the achievement tests, background questionnaires and attitude surveys of both Grades were translated into additional local instructional languages: Harari, Somali, Welaita, Keficho, Kembata, Hadiya, Gofa, Gamo, Dawaro, Sidama, Gedeo, and Himitigna. At last, the translated instruments were sent for printing in booklet forms. Two subjects were arranged in one booklet and packaged for the sample schools and the field routes.

Prior to adopting instruments from the National Baseline Learning Assessment and developing new ones for the Second National Learning Assessment, however, a thorough item analysis was carried out for item difficulty and discrimination power. Based on the analysis, certain items were modified or totally replaced by new ones and more new sets of items were prepared for each subject area.

Similarly, the study team revised the instruments for background information which included the previous attitude survey instruments for grades 4 and 8. The questionnaires for directors and teachers and the school checklist used in the EBNLA were also improved in line with the identified variables. A new background questionnaire for Grade 4 and an attitude survey questionnaire for Grade 8 were also developed based on suggestions made by the international consultant.

3.6. Data Collection and Administration of Instruments

A systematic and planned approach was used for the field work. The collection of data was organized to take place in 4teen routes. Each route had two selected training centers in which the training of data collectors was carried out. One route leader or trainer was assigned from the center (NOE) to manage the activities of each route. Prior to data collection or the field work, a consultative workshop was held with regional educational representatives to discuss the program of data collection, the sample schools and training centers, the assignment of centre coordinators, and recruiting data collectors in each region. In line with this, data collectors were selected by regions from the respective Woreda education offices and from non-sampled schools. One data collector was assigned to each Grade level (4 and 8) in each sample school. Based on the prepared guideline by the study team, a two day intensive training was given for data collectors by the route leader /trainer in each of the 27 centers. After the training, the data collectors were provided with the list of sample schools, instruments and working schedule at each school. The data collection was conducted at the same time in all sample schools from April 15-30/2004. Table 5 below summarizes participants of the field work by routes.

Table 5. Number of participants involved in field work by route

Routes	Route leaders	Centre facilitators	Data collectors
Mekele and Axum	1	2	43
D/Tabor and Gonder	1	2	17
D/markos and G/Beles	1	2	44
Ataye and Woldia	1	2	31
Nekemete and Assossa	1	2	45
Wolkite and Ambo	1	2	54
Bonga and Bedele	1	2	36
Shashemene and Adama	1	2	51
Sodo and Arbaminch	1	2	70
Awassa and Negele	1	2	43
Asebeteferi and Diredawa	1	2	51
Harari and Jigiga	1	2	67
Addis Ababa	1	1	46
Awash and Dubti	1	2	40
Total	14	27	638

Before students sat for the examinations, they were given a short-training on how to write or fill in the answers. They were also sensitized on the goal of the Ethiopian Second National Learning Assessment and how significant their participations could be for the successful accomplishment of the task. In order to reduce boredom in the course of taking long examinations, a limited form of motivation was administered to students. At the end or return from the field, reports were made by various groups; the field workers to the route leader, the centre coordinators to the respective region, and the route leaders to the centre or the study team. Data for the qualitative analysis were collected by center coordinators. There were center representatives from region and zone education bureaux who assisted in discussions, selecting participants and in translations.

3.7 Methods of Data Analysis

Two separate data files were created for grades 4 and 8, and the data were organized into these data bases at school and student levels. Before encoding the collected data into the computer, the instruments were first organized by region, type of instruments, grade level, subject area, school and respondents' code number. The data were first entered to access and then transformed to SPSS. Twenty five encoders from different departments of NOE were involved in entering the data into the computer. Parallel to this, other 20 persons were assigned to check and re-check the encoded printouts of each instrument. The encoding and checking task took beyond the expected time due to micro-management problems. The encoding was followed by the process of data cleaning, analysis and interpretation. The common statistical procedures applied for data analysis included descriptive statistics, t-tests, ANOVA, homogeneity tests, correlation and regression analyses.

CHAPTER 4: FINDINGS OF THE STUDY

4.1 Findings on Grade 4 Achievement Tests

The following section presents the analysis on the 4 achievement tests given to Grade 4 students across the nation. The tests were reading comprehension, English, mathematics, and environmental science. The composite average score of the 4 tests was also analyzed separately. The raw scores of each subject were converted into percentages. Each test was analyzed primarily at national and regional levels and then by gender of the students, and the location of the schools. The analysis of each test is accompanied by up to seven tables which show summary of the descriptive statistics, one-way analysis of variance and homogeneity subsets grouping whenever appropriate.

4.1.1 Overall Achievement of the Students at National Level

The overall achievement of the students at national level was presented initially prior to detailed discussion at different levels of analysis. Table 6 shows the summary of the descriptive statistics for all the subjects and the composite mean score at national level. The overall performances in all subjects were low with only reading comprehension and environmental science average reaching 50%. The mean percent score for mathematics (39.70) and English (38.68%) are notably lower than the national mean (48.48%) and also from the mean of the other two subjects.

Table 6. Summary of the descriptive statistics for all subjects at national level

Subjects	N	Mean	Std. Error	Std. Dev
Reading	12886	64.49	.19	21.97
English	13028	38.68	.15	16.88
Mathematics	13143	39.70	.15	17.51
Env. Science	13005	51.74	.15	17.42
Composite	12507	48.48	.13	15.07

Seen from the perspectives of the Ethiopian Education and Training policy, the overall achievement of Grade 4 students is less than the minimum expectations. Particularly, the mean percentage scores for Mathematics and English are by far less than the minimum requirements of 50%. Ironically, the two subjects are considered as core subjects in the primary school curricula. The fact that test scores are less than expectations implies that the quality of schooling is low in Ethiopia. It is now clear from good amount of evidence that the quality of schooling has an influence upon the speed with which societies can become richer and the extent to which individuals can improve their own productivity and incomes (UNESCO, 2004). Thus, low quality schooling in Ethiopia also implies immense influence on the country's efforts to poverty reduction and national prosperity.

4.1.2 Overall Achievement of the Students by Gender at National Level

Table 7 shows that the mean differences between males and females in all subjects were statistically significant in favor of males. The mean difference in mathematics was the highest. This is an indication that gender gap still exists.

Table 7. T-test for equality of means by gender at national level

Subjects	Gender	N	Mean	Std. Error Mean	Std. Deviation	t	Sig.	MD
Reading	MALE	7547	65.71	.25	21.78	6.858	0.000	2.71
	FEMALE	5221	63.00	.31	22.18			
English	MALE	7628	40.55	.20	17.39	16.398	0.000	4.90
	FEMALE	5282	35.66	.21	15.59			
Mathematics	MALE	7689	41.68	.21	18.15	16.244	0.000	5.02
	FEMALE	5350	36.67	.22	16.13			
Env. Science	MALE	7590	53.43	.21	17.89	12.376	0.000	3.82
	FEMALE	5313	49.61	.22	16.32			
Composite	MALE	7324	50.14	.18	15.43	14.690	0.000	4.02
	FEMALE	5082	46.12	.20	14.31			

The above findings in Table 7 are not surprising as many studies in developing countries show similar trends. For instance, Duncan (1989) refers to unpublished IEA studies which indicate that girls performed less than boys in science in three African countries namely, Ghana, Nigeria and Zimbabwe. She further refers to a Zambian study where examination results show that girls scored considerably below boys in all 4 subjects examined (English, social studies, mathematics and science). These results even concur with previous findings in Ethiopia for primary science and mathematics by Biazen and Junge (1988). Reasons and appropriate solutions for such disparity have been related to numerous factors including socio-economic, cultural, and institutional/policy, and school situations (see Odaga and Heneveld, 1995). Very recently, Kellagan and Greany (2004) referred to studies from MLA and SAQMEC projects in Africa where significant gender differences were not found in academic performance among female and male students. This shows that it is possible to overcome the gap in female and male student academic performances provided that the right approaches are adopted in Ethiopia.

4.1.3 Overall Achievement of the Students by Location at National Level

Table 8 shows that the mean differences between students from urban and rural schools in the composite score, reading comprehension and mathematics were statistically significant in favor of urban schools. The mean difference in reading comprehension (10.87) was the highest. Though not statistically significant, the mean difference in environmental science was in favor of rural schools.

Table 8. T-test for equality of means by location at national level

Subjects	Location	N	Mean	Std. Error	Std. Deviation	t	Sig.	MD
Reading	Urban	4329	67.44	.32	20.73	10.871	0.000	4.43
	Rural	8557	63.01	.24	22.42			
English	Urban	4365	38.87	.25	16.84	.907	0.364	.28
	Rural	8663	38.58	.18	16.90			
Mathematics	Urban	4427	40.38	.26	17.28	3.165	0.002	1.02
	Rural	8716	39.35	.19	17.62			
Env. Science	Urban	4304	51.49	.25	16.64	-1.176	0.240	-.38
	Rural	8701	51.87	.19	17.79			
Composite	Urban	4149	49.28	.22	14.41	4.165	0.000	1.19
	Rural	8358	48.09	.17	15.37			

Table 8 above indicates that the Ethiopian first cycle primary schooling is characterized not only by low level of academic achievement, but also the distribution of achievement levels favors urban schools over their rural counter parts. This finding is consistent with many

studies. For instance, Fuller (1986) reports several studies from developing countries whereby rural school students perform less than their urban counterparts. For Ethiopia which has been working for equality of educational opportunity between urban and rural schools over the last decade, the current finding is an alarm that primary school student performance in rural first cycle primary education requires improvement not only in comparison to urban primary education, but also in relation to the national standards. The above table also indicates that the highest difference between rural and urban schools was in reading comprehension. Given the key role of reading comprehension for learning other subjects and the fact that most primary schools are rural in Ethiopia, the finding suggests that strengthening the reading ability of rural students is a precondition for improved quality of primary schooling at the lower level for the nation.

4.2 Overall Achievement of the Students by Subject

4.2.1 Reading Comprehension

Table 9 shows the summary of the descriptive statistics for reading comprehension. A total of 12,886 students took the test. The mean score at national level was 64.5 with a standard deviation of 21.97. It can be seen in the table that overall performances were in the minimum acceptable range with all regions surpassing a 50% correct, on the average. Addis Ababa, Afar and Amhara scored well above the national mean 76.90, 71.92, and 70.92 respectively. The common characteristic of these three regions was that they took the test in Amharic which is a language that has been used as medium of instruction for long time. Somali scored the least with mean score of 50.53. In comparison to those three regions which scored highest, the introduction of Somali as a medium of instruction is recent.

Table 9. Summary of the descriptive statistics for reading comprehension by region

REGIONS	N	MEAN	SD
TIGRAY	947	63.37	16.77
AFAR	643	71.92	16.20
AMHARA	1641	70.92	22.08
OROMIA	2773	63.73	24.07
SOMALI	672	50.53	20.07
BEN-GUMUZ	771	60.34	21.89
SNNPR	2980	60.31	21.28
HARARI	722	68.60	19.61
ADDIS ABABA	946	76.90	17.19
DIRE DAWA	791	62.24	21.61
NATION	12886	64.49	21.97

One-way analysis of variance (Table 10) showed that there was a statistically significant mean difference between regions for reading comprehension ($F = 115.481$, $p < 0.001$). Multiple comparisons using Scheffe Post Hoc Test was conducted to identify homogenous subset groupings across regions.

Table 10. One-way analysis of variance on reading comprehension by region

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	464413.18	9	51601.46	115.481	0.00
Within Groups	5753481.65	12876	446.84		
Total	6217894.8	12885			

The homogenous subset grouping for reading comprehension resulted in 4 distinct groups (Table 11). The mean difference between the groups was statistically significant but members of each group did not differ significantly from each other. The highest performing region was Addis Ababa followed by Afar, Amhara and Harari. Somali's performance was way behind and all the other regions fell in group two.

Table 11. Homogenous subset groups for reading comprehension

Region	Number	Subset for alpha = .05			
		1	2	3	4
SOMALI	672	50.53			
SNNPR	2980		60.31		
BEN-GUMUZ	771		60.34		
DIRE DAWA	791		62.24		
TIGRAY	947		63.37		
OROMIA	2773		63.73		
HARARI	722			68.60	
AMHARA	1641			70.92	
AFAR	643			71.92	
ADDIS ABABA	946				76.90
Sig.		1	0.182	0.222	1

Table 12 shows the independent t-test for equality of mean scores for reading comprehension between males and females across regions. There are statistically significant mean differences in Tigray, Amhara, Oromia, and SNNPR in favor of males. Though not statistically significant, females in Afar, Harari and Addis Ababa performed better than males. The highest score for females was in Addis Ababa (77.04) and the least in Somali (49.45).

Table 12. T-test for equality of means for reading comprehension by gender within regions

Region	Gender	N	Mean	Std.	t	Sig.	Mean
TIGRAY	M	449	66.41	15.85	5.375	0.00	5.78
	F	498	60.63	17.12			
AFAR	M	401	71.32	16.38	-1.208	0.228	-1.59
	F	242	72.91	15.87			
AMHARA	M	904	72.52	22.92	3.257	0.001	3.57
	F	733	68.95	20.88			
OROMIA	M	1719	66.31	23.45	7.267	0.00	6.78
	F	1054	59.53	24.50			
SOMALI	M	426	50.50	20.72	0.517	0.605	1.06
	F	136	49.45	20.84			
BENISHA/GUMUZ	M	476	60.43	21.90	0.14	0.889	0.23
	F	295	60.20	21.92			
SNNPR	M	1809	62.79	20.57	7.962	0.00	6.29
	F	1169	56.49	21.80			
HARARI	M	425	68.36	19.46	-0.35	0.726	-0.52
	F	296	68.89	19.85			
ADDIS ABABA	M	433	76.74	17.68	-0.261	0.794	-0.29
	F	513	77.04	16.79			
DIRE DAWA	M	505	62.97	21.25	1.274	0.203	2.04
	F	285	60.93	22.25			
NATION	M	7547	65.71	21.78	0.000	.000	2.71
	F	5221	62.99	22.18			

The mean difference in reading comprehension between urban and rural schools was statistically significant in Amhara, Oromia, Somali, Harari, and Dire Dawa (Table 13). In all these regions except Oromia, urban schools performed better than rural ones. Though not statistically significant, rural schools performed better in Tigray and Ben-Gumuz than urban schools.

Table 13. T - test for equality of means for reading comprehension by location within regions

Region	Location	N	Mean	Std.	T	Sig.	Mean Difference
TIGRAY	U	396	62.16	15.36	-1.892	0.059	-2.09
	R	551	64.25	17.67			
AFAR	U	369	72.06	14.16	0.252	0.801	0.33
	R	274	71.73	18.61			
AMHARA	U	283	75.72	14.77	4.043	0.000	5.81
	R	1358	69.92	23.19			
OROMIA	U	548	60.21	26.85	-3.833	0.000	-4.39
	R	2225	64.60	23.26			
SOMALI	U	399	54.67	18.72	6.681	0.000	10.21
	R	273	44.47	20.47			
BEN-GUMUZ	U	277	60.11	21.08	-0.223	0.823	-0.37
	R	494	60.48	22.35			
SNNPR	U	298	61.06	19.76	0.639	0.523	0.83
	R	2682	60.23	21.44			
HARARI	U	428	74.05	18.16	9.565	0.000	13.39
	R	294	60.66	18.94			
DIRE DAWA	U	385	65.48	21.32	4.151	0.000	6.32
	R	406	59.16	21.46			
NATION	U	4329	67.44	20.73	10.871	0.000	4.43
	R	8557	63.01	22.42			

4.2.2 English

Table 14 shows the summary of the descriptive statistics for English. A total of 13,028 students took the test. The mean score at national level was 38.68 with a standard deviation of 16.88. Somali scored the highest (50.36) and Benishangul Gumuz and Tigray scored the least with mean scores of 32.44 and 32.95 respectively. The overall performance in all regions was very low for English with only Somali reaching 50% correct, on the average.

Table 14. Summary of the descriptive statistics for English by region

Regions	N	Mean	SD
TIGRAY	946	32.95	12.78
AFAR	641	35.30	13.75
AMHARA	1654	37.87	17.49
OROMIA	2962	39.49	17.94
SOMALI	670	50.36	20.98
BEN-GUMUZ	771	32.44	13.65
SNNPR	2928	38.93	15.28
HARARI	721	42.46	16.74
ADDIS ABABA	945	39.94	17.33
DIRE DAWA	790	37.19	15.75
NATION	13028	38.68	16.88

One-way analysis of variance (Table 15) shows that there is a statistically significant difference in the mean score for reading comprehension between regions ($F = 72.204$, $p < .001$). Multiple comparisons using Scheffe Post Hoc Test was conducted to identify homogenous groups across regions.

Table 15. One way analysis of variance of English across regions

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	176544.53	9	19616.06	72.204	0.000
Within Groups	3536653.48	13018	271.67		
Total	3713198	13027			

Table 16 shows that there are five homogenous subset groups. The highest performing group contains Somali followed by Harari, Addis Ababa and Oromia. From the remaining ones Beninshangul Gumuz, Tigray and Afar performed the least and fell under Group 1. When the alpha level is reduced from .05 to .01, the result of multiple comparison table shows only two distinguished groups. In this case, the five lower performing regions were significantly different from Somali Region.

Table 16. Homogeneous subset groups of English by region

Regions	N	Subset for alpha = .05				
		1	2	3	4	5
BEN-GUMUZ	771	32.44				
TIGRAY	946	32.95				
AFAR	641	35.30	35.30			
DIRE DAWA	790		37.19	37.19		
AMHARA	1654		37.87	37.87		
SNNPR	2928			38.93		
OROMIA	2962			39.49	39.49	
ADDIS ABABA	945			39.94	39.94	
HARARI	721				42.46	
SOMALI	670					50.36
Sig.		0.108	0.229	0.145	0.074	1

The mean difference between males and females in English was statistically significant in Tigray, Afar, Amhara, Oromia, Benshangul Gumuz, SNNPR, and Dire Dawa in favor of males. Though not statistically significant females in Somali and Harari performed better than males, and males in Addis Ababa performed better than females (Table 17).

Table 17. T-test for equality of means for English by gender within regions

Regions	Gender	N	Mean	Std.	t	Sig.	Mean Difference
TIGRAY	M	449	35.12	13.40	5.008	0.000	4.12
	F	497	31.00	11.88			
AFAR	M	400	36.49	13.59	2.841	0.005	3.17
	F	241	33.32	13.82			
AMHARA	M	905	40.41	18.82	6.511	0.000	5.57
	F	745	34.84	15.23			
OROMIA	M	1824	42.04	18.42	9.99	0.000	6.66
	F	1138	35.38	16.33			
SOMALI	M	425	49.11	22.20	-1.099	0.272	-2.38
	F	136	51.49	21.41			
BEN-GUMUZ	M	476	33.67	14.21	3.197	0.001	3.21
	F	295	30.46	12.46			
SNNPR	M	1788	41.37	15.55	10.993	0.000	6.25
	F	1137	35.12	14.04			
HARARI	M	424	42.35	16.66	-0.196	0.845	-0.25
	F	296	42.60	16.92			
ADDIS ABABA	M	432	40.62	18.20	1.107	0.269	1.25
	F	513	39.37	16.56			
DIRE DAWA	M	505	38.34	16.32	2.796	0.005	3.25
	F	284	35.09	14.49			
NATION	M	7628	40.55	17.39	16.398	0.000	4.89
	F	5282	35.66	15.59			

The mean difference between urban and rural schools in English was statistically significant in Tigray, Afar and Harari (Table 18). In Tigray and Afar, rural schools performed better than urban ones while in Harari, urban schools performed better than rural ones. The trend shows that there was no big gap between urban and rural schools.

Table 18. T-test for equality of means in English by location within regions

Regions	Location	N	Mean	Std. Deviation	t	Sig. (2-tailed)	Mean Difference
TIGRAY	U	396	30.13	11.92	-5.876	0.000	-4.86
	R	550	34.99	13.00			
AFAR	U	370	33.62	13.09	-3.638	0.000	-3.96
	R	271	37.58	14.31			
AMHARA	U	284	36.92	14.13	-1.008	0.313	-1.15
	R	1370	38.07	18.11			
OROMIA	U	588	40.25	19.16	1.156	0.248	0.96
	R	2374	39.30	17.62			
SOMALI	U	397	50.23	19.12	-0.192	0.848	-0.32
	R	273	50.55	23.46			
BEN-GUMUZ	U	277	32.59	14.09	0.226	0.821	0.23
	R	494	32.36	13.41			
SNNPR	U	296	38.23	12.94	-0.827	0.408	-0.77
	R	2632	39.01	15.52			
HARARI	U	428	44.83	16.64	4.656	0.000	5.83
	R	293	39.00	16.31			
DIRE DAWA	U	384	36.22	13.66	-1.684	0.093	-1.89
	R	406	38.10	17.47			
NATION	U	4365	38.87	16.84	.907	.364	.28
	R	8663	38.58	16.90			

4.2.3 Mathematics

Table 19 shows the summary of the descriptive statistics for mathematics. A total of 13,143 students took the test. The mean score at national level was 39.70 with a standard deviation of 17.51. Addis Ababa scored the highest (44.92). Overall performances were very low in mathematics, with no region reaching a 50% correct level on the average.

Table 19. Summary of the descriptive statistics for mathematics by region

Regions	N	Mean	SD
TIGRAY	945	34.50	13.92
AFAR	643	40.29	16.80
AMHARA	1711	43.27	20.13
OROMIA	2932	38.30	17.33
SOMALI	670	43.09	18.98
BEN-GUMUZ	771	34.34	15.00
SNNPR	2986	39.70	16.87
HARARI	719	40.71	16.47
ADDIS ABABA	981	44.92	18.28
DIRE DAWA	785	37.83	15.77
NATION	13143	39.70	17.51

One-way analysis of variance (see Table 20 below) shows that there was a statistically significant difference in the mean score for reading comprehension between regions ($F = 42.22$, $p < .001$). Multiple comparisons using Scheffe Post Hoc Test was conducted to identify homogenous groups across regions. The following table shows the results.

Table 20. One way analysis of variance on mathematics by region

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	113367.12	9	12596.35	42.221	0.00
Within Groups	3918117.3	13133	298.34		
Total	4031484.4	13142			

Table 21 below shows that there are 4 groups. Addis Ababa, Amhara and Somali fell in the upper most group followed by Afar and Harari. Benshangul Gumuz and Tigray fell in the lowest group.

Table 21. Homogenous subset groups of mathematics by region

Region	N	Subset for alpha = .05			
		1	2	3	4
BEN-GUMUZ	771	34.34			
TIGRAY	945	34.50			
DIRE DAWA	785		37.83		
OROMIA	2932		38.30		
SNNPR	2986		39.70		
AFAR	643		40.29	40.29	
HARARI	719		40.71	40.71	
SOMALI	670			43.09	43.09
AMHARA	1711			43.27	43.27
ADDIS ABABA	981				44.92
Sig.		1	0.143	0.109	0.796

The mean difference in mathematics between males and females in Tigray, Afar, Amhara, Oromia, Benshangul Gumuz, SNNPR and Dire Dawa was statistically significant in favor of males (see Table 22). Though not statistically significant, females in Somali and Harari performed better than males.

Table 22. T-test of mean difference for mathematics by gender within regions

Region	Gender	N	Mean	Std. Deviation	t	Sig. (2-tailed)	Mean Difference
TIGRAY	M	449	35.12	13.40	5.008	0.000	4.12
	F	497	31.00	11.88			
AFAR	M	400	36.49	13.59	2.841	0.005	3.17
	F	241	33.32	13.82			
AMHARA	M	905	40.41	18.82	6.511	0.000	5.57
	F	745	34.84	15.23			
OROMIA	M	1824	42.04	18.42	9.99	0.000	6.66
	F	1138	35.38	16.33			
SOMALI	M	425	49.11	22.20	-1.099	0.272	-2.38
	F	136	51.49	21.41			
BEN-GUMUZ	M	476	33.67	14.21	3.197	0.001	3.21
	F	295	30.46	12.46			
SNNPR	M	1788	41.37	15.55	10.993	0.000	6.25
	F	1137	35.12	14.04			
HARARI	M	424	42.35	16.66	-0.196	0.845	-0.25
	F	296	42.60	16.92			
ADDIS ABABA	M	432	40.62	18.20	1.107	0.269	1.25
	F	513	39.37	16.56			
DIRE DAWA	M	505	38.34	16.32	2.796	0.005	3.25
	F	284	35.09	14.49			
Nation	M	7689	41.68	18.15	16.24	0.000	5.02
	F	5350	36.67	16.13			

The mean difference of mathematics between urban and rural schools was statistically significant in Tigray, Afar, Amhara, Somali, SNNPR, Harari and Dire Dawa (Table 23). In Tigray, Afar, Amhara and SNNPR, rural schools performed better than urban ones. In Somali, Harari and Dire Dawa, urban schools performed better than rural schools.

Table 23. T-test of mean difference for mathematics by location within regions

Region	Location	N	Mean	Std. Deviation	t	Sig. (2-tailed)	Mean Difference
TIGRAY	U	397	32.66	12.92	-3.483	0.001	-3.18
	R	548	35.83	14.47			
AFAR	U	367	38.48	15.88	-3.164	0.002	-4.21
	R	276	42.69	17.70			
AMHARA	U	317	40.55	16.48	-2.673	0.008	-3.34
	R	1394	43.89	20.82			
OROMIA	U	581	38.69	17.99	0.602	0.547	0.48
	R	2351	38.20	17.17			
SOMALI	U	396	45.63	19.63	4.217	0.000	6.21
	R	274	39.42	17.40			
BEN-GUMUZ	U	277	34.78	15.07	0.6	0.549	0.68
	R	494	34.10	14.98			
SNNPR	U	300	34.44	13.07	-5.724	0.000	-5.85
	R	2686	40.28	17.14			
HARARI	U	429	44.64	16.80	8.128	0.000	9.74
	R	290	34.90	14.10			
DIRE DAWA	U	382	39.47	15.42	2.85	0.004	3.20
	R	403	36.27	15.96			
NATION	U	4427	40.38	17.28	3.165	0.002	1.02
	R	8716	39.35	17.62			

4.2.4 Environmental Science

Table 24 shows the summary of the descriptive statistics for environmental science. A total of 13005 students took the test. The mean score at national level was 51.74 with a standard deviation of 17.42. Amhara scored the highest (60.05), well above the national mean. The overall regional performances in environmental science were average, with most regions approximating 50% correct, on the average.

Table 24. Summary of the descriptive statistics for environmental science by region

Regions	N	Mean	SD
TIGRAY	938	45.87	11.88
AFAR	604	51.60	13.77
AMHARA	1708	60.05	17.19
OROMIA	2932	50.73	18.63
SOMALI	632	45.83	22.30
BEN-GUMUZ	768	47.84	16.75
SNNPR	2977	49.75	16.14
HARARI	720	53.35	16.86
ADDIS ABABA	929	54.64	14.02
DIRE DAWA	789	47.86	16.17
NATION	13005	51.74	17.42

One-way analysis of variance (Table 25) shows that there is a statistically significant difference in the mean score for reading comprehension between regions ($F = 72.44$, $p < .001$). Multiple comparisons using Scheffe Post Hoc Test was conducted to identify homogenous groups across regions.

Table 25. One way analysis of variance on environmental science by region

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	188474.64	9	20941.63	72.435	0.00
Within Groups	3756994.7	12995	289.11		
Total	3945469.3	13004			

Table 26 shows that there were six groups. Amhara was in the upper most group followed by Addis Ababa, Harari, Tigray, and Afar. Somali, Beninshangul Gumuz, and Dire Dawa fell in the lowest group. When the alpha level was reduced to .01, the result of multiple comparison tests showed Somali as performing significantly lower than the higher performing region of Amhara.

Table 26. Homogenous subset groups of environmental science by region

Region	N	Subset for alpha = .05					
		1	2	3	4	5	6
SOMALI	632	45.83					
BEN-GUMUZ	768	47.84	47.84				
DIRE DAWA	789	47.86	47.86				
SNNPR	2977		49.75	49.75			
OROMIA	2932		50.73	50.73	50.73		
AFAR	604			51.60	51.60	51.60	
TIGRAY	946			52.50	52.50	52.50	
HARARI	720				53.35	53.35	
ADDIS ABABA	929					54.64	
AMHARA	1708						60.05
Sig.		0.668	0.134	0.195	0.259	0.087	1

The mean difference in environmental science between males and females was statistically significant in Tigray, Amhara, Oromia, SNNPR, and Dire Dawa in favor of males (see Table 27). Though not statistically significant, females performed better than males in Somali and Harari.

Table 27. T -test of mean difference for environmental science by gender within regions

Region	Gender	N	Mean	Std. Dev.	t	Sig. (2-tailed)	Mean Difference
TIGRAY	M	448	55.80	15.70	6.377	0.000	6.26
	F	498	49.54	14.49			
AFAR	M	375	52.22	14.84	1.423	0.155	1.64
	F	229	50.58	11.77			
AMHARA	M	938	63.24	17.47	8.669	0.000	7.11
	F	766	56.12	16.05			
OROMIA	M	1804	53.77	18.84	11.419	0.000	7.90
	F	1128	45.87	17.22			
SOMALI	M	410	46.91	22.69	-1.268	0.205	-2.91
	F	126	49.82	21.93			
BENISHA/GUMUZ	M	473	47.90	17.47	0.139	0.889	0.17
	F	295	47.73	15.56			
SNNPR	M	1804	51.62	16.18	7.881	0.000	4.72
	F	1172	46.89	15.65			
HARARI	M	425	52.55	16.91	-1.544	0.123	-1.97
	F	295	54.52	16.74			
ADDIS ABABA	M	409	54.91	15.04	0.513	0.608	0.48
	F	520	54.43	13.17			
DIRE DAWA	M	504	49.27	15.95	3.27	0.001	3.90
	F	284	45.37	16.32			
NATION	M	7590	53.43	17.89	12.38	0.000	3.82
	F	5313	49.61	16.33			

The mean difference in environmental science between urban and rural schools was statistically significant in Tigray, Afar, Amhara, Somali, and SNNPR (see Table 28). In Tigray, Afar, Amhara and SNNPR rural schools performed better than urban ones. In Somali and Harari, urban schools performed better than rural schools.

Table 28. T-test for mean difference in environmental science by location within regions

Region	Location	N	Mean	Std. Dev.	T	Sig. (2-tailed)	Mean Difference
TIGRAY	U	397	47.39	14.42	-9.061	0.000	-8.82
	R	549	56.20	15.01			
AFAR	U	330	49.43	12.04	-4.304	0.000	-4.77
	R	274	54.21	15.22			
AMHARA	U	316	57.48	14.38	-2.942	0.000	-3.14
	R	1392	60.63	17.72			
OROMIA	U	583	51.40	19.62	0.973	0.33	0.84
	R	2349	50.56	18.38			
SOMALI	U	358	48.97	23.54	4.094	0.000	7.24
	R	274	41.73	19.89			
BEN-GUMUZ	U	277	47.70	15.64	-0.17	0.865	-0.21
	R	491	47.91	17.36			
SNNPR	U	300	45.21	14.14	-5.167	0.000	-5.06
	R	2677	50.26	16.27			
HARARI	U	429	57.60	16.27	8.626	0.000	10.52
	R	291	47.09	15.73			
DIRE DAWA	U	385	48.19	14.84	0.565	0.572	0.65
	R	404	47.54	17.36			
NATION	U	4304	51.49	16.64	-1.176	0.240	-.38
	R	8701	51.87	17.79			

4.2.5 Composite Score

Table 29 shows the summary of the descriptive statistics for the composite score. A total of 12, 507 students had mean composite score. Only those students who sat for all the 4 tests had mean composite score. The mean score at national level was 48.48 % with a standard deviation of 15.07. Addis Ababa, Amhara, and Harari scored well above the national mean 53.69%, 52.67%, and 51.29% respectively. Benshangul Gumuz scored the least, 43.69%.

Table 29. Summary of the descriptive statistics for composite mean by region

Regions	N	Mean	SD
TIGRAY	938	45.87	11.88
AFAR	594	49.64	12.01
AMHARA	1597	52.67	15.86
OROMIA	2724	47.83	16.59
SOMALI	613	47.18	15.80
BEN-GUMUZ	768	43.69	13.67
SNNPR	2886	47.27	14.40
HARARI	716	51.29	14.53
ADDIS ABABA	888	53.69	13.65
DIRE DAWA	783	46.20	14.58
NATION	12507	48.48	15.07

One-way analysis of variance (see Table 30 below) shows that there was statistically significant difference in the mean composite score between regions in the mean composite score ($F = 47.07$ at $p < .001$). Multiple comparisons using Scheffe Post Hoc Test was applied to identify homogenous groupings across regions.

Table 30. One way analysis of variance on composite mean by region

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	93063.104	9	10340.35	47.07	0.000
Within Groups	2745642.9	12497	219.70		
Total	2838706	12506			

Table 31 shows that there were five groups. Addis Ababa, Amhara and Harari fell in the upper most group. Benshangul Gumuz, Tigray and Dire Dawa fell in the lowest group.

Table 31. Homogenous subset groups of mean composite score by region

Region	N	Subset for alpha = .05				
		1	2	3	4	5
BEN-GUMUZ	768	43.69				
TIGRAY	938	45.87	45.87			
DIRE DAWA	783	46.20	46.20			
SOMALI	613		47.18	47.18		
SNNPR	2886		47.27	47.27		
OROMIA	2724		47.83	47.83		
AFAR	594			49.64	49.64	
HARARI	716				51.29	51.29
AMHARA	1597					52.67
ADDIS ABABA	888					53.69
Sig.		0.155	0.521	0.174	0.767	0.21

The mean difference in composite score between males and females was statistically significant in Tigray, Amhara, Oromia, SNNPR, Benshangul Gumuz and Dire Dawa in favor of males (Table 32).

Table 32. T -test of mean difference for composite mean by gender within regions

Region	Gender	N	Mean	Std. Dev	T	Sig. (2-tailed)	Mean Difference
TIGRAY	M	446	48.82	11.93	7.46	0.000	5.63
	F	492	43.19	11.20			
AFAR	M	368	50.31	12.58	1.74	0.082	1.76
	F	226	48.55	10.96			
AMHARA	M	876	55.41	16.39	7.77	0.000	6.10
	F	717	49.32	14.56			
OROMIA	M	1693	50.37	16.89	10.42	0.000	6.70
	F	1031	43.67	15.20			
SOMALI	M	398	46.60	16.63	-1.37	0.172	-2.39
	F	120	48.99	17.27			
BEN-GUMUZ	M	473	44.49	14.27	2.03	0.042	2.10
	F	295	42.43	12.57			
SNNPR	M	1759	49.62	14.48	11.18	0.000	6.02
	F	1126	43.60	13.50			
HARARI	M	421	51.05	14.43	-0.54	0.587	-0.60
	F	295	51.65	14.69			
ADDIS ABABA	M	390	54.09	14.25	0.77	0.440	0.71
	F	498	53.37	13.16			
DIRE DAWA	M	500	47.30	14.59	2.83	0.005	3.10
	F	282	44.24	14.39			
Nation	M	7324	50.13	15.43	14.69	0.000	4.02
	F	5082	46.12	14.31			

The mean difference in composite score between urban and rural schools was statistically significant in Tigray, Afar, Somali, SNNPR, Harari, and Dire Dawa (Table 33). In Tigray, Afar, and SNNPR rural schools performed better than urban schools. In Somali, Harari and Dire Dawa urban schools performed better.

Table 33. T-test of mean difference for composite mean by location within regions

Region	Location	N	Mean	Std. Dev	T	Sig. (2-tailed)	Mean Difference
TIGRAY	U	392	43.10	10.52	-6.16	0.000	-4.76
	R	546	47.86	12.40			
AFAR	U	327	48.09	10.45	-3.52	0.000	-3.45
	R	267	51.54	13.47			
AMHARA	U	279	52.44	12.29	-0.26	0.792	-0.28
	R	1318	52.72	16.52			
OROMIA	U	540	47.30	17.88	-0.84	0.403	-0.67
	R	2184	47.97	16.25			
SOMALI	U	342	49.72	16.89	4.54	0.000	5.75
	R	271	43.97	13.68			
BEN-GUMUZ	U	277	43.79	13.07	0.15	0.88	0.15
	R	491	43.64	14.02			
SNNPR	U	295	44.72	11.25	-3.21	0.001	-2.83
	R	2591	47.56	14.69			
HARARI	U	428	55.28	14.17	9.49	0.000	9.91
	R	288	45.37	12.97			
DIRE DAWA	U	381	47.26	13.16	2.00	0.046	2.08
	R	402	45.19	15.75			
NATION	U	4149	49.28	14.41	4.17	0.000	0.29
	R	8358	48.09				

4.3 Grade 4 Pupils' Background and Academic Achievement

A student background questionnaire was used as an instrument to gather the student background information. It contained a number of different items that helped the study team to collect general information about distance to school, watching television, listening to the radio, doing home work, interests in school subjects, etc. Findings of the study substantiate but sometimes challenge the conventional wisdom concerning the relationships between student background characteristics and academic achievement. For example, parents' level of education, tutorial support, teachers' qualification, teachers' teaching experiences and many other characteristics had either negative or no relationships with students' academic achievement. Some detailed analyses of such relationships, therefore, are indicated in this part of the study.

Table 34. Time needed to reach school and academic achievement.

	N	Composite Mean	Std. Deviation	% of Total N
up to 60 min.	10370	49.21	.1471	85.6%
Above 60 min.	1745	45.72	.1537	14.4%
Total	12115	48.70	.1485	100.0%

As Table 34 summarizes, students who travel up to 60 minutes scored a better mean (49) in the Grade 4 composite mean score than those who travel above 60 minutes (45.72). This implies that the longer the distance a student traveled from home to school the lower the score she/he achieved in academic subjects. This finding is not surprising as it is obvious that students who travel long distances spend their study time on the way. In fact upon arriving at schools, they may find it difficult to participate properly and attend class lessons as they are already tired.

Table 35. ANOVA on time required to reach school

Source of variations	SS	Df	MS	F	Sig.
Between Groups	1.819	1	1.819	82.971	0.001
Within Groups	265.495	12113	.02192		
Total	267.313	12114			

In Table 35, it is indicated that the F-ratio calculated on the observed mean difference between the two groups ($F_{(1, 12113)} = 82.971$, $P < .001$) was statistically significant. This implies that distance between home and school can influence Grade 4 pupils' academic achievement in Ethiopia. Moreover, the study identified the correlation between Grade 4 composite mean and distance from home to school ($r = -.082$) to be significant at the 0.01 level (2-tailed).

Table 36. Watching non-educational television and student achievement

Number of Days Devoted to Watching TV	Grade 4 Composite Mean	N	Std. Deviation	% of Total N
Not at all	50.94	6348	.1478	55.6%
1-2 days a week	48.68	1977	.1404	17.3%
3-4 days a week	44.84	801	.1400	7.0%
5-7 days a week	44.65	2284	.1474	20.0%
Total	48.86	11410	.1483	100.0%

Students were asked how frequently they watch television programs per week. The data indicated in Table 36 show that 55.6% of the sampled students selected the “Not at all” alternative, while 20% of them said they watch TV programs for 5 to 7 days a week. The finding regarding this item is negatively correlated ($r = -.177$) with the Grade 4 composite achievement. This correlation is significant at the 0.01 level (2-tailed) implying that as the number of days for watching TV per week increases, the level of students’ performance is likely to decrease (see Table 36 above). The negative relationship between TV viewing and student achievement is well recorded in many studies. As indicated in the review of literature in this report, it has been found out that TV viewing for leisure, as the other student characteristics, is inversely associated with student achievement. Obviously, spending time in front of a television screen takes away the time a student may use for studying or doing home work.

Table 37. Listening to the radio (other than educational radio) per week

Number of Days Devoted to Listening to the Radio.	Composite Mean	N	Std. Deviation	% of Total N
Not at all	49.22	2743	.1488	23.5%
1-2 days a week	50.21	2539	.1419	21.8%
3-4 days a week	49.99	1964	.1508	16.8%
5-6 days a week	47.30	4423	.1479	37.9%
Total	48.84	11669	.1479	100.0%

Students were also asked to rate how frequently they listen to radio programs per week at home. Out of the total number, 23.5% of the sampled students replied that they did not waste any single day listening to the radio; nevertheless, their achievement level (49.22%) was lower than the score (50.21%) of those students who reported that they devote 1-2 days per week to listening to the radio. In general, when the findings summarized in Table 37 are related to performance, for every group except the “Not at all” option raters, the greater the amount of time spent on listening to the radio per week, the poorer the performance in 4 of the Grade 4 subjects.

Table 38. Absenteeism and achievement

Frequency	Mean	N	Std. Dev	% of Total N
None	50.24	5157	.1419	42.8%
1-5 days	49.78	4546	.1449	37.7%
6-10 days	46.66	1014	.1597	8.4%
11-15 and more days	41.28	1338	.1477	11.1%
Total	48.77	12055	.1479	100.0%

Students' absenteeism is directly related to the level of their achievement as shown in Table 38. In this particular case, the relationship between absenteeism and performance was ($r = -.172$) significant but negatively correlated. Unless a student attends classroom instruction, it is obvious that not much can be expected in terms of academic achievement. In Ethiopia, students may miss classes due to a number of reasons. For instance, the qualitative study of this report indicates that parental demands for children's labor at home plays an important role. Students from economically low families may quit classes for doing some business on local market days to support themselves in school. Health and other factors may also contribute to student absenteeism in Ethiopia.

Table 39. Frequency at which attendance is taken

Frequency	Mean	N	Std. Deviation	% of Total N
Never	35.83	619	.1247	5.1%
Sometimes	44.49	1961	.1527	16.2%
Always	50.40	9506	.1432	78.7%
Total	48.70	12086	.1486	100.0%

To minimize the negative impact of pupils' absenteeism, taking frequent attendance by teachers is identified as a solution to the problem. Table 39 indicates that as the frequency of checking students' attendance increases, performance is likely to increase. These two variables correlate significantly in the positive direction with $r = .247$. This implies that taking attendance regularly has a positive contribution to students' achievement.

Table 40. Daily meals and achievement

Daily meals	Mean	N	Std. Deviation	% of Total N
Once	41.97	1007	.1487	8.4%
Twice	49.01	4175	.1479	34.6%
Three or more times	49.60	6877	.1458	57.0%
Total	48.76	12059	.1482	100.0%

From the Table above, it can be observed that a considerable number (43%) of the sampled students did not get sufficient number of meals. However, more than half of the students reported fed themselves three times a day. These responses are quite associated with the levels of students' overall performance in school subjects. The correlation between these two variables was ($r = .109$) significant at .01 level of P value, indicating that getting sufficient number of meals is a contributing factor for children to achieve well in school. Fuller (1986) refers to studies in Guatemala, Egypt and Uganda to show how malnutrition brought about lower student achievement (in verbal skills, reading, and mathematics and comprehensive exams) by causing illness in students. In Ethiopia, the current finding is extremely relevant because schools do not provide feeding services under normal circumstances. School feeding

is significant because little children can not grow physically and intellectually to their full potential unless they eat well. In this respect, gaining experiences from other developing countries seems essential to learn how they addressed this important dimension in lower primary education.

Table 41. Interest in learning English

Level of Interest	Mean	N	Std. Deviation	Min	Max	% of Total N
Low	38.20	738	.1251	.06	.82	6.1%
Average	46.05	3060	.1358	.03	.89	25.5%
High	50.80	8203	.1486	.01	.92	68.4%
Total	48.81	12001	.1480	.01	.92	100.0%

With regard to the level of students' interest towards learning the English language, Table 41 provides the responses of three different groups; "Low", "Average" and "High". Of the respondents, 68.4% said that they had "High" interest in the subject while 25.5% rated their interest as "Average"; 6.1% put their interest as "Low".

These different levels of students' interest positively correlated ($r = .227$) with their achievements in the subject, and the correlation was significant at the 0.01 level (2-tailed). That is, those who showed high interest in the subject achieved a high composite mean score of 50.80%, while those with an "Average" level of interest achieved the next higher mean score (46.05%). Those who reported the level of their interest in English as "Low" achieved the least mean score (38.20%). This implies that achievement in a particular subject can, to some extent, be gauged by the degree of interest in that particular subject.

Table 42. Interest in learning mathematics

Level of Interest	Composite Mean	N	Std. Deviation	% of Total N
Low	37.30	667	.1225	5.6%
Average	45.57	2652	.1432	22.1%
High	50.76	8679	.1449	72.3%
Total	48.87	11998	.1477	100.0%

Background information gathered from the sampled students with regard to their attitudes toward learning mathematics, indicated that 72.3% had a high degree of interest. However, 22.1% of them rated their interest as "average" and 5.6% as "low". These responses positively correlated ($r = .236$) with performances in the subject, and the correlation was significant at the 0.01 level (2-tailed). That is, those who reported to have high interest achieved a composite mean score of 51%, which is higher than the scores of the other two groups. This implies that performance in a particular subject can meaningfully be influenced by the level of interest in learning that subject.

Table 43. Interest in learning environmental science

Level of Interest	composite Mean	N	Std. Deviation	% of Total N
1 Low	37.43	622	.1244	5.2%
2 Average	45.76	2202	.1524	18.5%
3 High	50.35	9062	.1441	76.2%
Total	48.82	11886	.1482	100.0%

Table 43 shows the interest of Grade 4 sampled students in the learning of environmental science. Of the total number of respondents, 76.2% exhibited “High” interest in the subject while 18.5% showed “Average” interest. 5.2% of the sampled students rated their interest as “Low.” The levels of interest positively correlated ($r = .213$) with performances in the subject, and the correlation was significant at the 0.01 level (2-tailed). Students who reported to have high interest in the subject achieved a high composite mean score of 50.35% while, those with an “Average” level of interest achieved the next higher mean score (45.76). Students with “Low” interest achieved the least mean score (37.43%).

Table 44. English textbook distribution

Textbook Ratio	Composite Mean	N	Std. Deviation	% of Total N
1:1	49.65	5445	.1447	45.3%
1:2	48.29	1887	.1476	15.7%
1:3	48.01	863	.1655	7.2%
1:4	45.89	1251	.1543	10.4%
No textbook available	48.78	2587	.1439	21.5%
Total	48.74	12033	.1480	100.0%

Table 44 indicates that 45.3% of the students had an English textbook each. This group had the highest mean scores in relation to others. The majority of students, however, had one textbook for two or more and they scored lower than those who had a textbook each.

Table 45. Mathematics textbook distribution

Mathematics textbook ratio	composite Mean	N	Std. Deviation	% of Total N
1:1	50.25	6064	.1447	50.6%
1:2	47.78	2157	.1479	18.0%
1:3	43.77	673	.1532	5.6%
1:4	44.02	901	.1504	7.5%
No textbook available	48.84	2182	.1465	18.2%
Total	48.71	11977	.1480	100.0%

In Table 45, slightly over half of the students said that mathematics textbook- student ratio was 1:1, and this group scored higher than others. Yet a considerable proportion uses a textbook for two or more. These ones scored lower than their privileged peers.

Table 46. Environmental science textbook distribution

Envi Science Textbook ratio	composite Mean	N	Std. Deviation	% of Total N
1:1	50.05	5957	.1453	49.4%
1:2	48.40	2242	.1483	18.6%
1:3	44.68	772	.1533	6.4%
1:4	45.80	902	.1532	7.5%
No textbook available	48.02	2188	.1442	18.1%
Total	48.71	12061	.1477	100.0%

The above table shows that about half of Grade 4 students do not have their own environmental science textbooks. In fact, these students are also the ones who are at the disadvantage when it comes to academic achievement.

In summary, it can be concluded from the three tables that students' performance increases as the student-textbook ratio increases. The correlation (r) between Grade 4 English, mathematics, and environmental science textbook-student ratio and overall average score was .043, .073, .075 respectively, indicating that the better access to learning materials makes a difference in pupils' achievement. However, a considerable proportion of students do not have sufficient textbooks although some reports indicate that much work has been done in the sphere in Ethiopia. Referring to different studies, UNESCO (2004) indicates that lack of textbooks can result from an inefficient distribution system, and malpractice or corruption. In Zambia, for instance, it notes that not even 10% of the books procured had reached classrooms. UNESCO's suggestion is that liberalizing textbook markets and strong government coordination may contribute to drop textbook prices and higher availability.

4.4 Grade 4 Pupils' Attitudes in Socially Relevant Issues

The following shows the response frequency distributions for a health care, environmental protection, civics and ethics, cultural conditions, and education subscales. For analyzing mean differences, student responses depicting favorable attitude were assigned a score of 3 those depicting unfavorable attitude 1, and neutral responses 2. Negatively phrased items were shaded. Overall, attitudes tend to be in the favorable direction, suggesting that Ethiopia's social development curriculum is making a difference in positively shaping students' attitudes around health, environmental protection, civics and ethics, cultural issues and education values. Total number of students that responded to each of the attitude questions varied, ranging from 12616 to 12843.

4.4.1 Grade 4 Pupils' Attitudes in Socially Relevant Issues at National Level

Table 47. Attitude towards health care

Items	Percentage of Respondents		
	Agree	Neutral	Disagree
We can prevent the spread of diseases through personal and environmental hygiene.	79.27	10.52	10.21
I would like my parents to give birth over and over again so that I can have many brothers and sisters.	45.46	12.17	42.37
As far as we are aware of the modes of transmission of HIV/AIDS we will not be infected by the virus.	68.83	10.79	20.38
Blood tests cannot help identify HIV/AIDS infection.	34.62	16.67	48.71
Helping HIV/AIDS infected people is a social obligation.	63.09	11.54	25.37
Learning with those with HIV/AIDS causes no problem.	50.17	12.21	37.62
Eating meat regularly is a proper dietary practice.	35.79	12.40	51.81

Majority of the students (79.27%) agreed that with sanitation diseases could be prevented. The second highest percentage (68.83) showed agreement that if they knew how HIV/AIDS is transmitted and took care, they would not be infected. On the other hand, 51.81 % disagreed with the statement 'Eating meat regularly is a proper dietary practice' and 45.46 % agreed with the statement 'I would like my parents to give birth over and over again so that I have many brothers and sisters'; 48.71% disagreed with the statement 'Blood test can not help identify HIV/AIDS infection' where the neutrals (16.67%) and the agreed (34.62 %) made the 51.29%.

Table 48. Attitude towards environmental protection

Items	Agree	Neutral	Disagree
Drinking river water can expose one to water borne diseases.	58.46	10.76	30.78
Visiting historical places is wasting both money and time.	42.73	11.98	45.29
Planting and taking care of trees is pleasurable.	84.97	6.25	8.78
One should plant trees to prevent erosion.	80.71	6.46	12.83
Keeping the environment clean is everyone's duty.	77.78	7.81	14.41
It is possible to prevent the consequences of drought.	65.82	11.66	22.52
It is necessary to burn forests to drive away wild animals that destroy crops.	29.59	10.05	60.36

84.97 % agreed that "Planting and taking care of trees is pleasurable," 80.71% agreed that one should plant trees to prevent erosion and 77.78 % agreed that "keeping the environment clean is every one's duty"; 45.29 % disagreed with the statement "visiting historical places is wasting time and money" as opposed to the 42.73% who agreed on the same statement.

Table 49. Attitude towards civics and ethics

Items	Agree	Neutral	Disagree
Women are equal to men in knowledge and thinking.	57.03	11.00	31.97
It is appropriate to return property found lost either to its owner or to the school administration.	79.55	7.41	13.03
I believe that group work is important for academic success.	79.66	7.44	12.90
Gathering information about our surrounding is necessary to have knowledge about it.	77.17	8.76	14.07
Solving differences through dialogue is good culture.	79.52	7.38	13.10
Boys and girls should equally participate in household chores after school.	79.19	7.08	13.73

While 57.03% agreed to the statement “women are equal to men in knowledge and thinking, 31.97% disagreed. The statements “Lost property should be returned” (79.55%), “Group work is important for academic success” (79.66%), “solving differences through dialogue is good (79.52%) and “Boys and girls should equally participate in household chores” (79.19%) won high percentages.

Table 50. Attitudes towards harmful cultural practices

Items	Agree	Neutral	Disagree
Early marriage is a practice that should be abolished.	56.38	9.25	34.38
Female circumcision is a harmful traditional practice that should be abolished.	62.55	9.61	27.84

56.38% of the students agreed to the statement “Early marriage should be abolished” while 34.38% disagreed. “Female circumcision is a harmful traditional practice that should be abolished” won the agreement of 62.55% of the respondents.

Table 51. Attitude towards education

Items	Agree	Neutral	Disagree
Absenteeism and coming late to school are signs of laziness.	71.04	8.05	20.91
Student participation in teachers’ evaluation is appropriate.	79.41	6.80	13.80
Going to school and staying at home are the same to me	29.77	10.45	59.78
Learning is a means to earning a higher status.	84.85	5.37	9.79
Because I am a student, I have more knowledge and maturity than those not in school.	81.28	6.70	12.03

“Learning is a means to earn a higher status” (84.85%), “Students have more knowledge and maturity than those not in schools” (81.28%) and “Students should be involved in teacher evaluations “(79.41%) were statements that won extensive agreement. Those who disagreed with the statement “Going to school and staying home are the same to me” (59.78%) were relatively fewer than those in the former groups.

4.4.2 Grade 4 Pupils’ Attitudes towards Socially Relevant Issues at Regional Level

Analysis of subscale mean differences and region-by-response cross-tabulations provided insight into social awareness differences across regions. For analyzing mean differences, student responses depicting favorable attitude were assigned a score of 3 while those depicting unfavorable attitude were assigned a score of 1. Neutral responses were assigned a score of 2. The Scheffe paired comparisons procedure (ANOVA post-hoc tests) was applied to identify homogeneous groupings across regions using an alpha level of 0.005. The alpha

level was set at 0.005 to compensate for inflation of Type I error that can be expected when conducting more than one statistical test on the same population.

The following tables show the within region distribution of responses for individual attitude survey questions.

Table 52. We can prevent the spread of diseases through personal and environmental hygiene

Region	Agree	Neutral	Disagree
Tigray	66.63	19.79	13.58
Afar	85.71	7.54	6.75
Amhara	91.28	4.30	4.42
Oromia	73.42	12.38	14.20
Somali	81.95	5.30	12.75
Ben-Gumuz	81.72	9.01	9.27
SNNPR	76.11	13.67	10.22
Harari	81.69	9.53	8.77
Addis Ababa	90.96	4.21	4.83
Dire Dawa	75.06	11.05	13.88
NATION	79.27	10.52	10.21

Table 52 shows that big majority of respondents in Amhara region (91.28%) and Addis Ababa (90.96%) agreed that personal and environmental hygiene could help prevent the spread of diseases. Respondents from Tigray stood last in order of agreement (66.63).

Table 53. I would like my parents to give birth over and over again so that I can have many brothers and sisters.

Region	Agree	Neutral	Disagree
Tigray	50.05	10.96	38.99
Afar	64.77	7.42	27.80
Amhara	44.87	5.37	49.76
Oromia	43.70	19.56	36.74
Somali	76.49	7.62	15.89
Ben-Gumuz	56.61	7.21	36.18
SNNPR	40.91	12.42	46.67
Harari	39.03	13.01	47.96
Addis Ababa	27.26	9.77	62.96
Dire Dawa	42.45	14.71	42.84
NATION	45.46	12.17	42.37

Table 53 shows the between region differences of students' attitude regarding family planning. The overall national percentage of those who agreed to the statement and those who disagreed was very close (45.46% and 42.37% respectively). For most of the regions there was no great difference in the percentage between those who agreed and those who disagreed. The highest percentage difference was between respondents from Somali region (76.49% agreed, and 15.89% disagreed). In the Afar region too, the gap between those who agreed and the two groups was considerable (64.77% agreed and 27.80% disagreed). A relatively smaller percentage of respondents in Addis Ababa (27.26%) agreed with the statement.

Table 54. As far as we are aware of the modes of transmission of HIV/AIDS and take care, we will not be infected by the virus.

Region	Agree	Neutral	Disagree
Tigray	58.49	14.52	26.99
Afar	69.50	8.99	21.51
Amhara	71.03	6.93	22.04
Oromia	67.02	13.04	19.94
Somali	74.96	6.84	18.20
Ben-Gumuz	70.03	9.41	20.56
SNNPR	66.91	12.50	20.59
Harari	74.32	8.36	17.33
Addis Ababa	78.04	5.41	16.55
Dire Dawa	67.79	14.55	17.66
NATION	68.83	10.79	20.38

Table 55. Blood tests cannot help identify HIV/AIDS infection

Region	Agree	Neutral	Disagree
Tigray	30.78	20.04	49.18
Afar	32.10	15.32	52.58
Amhara	32.47	13.00	54.53
Oromia	36.73	18.19	45.08
Somali	47.83	12.54	39.63
Ben-Gumuz	37.23	14.52	48.25
SNNPR	32.69	21.30	46.00
Harari	33.54	10.87	55.59
Addis Ababa	26.95	11.24	61.81
Dire Dawa	43.27	15.56	41.18
NATION	34.62	16.67	48.71

Table 56. Helping HIV/AIDS infected people is a social obligation

	Agree	Neutral	Disagree
Tigray	50.81	14.70	34.49
Afar	75.12	8.87	16.01
Amhara	65.63	6.61	27.75
Oromia	60.30	15.81	23.89
Somali	64.74	8.24	27.02
Ben-Gumuz	62.92	9.50	27.58
SNNPR	52.64	13.86	33.51
Harari	74.55	9.55	15.91
Addis Ababa	89.98	3.20	6.82
Dire Dawa	67.36	12.95	19.69
NATION	63.09	11.54	25.37

Table 57. Learning with those with HIV/AIDS causes no problem

Region	Agree	Neutral	Disagree
Tigray	42.75	13.65	43.60
Afar	61.07	6.91	32.03
Amhara	53.08	6.58	40.34
Oromia	43.04	18.79	38.17
Somali	49.75	9.09	41.16
Ben-Gumuz	54.65	8.38	36.97
SNNPR	42.38	13.43	44.19
Harari	62.97	9.71	27.31
Addis Ababa	74.77	5.74	19.49
Dire Dawa	53.49	14.73	31.78
NATION	50.17	12.21	37.62

Tables 54, 55, 56 and 57 show the attitude of the respondents regarding HIV/AIDS issues. These tables show that respondents from Addis Ababa have better awareness of the issue. The number of respondents undecided about the issue was, however, not insignificant.

Table 58. Eating meat regularly is a proper dietary practice.

Region	Agree	Neutral	Disagree
Tigray	28.48	14.45	57.07
Afar	44.69	8.68	46.62
Amhara	34.36	6.21	59.43
Oromia	35.02	16.89	48.09
Somali	54.93	6.74	38.32
Ben-Gumuz	40.38	9.96	49.66
SNNPR	35.63	15.05	49.31
Harari	35.66	8.80	55.54
Addis Ababa	21.16	8.38	70.46
Dire Dawa	42.34	14.29	43.38
NATION	35.79	12.40	51.81

Table 58 shows that 70.46% of the respondents from Addis Ababa disagreed with the statement while only 38.32% disagreed from Somali Region. Those from the other regions varied between the two limits.

Table 59. Drinking river water can expose one to water borne diseases.

Region	Agree	Neutral	Disagree
Tigray	49.95	11.26	38.79
Afar	58.15	10.19	31.66
Amhara	60.01	6.81	33.18
Oromia	59.80	14.23	25.97
Somali	65.07	8.90	26.03
Ben-Gumuz	63.13	6.76	30.11
SNNPR	51.28	12.89	35.83
Harari	58.01	9.82	32.18
Addis Ababa	70.21	6.08	23.71
Dire Dawa	63.86	10.47	25.67
NATION	58.46	10.76	30.78

According to Table 59, 70.21% of the respondents from Addis Ababa agreed with the idea that drinking river water causes water borne diseases while the figure was 49.95% from Tigray. Others spread between the two extremes.

Table 60. Visiting historical places is wasting both money and time.

Region	Agree	Neutral	Disagree
Tigray	41.02	13.85	45.13
Afar	45.58	8.83	45.58
Amhara	42.65	7.79	49.55
Oromia	47.39	16.10	36.50
Somali	44.26	8.49	47.25
Ben-Gumuz	45.60	8.67	45.73
SNNPR	41.12	13.39	45.50
Harari	43.47	11.25	45.29
Addis Ababa	28.57	7.56	63.87
Dire Dawa	44.69	13.08	42.23
NATION	42.73	11.98	45.29

Table 60 shows that 63.87% of respondents from Addis Ababa and 36.50% from Oromia said that visiting historical places is not a waste of time and money. The other regions fall differently between the two.

Table 61. Gardening gives me pleasure.

	Agree	Neutral	Disagree
Tigray	77.55	9.76	12.69
Afar	90.75	4.08	5.17
Amhara	92.54	2.60	4.85
Oromia	80.27	7.88	11.85
Somali	83.93	4.59	11.48
Ben-Gumuz	84.71	6.78	8.51
SNNPR	84.36	7.48	8.16
Harari	84.85	5.61	9.55
Addis Ababa	94.22	2.48	3.30
Dire Dawa	81.75	7.20	11.05
NATION	84.97	6.25	8.78

Table 61 shows that the majority of students in all regions get pleasure from gardening. Significantly high in descending order were Addis Ababa (94.22%), Amhara (92.54%) and Afar (90.75%). Tigray, in this regard, was the lowest (77.55%).

Table 62. One should plant trees to prevent erosion.

Region	Agree	Neutral	Disagree
Tigray	62.98	8.69	28.33
Afar	87.15	4.23	8.62
Amhara	90.37	2.95	6.67
Oromia	75.23	8.92	15.85
Somali	79.93	5.22	14.85
Ben-Gumuz	83.31	5.07	11.62
SNNPR	81.78	7.59	10.63
Harari	81.21	6.21	12.58
Addis Ababa	88.62	3.00	8.38
Dire Dawa	79.43	6.86	13.71
NATION	80.71	6.46	12.83

Table 62 shows that majority of the respondents believe that planting trees could prevent erosion. 90.37% those from Amhara agreed with this statement while only 62.98% from Tigray did the same. The percentage for respondents from other regions was between these two points.

Table 63 below indicates the attitude of the sampled students towards environmental protection. 91.19% of the respondents from Addis Ababa agreed that protecting the environment from pollution is the duty of everybody, while pertinent figure for Oromia was only 64.85%. For the rest of the regions the percentage lay between the two boundaries.

Table 63. Keeping the environment clean is everyone's duty.

Region	Agree	Neutral	Disagree
Tigray	74.81	8.65	16.54
Afar	85.17	4.57	10.25
Amhara	86.28	3.03	10.69
Oromia	64.85	13.20	21.95
Somali	84.62	3.76	11.62
Ben-Gumuz	81.65	4.26	14.10
SNNPR	78.34	8.63	13.03
Harari	81.97	6.21	11.82
Addis Ababa	91.19	2.80	6.01
Dire Dawa	72.47	11.30	16.23
NATION	77.78	7.81	14.41

Table 64. It is possible to prevent the consequences of drought.

Region	Agree	Neutral	Disagree
Tigray	58.36	12.94	28.69
Afar	68.76	7.54	23.70
Amhara	71.68	6.90	21.42
Oromia	58.85	17.17	23.98
Somali	71.62	8.25	20.13
Ben-Gumuz	68.21	9.93	21.85
SNNPR	61.01	12.00	26.98
Harari	75.04	9.53	15.43
Addis Ababa	81.11	7.43	11.46
Dire Dawa	69.51	13.82	16.67
NATION	65.82	11.66	22.52

Table 64 shows that the majority of respondents agree that it is possible to prevent the consequences of drought. The percentage of students who agree, however, differs from region to region. For instance, 81.11% of the respondents from Addis Ababa agreed with the statement while only 58.36% agreed from Tigray. Others lay between these two points.

Table 65. It is necessary to burn forests to drive away wild animals that destroy crops.

Region	Agree	Neutral	Disagree
Tigray	30.36	11.48	58.15
Afar	31.71	7.22	61.07
Amhara	27.20	5.64	67.16
Oromia	26.36	14.76	58.89
Somali	34.60	8.11	57.28
Ben-Gumuz	32.62	7.59	59.79
SNNPR	35.07	10.75	54.18
Harari	29.51	7.95	62.54
Addis Ababa	19.86	5.35	74.79
Dire Dawa	28.33	12.29	59.38
NATION	29.59	10.05	60.36

Table 65 shows the percentage of respondents who agreed, disagreed remained neutral in each region with regard to the statement. 74.79% of the respondents from Addis Ababa disagreed while the percentage for SNNPR was only 54.18%. The percentage of those who disagreed with the statement in other regions is between these two.

Table 66. Women are equal to men in knowledge and thinking.

Region	Agree	Neutral	Disagree
Tigray	48.26	11.41	40.33
Afar	55.05	8.68	36.28
Amhara	68.58	5.15	26.27
Oromia	45.86	18.35	35.79
Somali	70.02	5.77	24.22
Ben-Gumuz	51.33	5.59	43.09
SNNPR	58.51	10.67	30.82
Harari	61.52	10.76	27.73
Addis Ababa	69.88	6.97	23.16
Dire Dawa	54.73	14.07	31.20
NATION	57.03	11.00	31.97

As seen in Table 66, majority of the respondents agreed that women are equal to men. The percentage of students who agreed or disagreed, however, differs from region to region. For instance, 70.02% of the respondents in Somali region agreed with the statement, while only 45.86% agreed in Oromia. The percentage of students who agreed in other regions spreads between these two points.

Table 67. It is appropriate to return property found lost either to its owner or to the school administration.

Region	Agree	Neutral	Disagree
Tigray	72.60	10.36	17.04
Afar	87.44	4.08	8.48
Amhara	88.46	3.47	8.07
Oromia	67.03	12.64	20.34
Somali	89.02	3.33	7.65
Ben-Gumuz	83.95	4.08	11.97
SNNPR	80.56	7.68	11.76
Harari	81.39	6.20	12.41
Addis Ababa	90.57	3.28	6.15
Dire Dawa	77.17	7.65	15.18
NATION	79.55	7.41	13.03

Table 67 shows that respondents from Addis Ababa stood at the top in terms of agreement (90.57%) agreed to the statement above. The percentage of those who expressed agreement from Oromia region was lower (67.03%) than that of all the other sampled students.

Table 68 (below) shows that the majority of students in all regions believe that group work is important for academic success. 90.08% of respondents from the Somali region agreed to the statement while only 72.41% did from Dire Dawa. The percentage of respondents in other regions who agreed to the statement spread between the two limits.

Table 68. I believe that group work is important for academic success.

Region	Agree	Neutral	Disagree
Tigray	77.88	7.26	14.85
Afar	82.78	6.16	11.06
Amhara	87.33	3.61	9.06
Oromia	72.93	11.08	16.00
Somali	90.08	4.63	5.29
Ben-Gumuz	81.18	5.13	13.68
SNNPR	78.41	8.31	13.28
Harari	83.82	5.65	10.53
Addis Ababa	84.79	4.52	10.69
Dire Dawa	72.41	10.09	17.50
NATION	79.66	7.44	12.90

Table 69. Gathering information about our surrounding is necessary to have knowledge about it.

Region	Agree	Neutral	Disagree
Tigray	77.52	9.85	12.63
Afar	81.42	6.46	12.13
Amhara	84.08	4.93	10.99
Oromia	66.36	14.22	19.42
Somali	86.14	5.61	8.25
Ben-Gumuz	79.00	6.30	14.70
SNNPR	77.69	8.84	13.48
Harari	79.20	5.81	14.98
Addis Ababa	88.33	4.44	7.23
Dire Dawa	71.47	10.15	18.38
NATION	77.17	8.76	14.07

In Table 69 we see that the percentage of respondents who agreed that gathering information is important for acquiring adequate knowledge about their surrounding lay between 88.33% (Addis Ababa) and 66.36% (Oromia). Even at the national level, it could be seen that the majority of students agreed to this idea (77.17%).

Table 70. Solving differences through dialogue is good culture.

Region	Agree	Neutral	Disagree
Tigray	75.84	8.85	15.32
Afar	84.91	4.87	10.22
Amhara	87.46	3.09	9.45
Oromia	70.54	11.72	17.74
Somali	87.19	4.93	7.88
Ben-Gumuz	82.12	4.64	13.25
SNNPR	80.08	7.55	12.37
Harari	81.89	6.09	12.02
Addis Ababa	86.96	4.24	8.80
Dire Dawa	73.19	10.10	16.71
NATION	79.52	7.38	13.10

As can be seen from Table 70, the percentage of respondents who agreed to the idea in the statement came between 87.40% for Amhara region and 70.54% for Oromia. The figure for the national level is 79.52% which means the majority agreed to the idea.

Table 71. Boys and girls should equally participate in household chores after school.

Region	Agree	Neutral	Disagree
Tigray	71.54	10.10	18.37
Afar	85.96	4.37	9.67
Amhara	86.40	2.66	10.94
Oromia	72.15	10.72	17.13
Somali	74.51	6.09	19.41
Ben-Gumuz	81.53	5.54	12.93
SNNPR	80.29	6.45	13.27
Harari	80.18	7.87	11.95
Addis Ababa	89.72	3.29	6.99
Dire Dawa	76.12	10.73	13.15
NATION	79.19	7.08	13.73

Table 71 shows that the majority of students in all regions agreed to the statement that boys and girls should participate in household activities equally. Highest agreement was secured in Addis Ababa (89.72%) while the lowest was in Tigray (71.54). The rest spread between these extremes.

Table 72. Early marriage is a practice that should be abolished.

Region	Agree	Neutral	Disagree
Tigray	45.75	10.87	43.38
Afar	55.13	4.74	40.13
Amhara	63.87	4.40	31.73
Oromia	48.73	15.06	36.21
Somali	55.39	7.63	36.98
Ben-Gumuz	60.40	6.36	33.25
SNNPR	58.27	9.23	32.50
Harari	55.91	7.27	36.82
Addis Ababa	69.16	4.92	25.92
Dire Dawa	55.90	11.41	32.69
NATION	56.38	9.25	34.38

It can be seen from Table 72 that although the majority of respondents in many regions agree that early marriage is something that should be abolished, substantial percentage of respondents in other regions disagree to its abolition. It is only 45.74% of the respondents in Tigray and 48.73% in Oromia that agreed to this statement. 69.16% of respondents from Addis Ababa who agreed to the idea make comparatively the highest group.

Table 73. Female circumcision is a harmful traditional practice that should be abolished

Region	Agree	Neutral	Disagree
Tigray	56.09	10.22	33.70
Afar	62.30	6.78	30.91
Amhara	74.93	3.39	21.69
Oromia	51.17	17.48	31.35
Somali	74.88	7.88	17.24
Ben-Gumuz	67.77	5.15	27.08
SNNPR	60.70	8.81	30.49
Harari	63.53	8.97	27.51
Addis Ababa	76.72	4.82	18.46
Dire Dawa	58.61	11.31	30.08
NATION	62.55	9.61	27.84

Table 73 shows the majority of students in all regions believe that female circumcision is a harmful traditional practice. Large percentages of respondents from Addis Ababa (76.72%) agreed to the statement making the highest in the sample. The smallest percentage of respondents who agreed to this statement, in comparison with other regions, was seen in the Oromia region (51.17%).

Table 74. Absenteeism and coming late to school are signs of laziness.

Region	Agree	Neutral	Disagree
Tigray	65.38	8.87	25.75
Afar	83.20	3.45	13.34
Amhara	82.98	3.50	13.52
Oromia	59.08	14.67	26.25
Somali	77.72	3.63	18.65
Ben-Gumuz	76.19	5.16	18.65
SNNPR	67.54	8.81	23.65
Harari	69.53	5.58	24.89
Addis Ababa	87.90	2.15	9.95
Dire Dawa	68.43	9.66	21.91
NATION	71.04	8.05	20.91

We understand from Table 74 that Addis Ababa agreed the most (87.90%) and Oromia agreed the least (59.08). Even at the national level, it could be seen that the majority of students agreed to the suggestion that laziness of a student is manifested by being absent from or coming regularly late to school (71.04%).

Table 75. Student participation in teachers' evaluation is appropriate.

Region	Agree	Neutral	Disagree
Tigray	73.02	10.17	16.81
Afar	83.33	4.76	11.90
Amhara	87.90	2.73	9.37
Oromia	74.44	9.42	16.15
Somali	85.09	6.03	8.88
Ben-Gumuz	81.73	3.47	14.80
SNNPR	76.73	7.05	16.22
Harari	80.64	6.40	12.96
Addis Ababa	87.32	3.71	8.97
Dire Dawa	74.68	10.98	14.34
NATION	79.41	6.80	13.80

Table 75 shows the majority of the students in all regions believe that they should participate in the evaluation of teachers. 87.90% of respondents from the Amhara Region and 73.02% from Tigray agreed to the statement making the highest and the lowest respectively. The percentage for respondents from the other regions was between these two lines.

Table 76. Going to school and staying at home are the same to me.

Region	Agree	Neutral	Disagree
Tigray	26.20	11.98	61.82
Afar	28.64	6.01	65.35
Amhara	27.74	6.25	66.01
Oromia	29.78	14.89	55.34
Somali	31.51	5.80	62.69
Ben-Gumuz	30.69	7.28	62.04
SNNPR	28.95	11.86	59.20
Harari	39.42	11.57	49.01
Addis Ababa	27.09	5.89	67.01
Dire Dawa	35.10	12.00	52.90
NATION	29.77	10.45	59.78

In Table 76 we see that the percentage of the national average for respondents who disagreed with the idea that spending time in school and at home are the same was 59.78%. The highest percentage of disagreement was 67.01% (Addis Ababa) and the lowest was 49.01% (Harari). The rest of the regions come between the two.

Table 77 below shows that the majority of students in all the regions believe that learning enables people to attain higher status. 94.64% of respondents from Addis Ababa, 92.36% from Amhara, and 91.97% from Afar have agreed to the statement. The lowest percentage of respondents who agreed to this statement, in comparison with other regions, was from Oromia (76.25%).

Table 77. Learning is a means to earn a higher status.

Region	Agree	Neutral	Disagree
Tigray	82.53	7.35	10.12
Afar	91.97	3.15	4.88
Amhara	92.36	2.25	5.39
Oromia	76.25	8.50	15.26
Somali	88.63	2.64	8.73
Ben-Gumuz	86.02	3.83	10.16
SNNPR	87.10	5.43	7.47
Harari	80.45	6.02	13.53
Addis Ababa	94.65	1.23	4.12
Dire Dawa	76.70	8.19	15.11
NATION	84.85	5.37	9.79

Table 78. Because I am a student, I am more knowledgeable and mature than those not in school.

Region	Agree	Neutral	Disagree
Tigray	80.17	8.42	11.41
Afar	88.21	4.56	7.23
Amhara	91.85	2.97	5.18
Oromia	69.77	10.54	19.70
Somali	82.83	4.67	12.50
Ben-Gumuz	84.66	4.50	10.85
SNNPR	82.85	6.20	10.94
Harari	80.66	7.25	12.08
Addis Ababa	90.73	3.81	5.46
Dire Dawa	75.06	8.87	16.07
NATION	81.28	6.70	12.03

Table 78 shows the majority of the students in all the regions believe that they are better than their peers who do not go to school. Significant percentages of respondents of Amhara (91.85%) and Addis Ababa (90.73%) agreed to the statement. Comparatively, the least agreement was observed with respondents from Oromia (60.77%).

4.5 Relationships between Grade 4 Pupils' Achievement and Factors

4.5.1 Correlations with Grade 4 Composite Score

Table 79 shows correlates of factors that affect achievement at different levels. The variables are organized into five blocks: School Structure and Supplies, Teacher Behaviour, School Management, Student Behaviour and Instruction / Support. Those factors which show statistically significant relationships ($p < .05$) with the composite score are highlighted.

Table 79 Spearman Rho Correlation with Grade 4 Composite Score

Group of Variables	Correlation with G4 Composite	Sig.
Block 1: School Supplies		
Availability of Mother Tongue Syllabus	0.065	0.22
Availability of the English Syllabus	0.046	0.38
Availability of Mathematics Syllabus	0.125	0.02*
Availability of Environmental Science Syllabus	0.155	0.003**
Pupil-Text Ratio in Mother Tongue Learning	-0.038	0.25
Pupil-Text Ratio in English Learning	-0.020	0.70
Pupil-Text Ratio in Environmental Science Learning	-0.060	0.25
Average of Syllabi Availability	0.122	0.02*
Availability of Mother Tongue Teacher's Guide	0.107	0.04*
Availability of English Teacher's Guide	0.096	0.07
Availability of Mathematics Teacher's Guide	0.064	0.22
Availability of Environmental Science Teacher's Guide	0.087	0.10
Total Teacher's Guide	0.128	0.01*
Teacher's Perception of Student's Learning Attitudes	0.206	<.001***
Block 2: Teachers Quality and Behavior		
Teacher Workload	0.017	0.74
Teacher Training on new teaching techniques	0.127	0.02*
Teacher Training on new assessment techniques	0.129	0.02*
Teacher Training average	0.106	0.05*
Teacher's Age	-0.039	0.45
Teacher Qualifications	-0.049	0.35
Total teaching experience	-0.059	0.26
Teacher Distance to school	-0.081	0.12
Other social commitments	-0.015	0.78
Block 3: School Management		
Students' Reporting Teachers taking Attendance	0.392	<.001***
Director Qualifications	0.069	0.20
Social obligations outside of job	-0.020	0.68
Director experience	0.104	0.05*
How often are teachers supervised?	0.104	0.05*
Block 4: Students Behavior		
Interest in subjects	0.298	<.001***
Meals a day	0.118	0.02*
Agreement between home and instructional language	0.160	0.002**
Students' sense of ownership in school property	0.204	<.001***
Time spent listening to the radio	-0.171	0.001**
Time spent watching TV	-0.134	0.01*
Distance from Home to school	-0.082	<.001**
Absences from school	-0.223	<.001***

Block 5: Instruction		
Director's perception of English instruction quality	0.101	0.05*
Director's perception of Mathematics teacher's instruction quality	0.108	0.04*
Director's perception of Science instructional quality	0.094	0.08
Director's perception of M/tongue instruction quality	0.121	0.03*
Director's perception of all instruction quality	0.140	0.007**

As to the availability and conditions of curriculum materials in schools under the “*school supplies block*”, there are 12 items considered important in bringing impact on the overall level of student achievement. Among these, Mathematics Syllabus, Environmental Science Syllabus, average of all syllabi and Teacher's Guide for Mother Tongue have significant simple correlation with composite score. These curriculum materials are associated with higher achievement.

With regard to *teachers' behavior*, the variables having positive correlation with composite score are teacher training on new teaching techniques, training on assessment techniques, the use of teacher's guide, and teacher's perception or expectation about the learning of their students.

On the other hand, some characteristics of teachers behavior such as “teacher age”, “teacher qualifications”, “total teaching experience”, “distance teachers travel from home to school” and “other social commitments of teachers” have negative but small relationships with levels of achievement as shown in Table 4.4.1

Concerning the *school organization and management block*, the degree to which head teachers focus on school matters, taking students' attendance frequently, directors' experience, and supervising teachers are considered as elements of the overall quality of management having positive relationships with achievement. Whereas, some factors like “directors' qualifications and social obligations outside of school” have weak or negative relationships with achievement.

Factors grouped under students' background result in high association with achievement. Some of these factors such as “students' interest in school subjects, the amount of meals students get in a day, similarity between home and instructional language, and students' sense of ownership in school property have strong positive correlation with achievement, while other factors like “time spent listening to the radio and watching TV programs, distance from home to school, and being absent from school” are associated with low level of performance having strong negative correlations with achievement.

4.5.2. Results of Multiple Regression Analysis

This section deals with the exploration of the overall influence of the factors that showed strong and significant correlation with the composite score by using multiple regression analysis at the school level. The variables were organized into five blocks. Initially each block was analyzed separately using the standardized method (ENTER) of multiple regression analysis. The main purpose of this was to examine the extent to which each block and each variable in each block influences or explains the variability of student achievement. This means not all variables examined in the correlation analysis strongly and significantly determine students overall achievements. Table 80 summarizes the result.

Table 80. School-level multiple regression results on the Grade 4 composite achievement score details on reduced models

Predictor models and Variables	Std. coefficient	Sig.	R	R-sq	F	df.
School Structure and curricular material supplies			0.266	0.071	10.215***	2,268
Availability and making use of syllabuses	0.337	0.002				
Availability and making use of teacher's guides	0.385	0.004				
Teachers' Behavior			0.253	0.064	8.147***	3,356
Teachers' perception of student learning	0.532	<.001				
Teachers' traveling long distances from home to school	-0.102	0.054				
Teachers' use of textbooks	0.095	0.022				
School Management			0.414	0.171	35.992***	2,349
Teachers taking attendance	0.16	<.001				
Directors' overall managerial experience	0.086	0.041				
Students' Home Background and Behavior			0.435	0.189	16.533***	5,355
Average interest in subjects	0.316	0.009				
Students' sense of ownership of school property	0.233	0.007				
Time spent listening to the radio	-0.284	0.008				
Absence from school	-0.345	0.01				
Similarity of home and school language	0.0041	<.001				
Instruction Support			0.162	0.026	8.361**	1,312
All teachers' instructional quality	0.342	0.007				

The variables under school structure and curriculum materials have explained 7.1% of the variation in learners' achievement. There are only two variables under this block, availability and making use of syllabuses and teacher's guides. Both variables showed strong positive contribution to achievement.

The variables under the block teacher's characteristics and behavior accounted for 6.4% of the variation in achievement. Under this block 'teachers' perception of student learning' showed very strong relation with achievement. Teachers' use of textbooks also contributed positively.

On the other hand the variable 'teachers' traveling long distance from home to school' showed negative relation.

The school management block explained 17.1% of the variation observed in achievement score. Taking attendance and the director's managerial experience both contributed positively.

The variables under students' home background and behavior explained 18.9% of the variation observed in achievement. Schools in which students demonstrate a sense of ownership towards school properties and where high level of interest in their learning is developed tend to exhibit higher achievement. The alignment of home and instructional language also contributed positively. On the other hand absenteeism and time spent listening to non-educational radio programs contributed negatively.

Finally the director's perception of instructional quality which is the only variable under the block instruction and support explained 2.6% of the variation observed in achievement.

Table 81 shows the Model Summary which resulted from the Hierarchical Multiple Regression Analysis. In the first column the numbers in bracket show the number of variables each model contains. The maximum total variance explained (shown under the column R Square) by a single model is 27.8 % (Model 5) which contains all the variables listed in Table 80 lumped together.

Table 81 School level multiple regression results on the composite score model summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1 (2)	.157	.025	.018	9.38	.025	3.790	2	300	.024
2 (5)	.287	.083	.067	9.14	.058	6.255	3	297	.000
3 (7)	.465	.216	.198	8.48	.134	25.189	2	295	.000
4 (12)	.525	.275	.245	8.23	.059	4.704	5	290	.000
5 (13)	.527	.278	.246	8.22	.003	1.122	1	289	.290

Table 82 shows the contribution of each block to the final model when entered in different orders. It can be seen from the table that all the blocks have significantly contributed to the explanation of the variation in achievement of learners when they are entered first. Their contribution, however, dwindles and becomes insignificant in the case of blocks one and five when they are entered last. This fact shows that the other three blocks explained most of the variations observed in achievement scores.

Table 82 Contributions of each block to R-square of completely reduced model

Predictor models and Variables	R square contribution if entered first	R square contribution if entered last
1. School Structure and Supplies	.021*	.004 (n.s.)
Availability and making use of syllabuses		
Availability and making use of teacher's guides		
2. Teacher Variables	.068***	.029**
Teachers' perception of student learning		
Teachers' traveling long distances from home to school		
Teachers' use of textbooks		
3. School Management	.164***	.060***
Student perception of the role of teachers taking attendance		
Director overall managerial experience		
4. Student Home Background and Behavior	.172***	.057***
Average interest in subjects		
Students' sense of ownership of school property		
Time spent listening to the radio		
Absence from school		
Similarity of home language and school language		
5. Instruction	.029**	.002 (n. s.)
Director's perception of all teachers' instructional quality		

The variables grouped under *Student Background and Behavior* explained the highest portion (17.2%) of the observed variation in student achievement at school level before other factors were considered and 5.7% after controlling or considering all other factors in the model. This means, of the 28.7% variations explained by the factors, *Student Background and Behavior* ranked highest in both cases of entry order for analyses. This block contained five variables. Among these, interest in the subject matters, students' sense of ownership of school property and alignment of home and instructional language showed positive relationship with achievement while absenteeism and time spent listening to the radio showed negative relationship.

School Organization and Management block explained 16.4% of the observed variation in student achievement before controlling other factors and still a significant 6.0% with other factors controlled. This means, of the 27.8% variations explained by the factors in the current study, *School Organization and Management* ranked as the second highest in both cases of entry order for analyses. This block contained two variables, namely frequency of taking attendance by the teachers and director's managerial experience. Both factors showed positive relationship with the student achievement.

Teachers' Variables explained the observed variation in achievement by 6.8% before other factors were controlled and 2.9% after controlling all other factors. This means of the 27.8% variations explained by the factors in the current study, *Teachers' Variables* ranked as the third highest in both cases of entry order for analyses. This block contained variables which asked about teachers' perception (judgment) of students, the distance they traveled from home to school, and the use of the textbooks. The distance the teacher traveled from home to school affected student achievement negatively while the other two variables were related positively with achievement.

School Supplies and Structures explained the observed variation in achievement by 2.1% before other factors were considered and by 0.4% after controlling all other factors in the model. This block contained two variables, namely availability and use of syllabuses and teacher's guides. Both factors showed significant positive relationships with achievement.

Finally the *Instruction and Support* block, which includes factors concerning the role of the director in supervising instructional activities, accounted for 2.6% of the total variation in student achievement before other factors were controlled and 0.2% after controlling all the other factors.

4.6 Comparative Analysis between EBNLA and ESNLA

4.6.1 General Comparison

One of the major purposes of this study was to make comparative analysis on the findings of EBNLA and ESNLA. In this part, an attempt is made to discuss the major findings obtained by the study. The following tables show the mean achievements of both EBNLA and ESNLA assessment measures followed by a short analysis.

Table 83. EBNLA mean scores at national level

	Reading	English	Mathematics	Envi. Science	Composite
Mean	64.25	40.46	39.31	48.14	47.94
N	10411	10297	10388	10495	10281

Table 84. ESNLA mean scores at national level

	Reading	English	Mathematics	Envi. Science	Composite
Mean	64.49	38.67	39.69	51.74	48.48
N	12886	13028	13143	13005	12507

Table 83 and Table 84 contain the summary report of the national mean scores on the baseline and the second national learning assessments of the country respectively.

The tables show that the sampled pupils performed better in the two basic reading comprehension tests than in the remaining three subjects in both cases. It is also true that their performances in the two environmental science tests were slightly higher as compared to the results of both English and mathematics tests.

Moreover, there is an acute shortage of textbooks since only less than half of the sampled students reported that the textbook-pupil ratio was one to one while more than half of them used either one textbook for many or did not have any. The correlations between textbook-student ratios in English, mathematics, and environmental science and the Grade 4 composite mean score are 0.043, 0.073, and 0.075 respectively. This indicates that the level of textbook distribution can bring about meaningful differences in pupils' performance.

Grade 4 students have also reported that the levels of their interests towards the instruction of these subjects vary considerably. Based on the reported levels of students' interest in their learning (HIGH, AVERAGE and LOW), a correlation matrix was produced and relationships between Grade 4 composite mean score and interest in the environmental science ($r = .213$), English ($r = .227$) and mathematics ($r = .236$) have been determined. This finding depicts that the level of one's interest in learning a subject determines one's performance more in Grade 4 mathematics and English than in environmental science. This might be partly due to the typical nature of the subjects and partly due to the inadequacy of human and material inputs made available to schools. There were also problems of large class size and lack of instructional facilities as reported by the samples.

Therefore, these and other related problems in schools might have contributed to the low performance of pupils, especially, in the English and mathematics subjects. In addition to this, the finding with reference the mathematics for all sampled pupils is similar to those reported by other researchers.

Finally, it is quite imperative to note that in both EBNLA and ESNLA, the obtained average mean scores in Grade 4 English and mathematics were far below the 50% minimum requirement to pass from one level to the next higher level as stipulated by the Education and Training Policy (ETP) of our country. Another point, quite worth noticing, is that ETP from the start claims much emphasis to be given to the instruction of these subjects at all levels of the education system.

Table 85. Learning Progress: Comparison between the two national learning assessments

Subject	Reading		English		Mathematics		Envi. Science		Composite	
	EBNLA	ESNLA	EBNLA	ESNLA	EBNLA	ESNLA	EBNLA	ESNLA	EBNLA	ESNLA
Time										
Means	64.25	64.49	40.46	38.67	39.31	39.69	48.14	51.74	47.94	48.48
Progress (T2-T1)	0.24		-1.77		0.38		3.60		0.54	

Table 85 compares the levels of pupils' achievement obtained in 4 academic subjects on the two NLAs conducted in 1992 E.C. and 1996 E.C. respectively.

The overall composite mean scores of EBNLA (47.94) and ESNLA (48.48) are almost similar with an increment of 0.54 points in the ESNLA, which is negligible. This implies that there is no gain in students' level of achievement between the baseline and the current assessment measures. In other words, it can be concluded that during the last 4 years, school instructions in Ethiopia didn't register progress in the academic performances of Grade 4 students.

When we consider the specific results gained in the individual subjects, the current status of student performance is not satisfactory. For example, taking the two English tests, there is a considerable drop in gain (-1.86%) from the baseline. And there is no significant change obtained in the results of the two readings and mathematics tests from EBNLA. Nevertheless, a 3.6% gain was achieved in the environmental science tests from the baseline assessment.

4.6.2 Specific Comparison

The sampling strategy called for carry over of 10 schools for each region that had participated both in the Ethiopian Baseline National Learning Assessment (EBNLA, 1992 E.C.) and the Ethiopian Second National Learning Assessment (ESNLA, 1996 E.C.). This strategy would provide for a comparison of baseline and current performance levels for the nation overall and, with substantially less power for detecting statistical significance, within regions. A logical alternative for evaluating change in performance from baseline would be to conduct statistical tests that are designed to detect mean differences from independent samples. In the latter case, the EBNLA sample and the ESNLA sample would be considered independent samples of schools taken at different years. This alternative to the paired comparisons analysis was, however, problematic due to the difference in the sampling strategies that were applied for the EBNLA and the ESNLA. Therefore, it was decided to include a set of schools in the current sample that had participated in the EBNLA and to conduct a repeated measures analysis to investigate change from baseline.

Though the paired comparisons analysis is, in theory, a more powerful test and can often yield very meaningful results because the subject (i.e., the school in this situation) serves as its own control, the limited number of matched schools in combination with the large variances in difference of scores from the Baseline to the ESNLA reduced the power for detecting statistically significant differences. In addition, since five statistical tests were conducted, statistical significance was tested at the 0.01 alpha level in order to compensate

for the inflation of Type I error that can be expected when conducting multiple test statistics on the same sample population.

The statistical results presented below should be considered only in conjunction with careful review and analysis of the descriptive results. It is often the case that when there are limited numbers of schools sampled (e.g., less than the 10 matched schools originally planned for the paired comparisons analysis) the sample population tends to be very skewed and precludes the valid conduct of parametric statistical tests (e.g., the dependent t-test). However, analysis of the normality of the data indicated that the samples, even within regions, were relatively normal. Using the standard error of skewness as a diagnostic statistical tool, it was determined that there was no marked deviation in spread that would justify using analysis alternatives to the paired t-test. This is confirmed in the relatively similar mean and median measures of central tendency given in the tables below.

The compiled descriptive statistics for the difference scores are given below for the nation overall and for each region. The difference score was calculated for each school by subtracting average school (percent correct) score obtained at Baseline from the average school (percent correct) score obtained at the ESNLA. These difference scores were then summarized for the Nation and for each Region as shown below. To support the interpretation of results, schools were grouped into high and low groups and these groupings were then crossed with responses on the six school reform variables from the director's questionnaires.

Progress in the Basic Reading Subject

Results below show that for the 83 matched schools sampled across the nation there was considerable variability in change from baseline. The worst performing school fell 24 percentage points while the highest showed a gain of approximately 20 percentage points. Though results were not statistically significant at the 0.01 alpha level, the performance of Oromia is noteworthy in that on the average, schools in this region fell in their performance by 7 percentage points.

Table 86 Reading: average change in percent correct scores from EBNLA to ESNLA

REGION	Number of Matched Schools	Mean	Standard Deviation	Median	Minimum	Maximum	Test of Mean Differences from Baseline to Second NLA (Paired t-test)	
							T Statistic	P Value
NATION	83	-1.56	9.04	-0.25	-24.78	19.99	-1.58	0.119
TIGRAI	9	1.25	7.20	4.25	-11.00	13.25	0.52	0.616
AFAR	7	4.02	6.48	6.00	-6.38	13.13	1.64	0.152
AMHARA	13	-1.67	5.32	-1.75	-10.88	8.17	-1.13	0.279
OROMIA	8	-7.09	6.51	-8.19	-16.38	2.50	-3.08	0.018
BEN-GUMUZ	6	-4.61	8.49	-3.63	-17.34	4.50	-1.33	0.241
SNNPR	17	0.10	11.46	2.91	-18.83	19.99	0.04	0.972
HARARI	6	-0.30	13.38	2.03	-24.78	13.60	-0.05	0.959
ADDIS ABABA	9	-0.27	6.74	2.25	-15.57	6.13	-0.12	0.907
DIRE DAWA	8	-7.55	10.16	-6.50	-20.43	7.00	-2.10	0.074

Progress in ENGLISH

Results below show that for the 83 matched schools sampled across the nation, there was a significant drop in performance from Baseline to the ESNLA.

Table 87. English: average change in percent correct scores from EBNLA to ESNLA

REGION	Number of Matched Schools	Mean	Standard Deviation	Median	Minimum	Maximum	Test of Mean Differences from Baseline to Second NLA (Paired t-test)	
							T Statistic	P Value
NATION	83	-3.35	8.01	-3.26	-25.59	16.23	-3.81	0.000*
TIGRAI	9	-1.49	6.52	-1.63	-11.94	7.93	-0.69	0.511
AFAR	7	-4.58	2.72	-4.69	-8.87	-0.06	-4.45	0.004*
AMHARA	13	-3.84	5.76	-3.88	-13.38	8.46	-2.41	0.033
OROMIA	8	-13.02	8.32	-14.38	-24.25	3.31	-4.43	0.003*
BEN-GUMUZ	6	-4.95	5.98	-3.72	-13.57	2.22	-2.03	0.099
SNNPR	17	-1.40	6.89	0.53	-21.81	7.94	-0.84	0.415
HARARI	6	-0.92	10.20	-0.34	-13.61	10.69	-0.22	0.833
ADDIS ABABA	9	-2.49	12.22	-2.63	-25.59	16.23	-0.61	0.558
DIRE DAWA	8	0.36	6.76	0.65	-7.97	9.04	0.15	0.884

*Significant negative change from baseline (p<0.01)

It can be seen in the above table that though there was considerable variability in the magnitude of the negative change, on the average, performances fell in every region for Grade 4 English. Attention is drawn to the maximum difference scores given in the table which underscores the fact that some schools did improve in English performance. Further research investigating the reasons for performance regression and for progress in English is recommended. Two regions stand out as having significant (p< 0.01) drops in English performance: Afar and Oromia. It can be seen that in these regions even the maximum scores fail to reflect significant gains.

Progress in Mathematics

The results below show that for the 83 matched schools sampled across the nation, there was no significant change in performance from Baseline to the SNLA. Performance change varied considerably with a minimum of drop in performance by approximately 15 percentage points to a 20 percentage points gain. As in English and environmental science, Oromia showed a significant drop (p<0.01) in performance from Baseline. All of the matched schools in Oromia fell in performance on Grade 4 mathematics.

Table 88. Mathematics: average change in percent correct scores from Baseline to SNLA

REGION	Number of Matched Schools	Mean	Standard Deviation	Median	Minimum	Maximum	Test of Mean Differences from Baseline to Second NLA (Paired t-test)	
							T Statistic	P Value
Total	83	-1.57	7.72	-1.84	-14.77	20.86	-1.85	0.068
TIGRAI	9	-1.73	6.54	-1.84	-13.98	6.60	-0.79	0.450
AFAR	7	-3.30	5.39	-3.43	-12.31	4.96	-1.62	0.156
AMHARA	13	-3.51	6.52	-5.29	-12.38	9.28	-1.94	0.076
OROMIA	8	-8.29	5.71	-8.77	-14.54	-2.14	-4.10	0.005*
BEN-GUMUZ	6	-1.69	8.35	-3.14	-10.21	9.69	-0.50	0.641
SNNPR	17	-0.29	7.80	0.21	-14.77	15.58	-0.15	0.879
HARARI	6	0.16	10.95	3.74	-14.00	9.98	0.04	0.973
ADDIS ABABA	9	2.78	9.80	0.86	-8.80	20.86	0.85	0.420
DIRE DAWA	8	1.22	5.78	2.39	-7.97	9.16	0.60	0.569

*Significant negative change from Baseline (p<0.01)

Progress in the Environmental Science Subject

Results below show that for the 83 matched schools sampled across the nation, there was no significant change in performance from Baseline to the ESNLA. Performance change varied considerably with a minimum of 14 percentage point drop to a high of approximately 16 percentage point improvement. On region, Oromia, showed a significant drop ($p < 0.01$) in performance from the baseline. All of the matched schools selected from Oromia fell in performance in Grade 4 environmental science.

Table 89. Environmental science: average change in percent correct scores from Baseline to SNLA

REGION	Number of Matched Schools	Mean	Standard Deviation	Median	Minimum	Maximum	Test of Mean Differences from Baseline to Second NLA (Paired t-test)	
							T Statistic	P Value
NATION	82	0.52	6.47	1.83	-14.44	15.95	0.73	0.468
TIGRAI	9	1.96	6.06	3.70	-7.99	11.47	0.97	0.361
AFAR	7	3.59	4.86	3.45	-5.78	9.52	1.96	0.098
AMHARA	13	0.17	5.96	1.79	-9.35	9.81	0.10	0.919
OROMIA	8	-5.55	3.71	-4.51	-11.90	-1.93	-4.23	0.004*
BEN-GUMUZ	6	-0.99	6.38	0.85	-10.72	5.63	-0.38	0.719
SNNPR	17	1.28	7.40	1.86	-13.74	15.95	0.72	0.485
HARARI	6	2.60	7.16	5.37	-8.09	8.55	0.89	0.415
ADDIS ABABA	8	1.91	7.01	3.13	-10.57	12.08	0.77	0.466
DIRE DAWA	8	-0.58	6.49	-0.05	-14.44	7.49	-0.25	0.807

*Significant negative change from baseline $p < 0.01$)

Compiled Performance Composite Score across Regions

The following table presents the descriptive statistics for the average difference score from EBNLA to the ESNLA. Only Oromia showed a significant change, a significant ($p < 0.01$) drop in the compiled achievement change score. With the exception of Oromia some schools in all regions made noteworthy improvements, with maximum gains ranging from 4.3 to 14.22 percentage points.

Table 90. Average difference in percent correct scores across Grade 4 achievement tests

REGION	Number of Matched Schools	Mean	Standard Deviation	Median	Minimum	Maximum	Test of Mean Differences from Baseline to Second NLA (Paired t-test)	
							T Statistic	P Value
NATION	83	-1.35	6.82	-0.05	-15.99	14.22	-2.45	0.017
TIGRAI	9	0.09	5.81	-0.16	-11.23	7.86	0.04	0.966
AFAR	7	0.02	3.37	1.35	-4.12	4.33	0.01	0.999
AMHARA	13	-1.87	5.75	-0.05	-9.01	10.56	-1.17	0.265
OROMIA	8	-8.28	5.18	-9.21	-15.99	-0.05	-4.52	0.003
BEN-GUMUZ	6	-3.06	6.88	-2.93	-12.96	4.30	-1.09	0.326
SNNPR	17	-0.07	6.89	2.72	-12.44	8.55	-0.04	0.969
HARARI	6	0.44	9.87	2.92	-14.11	10.35	0.11	0.918
ADDIS ABABA	9	0.96	8.55	1.39	-15.13	14.22	0.34	0.744
DIRE DAWA	8	-1.73	6.47	-1.13	-12.07	6.63	-0.75	0.474

4.7 Analysis of Qualitative Data

4.7.1 Overview

This study was set to complement the quantitative study, which used statistical tools to determine the extent to which learning takes place in primary schools, using interviews and focus group discussions. It also explored the main explanatory conditions of student learning using basically similar questions used in the quantitative study from different sources.

The qualitative study was carried out in all the regions where quantitative data has been collected. The source of information included teachers, Grade 4 students and parents. Each of these respondents was asked relatively similar questions so that the acquired information could be cross-checked using the triangulation of sources of data. Large numbers of languages were used for interviews and focus group discussions. Findings from Grade 4 students' qualitative study were included in this report because their views are considered valid as they could provide cumulative information acquired in the entire primary education system. The analysis of the collected data has been carried out at national and regional levels. Student learning was gauged by asking respondents what students learn in schools and if the learning was relevant to their needs. Explanatory conditions were conceptualized around such categories as school supplies, teacher behaviors, curricular and system level issues, school administration, parental behavior and student personal characteristics. In the following section the national and regional level of analyses and interpretation of data are presented.

4.7.2 Characteristics of Student Learning

Teachers believe that primary school children gain basic knowledge, skills and attitudes. However, in many regions teachers expressed dissatisfaction with the extent to which students learn. As it can be seen from the summary given in Table 90, teachers used such metaphors as "schooling is better than remaining at home" and "they learn how to read and write" to show the minimal learning taking place in schools. Repeatedly, teachers expressed "the lack of student interest or vision" in learning to show how little is going on in academic achievement. In few cases, it has been expressed that students gain useful knowledge, particularly in environmental understanding and personal hygiene.

Parents perceive that students' learning did not make a significant change. They underline that their children come from school without sufficient and appropriate knowledge. There is still an agreement with teachers in many cases that "students lack interest and motivation to learn". Some even do not acknowledge that learning takes place while others say that "students lack normal development". Marginal or minimal knowledge acquired is also expressed as "those who go to school are better than those who remain at home". In some circumstances, there are feelings showing that the "time spent on learning and learning gains do not commensurate". In a very particular case, parents acknowledged that students learn useful knowledge and they "developed communication skills". The emphasis on the development of communication skills is interesting, because this particular case is in one of those regions where the medium of instruction is the mother tongue. Parents expressed concerns about "pooling factors" that contribute to low level learning in schools.

In most cases, students believe that they have satisfaction with their learning. They emphasized that they get sufficient knowledge and skills in environmental issues, hygiene, civics and social skills. Few students perceived "lack of interest of students in learning".

A comparison of the results of qualitative study with quantitative study indicates that the findings are complementary to each other. A composite scores analysis of test results for different subjects and regions at the national level was 48.8% for Grade 4. What this finding says is that learning indeed takes place in the Ethiopian Grade 4 Primary Education, but it is

not satisfactory. This is so because the composite scores of students is less than the expected average for the nation. In the table below, one of the consistent positive evaluations was expressed in Harari. If one looks into the Harari quantitative achievement results, it becomes vivid that the average achievement score is the third highest. This implies that findings are consistent with each other although the methods used for data collection and analyses are different.

Teachers, students' and parents views are almost similar in almost all regions except for Harari. This case requires a special attention and continued research as to why the region continues to perform more than others. Table 91 summarizes the findings of this part of the study below.

Table 91. Student learning as viewed by teachers, students and parents by region.

Regions	Respondent views on what students learn and its relevance		
	Teachers	Students	Parents
Oromia	Basic valuable skills and attitudes. Some students have no interest to learn. Schooling is better than non-schooling	Basic knowledge about environment. Have interest to learn. Schooling is better than non-schooling	Basic knowledge about environment Learning is unsatisfactory and weak. Schooling is better than non-schooling
Amhara	Lack of satisfaction and confidence in students to learn. Only the ability to read and do some writing	Acquire important knowledge.	Failure and no adequate learning Better than those who remain at home.
Tigray	Students do not show the required competence. Lack of vision.	Lack of motivation to learn	Lack of motivation to learn Differentiate between good and bad, Better than those who remain at home.
SNNPR	No satisfaction with student learning	Satisfied with their learning Basic knowledge about the environment, and hygiene Understanding of rights and responsibilities, social skills	Do not know that their children learn.
Addis Ababa	Low interest to learn Good moral make-up, but spoiled by environment	Low interest to learn	Sufficient learning does not take place Student lack normal development
Dire Dawa	Interest to learn is high Insufficient/inadequate knowledge is gained by students	Interest to learn is high	Children acquire inadequate/insufficient knowledge
Somali	Student understanding of their environment is better. Keep personal hygiene	Learned reading and writing	Better than those who remain at home.
Afar	Basic learning takes place, but not satisfactory. Low or no interest to learn. Students have no clear visions.	Low interest to learn	Low interest to learn
Harari	Grasped useful knowledge	Grasped useful knowledge Believe in their changed social attitudes	Grasped useful knowledge Developed communication skills
Ben-Gumuz	Basic knowledge, skills and attitudes Students do not have interest to learn No satisfaction with what they learn	Basic knowledge, skills and attitudes	Basic knowledge, skills and attitudes Students do not have interest to learn No satisfaction with what they learn Time spent and learning gains do not commensurate.

4.7.3 Conditions Affecting Student Learning

The analyses of explanatory conditions as expressed by respondents were conceptualized into issues at school and outside school levels. School level issues consistently revolved around such factors as school facilities and supplies, teacher behaviors, availability and competence. Issues related to the curriculum, system characteristics, parents and students were considered as outside school factors that influence student performance. The analysis of these discussions are presented in the following sections both using narrations and tables.

1. School Supplies and School Facilities

Table 91 below summarizes that teachers, students and parents have the opinion that lack of school supplies and facilities impede student learning. In almost all regions, shortage of textbooks, lack of reading rooms, lack of classrooms, lack of laboratories, insufficient utilization of the available support services, and malfunctioning of the available resources were some of the main issues raised in the discussions.

In addition to shortages of supplies and facilities, the main concern of teachers was the poor utilization of the available support services including tutorials and student counseling. Reasons for this were related to student involvement in business and other activities for self sufficiency. This in turn led to the shortage of time to participate in extra-school activities required for effective learning.

Table 91 also shows that students and parents in all regions shared almost all of the concerns of schools teachers. If one goes back to the quantitative study, it is obvious that school supplies have a great influence in student achievement. The regression analysis indicated that school supplies explain about 7% of student achievement. The correlation analysis also showed that some school supplies like the availability of the syllabi and teacher's guides have significant positive relations with student achievement. These findings lead to the fact that the more schools are supplied with the necessary facilities the better will be student achievement.

The concern with the shortage of textbooks is a nationwide phenomenon. Given the dependence of the school system on textbooks the issue has to be taken seriously. What the quantitative study tells in this respect is even more alarming. The correlation analysis indicates that the pupil-text book ratio negatively correlated with student achievement. This means the preparation and distribution of textbooks of the current type makes negative difference in learning rather than the positive development expected in society. In this sense, what the finding tells is not the insignificance of textbooks in student learning. Rather, it notes that if the quality of textbooks that students use is not improved and standardized, it causes student failure to learn and achieve.

Table 92. School supplies and school facilities.

Regions	Respondents' views on the influence of school supplies on student learning		
	Teachers	Students	Parents
Oromia	Shortage of text books Lack of library and reading room Lack of laboratories Large number of students in the classroom	Shortage of text books Lack of library and reading room Lack of laboratories	Large number of students in the classroom
Amhara	Inadequate utilization by students of available services (e.g. tutorials) Shortage of textbooks, blackboards Lack of laboratories Shortage of classrooms.	Large number of students in the classroom. Wastage of school tutorial services due to lack of time.	Inadequate utilization by students of available services (e.g. tutorials) Shortage of textbooks, blackboards Lack of laboratories Shortage of classrooms
Tigray	Lack of textbooks Shortage of laboratories and libraries	Lack of textbooks Shortage of laboratories and libraries	Lack of textbooks Shortage of laboratories and libraries
SNNPR	Minimal use by students of materials and tutorials provided by teachers, Lack of textbooks, laboratories and reading rooms, SPCs Shortage of classrooms	Lack of textbooks Shortage of seating places in the classroom	Lack of adequate services like adequate classrooms, text books
Addis Ababa	Shortage of laboratories and libraries	Shortage of laboratories and libraries	Shortage of classrooms Shortage of laboratories and libraries
Dire Dawa	Shortage of textbooks Lack of library, reading room, laboratory	Shortage of textbooks Lack of library, reading room, laboratory	Shortage of textbooks Lack of library, reading room, laboratory
Somali	Shortage of basic learning materials from parents Inadequate use of tutorial services for students are engaged in business Shortage of libraries and laboratories	Shortage of basic learning materials from parents Inadequate use of tutorial services for students are engaged in business	Shortage of basic learning materials from parents Inadequate use of tutorial services for students are engaged in business Shortage of libraries and laboratories
Afar	Shortage of textbooks and lack of teaching materials	Shortage of laboratories and libraries	Shortage of laboratories and libraries
Harari	Lack of reading rooms Malfunctioning of available facilities Lack of laboratories Instructional services were not properly used for students spend time doing business	Lack of reading rooms and malfunctioning of those that are available Lack of laboratories	Available school services are not properly used by students
Ben-Gumuz	Shortage of textbooks Unsatisfactory use of tutorial and counselling services	Lack of textbooks for arts and music education	Lack of teaching materials including textbooks

2. Teacher Behaviour, Availability and Competence

Teachers, parents and students showed the influence of the teacher in learning achievement in many ways. Teachers in almost all regions feel that teachers lack the required qualification and competence to deal with the curriculum they are supposed to teach. Except in Amhara, Somali and Harari regions where this issue was not directly reflected, teachers in all other regions pointed out that there is a gap between their current professional capability and the demands of the curriculum. Late assignment of teachers in some regions has been identified to cause a problem for student learning.

Table 93. Teacher behavior, availability and competence

Regions	Respondents' views on the influence of teacher behaviour, availability and competence on student learning		
	Teachers	Students	Parents
Oromia	Shortage of qualified teachers Lack of teachers' competence to teach subjects Teachers give tutorial services, but student use is inadequate	Teachers lack adequate preparation to teach in the classroom	Failure to cover yearly portions
Amhara	Teacher support exists, but used inadequately by students Loose relationship between teachers and parents	Teachers do not cover all portions in the syllabi	Lack of qualified teachers Loose relationships between the schools and community Insufficient imparting of knowledge
Tigray	Shortage of qualified teachers Late assignment of teachers in schools	Teachers do not care to help Teachers do not give feedback for their learning	Shortage of qualified teachers Late assignment of teachers in schools
SNNPR	Teacher support is available, but not used adequately. Lack of qualified teachers Assignment of incompetent teachers	Failure to cover the syllabi Inadequate use of teacher support	Instruction is impeded by students participation in business
Addis Ababa	Teachers do not cover large number of subjects and portions Teachers lack competence to teach most subjects	Teachers are unable to cover subjects and portions	Teachers do not cover large number of subjects and portions Teachers lack competence to teach most subjects
Dire Dawa	Shortage of qualified teachers	Shortage of qualified teachers, particularly in English	Insufficient imparting of knowledge in schools
Somali	Inadequate use teachers' support like tutorial	—	—
Afar	Lack of qualified teachers Teachers' dissatisfaction in their work	Teachers are absent frequently	Teachers are absent frequently
Harari	Insufficient use of tutorials and make up classes	Participation in business does not allow to use teacher efforts	Self-help activities by students limits teachers' efforts
B. Gumuz	Tutorial services Counselling and competition among students Shortage of qualified teachers	Inadequate use of tutorial services Teachers' lack of competence or qualification	Inadequate teaching practices

Teachers also commonly showed that they are not capable of covering all the portions of the curriculum given for the year. They attribute this to large amount of contents prescribed in the curricula. In some cases, job dissatisfaction of teachers has been reported to influence learning. On the positive side, teachers in many regions indicated that they organize student support services, although it has been insignificantly utilized.

Students agree with teachers on many points. They recognize the lack of qualification and competence on the part of teachers. Students are aware of the fact that the contents of the curricula are too much to cover. At the same time, students acknowledge teachers lack of preparation, absenteeism and lack of willingness to provide feedback in their learning as impediments to achievement. Students also acknowledge the efforts of teachers in the provision of support services. However, it has been underutilized due to student engagement in other activities. Apart from the above, parents and teachers mention that there is a loose relationship between the school and the community.

A comparison of the findings of the quantitative study with the above findings ascertains that teacher's behavior and competence is a strong factor influencing student achievement. In the multiple regression analysis, it has been found out that teachers' behavior explains 17.1% of the variations in student learning. In the correlation analysis, teacher training on the average has shown a strong significant relation with student achievement. For summary of teacher behavior that influences student learning, refer to Table 92 above.

In summary, the most important teacher qualities that were identified to influence student learning are lack of teachers' qualification and competence, late assignment of teachers to the work place, inability of teachers to cover portions given in the curriculum, absenteeism and lack of preparation for teaching.

3. Student Characteristics

Student behaviors that negatively influence academic achievement have been identified in almost all regions. From teachers' perspectives to absenteeism, less interest and motivation for learning, engagement in unwanted and wasteful environment, lack of discipline, lack of study time, lack of background in kindergarten education, and engagement in business have been mentioned as the most important influences on student learning.

Student absenteeism has been associated with parental labor demands on children. Since students achievement very much relates to the time spent on learning, absenteeism of students can be considered as a national problem demanding attention. Less interest in learning and lack of motivation of students have been related to the presence of undesirable environments around the schools as well as the consideration of employment perspectives. Teachers complain in almost every part of the nation about lack of student discipline. One explanation for such feeling may be related to the new approaches to teaching-learning in which teachers require to adopt new attitudes towards student-centered approach. Given the passive nature of the instruction process in previous circumstances such reflection from teachers is only normal. Capacity building in this area seems vital.

Table 94. Student behaviors influencing academic achievement

Regions	Respondents' views on the influence of student characteristics on student learning		
	Teachers	Students	Parents
Oromia	Less class attendance due to parental labour demands Undisciplined acts of students	Attraction towards TV, radio, movies, etc Low level of parental support	Lack of discipline due to large number of students in the classrooms
Amhara	Some students have less interest to learn Some waste time in unwanted engagements Some students lack discipline. Students lack study time due to labour work	Most students have low material support from home	Some students have less interest to learn Some students waste time in unwanted engagements
Tigray	Most students do not attend KG and that creates problems for later learning Students lack interest in learning Students have confused assumptions	Less time for study due to parental need for labour	Most students receive low material support from home
SNNPR	Engaged in trade for helping themselves Irregular classroom attendance	Engaged in trade for helping themselves.	Most students get low parental support. Do petty trade to help their schooling
Addis Ababa	Less interest to learn and study Student engagement in wasteful environment	Less interest to learn and study	Students get low parental support and follow-up
Dire Dawa	Students have interest to learn Less classroom attendance because of business	Less classroom attendance because of business	Students have interest to learn
Somali	Lack of study time due to engagement in trade	Lack of study time due to engagement in trade	Lack of study time due to engagement in trade
Afar	Less interest to learn Engagement in chewing chat Lack of discipline from students	Lack of study time due to labour demands at home	Low level of parental support
Harari	Students are engaged in trade to help themselves Students are discouraged by jobless school leavers	Students are engaged in trade to help themselves Students are discouraged by jobless school leavers	Students are engaged in trade to help themselves. Students are discouraged by jobless school leavers.
B. Gumuz	Students have less interest to learn Lack of discipline among students	Lack of confidence in education	Students have less interest to learn Arrogance of school children on the rise

From teachers explanations, it is also understood that students lack competence for primary schooling due to the fact that the curriculum requires attending pre-primary education. It is equally and significantly mentioned by teachers that students lack study time due to their engagement in business and other activities. The findings show that student achievement is not only conditioned by schools factors, but also by matters outside it as well. If students do business to help themselves, it is obvious that their living standards need to be improved to enable them concentrate on their studies. Since this requires a long-term strategy, however, schools may need to consider these situations of learners by arranging flexible programs to minimize lost learning.

Students and parents agree with teachers concerning student behaviors influencing their academic performance. A different expression from the students shows that there is lack of confidence on the part of students to learn. This can be explained in different ways, but as the next section shows the difficulty with the curriculum and lack of vision of the significance of learning may play their roles for such behaviors.

Like the qualitative data, the quantitative results summarized in the previous section indicated that most of the variations in student academic achievement in Ethiopia are explained by student behavior. In the correlation analysis, significant relations have been observed between student behavior and student achievement more than any other factor. In the regression analysis it has been demonstrated that student behavior explains 18.9% of the variation observed in the current study. This is by far the strongest factor explaining student learning in this country.

4. School curricula

Issues related to school curricula and the school system have strongly surfaced during the interviews and focus group discussions. As Table 95 below shows, teachers, students and parents were unanimous in pointing out some of the features of the curriculum and the system. These include the difficulty of the curriculum both to the students and teachers, difficulty to cover portions due to the mismatch between contents and time, the impracticality of self-contained classes, the bare advantage of continuous assessment and the discouraging influence of the free promotion of students from one Grade to the other.

Teachers indicate that the curriculum is difficult both for themselves and the students. This has been already indicated in the previous section where teachers' belief indicated that they lack the necessary qualifications and competence in relation to the new curricula. It is alarming that the curriculum is seen as a mismatch between time and content. Even more so is the perceived difficulty of the subjects for students. Unless students understand the subject they are supposed to learn, it threatens the whole efforts of schooling.

Table 95 The influence of the curriculum on students achievement.

Regions	Respondents' views on the influence of the curricula on student learning		
	Teachers	Students	Parents
Oromia	Curriculum is difficult both to the students and teachers Difficult to cover portions due to the mismatch between contents and time Self-contained classes are impractical Continuous assessment is a useful practice Free promotion increases failures repetition.	Curriculum is relevant. Learning in the mother tongue is relevant. Self-contained classes are impractical for a single teacher. Continuous assessment is a useful practice. Free promotion discourages hard work.	Curriculum does not prepare children for work. Curriculum misses some social values of society. Moral education is lacking. Learning in the mother tongue is relevant. Free promotion discourages hard work.
Amhara	The curriculum is difficult particularly for rural children. Curriculum demands unavailable facilities and conditions. Curriculum lacks cohesion. Teachers can not implement self-contained classes Free promotion discourages students Problems of translation and interpretation of books.	Curriculum demands unavailable facilities and conditions. Self-contained classes are impractical. Continuous assessment is important. English at higher levels may be difficult.	The curriculum is difficult particularly for rural children. Free promotion discourages students.
Tigray	Teachers were not involved in policy formulation. Self-contained classes are not supported by appropriate methods and materials.	The curriculum is difficult. Portions are not covered.	Self contained classes are impractical.
SNNPR	Curriculum is relevant, but difficult. Self-contained classes make learning difficult. Free promotion advances incompetence	Curriculum is relevant, but difficult. Free promotion advances incompetence.	Curriculum is relevant, but difficult. Self contained classes make learning difficult. Free promotion advances incompetence.
Addis Ababa	Curriculum is difficult to students. Self contained classes are inappropriate. Free promotion is useless.	Curriculum is difficult for teachers and students. Free promotion is useless.	Curriculum is irrelevant. Self contained classrooms are inappropriate.
Dire Dawa	Subjects are difficult. Self-contained classes are impractical Free promotion advances incompetence	Subjects are difficult.	Subjects are difficult. Self contained classes cause insufficient learning
Somali	The curriculum is difficult.	Period allotment is inappropriate.	The curriculum is difficult.
Afar	Regional promotion policy at lower Grades is discouraging.	Subjects are difficult	Self contained classes are not useful
Harari	Learning in the mother tongue is good. Subjects are difficult. Self-contained classes are better than non-self contained. Continuous assessment is good. Free promotion hurts learning.	Learning in the mother tongue is good. Subjects are difficult. Curriculum is relevant. Continuous assessment is good.	Learning in the mother tongue is good. Subjects are difficult. Continuous assessment is good.
Ben-Gumuz	Subjects are difficult. Curriculum is irrelevant to life conditions. Self contained classes are difficult. Free promotion discourages hard work.	Subjects are difficult. Curriculum helps to solve some problems.	Subjects are difficult. Curriculum does not prepare for work. Free promotion discourages hard work.

Self-contained classes were considered impractical for the reason that teachers are not competent to teach all subjects. It is also considered boring both for teachers and students. The use of continuous assessment is taken as a positive side of the curriculum by teachers. However, the practice of free promotion has been rejected almost in regions due to the perception that such policy encourages incompetence and the pushing through of children who do not deserve promotion. Since the policy to use continuous assessment is meant to follow the progress of students from time to time in order to determine student promotion, the automatic promotions in schools seem to indicate failure to use the continuous assessment approach for its intended purpose.

Students and parents reflected similar opinions concerning the curriculum. In this respect, parental special concerns require special attention. In addition to what has been commonly pointed out in the above analysis, they say that the curriculum is irrelevant since it does not prepare children for the world of work. Given the two primary cycles which also have terminal points in their structure, this concern is valid. Parents also say that some social values are not embedded in the curricula. There is a suggestion that moral education be included in the curricula. Parents appreciate that learning takes place in the mother tongue. Some complained, however, that the current curriculum is particularly difficult for rural children. This observation has to be understood in relation to the previous finding which indicated the negative influence of the absence of kindergarten education for children. Since most rural children have no access to these facilities, there is a high possibility that the curricula become difficult for learning.

5. Parental Support

All respondents perceive that parental support of students' learning is very low. Even though parents say that they provide learning materials, teachers pointed out that once parents send their children to school, they feel they are relieved and never go to school to follow the progress of their children. In almost all regions, except Addis Ababa, respondents confirm that students are self-supporting, and are engaged either in petty business or home labor. They feel that these have resulted in students' truancy and low achievement. Parents say that they do not let their school children attend tutorial classes when it is out of school time.

The fact that parents provide less academic support is not surprising. This is because parents are neither as educated as their children nor do they understand how schools operate. Follow up of children and the interest to discuss with teachers help find out what goes on between their child and the school or teachers. On the other hand, parents do participate in formal committees, raising funds, and classroom construction or school expansion. This is by itself a positive trend in school-community relations. Table 95 below is the summary of findings regarding parental support to school children nation wide.

Table 96. The influence of parental support on school achievement in Ethiopia.

Regions	Respondents' views on the influence of parental support on student learning		
	Teachers	Students	Parents
Oromia	Inadequacy of academic support and follow up.	Demand for children's labour Support limited to basic learning materials	Low perception of the importance of education Poverty of parents High Participation in construction of rooms and teacher- homes.
Amhara	High Participation in schools affairs especially in fund raising. Academic support is minimal.	No time for studying due to labour demands at home	No support for tutorial time
Tigray	Rare participation in school affairs.	Parents have no participation	Rare participation in student affairs
SNNPR	Low parental academic support High Participation in formal committees	Low parental academic support	Lack of awareness on the importance of education Participation in school expansion.
Addis Ababa	Parents make no follow-up and do not come to school for discussing children's affairs.	—————	Schools do not invite parents for discussing children's affairs.
Dire Dawa	Fund raising whenever requested	Less academic support by parents	Employing teachers Building classes
Somali	Parental participation and support is insignificant.	Parental participation and support is insignificant.	Parental participation and support is insignificant.
Afar	Parents' academic support is minimal. Demand for labour	Parents' academic support is minimal. Demand for labour	Parents' academic support is minimal. Demand for labour Participation in school affairs is insignificant.
Harari	Low parental academic support	Parental academic support is low. Teachers employed at home	Fund raising and construction of rooms and teachers residences
Ben-Gumuz	Minimal parental academic support	Participate in school committees	Participate in schools expansion Construct teacher residences

5. School administration

In most of the discussions, the influence of school administration on student achievement has been implied to have direct or indirect consequences. One of the most important findings is the fact that school administrations have been able to initiate student support services that were insignificantly used. Poor use has been ascribed to students' lack of time to make use of those services and the lack of parental will to let children to benefit from those support initiatives like tutorials and guidance and counseling. On the other hand, there are reports that schools assign less competent teachers to teach in lower primary Grades. In some regions there are indications that schools operate under severe shortage of budgetary resources.

Although schools were unable to initiate significant parental academic support for children, it is obvious that their initiatives in terms of generating community support for school expansion, classroom construction and the construction of residential houses for teachers are common in almost all regions. However, it is clear that the construction of classrooms and

student enrolment do not match since reports from all regions point to the problem of large classes.

Another interesting finding is the prevalence of student and teacher absenteeism in schools. In the quantitative study, it has been found out that in schools where the director supervises student and teacher attendance, student achievement has been higher than others. However, failure to deal with disciplinary issues has been found out to negatively affect student learning. In fact, it was one of the responsibilities of the school administration to maintain discipline in schools. It was also indicated that lack of school facilities and supplies are grounded in school operations all over the nation. The school administration was unable to tap community resources to alleviate this problem throughout the nation.

CHAPTER 5: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary

The main purposes of the Second National learning Assessment (for Grade 4) was to measure learning attainments of Grade 4 students, and identify the factors that determine those attainments. It was also aimed at providing comparative information on school improvement from the National Baseline Learning Assessment conducted in 1999/2000. In order to obtain the required information for the study, both quantitative and qualitative research approaches were used. All Grade 4 students in the country were targeted for the purpose of providing national and regional estimates of student achievement results. A total of 13,248 randomly selected sample students from 376 schools in all regions except Gambella have been included. For the purpose of generating data on factors determining scholastic achievement, 751 teachers and 375 school principals were included. In the qualitative study, 321 students, 326 teachers and 246 parents from 44 schools provided information. The instruments of data collection included achievement tests, attitude surveys, and school checklist forms, questionnaires for teachers and school directors, and focus group discussion guides. Findings from the study show the following main results.

1. *The summary of composite achievement results at national level indicates that Grade 4 students achieved less than the expected minimum standards by the Ministry of Education, but their results are slightly better than the National Baseline Learning Assessment.*

As it has been summarized in Table 96, the overall performances in all subjects were low with only reading comprehension and environmental science averages reaching 50% correct. The mean percent score for mathematics (39.70) and English (38.68%) were notably lower than the national mean (48.48%) and also from the mean of the other two subjects.

2. *A summary of composite scores by gender indicates that the achievement of boys was higher than that of girls.*

At the national level, the overall composite mean score found for girls was 46.12%, while it was 50.14% for boys, indicating a gap of 4.02 percentage points. This same general pattern appears in all the subject areas (basic reading 2.71%; English, 4.9%, mathematics, 5.0%, and environmental science 4.82%). These results indicate that concern for gender equity in schooling needs to extend beyond access, to the provision of quality education that supports and encourages girls' achievement as well. Though there is an indication of increment in the achievement of girls since the baseline study, the gap between male and female student achievement has widened from baseline.

3. *A summary of composite scores by location indicates that urban students achieved better than rural students.*

Students from urban schools achieved a better mean score (49.28%) as compared to the mean score (48.09%) of students from rural schools at the national level. This indicates that the urban/rural difference in students' achievement is 1.19%, which is statistically significant at $p < 0.001$ level.

4. *A summary of composite scores by regions indicates that there were disparities in Grade 4 student achievement results among regions.*

The findings for each achievement test and their composite mean are calculated across regions and summarized in Table 97.

Table 97. Summary of Grade 4 achievement scores by region

Region	Basic Reading	English	Mathematics	Environmental Science	Composite
TIGRAI	63.37	32.95	34.50	52.50	45.87
AFAR	71.92	35.30	40.29	51.60	49.64
AMHARA	70.92	37.87	43.27	60.05	52.67
OROMIA	63.73	39.49	38.30	50.73	47.83
SOMALI	50.53	50.36	43.09	45.83	47.18
BEN-GUMUZ	60.34	32.44	34.34	47.84	43.69
SNNPR	60.31	38.93	39.70	49.75	47.27
HARARI	68.60	42.46	40.71	53.35	51.29
ADDIS ABABA	76.90	39.94	44.92	54.64	53.69
DIRE DAWA	62.24	37.19	37.83	47.86	46.20

From the table above, one can infer that there are also considerable variations in regional mean scores in each subject. Variations within each region also tended to be high, with standard deviations ranging from 12 to 24 percentage points in every subject.

Some regions showed a similar pattern of achievement across subject areas and scored better composite means. These are Addis Ababa (53.69), Amhara (52.67), Harari (51.29), and Afar (49.64) regions where each one of them scored above the national mean percentage point. The first three regions have also achieved above the minimum requirement of 50% pass mark. The remaining seven regions have regional mean scores below the national mean. These regions are Afar, Oromia, SNNPR, Somali, Dire Dawa, Tigray and Benishangul-Gumuz.

5. *Student achievements in Grade 4 English were significantly less than the national composite average for all subjects, and there were disparities among regions and between sexes.*

The average mean score obtained at the national level in the Grade 4 English was 38.68 with a standard deviation of 16.88. This is indeed significantly less than the composite scores for all subjects at national level (i.e. 48.8%).

The Somali Region scored the highest mean average of 50.36% while Tigray and Benishangul-Gumuz Regions achieved the least average scores of 32.95% and 32.44% respectively. As to the performance levels of each region, five of the regions - Somali (50.36%), Harari (42.46%), Addis Ababa (39.93%), Oromia (39.48%), and SNNPR (38.93%) scored relatively better mean scores (above the national mean of 38.7%) while all the remaining regions scored below the national average score ranging from 37.87% to 32.44%. The one way analysis of variance showed that there is a significant difference ($F = (9, 13018), 72.204, P < .001$) among the mean scores of the sample regions.

Boys at the national level scored higher (40.55%) in the Grade 4 English test as compared to girls (35.65%) showing a gap of 4.9%. This difference between male and

female students' performance is significant at a P-value less than 0.001. The gender gap in student achievement has also been observed in all the regions except in Harari where female students scored slightly better than male students.

Nationally, students from urban and rural schools have achieved nearly the same mean score (38.87%, and 38.58% respectively) in the Grade 4 English test. This indicates that there was no statistically significant difference between the performance levels of urban and rural students.

6. *Student achievement results in Grade 4 mathematics was significantly less than the national composite average for all subjects and there were disparities among regions and between sexes.*

The mean score achieved in Grade 4 mathematics test at the national level was 39.70% with a standard deviation of 17.2. With regard to the performance levels of each region, five of the regions: Addis Ababa (44.92%), Amhara (43.27%), Somali (43.09%), Harari (40.71%), and Afar (40.29%) achieved relatively better mean scores (above the national mean of 39.70%), and students from the SNNP Region scored (39.7%) which is equal to the national average. The remaining 4 regions scored below the national average ranging from 37.3% (in Oromia) to 34.34% (in Tigray and Benishangul-Gumuz regions). The one way analysis of variance computed on the Grade 4 mathematics test showed that there was a significant difference ($F = (9, 13133), 42.22, P < .001$) between the mean scores of the sample regions.

With respect to gender parity, boys scored better (41.68%) than girls (36.67) in the Grade 4 mathematics test with a gap of 5.01%. This difference between performances of male and female students is significant at P-value less than 0.001. The gap has also been observed in eight regions, except for the Somali Region where female students performed better than their male counter parts while in Harari the observed gap was very small (0.49%).

Students from urban and rural schools have achieved differently in the Grade 4 mathematics test with average scores of 40.38% and 39.35% respectively. Though the observed mean difference between urban and rural scores seems to be very small (1.03%), the calculated statistical test shows that this difference was significant at $P < 0.001$ level.

7. *Student achievement results in Grade 4 environmental science indicate that students scored more than the national composite scores, but this was limited only to few regions and boys.*

The national mean average obtained in the Grade 4 environmental sciences was 51.74% with a standard deviation of 17.41. In this test, the Amhara Region scored the highest mean average of 60.05% while the Somali Region achieved the least average score of 45.83%.

Concerning regional performances, only 4 regions - Amhara (60.04%), Addis Ababa (54.64%), Harari (53.35%), and Tigray (52.50%) scored relatively better mean scores (above the national mean of 51.74%) while all the remaining six regions scored below the national average score ranging from 51.60% (Afar) to 45.83% (Somali).

As to the observed gender gap in achievement at the national level, boys (53.43%) scored better mean than girls (49.61%) with an achievement gap of 3.82% in the Grade 4 environmental science test. This level of difference in performance is tested ($F_{(1, 12901)} = 153.166, P < .001$) as a significant deviation.

With regard to the comparison between school location and achievement, students from both urban and rural areas achieved very similar mean scores of 51.49% and 51.87% relatively, with only a 0.38% in favour of the rural schools. From this, it can be concluded that there is no difference between the performance levels of urban and rural students in the Grade 4 environmental science test at the national level.

8. *Student achievement results in Grade 4 reading comprehension test indicate that students scored more than the national composite scores, but this was limited only to few regions and boys.*

The national mean score achieved in the basic reading comprehension test was 64.5% which is almost similar to the result of the Ethiopian Baseline National Learning Assessment in the same subject (64.3%) with very similar standard deviations of 21.97 and 21.36, respectively. There is a statistically significant and positive relationship ($r = .472$) between the scores of the basic reading comprehension test and the school based measures in the mother tongue subjects of the same Grade level.

Boys at the national level scored high (65.71%) in the basic reading comprehension test compared to what girls scored (63.00%). This difference between the performances of male and female students is significant at P-value of less than 0.001. This implies that there is a considerable amount (2.71%) of gender gap. This trend of gender gap in students' achievement is observed in the seven regions except for Afar, Harari and Addis Ababa where female students scored slightly better than male students. The gender gap observed in the reading comprehension test is most notable in 4 regions: Oromia (with gender gap of 6.8%), SNNPR (6.3%), Tigray (5.8%), and Amhara (3.6%).

As to the performance levels of each region, Addis Ababa (76.90), Afar (71.92), Amhara (70.92), and Harari (68.60) scored a relatively high mean score (above the national mean of 64.49) while all the remaining regions scored below the national average score ranging from 50.53 (lowest in Somali) to 63.73 (moderate in Oromia). It shall be remembered that in the Baseline Assessment, the highest score in the Grade 4 basic reading comprehension test was registered to the Amhara Region, but not this time.

9. *The analysis of students' background in relation to student achievement shows that there was a relationship between academic scores and the student background characteristics.*

Students who travel longer distance from home to school achieved lower scores in academic subjects than those who travel shorter distances. Those who spend more time watching TV and listening to radio for leisure achieved lower scores than those who do not spend their time with these media. Frequent absenteeism from school has been found to relate to less achievement while getting sufficient amount of food was related to higher achievement. Students who had their own textbooks were better achievers than those who shared with others or who do not have textbooks at all. Similarly those who have shown interests in learning subjects were also the ones who scored better than who did not have such interests.

10. *The analysis of Grade 4 students' general attitudes indicates that the country's social development curriculum is making a positive difference in shaping students' attitudes towards health care, environmental protection, civics and ethics, cultural issues and education values.*

In health care, the main issues were personal and environmental hygiene, spacing of child birth, mode of transmission of HIV/AIDS and proper nutrition. In environmental protection, students expressed positive inclinations towards drinking clean water, preservation of historical heritages, planting of trees, environmental hygiene, and

prevention of drought and preservation of forests. In civics and ethics, the emphasis was on equality, honesty, socialization, use of information and culture of dialogue. Items on early marriage and female circumcision were included to see attitudes concerning harmful cultural practices while absenteeism, student participation in school affairs, learning and utilisation of personal earnings were some of the issues in which students expressed positive attitudes.

11. *The availability of curriculum materials has influential relationships with student achievement.*

The correlation analysis indicates that the availability of teacher's guides and subject syllabi had significant positive relations with student achievement in Grade 4. That is the more these materials were available to students, the higher was their achievement. Though not statistically significant, there is an indication that student textbook had negative relations with achievements. That is the higher the number of students using textbooks, the lower is student achievement.

12. *Teacher's personal characteristics and behavior have influences on student achievement in different directions.*

In the correlation analysis, teacher variables that positively and significantly relate to student achievement were teacher's perception of students' learning ability/attitude, teacher training on new teaching methods, teacher training on new assessment techniques, and the average level of teacher training for teaching. On the other hand, teacher's age, qualifications, total teaching experience, distance from school and social commitment were negatively correlated with student achievement, although these relations were not statistically significant.

13. *School management is another important dimension that influences student achievement.*

It has been found out that teachers taking classroom attendance, director's experience in school leadership, and how often the directors supervise teachers had significant positive relations with student achievement in Grade 4.

14. *The multiple regression analysis indicates that student home background and behavior were the strongest determinants of academic achievement in Grade 4 in Ethiopia.*

The variables grouped under *Student Background and Behavior* explained the highest portion (17.2%) of the observed variation in student achievement at school level before other factors were considered and 5.7% after controlling or considering all other factors in the model. This means, of the 28.7% variations explained by the final model, *Student Background and Behavior* ranked highest in both cases of entry order for analyses. This block contained five variables; among these, interest in the subject matters, students' sense of ownership of school property and alignment of home and instructional language showed positive relationship with achievement while absenteeism and time spent listening to the radio showed negative relationship.

15. *School Organization and Management is the second strongest determinant for student achievement in Grade 4.*

School Organization and Management block explained 16.4% of the observed variation in student achievement before controlling other factors and still a significant 6.0% with other factors controlled. This means, of the 27.8% variations explained by the factors in the current study, *School Organization and Management* ranked as the

second higher in both cases of entry order for analysis. This block contained two variables; namely, frequency of taking attendance by the teachers and directors' managerial experience. Both factors showed positive relationship with student achievement.

16. *Teacher variables are the third most important determinants for student achievement in Grade 4.*

Teachers' Variables explained the observed variation in achievement by 6.8% before other factors were controlled and 2.9% after controlling all other factors. This means of the 27.8% variations explained by the factors in the current study, *teacher variables* ranked as the third higher in both cases of entry order for analysis. This block contained variables which asked about teachers' perception (judgment) of students, the distance they traveled from home to school, and the use of the text books. The distance the teacher traveled from home to school affected student achievement negatively while the other two variables were related positively with achievement.

17. *School supply with reference to curriculum materials appeared as the fourth most important determinant for student achievement.*

School Supplies and Structures explained the observed variation in achievement by 2.1% before other factors were considered and by 0.4% after controlling all other factors in the model. This block contained two variables namely availability and use of syllabi and teacher's guides. Both factors showed significant positive relationships with achievement.

18. *Instruction and support is the least important influential factor that determines student achievement in Grade 4.*

Finally the *Instruction and Support* block, which included factors concerning the role of the director in supervising instructional activities, accounted for 2.6% of the total variation in student achievement before other factors were controlled and 0.2% after controlling all the other factors.

5.2 Recommendations

Based on the findings and implications of this study, the following specific recommendations are given.

1. ***The overall performance of the education system as measured by student achievement at Grade 4 needs improvement.*** The low level of overall achievement of students requires to be improved at least to a reasonable level of 50% as planned by MOE. Schools where students attain less than half of what they are expected to learn are in serious trouble.
2. ***The overall performance of girls in Grade 4 needs improvement.*** The overall achievement of girls requires to be improved in relation to that of boys. This is required in all subjects. The current results indicate that concern for gender equity in schooling needs to extend beyond access to the provision of quality education that supports and encourages girls' achievement as well. Eventhough there is an indication of increment in girls' achievement since the baseline study, the gap between male and female students achievement has widened from the baseline.
3. ***The performance of rural students in Grade 4 needs improvement.*** The performance of rural schools in comparison to urban schools requires improvement. Given that the majority of primary schools are rural, it is an alarming condition that should be addressed since achievements are already low at the national level.
4. ***Performance of students in Grade 4 core curricula (English and Mathematics) requires special attention.*** Mathematics and English achievement results were extremely lower than the other subjects. Some regions and girls achieved less than other regions and boys. These problems require solutions with reference to quality and equity.
5. ***Disparity of performance by region and sex of students in other Grade 4 subjects needs improvement.*** Although students' achievement scores were more in the national mean for environmental science and reading comprehension tests, they were significantly favoring boys and some regions. Attention must be given to bringing the equity among both sexes and regions.
6. ***The improvement of student achievement in Grade 4, first and foremost requires improvement in school management.*** School management with particular reference to teachers taking class attendance, director's experience in school leadership, and how often the directors supervise teachers showed significant positive relations with student achievement in Grade 4. Therefore it is required to encourage teachers to take class attendance, appoint and retain experienced school principals, and encourage principal's supervision of teachers.
7. ***Improvement of student achievement also depends on the improvement of student home background and personal behavior.*** Student interest in subjects, meals per day, agreement between home and instructional language, and students' sense of ownership in school property had significant positive relations with student achievement while time spent by listening to radio and watching television, distance from home to school and absence from school had significant negative relations with student achievement. Therefore, improving student interest towards subjects, improving the nourishment conditions, the use of home language in schools, and student participation in school affairs need to be improved. Moreover, the construction of nearby schools to reduce distance and discouraging wastage of time

- by listening to radios and watching TV, and encouraging students' classroom attendance need special attention.
8. ***Better supply of school curriculum materials for Grade 4 needs attention to improve student scholastic performance.*** This factor is one of strongest predictors of student achievement in Grade 4. Among these, the availability of teacher's guides and the availability of the syllabus to teachers stands top in the degree of order. Therefore, it is vital the teacher's guides and subject syllabi are made available to schools.
 9. ***special attention to teacher quality and behavior is the other priority area for improvement of students' learning in Grade 4.*** Teacher training on new teaching methods and assessment techniques, average teacher training level, and teacher's perception of student learning attitudes are important conditions for the improvement of student learning. Thus, organizing training programs on new methods of teaching, conducting teacher training on new assessment techniques, improving the average level of teacher training (pre-service), and encouraging teachers to learn about their students are some of the measures required.
 10. ***Strengthening the instructional practice of teachers and instructional support in the classroom are vital to improve student learning in Grade 4.*** The director's perception of the overall instructional quality or his/her knowledge and judgment of how teachers instruct their students has significant positive relation with student achievement. This means both the quality of the teachers as instructional leaders and the director's position to acknowledge performance as an instructional leader have to be strengthened. This can happen if the capacity of the principal and the teachers is developed using different trainings or other programs.
 11. ***Regional disparities in student achievements in Grade 4 need to be improved.*** There are considerable variations in regional mean scores in each subject. Variation within each region also tended to be high, with standard deviations ranging from 12 to 24 in every subject. This means that regions that achieved less on the test scores need to improve their learning outcomes as the responsibility to plan and implement primary educations is theirs.
 12. ***Learning from national learning assessment is required.*** Regular, reliable and timely assessments are key to improve learning achievement. Such assessments provide information about the level of performance of the education system at different levels and the main factors that influence learning. Therefore, it is a major requirement for all concerned bodies to learn from the current national learning assessment and make sure that the recommended improvements are realized in the years to come.

REFERENCES

- Anbessu Biazen and Barbara Junge (1988). *Problems in Primary School Participation and Performance in Bahir Dar Awraja*. Addis Ababa: Ministry of Education and UNICEF.
- Angelo, Tom. (May, 1999). *Doing Assessment as if Learning Matters Most*. American Association for Higher Education Assessment Bulletin.
- Brookover, W.B. and Lezotte, L.W. (1979). *Changes in School Characteristics Coincident with Changes in Student Achievement*. East Lansing: Michigan University.
- Duncan, W. (1989). *Engendering School Learning Science Attitudes and Achievement among Girls and Boys in Botswana*. Stockholm: Stockholm University.
- Edmond, R.(1979). Effective Schools of the Urban Poor: Educational Leadership (35): 15-24
- Ellet, C.D and Walber, H.J. (1979). *Principals' Competency, Environment and Outcomes in Educational Environments and Effects*. Berkeley, CA:McCutchan.
- Fagerlind, I. and Saha, L. (1989). *Education and National Development: A Comparative Perspective*. Oxford: Pergamon.
- Fullan, M.G. (1990). *Change Processes in Secondary Schools: Toward a More Fundamental Agenda*. In M.W. in McLaughlin, J.E. Talbert and N. Basica (eds.). The context of Teacher in Secondary schools: Teacher Realities. New York:Teachers College Press.
- Fuller, B. (1986). *Raising School Quality in Developing Countries: What Investments Boost Learning?* Report No. EDT7: Education and Training Dept. World Bank .
- Fuller, Bruce: (1987) Review of Educational Research *What School Factors Raise Achievement in Third World* . The World Bank.
- Greaney Vincent, and Thomas Kellaghan (1995) *Equity Issues in Public Examinations in Developing countries*_ world Bank, D,C.
- Heyneman, S.and Loxley.(1983). *The Effect of Primary School Quality on Academic Achievement across Twenty Nine High and Low Income Countries:*
- Himmel, E. (1984). *National Assessment in Chile: National Assessments Testing the System*. Washington ,D.C: Economic Development Institute of the World Bank
- Husen, T. (1967). *International Studies of Achievement in Mathematics: A Comparison of 12 Countries*. Stocklom: Almqvist and Wikshell International.
- Kellaghan, T. and Greaney, V. (2001). *Using Assessment to Improve the Quality of Education*. Paris: UNESCO

- Kellaghan, T. and Greaney, V. (2004). *Assessing Student Learning in Africa*. Washington, DC: World Bank.
- Keller, D. (2004). Ethiopia: Microsoft ® Encarta ® Online Encyclopaedia [http:// Encarta msn.com](http://Encarta.msn.com)
- Klitgard, R.E and Hall, G. (1973). *Are There Usually Effective Schools?* Santa Monica, CA:Rand. (ERIC.No.ED 085 402).
- McDill, E. L. et al. (1967). *Institutional Effects on the Academic Behavior of High School Students*, *Sociology of Education*, 40:181-99.
- McDill, E. L. and Rigsby, L.C. (1973). *Structure and Process in Secondary Schools*. Baltimore, MD: John Hopkins University Press.
- Ministry of Education (1986). *Evaluative Research of the General Education System in Ethiopia, Summary Report*. Addis Ababa: MOE.
- MOFED (2002). *Ethiopia: Sustainable Development and Poverty Reduction Program*. Addis Ababa: Ministry of Finance and Economic Development
- NCES (2003). *The Nation's Report Card: Mathematics Highlights USA*: National Centre for Education Statistics..
- NCES (2004). *The Nation's Report Card: The NAEP Guide*.USA: National Centre for Education Statistics.
- Prawalbruk, K. (1996). *National Assessment in Thailand: National Assessment Testing the System*. Washington,D.C: Economic Development Institute of the World Bank
- Parelius, R. J. and Parelius, A. P. (1987). *The Sociology of Education*. (2nd ed.). New Jersey: Prentice-Hall Inc.
- Puff, F. (1978). *Instructional Variables and Student Achievement in Reading and Mathematics: A Synthesis of Recent Process-Produce Research*. Philadelphia: Research for Better Schools, Inc. (ERIC, No.ED 189 135).
- Rosenshine, B.V. (1983). *Teaching Functions in Instructional Programs*. *Elementary School Journal* 83:335-51
- Seleshi Zeleke (1995). *Gender Differences in Mathematics Achievement as a Function of Attitudes in Grades 8 through 11* (in Northern Shewa Region). Masters thesis Addis Ababa University, School of Graduate studies.
- Serjiovanni, T.J. (1995). *The Principalship*. London: Allyn and Beacon.
- Stockard, J. and Mayberry, M. (1992). *Effective Educational Environment* London: Corwin Press, INC.
- Tekeste Negash (1990). *The Crises of Ethiopian Education: Some Implications for Nation Building*. Uppsala: Sweden

Thorndike,R.L.(1973). *Reading Comprehension in Fifteen Countries: An Empirical Study*. New York: John Wiley.

UNESCO, (1971). *Comparative Study of Secondary School Building Costs*.

World Bank Groups (2003). *Ethiopia Data Profile*. [http:// www.devdata world bank.org](http://www.devdata.worldbank.org)