National Assessment of Student Achievement (NASA) 2015

(Grade 3 and 5)

Government of Nepal
Ministry of Education
Education Review Office
Sanothimi, Bhaktapur
Webpage: www.ero.gov.np, email: eronasa@gmail.com
+977-1-6639556, +977-1-6639557 +977-1-6639556

National Assessment of Student Achievement (NASA) 2015
(Grade 3 and 5)

Government of Nepal
Ministry of Education
Education Review Office
Sanothimi, Bhaktapur, Nepal

National Assessment of Student Achievement (NASA) 2015 (Grade 3 and 5)

Published by: Education Review Office, Sanothimi, Bhaktapur, Nepal

Edited by:

Dr Lekha Nath Poudel
Dr Rishi Ram Rijal
Dr Binod Luintel

Data analysis and report writing:

Gopal Prasad Bhattarai; Sushil Babu Khanal
Shyam Prasad Acharya; Devi Ram Acharya; Sunita Shakya

Copyright: Education Review Office

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means without the permission in writing from the Education Review Office, Sanothimi except brief quotations in articles and review.

This publication can be cited as:
Education Review Office. (2016). *National Assessment of Student Achievement 2015 (Grade 3 and 5)*. Sanothimi: Author.

This report is also available at: www.ero.gov.np

Published in 2016 (2073 BS)

Copy: 500

Printed at:

Acknowledgements

After the publication of three reports on assessment of student achievement in Grade 8 twice and once in Grade 3 and 5 earlier, this is another report again on the assessment of student achievement in Grade 3 and 5 conducted second time in the year 2015. This report would not have been possible to produce in this form without contributions of numerous people and organizations. The Education Review Office acknowledges them for their supports and thanks them all for their time and effort.

During the process, a number of school teachers and subject specialists were involved in tests development and finalization without which test administration would not have been possible. They deserve special thanks for their contribution. Particular thanks go to the EDSC led by professor Dr. Tirtha Raj Khaniya for bearing the responsibility of tasks such as test booklet printing, transportation and delivery of the text booklets to the Districts, examining the answer sheets and accomplishing the preliminary data entry works. Due thanks go to the UNESCO Office in Kathmandu for its support provided to conduct an orientation program for the personnel of respective District Education Offices which eased the smooth test administration process. Special mention also goes to the personnel from the respective DEOs for the vital role they played in coordinating and monitoring the entire process during the test administration. Thanks also due to the students of respective Grades for taking part in the test and to teachers and the head teachers of respective schools for their cooperation in responding the relevant questionnaires that helped in generating required information for this report.

Special thanks go to the ERO personnel for their coordinating roles in item writing, moderation, conducting pretest, developing the final items, monitoring and coordinating the entire test administration process across the country, accomplishing data analysis and report writing. Among the personnel from ERO Sushil Babu Khanal, Devi Ram Acharya, Shyam Prasad Acharya and Sunita Shakya particularly deserve special thanks for their commendable works in coordinating not only the process of item development and finalization by analyzing their difficulty level, equating and standardizing them; but also devoting their full time in tabulating, analyzing the data and preparing the first write up of this report. Without their tireless effort during the process, the report could not come into this shape. Much thanks to Gopal Prasad Bhattarai, the Under Secretary, who coordinated the whole process of report writing, worked for proof reading and production of the report. This report would not have been possible to bring in this shape in time without his efforts. Due thanks also go to Prakash Kumar Kharel for his role in providing management and technical support to facilitate the entire process. We also appreciate the work done by Lav Dev Bhatta during the finalization of this report.

We appreciate the work done by the editors Dr Rishi Ram Rijal (Gauri Vigya Paramarsha Kendra) and Dr Binod Luintel (Mulung Foundation).

Finally, special gratitude and thanks to the Steering Committee headed by the Secretary for the MOE in taking policy decisions promptly to facilitate the assessment process. Similarly, due thanks also go to the Technical Committee for its support in the process of assessment program design and selection of the subject to assess. Equally the Subject Committees also deserve much thanks for their support provided not only in maintaining validity and reliability of the items but also in suggesting for the final form of the test.

This will be incomplete without extending thanks to many other names from the ERO not mentioned above who have directly or indirectly been involved in accomplishing the process and provided their support and assistance to bring the report in this shape.

ERO welcomes constructive suggestions and feedbacks from academia, professionals and practitioners for further improvement of the products in the days to come.

Dr Lekha Nath Poudel
Joint Secretary
Education Review Office

Table of Contents

Acknowledgements iii
Table of Contents v
List of Tables ix
List of Figures x
Abbriviations xi
Executive Summary 1
Chapter 1: Introducing NASA 2015 13
1.1 Introduction 13
1.2 Context of NASA 2015 13
1.3 Student Assessment as a Process 14
1.4 Characteristics of NASA 2015 18
1.5 Objectives of the Study 18
1.6 Method and Process Used for Assessment 18
1.7 Organization of Report 20
Chapter 2: Methodology 21
2.1 Introduction 21
2.2 Methodological Framework 21
2.3 Population and Sampling 22
Population Strata 22
Sample Size 23
Selection of Schools and Students 24
2.4 Item Preparation 26
Item Writing 26
Pre-testing of Item 27
Calibration and Adaptation of PIRLS and TIMSS Items 27
Selection of Linking Items 27
Selection of Items for Final Test 27
Final Tests 30
Preparation of Marking Schemes 37
2.5 Test Administration, Marking, and Data Entry 38
Test Administration 38
Marking and Data Entry 38
2.6 Background Questionnaires 38
A Model for Developing Background Questionnaires 38
Selection of Background Questionnaire 40
2.7 Specific Variables Used in the Analysis of Results 40
Equated Scores and IRT Modelling 40
Fennema-Sherman Attitude Scale 42
Socio-economic Status (SES) Eelated Variables 43
2.8 Statistical Methods Used in Analysis 44
Analytical Tools Used for the Statistical Analysis 44
Some Statistical Concepts Used for the Analysis of Results 45
2.9 Data Analysis and Report Preparation 47
2.10 Professional Standard and Ethical Consideration 48
Chapter 3: Analysis of Assessment Results for Grade 3 in Mathematics 49
3.1 Introduction 49

3.2 Basic Achievement Results in Mathematics....50
Overall Distribution of Achievement Scores....50
Achievement by Various Content Areas of Mathematics 51
Achievement in Various Levels of Cognitive Domain 52
Student Achievement by Type of Items 53
Comparison of Achievement Results between NASA 2015 and 2012....54
3.3 Achievement Score by Diversity Factors 54
Student Achievement by Districts 55
Student Achievement in Various Ecological Zones....56
Student Achievement by Development Regions 58
Student Achievement by School Type 59
Student Achievement in Rural and Urban Schools....60
Student Achievement in Various Language Groups 61
Student Achievement by Various Ethnic/Caste Groups....62
Gender and Student Achievement....63
3.4 Selected Explanatory Factors and Achievement 66
Achievement of Students from Various Socio-economic Status (SES) 67
Working Beyond School Hour and Achievement 77
Students' Attitude towards the Subject of Study and Achievement....79
Age of the Student and Achievement 80
Support for Study to the Student and Achievement....82
Homework and Achievement 82
Positive and Negative Activities in the School and Student Achievement 83
3.5 Summary of Findings 87
CHAPTER 4: ANALYSIS OF ASSESSMENT RESULTS FOR GRADE 5 IN MATHEMATICS 91
4.1 Introduction 91
4.2 Basic Achievement Results in Mathematics....91
Overall Distribution of Achievement Scores....91
Student Achievement by Various Content Areas in Mathematics 93
Achievement in Various Levels of Cognitive Domain 94
Student Achievement by Type of Items 95
Comparison of Achievement Results of NASA 2015 and 2012 96
4.3 Comparison with the International Standard....96
4.4 Achievement Score by Diversity Factors 97
Student Achievement by Districts 98
Student Achievement in Various Ecological Zones....99
Student Achievement by Development Regions 100
Student Achievement by School Type 102
Student Achievement in Rural and Urban Schools....103
Student Achievement in Various Language Groups 103
Achievement of the Students of Various Ethnic/Caste Groups....105
Gender and Student Achievement....106
4.5 Selected Explanatory Factors and Achievement 109
Achievement Variation of Students in Various Socio-Economic Status (SES)....110
Working Beyond School Hour and Achievement 120
Student Attitude Towards the Subject of Study and Achievement 122
Age of the Student and Achievement 123

Support for Study to the Student and Achievement ... 124
Availability of Textbook and Student Achievement ... 125
Homework and Achievement ... 126
Positive and Negative Activities in the School and Student Achievement ... 127
4.6 Summary of Findings ... 131
Chapter 5: Analysis of Assessment Results for Grade 3 in Nepali ... **135**
5.1 Introduction ... 135
5.2 Basic Results ... 136
Overall Distribution of Achievement Scores ... 136
Student Achievement in Various Content Areas in Nepali ... 137
Achievement in Various Levels of Cognitive Domain ... 139
Student Achievement by Type of Items ... 140
Comparison of NASA 2012 Achievement Results with NASA 2015 ... 141
5.3 Achievement Scores by Diversity Factors ... 142
Student Achievement by Districts ... 142
Student Achievement by Ecological Zone ... 144
Student Achievement by Development Region ... 145
Student Achievement by School Type ... 146
Student Achievement by Location of Schools ... 147
Language at Home and Student Achievement ... 148
Caste/ Ethnicity and Student Achievement ... 149
Achievement by Gender ... 151
5.4 Selected Explanatory Factors and Achievement ... 154
Socio Economic Status (SES) of Parents and Student Achievement ... 154
Working Beyond School Hour and Achievement ... 162
Student Attitude towards the Subject of Study and Achievement ... 164
Age of the Students and their Achievement ... 166
Support to the Students for Study and their Achievement ... 167
Homework and Achievement ... 168
Positive and Negative Activities at School and Student Achievement ... 169
5.5 Summary of Findings ... 172
Chapter 6: Analysis of Assessment Results for Grade 5 in Nepali ... **175**
6.1 Introduction ... 175
6.2 Basic Achievement Results in Nepali Grade 5 ... 176
Overall Distribution of Achievement Scores ... 176
Achievement in Various Levels of Cognitive Domain ... 179
Student Achievement by Type of Items ... 180
Comparison of NASA 2015 Results with NASA 2012 ... 181
6.3 Comparison of Results with the International Standard of PIRLS ... 181
6.4 Achievement Scores by Diversity Factors ... 183
Student Achievement by Districts ... 183
Student Achievement by Ecological Zone ... 185
Student Achievement by Development region ... 186
Student Achievement by School Type ... 187
Student Achievement by Location of Schools ... 189
Language at Home and Student Achievement ... 190
Ethnicity/Caste and Student Achievement ... 191
Achievement by Gender ... 193

6.5 Selected Explanatory Factors and Achievement 197
Socio Economic Status (SES) of Parents and Student Achievement 198
Working Beyond School Hour and Achievement 206
Students' Attitude towards the Subject and Achievement 208
Age of the Students and their Achievement 210
Support to the Students for Study and their Achievement 211
Availability of Textbook and Student Achievement 211
Homework and Achievement 212
Positive and Negative Activities at School and Student Achievement 213
6.6 Summary of Findings 216
Chapter 7: Analysis of Assessment Results for Grade 5 in English 219
7.1 Introduction 219
7.2 Basic Results 219
Distribution of Overall Results 220
Student Achievement in Various Content Areas 221
Achievement in Various Levels of Cognitive Domain 223
Student Achievement by Type of Items 223
Comparison of the results of NASA 2015 with NASA 2012 224
7.3 Achievement Scores by Diversity Factors 225
Student Achievement by Districts 225
Student Achievement by Ecological Zone 227
Student Achievement by Development Regions 228
Student Achievement by School Type 229
Student Achievement by Location of Schools 230
Language at Home and Student Achievement 231
Ethnicity and Student Achievement 232
Achievement by Gender 234
7.4 Selected Explanatory Factors and Achievement 237
Socio Economic Status (SES) of Parents and Student Achievement 237
Activities Beyond School Hour and Achievement 245
Students' Attitude towards the Subject of Study and Achievement 248
Age of the Students and their Achievement 249
Support to the Students for Study and their Achievement 250
Availability of Textbook and Student Achievement 251
Homework and Achievement 251
Positive and Negative Activities in School and Student Achievement 252
7.5 Summary of Findings 255
Chapter 8: Conclusions and Implications 259
8.1 Introduction 259
8.2. Summary of Context, Methods and Results 260
8.3 Implications 275
References 282

List of Tables

Table 2.1 Number of schools and students included in the sample Districts for Grade 3.....25
Table 2.2 Number of schools and students included in sample Districts for Grade 5 25
Table 2.3 Comparison of the characteristics of Mathematics test versions of Grade 3 31
Table 2.4 Comparison of the characteristics of Mathematics test versions in Grade 5.........31
Table 2.5 Characteristics of Mathematics tests of Grade 3 in various content areas 32
Table 2.6 Characteristics of Mathematics tests of Grade 5 in various content areas 32
Table 2.7 Characteristics of various domains in Mathematics tests of Grade 3....................33
Table 2.8 Characteristics of various domains in Mathematics tests of Grade 5....................33
Table 2.9 Comparison of the characteristics of Nepali test versions of Grade 3 34
Table 2.10 Comparison of the characteristics of Ne.pali test versions of Grade 5 34
Table 2.11 Characteristics of Nepali tests of Grade 3 in various content areas 34
Table 2.12 Characteristics of Nepali tests of Grade 5 in various content areas 35
Table 2.13 Characteristics of Nepali tests for Grade 3 in various cognitive domains 35
Table 2.14 Characteristics of Nepali tests for Grade 5 in various cognitive domains 36
Table 2.15 Comparison of the characteristics of English test versions..................................36
Table 2.16 Characteristics of English tests of Grade 5 in various content areas.....................37
Table 2.17 Characteristics of English tests of Grade 5 in various cognitive domains 37
Table 2.18 Factor structure analysis for Grade 5 in Mathematics: Component Matrix 43
Table 2.19 Indicators of SES in NASA 2015 (Mathematics in Grade 5)...............................44
Table 2.20 Rough boundaries of Effect Sizes 46
Table 3.1 Achievement score in various content areas in Mathematics 51
Table 3.2 Students' achievement by types of items in Mathematics...................................53
Table 3.3 Comparison of achievements between NASA 2012 and 2015 in various content areas 54
Table 3.4 Average achievement scores of sample Districts in Mathematics 55
Table 3.5. Achievement scores in Mathematics in various Ecological Zones 57
Table 3.6 Student achievement scores in Mathematics in various Development Regions ...58
Table 3.7 Student achievement according to school types in Mathematics..........................60
Table 3.8 Student achievement in rural and urban schools 60
Table 3.9 Achievement of Nepali and non-Nepali language groups in Mathematics..........61
Table 3.10 Student achievement by linguistic background in Mathematics 61
Table 3.11 Student achievement by ethnicity/caste background...62
Table 3.12 Lowest and highest achievement in various castes/ethnicities by Ecological Zones 62
Table 3.13 Student achievement of boys and girls in Mathematics 63
Table 3.14 Dichotomizing the indicators for home accessories...74
Table 3.15 Descriptive statistics of the students' achievement in different age groups..........81
Table 3.16 Achievement score in Mathematics in relation to the support provided to the students...82
Table 3.17 Bullying and the achievement...84
Table 3.18 Students' response towards teacher and school-related activities in Mathematics85
Table 3.19 Students' responses towards teacher and school-related activities 86
Table 4.1 Achievement score in various content areas of Mathematics 93

Table 4.2 Achievement by type of items in Mathematics 95
Table 4.3 Comparison of achievements between NASA 2012 and 2015 in various content areas (Mean score) 96
Table 4.4 Average achievement score of Mathematics in sample Districts 98
Table 4.5. Achievement scores in Mathematics in various Ecological Zones 99
Table 4.6 Student achievement scores in Mathematics by various Development Regions 101
Table 4.7 Student achievement by school type in Mathematics 103
Table 4.8 Student achievement in rural and urban schools 103
Table 4.9 Achievement of Nepali and non-Nepali home language groups in Mathematics 104
Table 4.10 Student achievement by linguistic background in Mathematics 104
Table 4.11 Student achievement by ethnicity/caste background in Mathematics 105
Table 4.12 Caste/ethnicity-wise minimum and maximum achievement in various Ecological Zones 106
Table 4.13 Student achievement of boys and girls in Mathematics 106
Table 4.14 Dichotomizing the indicators for home accessories 117
Table 4.15 Descriptive statistics of the students' achievement in different age groups 124
Table 4.16 Achievement score in Mathematics in relation to the support provided to the students 125
Table 4.17 Availability of textbook of Mathematics and the achievement 125
Table 4.18 Bullying and the achievement 127
Table 4.19 Students' response towards teacher and school-related activities in Mathematics at school 129
Table 4.20 Students' response towards teacher and school-related activities in schools 130
Table 5.1 Achievement in various content areas of Grade 3 in Nepali 138
Table 5.2 Achievement by type of items in Nepali 140
Table 5.3 Comparison of achievements in 2012 and 2015 in Nepali 141
Table 5.4 Average achievement score in sample Districts in Nepali 142
Table 5.5 Student achievement by type of school in Nepali 147
Table 5.6 Location of school and student achievement in Nepali 147
Table 5.7 Home language and Student achievement in Nepali 149
Table 5.8 Student achievement by ethnic/caste background 150
Table 5.9 Dalit students' achievement in different Ecological and Development Regions 150
Table 5.10 Dichotomizing the indicators of home accessories 160
Table 5.11 Students' attitude towards Nepali subject 165
Table 5.12 Students' response towards teacher and school related activities in schools 165
Table 5.13 Descriptive statistics of the students' achievement in different age groups in Nepali 166
Table 5.14 Descriptive statistics of support to the students and achievement level in Nepali 167
Table 5.15 Bullying and achievement in Nepali 170
Table 5.16 Students' responses towards teacher and school related activities in school 171
Table 5.17 Students' response towards teacher and school related activities in schools 172
Table 6.1 Achievement in various content areas in Nepali 178

Table 6.2 Achievement by types of items in Nepali 180
Table 6.3 Comparison of achievements in 2012 and 2015 in Nepali................................181
Table 6.4 Average achievement score of sample Districts for Grade 5 in Nepali..............183
Table 6.5 Student achievement by type of school in Nepali 188
Table 6.6 Comparison of results by content areas of two types of schools 188
Table 6.7 Comparison of results of two types schools by cognitive domain......................189
Table 6.8 Location of schools and student achievement in Nepali 189
Table 6.9 Home language and Student achievement in Nepali...191
Table 6.10 Student achievement by ethnic/caste background..191
Table 6.11 Dalit students' achievement in Nepali in different ecological and Development Regions 192
Table 6.12 Comparison of achievement scores in various contents areas in Nepali............194
Table 6.13 Variation in achievement by cognitive domains in Nepali................................194
Table 6.14 Comparison of achievement scores by various castes/ethnicities 195
Table 6.15 Comparison of girls and boys' achievement by Ecological Zones 196
Table 6.16 Comparison of girls and boys' achievement by Development regions . 197
Table 6.17 Dichotomizing the indicators for home accessories..204
Table 6.18 Students' response towards teacher and school related activities 209
Table 6.19 Attitude towards the subject and student achievement 209
Table 6.20 Descriptive statistics of the students' achievement in different age groups 210
Table 6.21 Descriptive statistics of support to the students and achievement level in Nepali 211
Table 6.22 Availability of Nepali textbook and achievement 212
Table 6.23 Bullying and achievement in Nepali 213
Table 6.24 Students' responses towards teacher and school related activities in school 214
Table 6.25 Students' response towards teacher and school related activities 215
Table 7.1 Achievement in various content areas in English language 222
Table 7.2 Achievement by type of items in the English language 224
Table 7.3 Comparison of achievements in 2012 and 2015 in English 224
Table 7.4 Average achievement scores of sample Districts in English..............................226
Table 7.5 Student achievement by type of school in the English language 230
Table 7.6 Location of school and student achievement in English 230
Table 7.7 Home Language and Student Achievement in English.....................................231
Table 7.8 Achievement of students from various language groups in English..................231
Table 7.9 Student achievement by ethnicity in English 233
Table 5.10 Dalit students' achievement in various Ecological and Development Regions ..233
Table 7.11 Dichotomizing the indicators for home accessories..242
Table 7.12 Students' achievement in different age groups in English 249
Table 7.13 Support to the students' study and achievement level in English.......................250
Table 7.14 Availability of English textbook and the achievement in English 251
Table 7.15 Bullying and achievement in English...252
Table 7.16 Students' responses towards teacher and school related activities at school 254
Table 7.17 Students' response towards teacher and school related activities in schools 254

List of Figures

Figure 1.1 Administrative cycle of NASA 2015 15
Figure 1.2 Phases of the assessment process of NASA 17
Figure 2.1 Sample Districts for NASA 2015 24
Figure 2.2 Conceptual framework for the background information 39
Figure 2.3 An example of DTA modelling on paid work and achievement in Mathematics Grade 5 45
Figure 3.1 Distribution of overall results in Mathematics 50
Figure 3.2 School mean achievement score in Mathematics 51
Figure 3.3 Comparison of achievement scores in various content areas in Mathematics 52
Figure 3.4 Comparison of achievement score in various levels of cognitive domain in Mathematics 53
Figure 3.5 Average achievement score of sample Districts in Mathematics 56
Figure 3.6 Comparison of achievement scores by Ecological Zones in Mathematics 57
Figure 3.7 Comparison of achievement scores by Development regions in Mathematics 58
Figure 3.8 Distribution of achievement scores of students in community and institutional schools in Mathematics 59
Figure 3.9 Comparison of achievement between boys and girls in different content areas 63
Figure 3.10 Comparison of achievement between boys and girls in various ethnic/caste groups 64
Figure 3.11 Comparison of achievement between boys and girls in various Ecological Zones 65
Figure 3.12 Comparison of achievement between boys and girls in various Development Regions 66
Figure 3.13 DTA of mother's education and students' achievement in Mathematics 68
Figure 3.14 DTA of father's education and students' achievement in Mathematics 69
Figure 3.15 Relation between mother's education and students' achievement in Mathematics70
Figure 3.16 DTA of mother's occupation and students' achievement in Mathematics 71
Figure 3.17 Comparison of mothers' occupation and students' achievement in Mathematics 71
Figure 3.18 Comparison of fathers' occupation and students' achievement in Mathematics 72
Figure 3.19 Comparison of fathers' occupation and students' achievement in Mathematics 72
Figure 3.20 Relation between home possessions and achievement in Mathematics....74
Figure 3.21 Relation between home accessories and achievement in Mathematics 75
Figure 3.22 Relation between the SES and achievement in Mathematics....76
Figure 3.23 Distribution of achievement by socio-economic status and type of schools in Mathematics 77
Figure 3.24 DTA of paid work and achievement in Mathematics 78
Figure 3.25 DTA of household work and achievement in Mathematics 79
Figure 3.26 Relation between students' attitude and achievement in Mathematics 80
Figure 3.27 Relation between the age of student and achievement in Mathematics....81
Figure 3.28 Effect of homework in students' achievement in Mathematics....83
Figure 3.29 Effect of bullying in the achievement of Mathematics 84

Figure 3.30 Relation between positive actions at school and achievement in Mathematics...86
Figure 4.1 Distribution of overall results in Mathematics ...92
Figure 4.2 School mean achievement score in Mathematics...93
Figure 4.3 Comparison of achievement scores in various content areas in Mathematics...94
Figure 4.4 Comparison of achievement scores in various levels of cognitive domain in Mathematics ...95
Figure 4.5 Student achievements in the international TIMSS Mathematics Scale...97
Figure 4.6 Average achievement score of sample Districts in Mathematics ...98
Figure 4.7 Comparison of achievement scores by Ecological Zones in Mathematics ...100
Figure 4.8 Comparison of achievement scores by Ecological Zones in Mathematics ...101
Figure 4.9 Distribution of students' achievement scores in community and institutional schools ...102
Figure 4.10 Comparison of achievement between boys and girls in different content areas ...107
Figure 4.11 Comparison of achievement between boys and girls in various ethnic/caste groups ...108
Figure 4.12 Comparison of achievement between boys and girls in various Ecological Zones ...108
Figure 4.13 Comparison of achievement between boys and girls in various Development Regions ...109
Figure 4.14(a) DTA of mother's education and students' achievement in Mathematics ...111
Figure 4.14(b) DTA of father's education and students' achievement in Mathematics...111
Figure 4.15 Relation between parent's education and students' achievement in Mathematics ...112
Figure 4.16 DTA of mother's occupation and students' achievement in Mathematics ...113
Figure 4.17 Comparison of mothers' occupation and students' achievement in Mathematics...114
Figure 4.18 DTA of fathers' occupation and students' achievement in Mathematics...115
Figure 4.19 Comparison of fathers' occupation and students' achievement in Mathematics ...115
Figure 4.20 Relation between home possessions and achievement in Mathematics...117
Figure 4.21 Relation between home accessories and achievement in Mathematics ...118
Figure 4.22 Relation between the SES and achievement in Mathematics...119
Figure 4.23 Distribution of achievement by SES and type of schools in Mathematics ...120
Figure 4.24 DTA of paid work and achievement in Mathematics...121
Figure 4.25 DTA of household work and achievement in Mathematics ...122
Figure 4.26 Relation between students' attitude and achievement in Mathematics ...123
Figure 4.27 Relation between the age of student and achievement in Mathematics...124
Figure 4.28 Effect of homework in students' achievement in Mathematics...126
Figure 4.29 Effect of bullying in the achievement of Mathematics ...128
Figure 4.30 Relation between positive actions at school and achievement in Mathematics ...130
Figure 5.1 Distribution of overall achievement scores of Grade 3 ...136
Figure 5.2 Distribution of schools' mean achievement scores of Grade 3 in Nepali ...137

Figure 5.3 Comparison of achievement scores in various content areas for Grade 3 in Nepali 138
Figure 5.4 Comparison of achievement score at various levels of cognitive domain in Nepali 139
Figure 5.5 Comparison of results of NASA 2012 and 2015 in various content areas in Nepali 141
Figure 5.6 Comparison of average achievement of sample Districts in Nepali 143
Figure 5.7 Comparison of achievement across various Ecological Zones in Nepali 144
Figure 5.8 Comparison of the achievement among various Development Regions in Nepali 145
Figure 5.9 Distribution of the students' mean score in community and institutional schools....................146
Figure 5.10 Achievement in Nepali by various language groups....................148
Figure 5.11 Madhesi students' achievement in various content areas in Nepali 151
Figure 5.12 Comparison of achievement of boys and girls in different content areas of Nepali....................152
Figure 5.13 Comparison of achievement between boys and girls in various castes/ethnicities in Nepali 152
Figure 5.14 Comparison of achievement by gender in various Ecological Zones in Nepali 153
Figure 5.15 Comparison of achievement of boys and girls in various Development 154
Figure 5.16 DTA of mother's education and students' achievement in Nepali 155
Figure 5.17 DTA of father's education and students' achievement in Nepali 156
Figure 5.18 Relation between parents' education and student achievement in Nepali 156
Figure 5.19 DTA of mother's occupation and students' achievement in Nepali 157
Figure 5.20 DTA of father's occupation and students' achievement in Nepali 158
Figure 5.21 Relation between home possessions and achievement in Nepali 159
Figure 5.22 Availability of home accessories and achievement in Nepali 160
Figure 5.23 Relation between the SES and achievement in Nepali 161
Figure 5.24 Achievement of schools with various levels of SES in Nepali 162
Figure: 5.25 Relationship between watching TV and achievement in Nepali....................163
Figure 5.26 Relationship between household chores and achievement in Nepali 163
Figure 5.27 Paid work and achievement in Nepali 164
Figure 5.28 Relation between attitude and achievement in Nepali 166
Figure 5.29 Relation between the age of students and achievement in Nepali 167
Figure 5.30 Support received and achievement....................168
Figure 5.31 Relation between assigning homework and achievement in Nepali 169
Figure 5.32 Effect of bullying in student achievement in Nepali 170
Figure 5.33 Relation between positive actions in school and achievement in Nepali 172
Figure 6.1 Distribution of overall achievement scores in Nepali 176
Figure 6.2 Distribution of schools' mean achievement scores in Nepali 177
Figure 6.3 Comparison of achievement scores in various content areas in Nepali 178
Figure 6.4 Comparison of achievement score in various levels of cognitive domain in Nepali 179

Figure 6.5 (a) Comparison of results of NASA 2012 and 2015 in various content areas in Nepali 181
Figure 6.5 (b) Comparing NASA 2015 results in Nepali with PIRLS reading scale.............182
Figure 6.6 Comparison of average achievement of sample Districts for Grade 5 in Nepali 184
Figure 6.7 Comparison of achievement of various Ecological Zones in Nepali 185
Figure 6.8 Comparison of the achievement of various Development Regions in Nepali ...186
Figure 6.9 Distribution of the students' mean score in community and institutional schools187
Figure 6.10 Language group-wise achievements in Nepali..190
Figure 6.11 Madhesi students' achievement in various content areas in Nepali 193
Figure 6.12 Comparison of achievement of boys and girls in different content areas in Nepali...193
Figure 6.13 Comparison of achievement of boys and girls in various castes/ethnicities in Nepali 195
Figure 6.14 Comparison of achievement by gender in various Ecological Zones in Nepali 196
Figure 6.15 Comparison of achievement of boys and girls in various Development Regions in Nepali 197
Figure 6.16 DTA of mother's education and students' achievement in Nepali 199
Figure 6.17 DTA of father's education and students' achievement in Nepali 199
Figure 6.18 Relation between parents' education and student achievement in Nepali 200
Figure 6.19 DTA of mother's occupation and students' achievement in Nepali 201
Figure 6.20 DTA of father's occupation and students' achievement in Nepali 202
Figure 6.21 Mother's and father's occupation and students' achievement in Nepali 202
Figure 6.22 Relation between home possessions and achievement in Nepali 203
Figure 6.23 Availability of home accessories and achievement in Nepali 204
Figure 6.24 Relation between SES and student achievement in Nepali 205
Figure 6.25 Achievement of schools with various levels of SES in Nepali 206
Figure: 6.26 Relationship between watching TV and achievement in Nepali........................207
Figure 6.27 Relationship between household chores and achievement in Nepali 207
Figure 6.28 DTA of paid work and achievement in Nepali 208
Figure 6.29 Relation between attitude and achievement in Nepali 209
Figure 6.30 Relation between the age of the students and achievement in Nepali 210
Figure 6.31 Relation between assigning homework and achievement in Nepali 212
Figure 6.32 Effect of bullying in student achievement in Nepali 213
Figure 6.33 Relation between positive actions in school and achievement in Nepali 215
Figure 7.1 Distribution of overall achievement scores in English.......................................220
Figure 7.2 Distribution of schools' mean achievement scores in English 221
Figure 7.3 Comparison of achievement scores in various content areas in English 222
Figure 7.4 Comparison of achievement scores in various levels of cognitive domain in English 223
Figure 7.5 Comparison of results of NASA 2012 and 2015 in various content areas in English 225
Figure 7.6 Comparison of average achievement of sample Districts in English 226
Figure 7.7 Comparison of achievement of various Ecological Zones in English................227

Figure 7.8 Comparison of the achievement of various Development Regions in English...228
Figure 7.9 Distribution of the students' scores in community and institutional schools in English ...229
Figure 7.10 Comparison of the achievement of students from various language groups in English ...232
Figure 7.11 Madhesi students' achievement in various content areas in English...234
Figure 7.12 Comparison of achievement of boys and girls in different content areas of English ...234
Figure 7.13 Comparison of boys and girls' achievement across ethnicities in English ...235
Figure 7.14 Comparison of achievement by gender in various Ecological Zones in English ...236
Figure 7.15 Comparison of achievement of boys and girls in various Development Regions ...236
Figure 7.16 Comparison of achievement of boys and girls by school types in English ...237
Figure 7.17 DTA of mother's education and students' achievement in English ...238
Figure 7.18 DTA of father's education and students' achievement in English ...239
Figure 7.19 Relation between parents' education and student achievement in English ...240
Figure 7.20 DTA of mother's occupation and students' achievement in English ...240
Figure 7.21 DTA of father's occupation and students' achievement in English...241
Figure 7.22 Relation between home possessions and achievement in English ...242
Figure 7.23 Availability of home accessories and achievement in English ...243
Figure 7.24 Relation between the SES and achievement in English ...244
Figure 7.25 Achievement of schools with various levels of SES in English ...244
Figure: 7.26 Relationship between watching TV and achievement in English ...245
Figure 7.27 Relationship between household chores and achievement in English ...246
Figure 7.28 Relationship between study/doing homework at home and achievement in English ...246
Figure 7.29 Relationship between playing games and achievement in English ...247
Figure 7.30 Relationship between paid job and achievement in English ...247
Figure 7.31 Relation between students' attitude and achievement in English ...248
Figure 7.32 Relation between the age of the students and achievement in English ...249
Figure 7.33 Relation between homework and feedback and achievement in English ...251
Figure 7.34 Effect of bullying in student achievement in English ...253
Figure 7.35 Relation between positive actions in school and achievement in English ...255

Abbriviations

ANOVA	Analysis of Variance
BPEP	Basic and Primary Education Programme
CBS	Central Bureau of Statistics
CERID	Research Centre for Educational Innovation and Development
CERSOD	Centre for Educational Research and Social Development
DEO	District Education Office
DOE	Department of Education
DTA	Decision Tree Analysis
EDSC	Education and Development Service Centre
ERO	Education Review Office
ES	Effect Size
FSAS	Fennema–Sherman Attitude Scales
GLM	General Linear Modelling
IRT	Item Response Theory
Max	Maximum
Min	Minimum
NASA	National Assessment of Student Achievement
OECD	Organization for Economic Co-operation and Development
OMR	Optical Mark Reader
OPLM	One Parametric Logistic Model
PEDP	Primary Education Development Project
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
P-SES	Percentage of Socio-Economic Status
SD	Standard Deviation
SE	Standard Error
SES	Socio-Economic Status
SPSS	Statistical Package/Programme for Social Science
TIMSS	Trends in International Mathematics and Science

Executive Summary

Introduction

Significance of formal schooling has long been recognized for a couple of reasons. First, it develops among youngsters the mastery over the content knowledge on core subjects, which helps them understanding the social, political, economic world. Second, through developing that subject specific knowledge and understanding, schooling also fosters improved literacy and numeracy skills in the form of reading, writing and computation ability that enable them for information retrieving, processing and communication skills to live an informed life. On the top of them, significance of schooling lies in developing basic cognitive skills in the form of literacy and numeracy, which in turn form the foundation of problem solving, reasoning, advance thinking including the ability to adopt, diffuse and innovate technology that have direct impacts on job market outcomes, economic productivity and social development. Additionally, these skills and knowledge acquired in early Grades of schooling lay strong foundation for individuals to learn throughout their life helping people adapt change to meet future challenges.

As prosperity and future development of a country and individuals largely depends upon these knowledge and skills learnt in schools by our youngsters; policy makers, educators, parents and society seek to have reliable information on how far those children are learning at school, whether they will be able to meet future challenge and whether there is major gaps in their learning that need to be addressed. They have reason to be confident that their children are acquiring the knowledge and skills that they will need to function as workers, family members and citizens in the 21st century. Besides, measurement of these knowledge and skills is also essential to tracking development indicators and assessing the effectiveness of educational policies, programs and practices (Boreman, Hews, Overman & Brown, 2003; Hanushek and Rivkin, 2010). With these motifs of seeking reliable information on students learning at schools and measuring the effectiveness of education system, assessing students' achievement by means of standardized test of a large-scale at various levels has evolved in the front of education. Nepal is no exception to adopt large-scale achievement test since the mid of 1990s to determine what students have learned because of their educational experiences.

Context

Various assessments of student achievement in Nepal were also conducted before 2011 for various Grades (in 1995, 1997, 2001 for Grade 3; in 1997 for Grade 4; in 1998, 1999; 2003, 2008 for Grade 5; in 1999 for Grades 6 and 8 and again in 2008 for Grade 8) aiming at assessing learning outcomes of students. The aim of such assessment of student achievement is also to determine the level of learning at respective Grades and to feed into policy in order to improve the education system. However, they all were conducted selecting relatively small sample and could be inadequate to nationally represent. Large-scale assessment with a nationally representative sample for the first time came into practice since 2011 after the establishment of the ERO under the MOE. During the implementation of SSRP 2009-2015, four assessments, two rounds in each Grade 3, 5 and 8, were planned to accomplish. Accordingly, two rounds of large-scale assessments for Grade 8 in 2011 and 2013 and one round for Grade 3 and 5 in 2012 have already been accomplished. Next to them, this is another, the second round of large-scale assessment accomplished in 2015 aiming to assess the learning outcomes of 3rd Graders in Mathematics and Nepali and of 5th Graders in Mathematics, Nepali and English.

Main Objectives

Like the earlier assessments, this study also aims to generate objective, accurate, and comparative information on learning achievement of students in Primary schools being motivated to evaluate the education system to support in obtaining results as nuanced in the curricula. The main objectives of this assessment were to determine the learning level of Grade 3 students in Mathematics and Nepali, and Grade 5 students in English, Mathematics and Nepali against the curricula goals; to create a reliable database on the learning level on those subjects for benchmarking; compare them with earlier achievements in order to monitor the progress over time; and to generate recommendations for policymaking to improve learning level of students.

Methodology

At the mid of 2015, the test was conducted to assess learning achievements of students in Mathematics and Nepali for Grade 3 and Mathematics, Nepali and English for Grade 5 in 1543 randomly selected schools from 23 sample Districts covering all Ecological Zones and Development regions. Altogether 74,078 students (33,863 in Grade 3 and 40,015 in Grade 5 from randomly stratified 1543 sampled schools) participated in the assessment. In the sample for Grade 3, 16927 students were boys and 16936 were girls. Similarly, out of the total sample for Grade 5, 19926 students were boys and 20089 were girls. Out of the 75 Districts of Nepal, the dataset represents a randomly selected 23 Districts covering all five Development regions (Eastern, Central, Western, Mid-Western, Far-Western), and the Kathmandu Valley, as well as all Ecological Zones (Mountain, Hill, Tarai). In the sample, both rural and urban schools as well as community and institutional (private) schools are proportionally represented so that the results of the assessment can credibly be extended to the entire student and school populations of Nepal.

Final items were standardized after pretesting six sets of test papers in each subject in 7724 students (ranging from 1514 to 1565 students in each subject) 153 schools from 11 Districts representing different strata. Only those items having high discriminatory power were selected for final test by analysing each of their difficulty level. Based on the pretest results of the items, the difficulty levels of the tests were set around 50–60%. Some linking items from international tests like PIRLS and TIMSS also were used to make the test results comparable to the international standard. All the items were analyzed and equated using IRT modelling.

Three versions of the items in each subject were administered and the final scores were equated by utilizing the IRT modelling. Reliability of the tests was found high and the validity was assured by applying specification grids of the national curriculum developed by the Curriculum Development Center (CDC). From a methodological standpoint, the process and practices of the inquiry has successfully followed the procedures as used in some international level test with some contextualization on them to reflect the reality of the Nepali context. Thus, this test is believed to fulfill the national and international ethical principles, criteria and standard to qualify it to be a credible assessment. The results were linked to the set of results from the 2012 assessments well as to the international item banks of TIMSS and PIRLS.

The tests were administered uniformly at a time in one shot in all the sample schools throughout the country in the scheduled day. Each selected school was assigned to conduct test in one of the selected subjects for each Grade. Thus, the students in a Grade were required to participate in one of the selected subjects assigned to the school. The answer sheets were marked and achievement scores were tabulated using Optical Mark Reading (OMR) machine.

The results are reported mainly as percentages of maximum marks where 100 (%) represents all tasks solved and 0 (%) for none. As a result of pre-testing of the items, the difficulty levels of the tests were set at 50—60%. For English and Nepali language assessment, the Common European Framework of Reference for Languages (CEFR) was also used to obtain the level of students from a language achievement point of view.

Main Results

In the test, the average achievement has remained at 45 to 52 varying by subjects and Grades. In Nepali, the 3rd Graders have achieved 52 percent which is 46 percent for 5th Graders. Similarly, in Mathematics, the higher Graders have achieved (48) lower Graders (45). However, comparison of two grades based on their average assessment scores may not be preferable, it indicates some trends about the achievement of students. The language proficiency level of 5th Graders is more or less at the same level in both English (47) and Nepali (46). To look at the mean achievement from equality point of view, gender parity is found to be maintained in both the Grades in all the subjects. It is notable that girls are ahead of boys by 2-3 percent point in Nepali in both Grades and by 2 percent points in Grade 5 Mathematics, whereas in English, they are seen equal (47). However, when seen from other perspectives going deep down into the results on different variables, low levels of learning with high discrepancies are noticed. Having analyzed the database for each subject, the main results have been summarized in the following points.

- **Distinctly divided school and student population into various groups in terms of achievement**

The population of schools participating in the test is seen to have divided into different distinct groups from the lowest, low, average to the high and highest achieving. Out of the total, 5 to 5.6

percent schools fall into the lowest achieving even less than 10 percent and 1.4 to 3.2 percent of them fall into the highest achieving more than 90 or above of the average score. From minimum 17.3 to maximum 24 percent schools fall into the low achieving group with average score from 10 to 30 percent. Maximum percentage of schools ranging from 37.3 to 49.1 belongs to the average performing achieving 31 to 59 percent of the score. Among the high achieving 26.3 to 38.1 percent schools with average score from 60 to 89 percent.

Not only the schools are divided into various groups, the population of student also is seen to be divided into two to three distinct groups in terms of the achievement score with remarkable variances: both low (15 percent or below) and high (63 percent) performing students in community schools and mainly high (above than 60 percent) and highest (above than 80 percent) performing students in the institutional schools. In comparison to Mathematics and Nepali for both Grades and English for Grade 5, the 5th Graders population for Mathematics is more normally distributed. In community schools, 50% students belong to the low performing groups below the average forming a large plot whereas in institutional schools more than 50% students lie at more than average achieving population forming its bigger plot with a small plot of low achieving group. It is notable that community schools also are found to have achieved as high as institutional schools but their number is too small, which is not enough to raise average achievement. Similarly, institutional schools are also there achieving as low as community schools whose number is small. Because of the low achieving large population and high achieving small population in community schools and mainly high to higher achieving population in institutional school, the entire system is seen to have shifted towards low performing one indicating that the system is not able to give sufficient support for those students who are lagging behind in the early Grades. From an equal opportunity viewpoint, such variations in achievement is not a positive sign that these 'low' and 'high' achiever students are concentrated in community schools in certain pockets in rural and outreach parts of the country.

- **Disproportionately distributed achievement level among Ecological, Development regions and Districts**

The achievement level is found disproportionately distributed among different Geographical regions and Development regions, and Districts showing a wide inequality in learning opportunity for children. Among the Ecological regions, the Valley is always performing the highest one achieving 51-64 percent in Grade 3 and 57-68 percent in Grade 5, whereas the Mountain records the lowest of all achieving 40-45 percent in Grade 3 and 37-42 percent in Grade 5. After the Valley, higher achievement is found in Hill in Nepali with average score 46-51 by 2-3 percent points above than the Tarai against 43-49 percent, whereas in Mathematics and English the Hill lags behind the Tarai by 3-7 percent points lower with average score 45-49 percent against 42-45 percent of the Hill. The widest difference between them is seen English with 7 percent points which is smallest in Mathematics with 3 percent points.

Significant variations in achievement are also notable between and among the Development regions. Among the five Development regions, the western region is found to be outperforming the rest except in Grade 3 Mathematics with average score 50-57 percent in Nepali and 48 percent in English. Only in Grade 3 Mathematics, it lags behind Eastern with just 1 percent points achieving 45 against the 46 of the Eastern. After the Western, Eastern region is seen performing better than the other with very small differences except in Mathematics achieving 43-45 percent in Nepali, 46-48 percent in Mathematics and 46 percent in English. In Case of Grade 5 Mathematic, Eastern region lags by 1 percent behind the Central achieving 47 percent in comparison to 48 percent of the later whereas in Grade 3, both the regions achieve equally 46 percent in this subject. The Central region is seen better achieving than Mid-Western with average score 43-45 percent in Nepali, 46-48 percent in Mathematics and 41 percent in English against 41-43,37-43 and 38 percent respectively of the Mid-Western. While compared the achievement between Mid-Western and Far Western, both of them are seen equal in English achieving 38 percent being the lowest of all. In Mathematics for both Grades and in Nepali for Grade 3, Mid-Western region is seen performing the lowest achieving 43 and 41 percent against the 45 and 40-45 percent respectively of the Far Western. In Nepali for Grade 5, the Far Western is seen the lowest achieving 37 percent.

In addition to the geographical and regional variations, wider inequality in achievement is found among the Districts too with notable differences (25-47 percent points in Nepali, 21 percent points in Mathematics and 44 percent points in English). In the lowest performing Districts, the average achievement goes down to as low as 33-34 percent in Nepali, 31-36 percent in Mathematics and 26 percent in English; whereas it rises up to 59-80 percent in Nepali, 57- 59 percent in Mathematics and 70 percent in English. Among the high achiever are Mustang (80 and 59), Kathmandu (66 and 56), Jhapa (64 and 58), Ilam (62 and 51), Nawalparasi (55) and Syangja (56) in Nepali; Jhapa (59-57), Kathmandu (56) and Ilam (52-54) in Mathematics; Kathmandu (70), Bhaktapur (69), Lalitpur (64) and Jhapa (61) in English. On the contrary, Routaht (33-37), Bhojpur (36-38) and Jajarkot (34) in Nepali; Shankhuwasava (35), Bhojpur (36-38), Jajarkot (33), Morang (36) in Mathematics; Bhojpur (26), Jajarkot (33),Pyuthan (33), Kalikot (35) in English. Once again, from the perspectives of equality in opportunity, such wider variation in achievement is not positive sign for the system.

- **Wider inequality in learning achievements among students by location and types of school**

Variance in achievement also is seen while analyzing the results in terms of location and types of schools. Students from urban schools are seen performing better than the rural ones. For instances, the achievement score for urban students has reached to 53 to 59.8 in Nepali Grade 3 and 5, 50.9 to 54.8 in Grade 3 and 5 Mathematics and 66 in English; whereas it is 44.4 to 48.9.4 in Nepali Grade 5 and 3,43.5 to 47.4 in Mathematics Grade 3 and 5 and 39.9 in English for rural. Such level of remarkable difference is not justifiable in any grounds from equality point of view.

Besides the variance in terms of location of schools, achievement gap is also seen wider between the types of schools students attended, as students' average achievement in institutional school is found notably higher than the others. The average achievement score in institutional schools reaches at 54.1 to 60.8 in Grade 3 and 5 Mathematics, 61.4 to 67.4 in Grade 5 and 3 Nepali and 73 in English which is 42.8 to 45.9 for Nepali Grade 5 and 3, 41.0 to 44.5 for Mathematics in Grade 3 and 5 and 38 in English in Community schools. One of the reasons behind such variances may be the higher socio-economic status of the family sending children to institutional schools as higher socioeconomic status and educational achievement are positively correlated.

It is noteworthy that there are also few number of community schools where the average results are at the same level as in the institutional schools even though the SES is remarkably lower. In these schools, either the processes are more effective than in the institutional schools or the students are of the same ability as those in the institutional schools and are not adversely affected by the processes within the school or their socio-economic status. It could be concluded that higher achievement of community schools is not due to the system but because of the individual effort of school, teacher or students themselves.

- **An unbalance learning across all curricula contents**

Against the expectation of curricula for similar level of learning in all content areas, the dataset is evident that certain contents of the curricula are learnt less effectively than others. For instances, in Grade 3 Mathematics, the achievement level in *Arithmetic* (45) and *Numeracy* (49) are remarkably lower than *Algebra* (57) and *Geometry* (57) whereas in Grade 5 it is lower in *Algebra* (46) and in *Numeracy* (47) than in *Arithmetic* (51) and in *Geometry*(49).Similarly, in Nepali and English, the *Writing* (42 in Grade 3,33 in Grade 5 for Nepali and 40 in Grade 5 English) followed by *Reading* skills (51 in Grade 3, 39 for Grade 5 Nepali and 46 in English) are notably poorer in comparison to the achievement in *Grammar* (58 for Grade 3 Nepali and 53 for Grade 5 English) and *Vocabulary* (76 for Grade 3,52 for Grade 5 in Nepali and 62 in English). From the data set, it is also evident that instead of improving, achievement level in these skills lowers down at the later Grade. Circumscribed with such unbalance learning on some domains, the entire system is shifted towards a low performing making it less effective to yield better results. Such glaring evidence raises an issue that why schooling at upper Grade fails to foster writing and reading skills.

- **Lower level of cognitive ability for the tasks requiring higher ability**

Due to poor writing and reading, students are found comparatively poorer in solving the *Subjective* types of items in which average score ranges from 33-48 in Nepali, remains 35 in Mathematics and 40 English in comparison to 55-63 in Nepali, 54-57 in Mathematics and 54 in English in *Objective* items. From this fact, it is evident that students are found poor in describing in detail the information, elaborating the ideas, explaining the concepts in structured form or to create short paragraphs in the relevant themes in an organized way. Similarly, they are also weaker in producing fluent texts, preparing synthesis and abstracts from the given text.

The dataset is evident that students in all subjects are seen performing comparatively better in *Knowledge* level tasks (56 to 65) whereas they are found very poor in the tasks requiring *Higher* ability which ranges from just 31 at the lowest to 39 at the highest. For instances, as many as 22.8 and 15 percent students from Grade 3 could not solve any of the tasks requiring higher ability in Mathematics and Nepali respectively, whereas such figures of students not solving higher ability tasks is seen to be rather big in Grade 5 which is 25 percent in Mathematics, 9 percent in Nepali and 29.7 percent in English. Likewise, having analyzed achievement scores in all levels, average achievement goes lowering down as order of cognitive level are higher up. From this fact it is evident that students are somehow good in the tasks requiring recalling simple facts, recognizing correct answers that are explicit in the texts, comprehending basic information, fundamental thinking, basic description or interpretation of given paragraphs, table, charts in few steps. However, they are seen very poor in demonstrating the ability to solve complex problems, to analyze, deduce logic, generalize, justify an argument or viewpoint, and in the ability to transfer learning from one context to another. Similarly, they are found weaker in solving problems requiring multistep functions or procedures. By types of school, the students in the institutional schools are seen more able to solve problems requiring higher ability in comparison to their peers in community school with a wide variations (for example 27.9-45.7 Nepali in Grade 5). More discouraging is the fact that higher the Grade, the more is the population with lesser higher ability tasks.

- **Influence of caste/ethnicity and home language backgrounds in achievement**

From the data sets, it is evident that language spoken at home and ethnicity/caste–the cultural background of the students, among others, has been one of the influential factors in determining student achievement. While lots of efforts have been put to reduce disparity in achievement, differences are still noticed in their achievement by ethnic/caste groups. For instances, Brahamans always are seen highest performing groups (61-54 in Nepali, 50-56 in Mathematics; and 59 in English) whereas Dalits and Madhesis are found performing the lowest which is just 47-44 in Nepali, 41-43 in Mathematics and 40 in English for Dalits, and 45 and 41 in Nepali and 41 in English for Madhesis. In Mathematics, Madhesis are seen comparatively better than the Minorities but lower than the group "Others". The Chhetris and Janajatis are above the Dalits and Madhesis but more or less at the similar level to the groups "Others". A positive sign from equity point of view is that the Dalit students have performed remarkably better than the national mean (48%) in the Western Mountain (62%) and Eastern (51%) and Western Tarai (49) as well. However, the results are much lower than the average in other regions and very poor in Far-Western and Mid-Western Tarai area (30%). The number of students in certain strata is small, and hence it may be wise not to make the too strong implications of the results. However, in the Kathmandu Valley Dalits performed high (50%).

While comparing the achievements of different language groups, a notable difference is found between the Nepali and other language groups mentioned as "Non-Nepali". In English, it is found that the Newari speaking have performing higher (63.9%) than the students from Nepali speaking community (51.0%). On the other hand, the students from Rai (34.8%), Magar (38.7%), "Others" (39.0%), Gurung (39.1%), Tharu (40.7%) and Tamang (44.4%) speaking performed much lower in comparison to the students from Nepali speaking communities. However, the results are different in other language groups as their performance is lower than that of students having Nepali as home language. The disparity is found widest (6.9) in Madhesi community, the lowest performance level also is found in the girls of the same community (38.3). The girls' performance is found better in Chhetri, Janajati, Minority and others communities, whereas the boys' performance is found better in Brahman and Dalit communities. The girls' performance is found better than the boys in Valley and Eastern region and the boys' performance is seen better in other remaining regions. The girls' performance also is found better than the boys in institutional schools (3.8%), whereas boys' performance is found slightly better in community schools.

- **Association of low socio-economic status with remarkably lower learning achievement**

Students learning level in Nepal is found dependent largely on the various components of socio-economic status of parents such as parent's education, occupation, home accessories and possessions, available for children's study at home. The difference in achievement due to the high and low SES is seen wide ranging from 27 percent (Grade 5) to 32 percent (Grade 3) in Nepali, 27 percent in Grade 5 Mathematics and 53 percent in English– the widest of all, for instances.

Datasets show that low achievement is persistent particularly among those children whose parents are illiterate. For instances, the students whose parents are illiterate achieve very low 42-43 in Grade 3 Nepali and 46-47 in Grade 5 Mathematics in comparison to above than 52 for those whose parents are literate. In English, the children of illiterate parents achieve 40, which go up to 62 for those whose parents are literate. From the data it is also evident that students' achievement increases along with rise in parents' educational level. It is also seen that mothers literacy and education level have contributed more than the father to raise achievement level.

Similar to the parents' education, availability of *home possession* and *accessories* are also correlated with achievement level. In Mathematics, the children with 0 to 3 home possessions out of the 8, the achievement level is lower than the national average and those with more than 3 home possessions the achievement is higher than the national average. The average achievement score of students from the families with 7 possessions is the highest (54- 55). The same pattern shows with home accessories too. Similarly, the results are very poor (36.5-39) when there is none or less than 3, whereas it is remarkably high (51.9-56.8) when there are all of the accessories. In Nepali, the achievement is < 37 to < 42 when none or two of the accessories are available whereas it is > 55 to > 51 when more than two are available and it rises further up to > 51-60 if all of them available. Same trend is noticed in English too. For instance, the achievement is 35.4 for the children with no or less than two accessories whereas it 69.6 if all of them are available.

Parents occupation is seen another detrimental factor for student achievement in Nepal. The achievement of the children is low (43-49) whose parents either mother or father works abroad, it is further low (41-47) for the children whose parents works for agriculture, daily wages or unspecified jobs in comparison to the children whose parents are either teacher or business persons or hold government job (52-60) 5 to 7 points advance than the agriculture related occupation or daily wages.

- **Effect of homework and support for study in higher achievement**

Homework is considered an integral part of teaching and learning for boosting up students achievement if assigned purposefully and feedback provided, which can be used as a means of assessment tools, exercise or drill. Based on the dataset, it is seen that the students receiving homework regularly achieve 47-54 and receiving sometimes which is higher (31-38) than those not receiving it. The difference is statistically significant ($p < 0.001$). Though the number of students not receiving homework is seen very small, and hence the effect size is also small (η^2=0.02 and 0.02)) which explains only 0.9 percent variance.

In order to raise the achievement, support for children for study in any form also is seen essential as the students receiving support from parents, their older siblings or teacher are found achieving higher score. In Nepali, the support from mother or from older siblings is seen beneficial (56) whereas teachers support is seen effective in Mathematics to achieve high score (49-53). In English, those achieve high (50-53) who receive support from their mother or older siblings. However, the children studying independently also are seen achieving as high as 49-54 above than the children receiving support from others. These children achieving better than others without support may be talented. As the ability to afford for support at home reflects better socio-economic condition of the family, the children from such family always achieve high.

- **Effects of availability of textbook in students' achievement**

In the context of Nepal, textbook has remained only the sole means for students to study and learning due to the limited availability of other reference materials for study. Besides, it has also been a reliable means for teachers as well to conduct teaching at classroom since entire teaching is dependent in textbooks. So, having or not having it in the hands of students and teachers is a matter of great importance to achieve better in schools. Being cognizant of its importance in teaching and learning, government has also been putting its concerted efforts to ensure timely availability for all students. However, the dataset reveals the fact that still it is not available for 3.7 to 4.4 students even by the end of academic session which has resulted in varied level of achievement between having it or not. For instance, the achievement in Nepali rises up 48 for the students having textbooks, which lowers down to 40 for those not having textbooks with a difference by 8 percent points. In English, the difference is found at 4 percent points (48 for having and 44 for not having textbooks). The great variance by 10 percent points is found in Mathematics in which students having textbooks achieve as much as 50 against 40 of the students lacking textbooks.

- **Effects of attending schooling at proper ages**

Although proper ages of children to study in Grade 3 and 5 are 8 and 10 respectively, the ages of students attending these Grades vary widely from below 7 years to above 13 and even 14 years above. As the ages vary, the achievement also vary from 40.7 (Grade 3 in Math) for the 7 years old children and to 42.9 for those who are 13 and above years-the lowest, whereas it is 44.7 to 46.8 for those who belong to the age group from 8 to 11. The students who are at the age 10 achieve the highest (46.8) followed by 9 (46.2) and 11 years (46.1). Similarly, in Grade 5 too, among the lowest achiever are 9 years (42.9) and above 14 years old whose achievement is 41.9 whereas higher achiever are the students belonging to the age groups from 10 to 13 whose achievements range from 44.5(for 10 years old) to 49 .2 (for 11 years old) followed by 48.6 (for 12 years old) and 46 (for 13 years old) in Nepali. The same is the trend in other subjects and Grade as well that is lower than average achievement for under and higher aged children and better achievement for proper aged children. The students studying at their overage means that either they started their schooling at later age or repeated the earlier or the same Grades many times. However, it is good that these over aged children (35.18 percent in Grade 3 and 10.05 percent in Grade 5) are retained at school and hence need additional support in their study. Given context implies that attending school at proper age leads children towards a higher possibility to achieve better.

- **Effect of utilizing beyond school hour time in achievement**

Achievement level of students varies widely depending upon the amount of time they spent beyond school hours for household chores, paid work and other activities like watching TV, playing games and chatting with friends. Dataset shows that the achievement score in Nepali rises to 57.8 and 57.4 (Grade 3), 51.0 and 51.8 (Grade 5), 53.8 and 54.4 (English) for those spending less than an hour or 1-2 hours a day in watching TV which lowers down to 53 and 42 (Grade 3), from 45.9 and 40.1 (Grade 5) and 49.5 and 39.5 (in English) for those who spend much time more than four hours a day in it. It is interesting that the achievement is remarkably lower (48.8 for Grade 3 Nepali, 44.6 for Grade 5 Nepali and 40.4 for English) for those who spend no time in watching TV. Similar trend is noticed in Mathematics too in both Grades.

Similarly, the achievement is lower (49 in Grade 3, 46 in Grade 5 Nepali and 40.4 in English) than the national average for the students who do not work at all at home in comparison to (57-53, Grade 3; 48- 50.6 Grade 5 in Nepali; 49- 54.4 in English) those who works for 1 to 2 or more hours a day for household chores. It is notable that as the amount of time for household chores increases to more than four hours, the achievement is seen to be falling down to 47.1 (Grade 3) and to 44.2 (Grade 5) in Nepali 39 .4 in English. From the given facts, one can draw that involvement of children in decent level of household chores for a minimum amount of time (1-2 hours a day) serves a part of socialization process that does not affect negatively in students learning.

Though child labor is prohibited by laws in Nepal, the children studying at primary Grades are reported to have been working for paid job of any form. It is reported that 11 to 13.1 percent students studying in Grade 3 and 12.7 to 16.5 students studying in Grade 5 are involved in paid job of any kinds working for more than an hours a day. Data shows that the students not involved in paid job achieve (49.3-53.1 in Math, 51.0-57.3 in Nepali) above the national average, which is higher than (41.5 in Mathematics, 48-48 in Nepali) for the students being involved in paid job of any kind even less than an hour. Those working for more than four hours a day achieve the lowest (23-40.1) of all. Working for paid job or involvement of them for more than four hours a day indicates poverty of the family requiring the children to work for earning for pocket money or for livelihood at a minimum subsistence level. So, apart from banning the child labor by laws, something more is also required to prevent children from working paid job to address this structural problem persisting within Nepalese socio-economy.

- **Stagnant learning achievement over the years**

To compare the average achievement in each subject for 2015 with that of previous assessment of 2012, it is seen to be stagnant over the years in all subjects. For instances, the average achievement score in Nepali was 63 (Grade 3) and 60 (Grade 5) in 2012, whereas it has fallen down to 52 (Grade 3) and 46 (Grade 5) by more than 11 to 14 in this assessment. In Mathemaics, it has fallen down by 5 to 15 percent points remaining at 45 (Grade 3) and 48 (Grade 5) in 2015 against 60 (Grade 3) and 53 (Grade 5) in comparison with 2012. The same is the trend in English too which has fallen down by 7 percent point from 54 percent in the earlier to 47 percent now.

To look at the content wise results of each subject, the achievement level in *Reading* has fallen down by 13 (Grade 3), 17 (Grade 5) in Nepali and 4 percent points in English coming down to 51, 39 and 46 in this assessment from 64, 56 and 50 in the earlier. Similarly, in *Writing*, it has fallen down to 42 (Grade 3), 33 (Grade 5) in Nepali and 40 in English from the earlier 54, 58 and 49 with difference by 12, 25 and 9 respectively. Likewise, in *Grammar* too, the same level of decrease is evident in both Grades and subjects which has come down to 58, 38 in Nepali and to 53 in English from 65, 64 in Nepali and 56.5 in English by 7.27 percent points in Nepali and 3.5 percent points in English. Exceptions are the 3rd Graders in *Vocabulary* (Nepali) who have improved their achievement by 16 percent points achieving 76 in this assessment against 60 in the earlier which has lowered among 5th Graders from 70 (Nepali) and 62 in 2012 below to the 52 and 57.8 in this assessment. With regards to Mathematics, the achievement level does not rise up to the level of earlier assessment, for example the earlier average score reaching to 61 (Grade 3) and 54 (Grade 5) in *Arithmetic* has come down to 45 and 51 respectively in this assessment. Likewise, it has fallen down to 51 and 49 in this assessment from 60 in the earlier. In *Algebra* the 3rd Graders have raised their average score (by 17 percent points) to 57 in this assessment from

40 in the earlier one whereas it has fallen down by 3 percent points from 49 below to 46 for 5th Graders in this assessment.

Achievement level is seen to be stagnant also in each level of cognitive domains. In Nepali, it has lowered down by 4 to 9 percent points for 3rd Graders and by 6 to 14 percent points for the 5th Graders Compared to the earlier assessment. In comparison to the previous assessment, the highest decline is noticed in *Application* and *Higher* ability by 14 and 11 percent points lower for 5th Graders 5 followed by in *Knowledge* and *Comprehension* by 7 and 9 percent points for 3rd Graders. In Mathematics, it has lowered down by 19 percent points in *Comprehension* and by 17 percent points in *Application* and *Higher* ability for Grade 3 whereas followed by 15 percent points in *Application* for Grade 5. In the rest, the decline is not so big which has lowered down by 3 percent points in *Higher* ability and by 8 and 7 percent points respectively in *Knowledge* and *Comprehension*.

Concluding up

Having analyzed the datasets, first, it is noticed that learning achievements vary a lot by subjects remaining the highest in Nepali and the lowest in Mathematics for Grade 3 with 7 percent point difference. Almost equal level achievement is seen in Grade 5 in all the subjects with difference 1 to 2 percent points. Gender disparity is not seen in both Grades and in all subjects where girls outperform the boys in Mathematics and Nepali remaining equal in English. It is positive sign from the equality point of view.

Regional and District variances are seen to be continuing though the Hill and Eastern region have improved their position from lowest and lower in earlier to the highest in this test after the Valley, which incomparable with other. Reducing such gaps between the lowest and the highest and achieving equality among all is seen challenging for the system. Another challenge is to improve our system to catch up with the international standards.

Chapter 1: Introducing NASA 2015

1.1 Introduction

The literacy and numeracy skills as well as subject specific knowledge and understanding acquired at school level largely determine the young people's prospects on succeeding in further education and prepare them for better adult life. Therefore, personal development of an individual and prosperity of a nation largely depend on knowledge, skills and understanding acquired by citizens from schools, which will enhance their productivity and enable them to adopt better to rapidly changing economy.

In this regard, understanding what and how students are learning at school is one of the important concerns for all parents, teachers and general public – which makes us know how well school education system has equipped the youngsters with knowledge and skills they need to better their lives and to be able to face future challenges. Besides, measurement of students' knowledge and skills is also essential for tracking the development of education and assessing the effectiveness of educational policies and practices (Hanushek and Rivkin, 2010). Measuring learning achievement is important not only for determining the existing level of learning and finding the gaps but also in providing feedback for improving quality of learning.

With these concerns, several initiations have been made worldwide for measuring and monitoring students' learning achievement by means of large-scale assessments during the last decade of 20th century. Nepal is no exception, as it has been measuring students' learning by means of large-scale assessment of student achievement for various Grades since 1995.

This report is the description of methods and process, and the analysis of major results of National Assessment of Student Achievement 2015 (hereafter, NASA 2015) conducted by Education Review Office (ERO) for Grade 3 in Mathematics and Nepali, and for Grade 5 in Mathematics, Nepali and English in 2015 for the students of academic year 2014/2015. This introductory chapter of this report briefly describes the context of the assessment, presents the history of the practice of National Assessment of Student Achievement in Nepal, deals with the assessment process, and points out the characteristics of NASA 2015. Similarly, it includes the objectives of this assessment as well as a brief indication of methods and process adopted in this assessment. Finally, it describes the structure of this report.

1.2 Context of NASA 2015

A number of national assessments of student achievement were accomplished for various Grades before 2011 (see BPEP, 1995; 1997; EDSC, 1997; BPEP 1998; PEDP, 1998; EDSC, 1999; CERID, 1999; EDSC, 2001; EDSC, 2003; CERSOD, 2001; EDSC, 2008; Fulbright, 2008), which were mostly carried out with small samples and were based on classical test theory. Large-scale assessment based on modern test theory, particularly the Item Response Theory (IRT) came into practice when Education Review Office (ERO) adopted it first in assessing the learning achievement of Grade 8 students in 2011. Since then, series of large-scale assessments following the same theory have been accomplished for Grade 3 and 5 in 2012 and for Grade 8 in 2013. Next to those, it is another large-scale assessment for Grade 3 and 5 conducted in 2015.

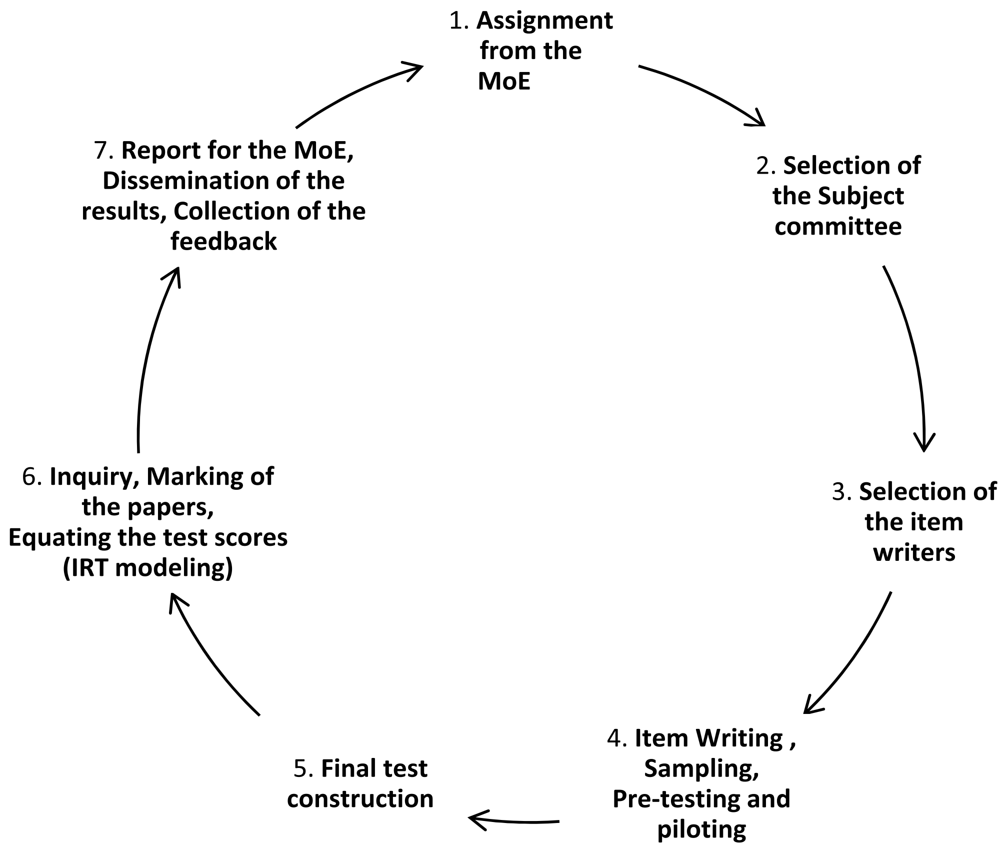
As planned in School Sector Reform Plan (SSRP) 2009-2015, two rounds of assessments for Grade 8 in 2011 and 2013 and one round in Grade 3 and 5 in 2012 have already been completed. This is the second round of assessment for Grade 3 and 5, which was conducted on 12th June 2015. Like the earlier one, the test also aims to assess the achievement in Mathematics and Nepali for 3rd Graders and in Mathematics, Nepali and English for 5th Graders. Though this achievement test was conducted in the year 2015 among students who had just completed their Grade 3 and 5 in the academic year 2014/2015, this report has been entitled as “NASA 2015” to reflect its specific timeframe.

As envisioned in SSRP, the main aim of establishing ERO is to inform regularly the educational stakeholders including the government, teachers, parents, schools, students and civil society about the effectiveness, efficiency, equity and quality of education so that equity and quality will be improved regularly. In order to provide feedback for policy formulation and programme implementation in education system, ERO is entrusted to assess student achievement regularly, carry out the performance audit of educational institutions and schools, and publicize the assessment and audit reports. Given this mandate since its establishment, ERO has been conducting National Assessment of Student Achievement (NASA) for Grade 3, 5 and 8 in various subjects.

1.3 Student Assessment as a Process

Like the previous National Assessments of Student Achievement (NASA) in Nepal, this assessment also includes several nested cycles. The wider and more general cycle revolves seven administrative steps in the assessment process: from the assignment from the MoE to the releasing of the final report to the MoE (fig. 1.1). Inside this general framework, there is another process of preparing the measurement instruments (stages 4 and 5 in figure 1.1) and still another process of analysing and interpreting the results (stages 6 and 7 in figure 1.1).

Figure 1.1 Administrative cycle of NASA 2015



The Steering Committee formed in the MoE ratified the need to continuing student's assessment program for several years. According to the plan, student achievement is assessed in every alternative year at Grade 8 and every another year at Grade 3 and 5 at least until 2016. Accordingly, NASA 2012 was the first large-scale assessment for Grade 3 and 5; and this is the second cycle of assessment for these Grades in Mathematics (Grade 3 and 5), Nepali (Grade 3 and 5) and English (Grade 5).

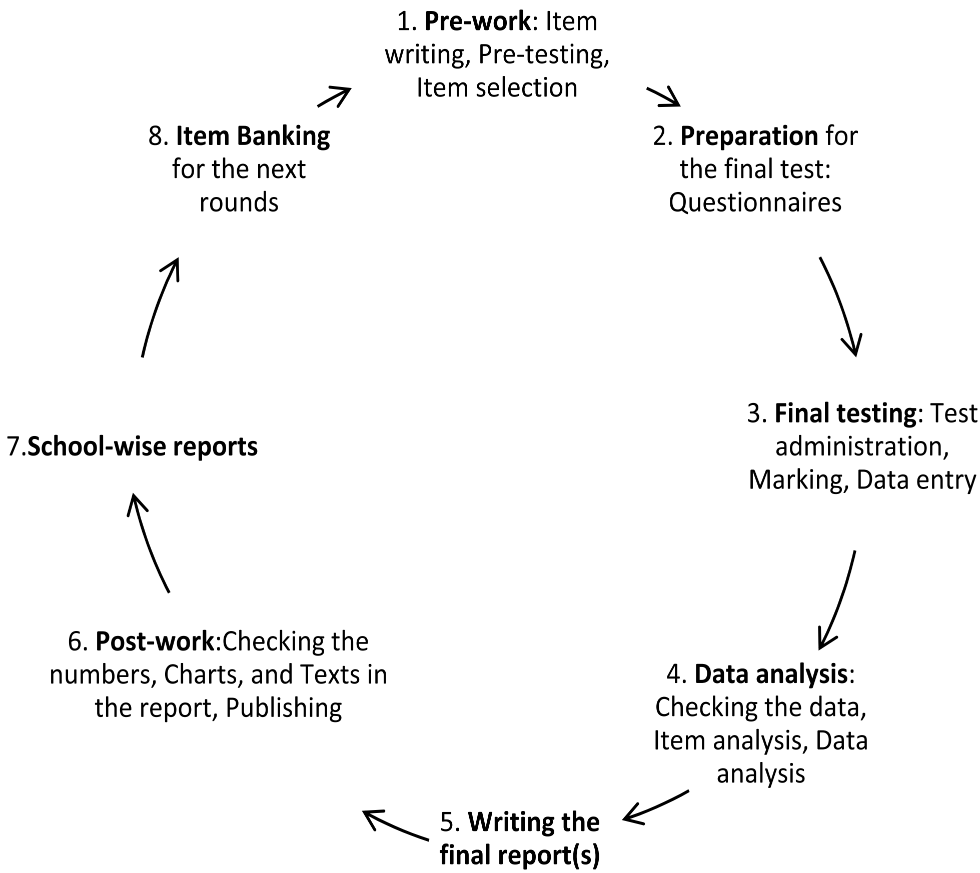
At the end of 2014 and the beginning of 2015, a number of experienced classroom teachers from schools, university, and subject experts from the Kathmandu Valley were assigned the task of test items development. The task was to create a sufficient number of items representing different sub-topics in each content area, representing various difficulty levels and cognitive levels in six versions for the pre-test – and ultimately in three versions for the final test– for each subject and Grade. The items were pre-tested in two layers: firstly, primarily to find out how much the language used in the tests effect the performance of Grade 3 and 5 students; and secondly, testing was done in 153 schools in several Districts to find the stable item parameters for the final test. Besides the newly created items, some of the items from the previous assessment were also used as linking items for making the results comparable to the earlier ones.

The final test versions were compiled by using the following six principles: (1) Contents based on the curriculum (construct validity), (2) Contents' coverage to be as wide as possible (content validity), (3) Proper structure of the levels of cognitive domain (ecological validity), (4) High test discrimination (reliability), (5) Proper difficulty level, and (6) Comparability of the results with the international results (Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS). All the tests showed up with high discriminative power (reliabilities-Cronbach α is higher than 0.90 for all versions of items in all subjects).

The final tests were conducted in both community and institutional schools from 23 Districts representing different ecological zones and development regions, rural and urban locations. The test in each school was administered by the respective DEOs with the assistance of head teachers and teachers of the respective schools. Marking of students' works and data entry were done centrally by outsourcing the task to a professional and trusted consulting company. Data tabulation and analysis – including the equating of the test scores by using the IRT modelling – was done in-house by the technical staff of ERO. The first draft of this report was completed in April 2016. The first results were disseminated to the MOE personnel prior to finalizing, wider sharing and publishing the report.

From methodological point of view, the assessment process comprises eight phases: 1) Pre-work phase, 2) Preparation phase, 3) Final testing phase, 4) Data analysis phase, 5) Report writing phase, 6) Post-work phase, 7) School-wise reporting phase, and 8) Item banking phase (see, fig. 1.2).

Figure 1.2 Phases of the assessment process of NASA



From methodological point of view, the Pre-work phase includes item writing, reviewing, pre-testing, item analysis and selection for the final tests. At the Preparation phase, the background questionnaires were prepared, relevant stakeholders like DEOs and schools were informed about the assessment day, date, processes in detail organizing a two-day orientation program at the center and one day program at each District. At the Final testing phase, the tests were administered at the same time in different parts of the country on the same day in a uniform way. The Data analysis phase started with the screening of the datasets to find possible errors in the data. At this phase, the final item analysis was done by omitting a few items with poor response rate. Finally, the data analysis was carried out by ERO using One Parameter Logistic Model (OPLM) together with SPSS software. During the data analysis, several discussions and inputs were received from relevant external agencies and persons. The Report writing phase took several weeks, seeking comments in several rounds. Before the report was made ready for printing, at the Post-work phase, it was deemed relevant to double check the numbers and interpretations of the results. At the School reporting phase, the schools brief reports of the assessment will be prepared for schools and other stakeholders. At the final phase – before next assessment round – the items and their parameters will be banked for the later use.

1.4 Characteristics of NASA 2015

In the report of NASA 2013, ERO (2015a) has discussed the general characteristics of NASA – especially in connection with NASA 2011. Similar to the earlier ones, this assessment also shares most of those features, such as large-scale and wider coverage, use of Item Response Theory (IRT), comparison with previous studies, international flavor, item banking, item-wise analysis and shared approach. However, there are some elements worth noting, which differ from earlier NASA process. One is that NASA 2015 was administered and led entirely without international support from the very beginning of item writing to data analysis and report producing. Hence, the process itself showed that the in-house team of the ERO is capable of doing such demanding tasks on its own.

1.5 Objectives of the Study

The aim of a large-scale national level student assessment is to produce objective, accurate, and comparative information on achievement of students in schools. However, assessment is not limited only to assessing students' achievement, but also serves as a tool to evaluate the education system to support in obtaining results as nuanced in the curricula. Practically, NASA 2015 is motivated to find out whether and to what extent the students in Grade 3 and 5 have reached the goals set in the national curricula. Specifically, the objectives of this assessment were to:

1. Determine the current national level of achievement of Grade 3 students in Mathematics and Nepali, and of Grade 5 students in Mathematics, Nepali, and English;
2. Determine variations in student achievement between and among different Ecological Zones, Development regions, Districts, school location (rural/urban), school type (community/institutional schools), caste/ethnicity and language groups and gender;
3. Examine the extent to which home background, socio-economic status of family and other student-related factors influence learning achievement;
4. Compare student learning achievement in the current study with that of the previous NASA studies;
5. Compare student learning achievement in Nepal with that shown by international studies: PIRLS (Reading) and TIMSS (Mathematics).
6. Generate evidence-based data for monitoring the trends in students' achievement for these subjects over the period.
7. Draw implications for policy making to improve educational quality and for concerned stakeholders to raise the standards of student learning at school level.

1.6 Method and Process Used for Assessment

This assessment was carried out using quantitative method based on a large-scale survey of student assessment using standardized tools – that is, a set of questions in each subject based on curriculum of the subject approved by the government of Nepal. Set of questions were prepared by a group of qualified subject teachers and subject specialists, pre-tested in some schools and revised based on pre-test results. During the process of item writing and selection, the major performance standards expected by the curriculum were analysed to ensure content validity of the set of questions. While revising the question set based on pre-test, each item was analysed by determining the difficulty level and the difficulty levels of the tests were set around 50 to 60%.

Item Response Theory (IRT) modelling was used from the beginning of item construction to the preparation of marking scheme for data analysis, which also helped to compare NASA 2015 results with some international assessments like TIMSS and PIRLS and with earlier assessments. Besides, a set of background questionnaires was also used among the students in order to identify the variables that influence the achievement of the students. Questionnaires were also used for teachers and head teachers regarding classroom and school management.

In this assessment, out of the total school population, 1542 schools for Grade 3, and 1543 schools for Grade 5 were taken as samples from 23 sample Districts with a small variation of number of schools in each subject. The total number of students in the sample of Grade 3 and 5 was 87,824, almost equally distributed in each subject. Sampling strategy used was, therefore, the random stratified method with random selection from different strata. The following are the strata considered while selecting the samples:

1. Ecological Zones (Mountain, Hill, Tarai, and Kathmandu Valley);
2. Development Regions (Eastern, Central, Western, Mid-Western, Far-Western, and the Kathmandu Valley);
3. Districts (75 altogether);
4. School type (Community and Institutional); and
5. School location (Rural and Urban).

However, some students in the sample schools were absent during the administration of test and the total number of students appeared in the test was 80,881. Again, due to the inadequate information, some of the students' answer sheet and students' background information questionnaires were not included in the analysis for this study. Therefore the total number of students included in the analysis was 73,878, which was the total number of sample students for this study, covering 33,863 students for Grade 3 and 40,015 students for Grade 5.

Each test paper was marked and scores in all item were tabulated in computer using OMR sheet and analysed using statistical methods, descriptive as well as inferential statistics as appropriate. As the descriptive statistics, univariate analysis was carried out including distribution of scores in various categories by calculating percentages and frequencies.

Calculation of mean score and dispersion of data by calculating standard deviation were carried out for the analysis. Similarly, Pearson's product moment correlation coefficients were calculated to correlate various results, and mean achievements were compared using t-test as well as inferential statistics like p-value and effect sizes.

1.7 Organization of Report

This report is organised in eight chapters. Chapter 1 introduces NASA 2015. This chapter mainly consists of a brief background, characteristics and objectives of NASA 2015. The second chapter presents the description of the methodology applied in NASA 2015. This chapter gives details on the methods that were applied in NASA 2015 including sample selection and determination of size of the sample, item writing and selection procedures, reliability and validity, test administration, marking of answer papers and data entry, analysis of the results, and statistical tools used. The analysis of achievement results of each subject are detailed out from the third to seven chapters. Chapters 3 and 4 present the analysis of results of Mathematics of Grade 3 and 5 respectively. Chapters 5 and 6 include the analysis of results of Nepali subject for Grade 3 and 5 respectively. Chapter 7 includes the analysis of results of English subject of Grade 5. Each chapter of this report describes basic results, diversity factors and achievement, and selected explanatory factors and achievement. The final chapter concludes the report by summarising entire processes of NASA 2015, major findings, implications and conclusions of the study.

Chapter 2: Methodology

2.1 Introduction

The National Assessment of Student Achievement 2015 (hereafter, NASA 2015) is the second round of national assessment for Grade 3 and 5 students conducted by Education Review Office (ERO) in the year 2015. This is the second large-scale national level assessment administered in Mathematics and Nepali of Grade 3 and Mathematics, Nepali and English of Grade 5. The students who participated in this assessment were from the academic year 2014/2015 and the assessment was conducted at the end of this academic year.

Some of the tasks of the assessment were outsourced and the rests were carried out by ERO. ERO carried out the tasks of tool development such as test items and questionnaires development, pre-testing and finalisation of test items, sampling and data analysis, whereas test administration, marking of answer papers and data entry were outsourced to a consulting company. The assessment was conducted in 23 sample Districts that were selected using stratified random sampling method covering each of the defined strata altogether in 1542 schools for Grade 3 and 1543 schools for Grade 5. The total number of schools that participated in assessment was 1704. Standardized tests were administered for the students at each sample school. Calibration of test items and equating of scores were done using Item Response Theory (IRT). The results of the assessment are presented in this report by analysing students' score using various statistical tools.

The aim of this chapter is to describe the methodology followed in NASA 2015. The methodology includes–sampling, item construction, background questionnaire preparation, test administration, scoring and data entry, equating of the test scores over three versions used in final testing, and the statistical methods used in analysis. This chapter is structured into seven sections. The first section presents the overall methodological framework; the second section presents sampling method and process. The third and fourth sections respectively deal with the process of item writing and test construction, and the process of developing background questionnaires. Similarly, the fifth section describes the variables used in the analysis; the sixth section presents the process of equating the test scores over three versions used in the final testing. The final section describes the statistical methods used to analyse the data.

2.2 Methodological Framework

The following points provide an overview of the overall methodological framework for NASA 2015.

- This is a sample-based assessment for Grade 3 in Mathematics and Nepali subjects, and for Grade 5 in Mathematics, Nepali and English subjects.
- Ecological and Regional representations were ensured using stratified random sampling method while selecting the samples.
- Test items were based on approved curricula of the government of Nepal for respective Grades.

- Test items, standardised through pre-testing and thorough analysis of difficulty level, were used.
- Using background information questionnaire to students, teachers and head teachers and this assessment has identified several factors affecting student achievement.
- In order to maximise the representation of contents of the curriculum and the reliability in test administration, three versions of test were used by equating the scores of those three versions.
- Some international items from TIMSS (in Science and Mathematics) and PIRLS (in Nepali and English) were calibrated and used to make results comparable to international assessment results.
- Data recording and entry were done using OMR sheet, and results were analysed using OPLM model.

The subsequent sections describe methodological process in detail.

2.3 Population and Sampling

The population for this assessment was the total students studying at Grade 3 and Grade 5 in the academic year 2014/2015 in Nepal. Population of school for this assessment was the total number of schools in Nepal with Grades 1 to 5 in the same academic year, which were 34,335 in numbers. To represent various strata of population, sample Districts for the assessment were selected using stratified random sampling method. Similarly, to represent institutional and community schools the stratified random sampling method was also used to select the schools from each District. Sample schools were selected randomly from each stratum. All students from each sample school participated in the assessment. This section describes the various strata identified for sampling, sample size and the process of selecting sample schools.

Population Strata

School was taken as the basic unit for sampling. Sample schools were selected in such a way that they should represent the country as widely as possible, and the selected students should represent the whole student population as widely as possible. The sampling strategy applied for this assessment was proportional stratified sampling with random selection. For sampling purpose, the following strata were considered:

- Ecological Zones (Mountain, Hill, Tarai, and the Kathmandu Valley);
- Development Regions (Eastern, Central, Western, Mid-Western, Far-Western, and the Kathmandu Valley);
- Districts (75 altogether);
- School type (Community and Institutional);
- School location (Rural and Urban).

Sample Districts were selected from each Ecological Zone and Development Region including the Kathmandu Valley. Although Nepal is categorised into three Ecological Zones, namely, Mountain, Hill and Tarai, for the purpose of this assessment Kathmandu Valley was also considered as a separate Zone. Therefore, Mountain, Hill, Tarai, and the Kathmandu Valley are taken as the four Ecological Zones. Similarly, Nepal is divided into five Development regions; however, for the purpose of this assessment, Kathmandu Valley was also considered a separate Development Region and therefore, Eastern, Central, Western, Mid-Western, Far-Western, and the Kathmandu Valley are taken as the sixth Development Region in this assessment. The Kathmandu Valley was taken as a separate stratum, as it is the most densely populated area in the country with more opportunities than other areas. Not only from the population point of view, but also from the mixed ethnicities, distinct weather condition, wide economic and development activities as well as the dense human capacity make the Valley a unique fourth geographical Zone as well as sixth Development Region in the analysis. Hence, there were 16 basic strata while selecting the sample Districts.

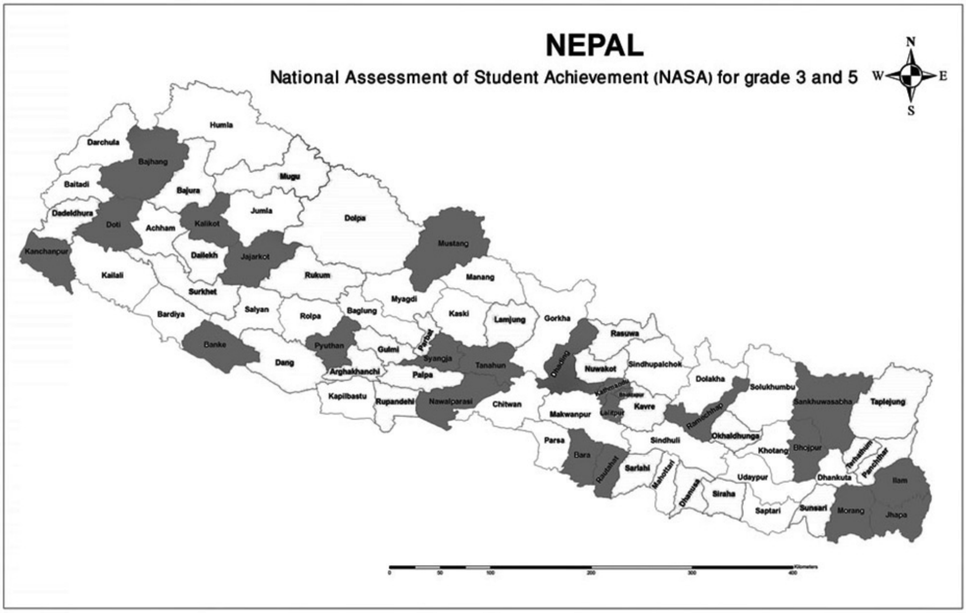
Sample Size

While selecting sample Districts, all the Districts were grouped into three categories each consisting 25 Districts – covering each of the above strata (three Ecological Zones and five Development Regions). The selection of Districts from each group was done randomly within the stratum. A group of 25 Districts together including the three Districts from Kathmandu Valley were taken as the sample Districts for this assessment. Number of schools in each District was fixed proportionally and then the names of the sample schools in each District were selected randomly provided that schools are selected proportionally from the two school types (community and institutional).

Due to the devastating earthquake of April 25, 2015, ERO postponed assessment date, which was already fixed for the first week of May 2015, and it was re-scheduled for June 12, 2015. However, there was still difficulty to conduct assessment in two sample Districts Gorkha and Ramechhap, which were most affected Districts from the earthquake. Therefore, these two Districts were excluded from the sample Districts and the rest 23 Districts were included in the sample for this assessment, provided that total number of sample students was the same as the number of students initially planned from left out two Districts was added to the rest 23 Districts. For this, it was planned to take all students of most of the schools as the sample. While doing so, some students in the sample schools were absent during the administration of test and the total number of students appeared in the test was 80,881. In addition, some of the students' answer sheet and students' background information questionnaires were not included in the analysis in this study because of inadequate information. Therefore, the total number of students' included in the analysis was 73,878.

The 23 sample Districts represent all strata including three Ecological Zones, five Development Regions and the Kathmandu Valley. Figure 2.1 shows the location of 23 sample Districts including Ecological Zones and Development Regions.

Figure 2.1 Sample Districts for NASA 2015



The number of schools and the students in each sample District for Grade 3 and 5 are presented in tables 2.1 and 2.2 respectively. Selection of 23 Districts among 75 Districts with a representative number of schools from each stratum can reasonably have a good coverage of all the Development Regions as well as Ecological Zones.

In the second phase of sampling, number of schools in each of the 23 Districts was determined proportionally based on the number of schools in each stratum. Based on the latest official list of schools, the number of schools running 3rd and 5th Grades was 34,335 (DOE, 2014). The number of sample schools per Grade was estimated to be 1650. Therefore, the number of schools per subject was 550 for Grade 5 and 825 for Grade 3. In each District, the numbers of community and institutional school were determined proportionally. However, the schools were not selected at this phase, but only the number of schools in each District was determined. Similarly, the number of students in each subject was estimated to be 17,500.

Selection of Schools and Students

The number of schools and students participated in this assessment for Grade 3 in Mathematics and Nepali, and for Grade 5 in Mathematics, Nepali and English subjects are presented in the tables 2.1 and 2.2.

Table 2.1 Number of schools and students included in the sample Districts for Grade 3

District	Development Region	Ecological Zone	N. of Schools		N. of the Students	
			Math	Nepali	Maths	Nepali
Illam	Eastern	Hill	33	32	543	478
Jhapa	Eastern	Tarai	44	44	1040	1083
Morang	Eastern	Tarai	46	47	1064	967
Sankhuwasabha	Eastern	Mountain	28	27	476	546
Bhojpur	Eastern	Hill	26	27	300	490
Ramechhap	Central	Hill	32	32	487	538
Lalitpur	Valley	Valley	32	31	964	630
Bhaktapur	Valley	Valley	23	23	496	466
Kathmandu	Valley	Valley	88	88	1864	1685
Dhading	Central	Hill	41	42	892	814
Rautahat	Central	Tarai	31	31	661	948
Bara	Central	Tarai	33	32	1133	1035
Tanahu	Western	Hill	44	44	965	903
Syangja	Western	Hill	41	41	738	761
Mustang	Western	Mountain	5	5	31	29
Nawalparasi	Western	Tarai	45	45	939	1019
Pyuthan	Mid-Western	Hill	25	25	696	733
Banke	Mid-Western	Tarai	29	29	847	816
Jajarkot	Mid-Western	Hill	27	28	703	506
Kalikot	Mid-Western	Mountain	19	19	481	426
Bajhang	Far-Western	Mountain	30	30	636	692
Doti	Far-Western	Hill	26	26	549	610
Kanchanpur	Far-Western	Tarai	23	23	569	614
Total			**771**	**771**	**17074**	**16789**

Table 2.2 Number of schools and students included in sample Districts for Grade 5

District	Development Region	Ecological Zone	Number of Schools			Number of Students		
			Maths	Nepali	English	Maths	Nepali	English
Illam	Eastern	Hill	22	22	21	429	444	435
Jhapa	Eastern	Tarai	29	29	30	724	752	740
Morang	Eastern	Tarai	31	31	31	788	790	820
Sankhuwasaa	Eastern	Mountain	19	18	18	469	409	381
Bhojpur	Eastern	Hill	17	18	18	432	492	454
Ramechhap	Central	Hill	22	21	21	576	493	448
Lalitpur	Valley	Valley	21	21	21	491	440	460
Bhaktapur	Valley	Valley	15	16	15	245	446	273
Kathmandu	Valley	Valley	59	59	58	1494	1172	1393
Dhading	Central	Hill	28	27	28	777	718	671
Rautahat	Central	Tarai	20	21	21	661	611	581
Bara	Central	Tarai	22	22	21	788	905	804
Tanahu	Western	Hill	29	29	30	791	802	892
Syangja	Western	Hill	29	27	27	607	680	673

District	Development Region	Ecological Zone	Number of Schools			Number of Students		
			Maths	Nepali	English	Maths	Nepali	English
Mustang	Western	Mountain	3	4	3	29	39	53
Nawalparasi	Western	Tarai	30	30	30	744	725	734
Pyuthan	Mid-Western	Hill	17	16	17	661	591	532
Banke	Mid-Western	Tarai	19	20	19	576	639	674
Jajarkot	Mid-Western	Hill	18	18	19	386	391	461
Kalikot	Mid-Western	Mountain	13	13	12	345	344	270
Bajhang	Far-Western	Mountain	20	20	20	522	517	490
Doti	Far-Western	Hill	17	17	18	418	488	434
Kanchanpur	Far-Western	Tarai	16	15	15	561	467	473
Total			**516**	**514**	**513**	**13514**	**13355**	**13146**

In the third phase of sampling, list of all schools was taken from the Department of Education (DOE) database. The list included information such as the number of students, school type (community or institutional), running Grades and address of each school. The information was used as the basis for random selection of schools in the District. To select school having both Grades 3 and 5; however, schools running only up to Grade 4 were excluded from the sample. Similarly, in some Districts a small number of schools were excluded from sample due to the long distances and lack of transportation. Student samples were taken to make it reasonably comparable with the other Districts to possible extent.

There were some variations in the estimated total number of students when the schools were selected randomly from the list of schools in each District. Selected number of students from the DOE database was altogether 87,824. During the test administration, some students were absent and only 80,881 students were participated in the test. Again, during the tabulation, using OMR and data cleaning some answer papers were found with inadequate information, which were excluded from the analysis.

Finally, the data of 771 schools and 17,074 students in Grade 3 Mathematics, and 771 schools and 16,789 students in Nepali were included in the analysis. Similarly, in Grade 5, the number of schools and students included in the analysis were 516 and 13,514 respectively in Mathematics, 514 and 13,355 in Nepali, and 513 and 13,146 in English. While selecting the students, students with some typical characteristics such as blind and mentally retarded were excluded from the sample.

2.4 Item Preparation

Item preparation includes item writing, pre-testing of items, adaptation of linking items from TIMSS and PIRLS released item banks, selection of linking items from previous NASA item banks and preparing three versions of final test booklets.

Item Writing

A team of item writers including schoolteachers teaching in Grades 3 and 5, curriculum officers, and university teachers was formed to accomplish the task of preparing sufficient number of items for pre-test. Item writing workshop was conducted for each subject in which more than six sets of items were prepared for each subjects. Further process and work done after developing test items through subject workshops were as follows: items were screened, edited and re-written by an expert team of each subject, and then subject committee in each subject worked to prepare six sets of items of each of the subjects for pre-testing. The final pre-test papers for pre-testing were printed in the confidential printing press.

Pre-testing of Item

Six versions of test items prepared for pre-testing were pre-tested in 153 schools of 11 Districts. Total students participated in pilot test were 7,724 in two subjects from Grade 3 (Mathematics and Nepali) and three subjects from Grade 5 (Mathematics, Nepali and English) of which the number of students participated in each subject ranged from 1,514 to 1,565. The pre-test was administered by the personnel in the District Education Offices (DEOs) after organizing an orientation session to the head teachers of respective schools. DEOs monitored the processes in the schools, collected the papers and sent them to ERO in Kathmandu for marking and data entry. Officers from District Education Offices were also oriented to the process of pre-testing in a one-day workshop whereby the objectives of pre-testing as well as NASA were shared. ERO personnel closely monitored the pre-test process. To avoid the leaking of the items, all the papers were counted before and after the process so that no papers were left in schools or at DEO.

Calibration and Adaptation of PIRLS and TIMSS Items

The released items from Trends in International Mathematics and Science Study (TIMSS) in Mathematics and Progress in International Reading Literacy Study (PIRLS) in Nepali and English languages for Grade 5 were calibrated. The items selected for Mathematics and Nepali were translated into the Nepali language and edited by the respective subject committee. Some of the items adapted from TIMSS were used as the linking items for three versions of test papers whereas in the case of Nepali and English, all items adapted from PIRLS were used as linking items. Parameters of all the selected items were fixed based on their released parameters and the other items were calibrated in the same parameters of adapted TIMSS and PIRLS items in respective subjects. Item Response Theory (IRT) modelling was used to calibrate the items.

Selection of Linking Items

Some items were selected from NASA 2012 for Grades 3 and 5 in respective subjects in order to make results comparable. Besides, some additional items were used as the linking items for three versions of each subject in each Grade so that the versions will be made comparable. Three versions of test papers were made comparable using IRT modelling for item calibration as well as equating three versions of tests.

Selection of Items for Final Test

Items for final test were selected from six sets of piloted test items. One of the three sets contained the items that were newly developed and piloted; the second set of test items was from the released items of TIMSS and PIRLS; and the third set of items was the test items used in NASA 2012. The subject committee of each subject selected final items from the above-mentioned three sets of items. While selecting the items for final test Subject Committees followed the following six basic principles of item selection.

(a) *Content dependence of the curriculum*

The basis of the construct and content validity of the final tests lies in the "theoretical framework", that is, in the national curricula. In the national assessment, the main idea is to test how well the objectives expressed in the national curricula are fulfilled. Specification grids for each subject and Grade were prepared based on the curricula. In the grids, the time spent on the tasks in the curricula was operationalized as percentage for each topic and sub-topic. This information was used as a basis for item writing and item selection where the marks on the tests are proportional to those percentages in the grid.

(b) *Content coverage as wide as possible*

To ensure content validity, items were selected from a broad range of topics to cover each as much as possible. However, the tests were not that long to make it possible to cover all sub-topics. A sub-test length of 3 to 4 items may be taken as a minimum length to discriminate the test takers from each other sufficiently. Thus, an attempt was made to include as many sub-topics as possible in the test. The selection was, however, proportional. When there were sub-topics of wider coverage in the curriculum than the others, more items were selected from those areas in test construction. Content coverage was widened by using three different versions in the final testing in each subject of both Grades. Several linking items linked all the versions to each other, and they were linked to some international standards by using linking items from TIMSS and PIRLS released item banks.

(c) *Proper structure of cognitive levels*

Bloom's taxonomy on the cognitive domain was used as the basis of cognitive levels. Bloom's original classification of Knowledge, Comprehension, Application, Analysis, Synthesis and Evaluation was shortened into four categories: Knowledge, Comprehension, Application, and Higher skills (Bloom *et al.* 1956; Metfesser, Michael and Kirsner 1969).

(d) *High discrimination power of the test (Reliability)*

Two technical areas related to high reliability of test were addressed: item discrimination and item difficulty. These main item parameters, that is, item difficulty and item discrimination, classically estimated by using the proportion of correct answers (p) and the item-total correlation (ρ_{gx}), are inter-related so that item discrimination is the highest when the difficulty level is around 0.50. Knowing that the variance of the dichotomous item is strictly related to item difficulty, that is, $\sigma^2 = p(1-p)$, the classical formula of alpha reliability can also be obtained using the following formula:

$$rel = \frac{k}{k-1}\left[1 - \frac{\sum_{i=1}^{k}\sigma_i^2}{\left(\sum_{i=1}^{k}\rho_{it}\sigma_i\right)^2}\right]$$

k = number of items
σ_i^2= variance of the scores on item i
σ_i= standard deviation of the scores on item i
$\rho_{it} = \rho_{gX}$ = item test correlation

It is noteworthy that there were only two sources of information needed for estimating the reliability of the test: the item discrimination (ρ_{it}) and item variance (σ_i^2). It is also noteworthy that the alpha reliability is maximized when the sum of the elements ρ_{it} and σ is the highest. When knowing that the variance is the highest when the proportion of the correct answer is p = 0.50, it makes sense why it is wise to select items with high item discrimination as far as possible and mediocre difficulty level in the test.

Generally, the values for reliability lower than α=0.70 are not taken as accurate enough to be accepted when comparing the scores of different groups with each other. On the other hand, the values higher than α = 0.60 can be accepted for a new instrument (see, Nunnally 1978; De Vellis 1991; Hair *et al.*, 1998).

(e) *Proper difficulty level*

The most balanced test for the national assessment is achieved when the items are selected from the whole range of abilities. Thus, there should be easy, mediocre and demanding items in the test. These kinds of items can discriminate the best and the poorest as well as the mediocre students. One possible solution, as used in NASA 2012, is to select the items so that of the total,

- 10 % items are very easy,
- 20 % items are quite easy,
- 40 % items are of medium difficulty level,
- 20 % items are quite demanding,
- 10 % items are very demanding.

Combining the principles, by selecting the easiest items from the pre-test, the aim was to raise the average difficulty level in the tests near $\overline{p}$ = 0.60, that is, 0.50 + 0.10 (from guessing in the multiple-choice questions) or even higher when possible. This is the way to construct a test, which could discriminate not only among the mediocre students but also among the highest and lowest performing students. The items were selected based on the classical item difficulty parameter p.

(f) *Comparability with the results of NASA 2012 and with the international results*

The sixth principle in item selection was that the results should be comparable with the results of NASA 2012 and with international test results. Better international comparison can be done when the Nepalese results are made comparable to the international assessments TIMSS and PIRLS. For this, there should be linking items in the test from the TIMSS and PIRLS item banks. Based on the released items and their parameters (TIMSS 2007; 2009a; 2009b; Adams and Wu, 2002; PISA, 2006; 2009), and by using IRT modelling, it is possible to find the baseline for the comparison.

The idea is that when one knows the difficulty parameter of the international items, those values are fixed in the datasets in Nepal, and thus the local items are calibrated onto the same scale as the international items. The linking items for international assessments used in NASA 2015 are from the TIMSS Mathematics items released in the year 2003 and 2007 and PIRLS reading comprehension based items released in the year 2011. Both international item banks are from Grade 4 students; the items were supposed to be somehow easy for Grade five students and somehow difficult for Grade three students. In Mathematics, the released parameters of TIMSS of the selected items were fixed at the same time in both Grades (3 and 5). In Nepali, the calibration was done first separately in Grade three and Grade five and later they were re-scaled to make the scores comparable. In both cases the test scores are equated (separately in Mathematics, Nepali and English) and, hence, the scores are made comparable in both the Grades but not over the subjects.

Final Tests

In each Grade, three versions of test items were administered simultaneously in the same classroom in each subject. Some standard procedures and principles were set using IRT modelling for the assessment of Mathematics in order to make the results comparable. The following principles were adapted from IRT modelling practices: First, no decimal number scores are allowed in the marking of answer sheet – i.e., the students' responses are always marked in whole numbers. If the students are not qualified to secure full score, 0.5 or any decimal score is not provided in any case. Second, the marking scheme has been prepared to make exactly the same judgment in each item. Third, marks of each of the items are recorded in the dataset. Finally, a linking procedure is set between the different versions of the test. All the versions were linked with each other using the identical linking items. The common items for each test version, the linking items, were carefully selected from the pre-tested items, TIMSS released items for Mathematics, PIRLS released items for Nepali and English, and from the NASA 2012 item banks. While selecting items, it was considered that the standards set by the national curriculum should be reflected by the test.

Mathematics test

All versions of Mathematics test (M_1, M_2 and M_3) were similar in content characteristics. Construct validity is quite high from curriculum point of view, as the number of items in each subject match well with curricular contents. Different versions of tests include a wider variety of sub-topics under each topic.

Classical item and test analysis methods were used in pre-test phase for finding the percentage of correct answers, that is, item discrimination power. IRT was used for item calibration, finding the latent ability (Theta,θ), and equating the versions M_1, M_2 and M_3 and TIMSS database. SPSS software was used for the classical analysis and One Parametric Logistic Model software (OPLM) (Verhelst, Glas, Verstralen, 1995) was used for IRT modelling. The original output is the latent ability (θ) which is a standardized normal score ranging usually from –3 to + 3. The parameters of the international items were fixed during the item calibration so that all the test items of the year 2015 were calibrated in the international TIMSS scale. After the calibration of the items, all the scores in the versions M_1, M_2 and M_3 were transformed into the same scale, that is, the scores were equated. The equating process of scores of different versions made three versions comparable to each other. The value of each test version was later transformed into equated scores and further the equated scores were converted into percentage. Table 2.3 shows the average original marks as well as equated score for Grade 3 in Mathematics calculated for three versions M_1, M_2 and M_3 separately. The number of linking items is high for obtaining stable calibration between the versions and sufficient for calibrating the scales over the Grades.

Table 2.3 Comparison of the characteristics of Mathematics test versions of Grade 3

Version	N	Max. Marks	Original score	Equated Mean score*	SD[1]	CV[2]
M_1	5635	53.0	26.0	47.05	23.33	49.59
M_2	5702	53,0	25.3	44.94	22.72	50.57
M_3	5737	53.0	23.1	41.96	21.60	51.48
Total	**17,074**	**53.0**		**44.63**		

1) SD = Standard Deviation 2) CV = Coefficient of Variation = SD/Mean*100

Table 2.3 shows that the maximum mark obtained by the students in each version is equal to 53 where the equated mean ranges from 41.96 to 47.05.

Table 2.4 shows the average marks together with equated score of Grade 5 calculated for three versions, M_1, M_2, and M_3, separately. As all the versions of test were linked with each other by several identical linking items, the equated mean score in all three versions vary with a small range from 47.10 to 49.54 percent.

Table 2.4 Comparison of the characteristics of Mathematics test versions in Grade 5

Version	N	Max. Marks	Original Score	Equated Mean Score*	SD	CV
M_1	4498	55	27.2	49.54	21.18	42.75
M_2	4590	55	25.4	47.10	20.48	43.49
M_3	4426	55	26.5	48.25	21.71	45.00
Total	**13,514**	**55**		**48.29**		

In Grade 3 Mathematics, five common items were used in all versions as linking items, which were borrowed from TIMSS released item banks; and other seven items were used as the linking items for any two versions among three versions. Among 12 linking items, six were from 2012 item banks. In Grade 5 Mathematics, six common items were used in all versions as linking items, which were borrowed from TIMSS released item banks; and other 10 items were used as the linking items for any two versions among three versions. Among 16 linking items, six were from NASA 2012 item banks.

As in the curricular contents, the test items were classified into four categories: Arithmetic, Algebra, Geometry and Numeracy (see tables 2.5 and 2.6). Arithmetic includes basic operations, time, money and measurement, fraction, decimal, percentage, unitary methods and simple interest, bill and budget, and statistics. Similarly, Algebra includes algebra and sets. Numeracy includes the knowledge of numbers whereas Geometry includes shape, size and their measurement.

Table 2.5 Characteristics of Mathematics tests of Grade 3 in various content areas

Topics	Marks			Percentages			Percentage in curriculum	Reliability		
	M_1	M_2	M_3	M_1	M_2	M_3		M_1	M_2	M_3
Arithmetic	36	36	34	66.7	66.7	61.8	65	0.904	0.902	0.856
Geometry	6	9	4	11.1	16.7	7.3	7	0.417	0.505	0.4
Algebra	5	3	6	9.3	5.6	10.9	9	0.63	0.38	0.75
Numeracy	7	6	11	13.0	11.1	20.0	19	0.689	0.605	0.501
Total	**54**	**54**	**55**	**100**	**100**	**100**	**100**	**0.902**	**0.921**	**0.932**

For Grade 3 in Mathematics, among the three versions of test items M_1 and M_2 versions were of the same length whereas M_3 was longer than M_1 and M_2. As the longer version, M_3 carried a maximum of 55 marks whereas the shorter versions M_1 and M_3 carried 54 marks each. In Grade 3, overall internal consistencies (given by Alpha-reliability) of the whole tests were very high (α = 0.90 to 0.93) in each version; however, some of the categories (Algebra and Geometry) contain a few items, hence reliability is somehow lower. The reliability of the score in the total sample cannot be given in a classical way because it can be estimated only version-wise.

Table 2.6 Characteristics of Mathematics tests of Grade 5 in various content areas

Topics	Marks			Percentages			Percentage in Curriculum	Reliability		
	M_1	M_2	M_3	M_1	M_2	M_3		M_1	M_2	M_3
Arithmetic	28	26	25	51.0	48.0	46.0	58	0.84	0.823	0.826
Geometry	7	10	9	13.0	18.0	16.0	9	0.498	0.633	0.651
Algebra	9	10	8	16.0	18.0	14.0	15	0.706	0.703	0.484
Numeracy	11	9	13	20.0	16.0	24.0	18	0.64	0.616	0.783
Total	**55**	**55**	**55**	**100**	**100**	**100**	**100**	**0.905**	**0.907**	**0.918**

For Grade 5, all the versions of Mathematics test were of equal length carrying a maximum of 55 marks each. In Grade 5, reliabilities of the total scores of the different versions (α) were over 0.90, that is, the test scores could discriminate the individual pupils with high accuracy. Because of the limited number of items, the reliability of Algebra, Geometry and Numeracy in Grade 5 remains relatively low compared to other content areas.

The items used in the tests varied from objective items (that is, the multiple choice items, fill in the blank, true or false, very short answer items) to subjective items, usually productive items (short answer type and long answer type items). The test also met all the cognitive levels of Bloom's taxonomy (see Bloom *et al.* 1956; Metfesser, Michael, and Kirsner, 1969). In Mathematics test only four categories of Bloom's taxonomy are used, i.e., Knowledge, Comprehension, Application and Higher ability.

Table 2.7 Characteristics of various domains in Mathematics tests of Grade 3

Cognitive level	Marks			Percentages			Reliability		
	M_1	M_2	M_3	M_1	M_2	M_3	V_1	V_2	V_3
Knowledge	17	17	13	31.5	31.5	23.6	0.842	0.789	0.726
Comprehension	19	15	22	35.2	27.8	40.0	0.833	0.783	0.764
Application	12	13	16	22.2	24.1	29.1	0.726	0.738	0.708
Higher ability	6	9	4	11.1	16.7	7.3	0.464	0.676	0.371
Total	**54**	**54**	**55**	**100**	**100**	**100**	**0.932**	**0.921**	**0.902**

Table 2.8 Characteristics of various domains in Mathematics tests of Grade 5

Cognitive level	Mark			Percentages			Reliability		
	M_1	M_2	M_3	M_1	M_2	M_3	V_1	V_2	V_3
Knowledge	12	15	16	22.2	27.8	29.1	0.647	0.756	0.71
Comprehension	18	19	14	33.3	35.2	25.5	0.788	0.801	0.803
Application	20	13	20	37.0	24.1	36.4	0.763	0.631	0.766
Higher ability	5	8	5	9.3	14.8	9.1	0.536	0.562	0.625
Total	**55**	**55**	**55**	**100**	**100**	**100**	**0.905**	**0.907**	**0.918**

In both the Grades in Mathematics (table 2.7 and 2.8), overall reliability of each version is high (> 0.90). However, when we see the reliability of various levels of cognitive domain separately, the reliability of items representing Higher ability and Application are low compared to Knowledge and Comprehension in both Grades.

Nepali language tests

As in Mathematics test, classical item and test analysis methods were used in the pre-test phase of Nepali subject to find the percentage of correct answers, item discrimination power, and the test reliability. IRT was used for item calibration, finding the latent ability (Theta, θ) as well as comparing and equating the versions N_1–N_3 and PIRLS database. SPSS software was used for classical analysis, and OPLM (Verhelst, Glas, Verstralen, 1995) was used for IRT modelling. The parameters of the international items were fixed during the item calibration so that all test items of this assessment were calibrated to the international PIRLS scale. After calibration of items, all the scores in versions N_1–N_3 were transformed into the same scale, that is, the scores were equated. This means that all the scores in each test version were made comparable. The original output is the latent ability (θ) which is a standardized Normal score ranging usually from –3 to +3. These values in each test versions were later transformed to equated scores and further the equated scores were converted into percentage.

Table 2.9 shows the mean marks of Grade 3 students calculated for three versions N_1, N_2 and N_3 separately. All the versions were linked with each other by using the identical linking items. There were five items from a PIRLS 2011 released Reading passage as the linking items in all the three versions. Similarly, 18 items were also used from NASA 2012 item banks. In addition to those items from PIRLS and NASA 2012 item banks, 8 items were common for all versions and additionally 19 items were linked in N_1 and N_2; 16 items were linking items of N_1 and N_3 and 5 items were used as the linking items of N_2 and N_3.

Table 2.9 Comparison of the characteristics of Nepali test versions of Grade 3

Version	N	Max. Marks	Original Mean score	Equated Mean score	SD	CV
N1	5644	56.0	28.6	51.1	25.74	50.39
N2	5656	54.0	28.6	51.5	23.39	45.38
N3	5489	55.0	31.1	52.0	24.66	47.43

Table 2.10 shows the average marks of Grade 5 calculated for three versions N_1, N_2 and N_3 separately. All versions were of equal length as each carries 62 marks. All the versions were linked with each other by using the identical linking items. There were 8 items from a Reading passage of PIRLS 2011 released items bank as well as 21 items from NASA 2012 item banks as the linking items over the three versions. In addition, 10 items from PIRLS and NASA 2012 item banks were common for all versions. 18 items were used as the linking items for N_1 and N_2; 22 items were used for N_1 and N_3 and 1 item for N_2 and N_3. In order to obtain stable calibration between the versions, linking items included in each test were in large number.

Table 2.10 Comparison of the characteristics of Ne.pali test versions of Grade 5

Version	N	Max. Marks	Original Score	Equated Mean Score*	SD[1]	CV[2]
N1	4526	62	29.6	47.9	21.72	45.39
N2	4517	59	22.2	45.2	20.67	45.77
N3	4312	61	23.0	45.8	19.97	43.61

As in curricular contents, the test items were classified into four categories: *Reading, Writing, Grammar and Vocabulary*. It is worth noting that *Grammar and Vocabulary* are not mentioned explicitly in the curriculum for Grade 3 though "*Functional Grammar*" is used as a term whereby these contents are included. However, *Listening* and *Speaking* are explicitly mentioned in the curriculum.

Table 2.11 Characteristics of Nepali tests of Grade 3 in various content areas

Topic	Score			Percentage			Weightage of Curriculum	Reliability			
	N_1	N_2	N_3	N_1	N_2	N_3		N_1	N_2	N_3	Weighted Mean of Reliability
Reading	18	18	18	32.1	33.3	32.7	25	0.885	0.888	0.884	0.923
Writing	19	17	18	33.9	31.5	32.7	25	0.760	0.777	0.753	0.886
Grammar	11	11	10	19.6	20.4	18.2	5	0.879	0.784	0.835	0.763
Vocabulary	8	8	9	14.3	14.8	16.4		0.810	0.726	0.801	0.833
Listening							20				
Speaking							25				
Total	**56**	**54**	**55**					**0.934**	**0.913**	**0.922**	

In Grade 3 Nepali, overall internal consistencies (given by reliability) of the whole tests in each version were very high ($\alpha = 0.91$). The reliability of the score in the total sample cannot be found in a classical way because it can be estimated only version-wise.

As in the curricular contents, the test items for Grade 5 were classified into four categories: *Reading, Writing, Grammar,* and *Vocabulary*. It is worth noting that the curriculum for Grade 5 students does not explicitly include *Grammar* and *Vocabulary* though weightage is allotted for the *Functional Grammar* integrating *Vocabulary* in it. Instead, *Listening* and *Speaking* are explicitly mentioned. For Grade 5, overall internal consistencies (given by reliability) of the whole tests in each version were very high ($\alpha > 0.91$). Due to the few items in some of the categories (*Grammar* and *Vocabulary*), however, reliability is somehow lower. The reliability of the score in the total sample cannot be found in a classical way because it can be estimated only version-wise.

Table 2.12 Characteristics of Nepali tests of Grade 5 in various content areas

Topic	Score			Percentage			Weightage of Curriculum	Reliability			
	N_1	N_2	N_3	N_1	N_2	N_3		N_1	N_2	N_3	Weighted Mean of Reliability
Reading	25	25	27	40.3	40.3	43.5	30	0.861	0.830	0.863	0.851
Writing	20	23	22	32.3	37.1	35.5	35	0.642	0.847	0.809	0.765
Grammar	5	10	6	8.1	16.1	9.7	5	0.704	0.792	0.629	0.710
Vocabulary	12	4	7	19.4	6.5	11.3		0.792	0.725	0.740	0.753
Listening							15				
Speaking							15				
Total	**62**	**62**	**62**					**0.915**	**0.908**	**0.906**	**0.910**

In Grade 5 Nepali, overall internal consistencies (given by reliability) of the whole tests in each version were very high ($\alpha = 0.90$ to 0.91). The reliability of the score in the total sample cannot be given in a classical way because it can be estimated only version-wise.

As in cognitive levels of Bloom's taxonomy (see Bloom et al. 1956; Metfesser, Michael, and Kirsner, 1969), the test items were from *Knowledge*, *Comprehension* and *Application* types in Grade 3, but Grade 5 test items represent four cognitive levels: *knowledge, Comprehension, Application* and *Higher Ability* (see tables 2.13 and 2.14).

Table 2.13 Characteristics of Nepali tests for Grade 3 in various cognitive domains

Topic	score			Percentage			Reliability			
	N_1	N_2	N_3	N_1	N_2	N_3	N_1	N_2	N_3	Weighted Mean of Reliability
Total	56	54	55				0.934	0.913	0.922	0.923
Knowledge	17	16	14	30.4	29.6	25.5	0.831	0.784	0.852	0.822
Comprehension	13	12	16	23.2	22.2	29.1	0.858	0.796	0.737	0.798
Application	14	14	19	25.0	25.9	34.5	0.823	0.744	0.840	0.802

Table 2.14 Characteristics of Nepali tests for Grade 5 in various cognitive domains

Topic	Score			Percentage			Reliability			
	N_1	N_2	N_3	N_1	N_2	N_3	N_1	N_2	N_3	Weighted Mean of Reliability
Knowledge	19	11	15	25.8	16.1	14.5	0.834	0.730	0.812	0.792
Comprehension	16	10	9	25.8	45.2	41.9	0.728	0.671	0.709	0.703
Application	16	28	26	17.7	21.0	19.4	0.730	0.812	0.782	0.775
Higher ability	11	13	12	30.6	17.7	24.2	0.553	0.634	0.571	0.586
Total	**62**	**62**	**62**	**30.6**	**17.7**	**24.2**	**0.915**	**0.908**	**0.906**	**0.910**

English test

Table 2.15 shows the average marks calculated for three versions of English test E_1, E_2 and E_3 separately for Grade 5. Based on the pre-tested items, the versions E_1 and E_3 were of equal length, whereas E_2 was longer one. All the versions were linked with each other by using identical linking items. Each of the shorter versions E_1 and E_3 scored a maximum of 61 marks and the maximum score for E_2 was 65. Eight items from a reading passage of PIRLS 2011 released items bank and 11 items from NASA 2012 item banks were used as the linking items over three versions. In addition to these 19 items linking all versions, some common items were used in E_1 and E_2, E_1 and E_3, and E_2 and E_3 as the linking items between the versions.

Table 2.15 Comparison of the characteristics of English test versions

Version	N	Maximum Marks	Original Mean Score	Equated Mean[1]	SD[2]	CV[3]
E_1	4462	61	29.3	29.7	15.51	52.14
E_2	4420	65	30.0	30.0	15.90	52.94
E_3	4264	61	29.9	31.5	16.25	51.60
Total	**13146**		**29.7**	**30.4**	**15.90**	**52.29**

Classical item and test analysis methods were used in the pre-test phase for finding the percentage of correct answers and the item discrimination power. IRT was used for item calibration and finding the latent ability (Theta, θ) as well as for comparing and equating the versions E_1–E_3 and PIRLS database. SPSS software was used for the classical analysis, and OPLM (Verhelst, Glas and Verstralen, 1995) was used for the IRT modelling. The parameters of the international items were fixed during item calibration so that all test items of this assessment were calibrated in PIRLS scale. After the calibration of items, all the scores in the versions E_1–E_3 were transformed into the same scale, equating the scores. This means that all the scores in each test version are comparable. The original output is the latent ability (θ) which is a standardized normal score ranging usually from –3 to +3. These values in each test versions were later transformed to equated scores; and the equated scores were further converted into percentage.

As in the curricular contents, the test items for Grade five English were classified into four categories: *Reading, Writing, Grammar* and *Vocabulary* (table 2.16). It is worth noting that the curriculum for Grade 5 students do not explicitly include Grammar and Vocabulary, though "*Functional Grammar*" is used as a term. Instead, *Listening* and *Speaking* are explicitly mentioned. From the beginning, it was thought that *Speaking* and *Listening* should be taken into account in the assessment. However, organizing these kinds of test items with objective and comparable manner would have been practically impossible in the situation where schools are lacking electricity and necessary equipments for organizing the listening test for all. Hence, they were omitted from the final test. However, *Vocabulary* and *Grammar* were included in the test items.

Table 2.16 Characteristics of English tests of Grade 5 in various content areas

Topic	Score			Percentage			Weightage of Curriculum	Reliability			
	E_1	E_2	E_3	E_1	E_2	E_3		E_1	E_2	E_3	Weighted Mean of Reliability
Reading	28	35	30	45.90%	53.85%	49.18%	25	0.868	0.914	0.917	0.901
Writing	14	14	14	22.95%	21.54%	22.95%	25	0.796	0.808	0.796	0.800
Grammar	11	12	15	18.03%	18.46%	24.59%		0.786	0.795	0.81	0.798
Vocabulary	8	4	2	13.11%	6.15%	3.28%		0.821	0.722	0.652	0.769
Listening							25				
Speaking							25				
Total	**61**	**65**	**61**	**100%**	**100%**	**100%**	**100**	**0.922**	**0.931**	**0.928**	**0.927**

Overall internal consistencies (given by Alpha-reliability) of the whole tests on each version were very high ($\alpha = 0.91$ to 0.93), however, some of the categories (*Grammar* and *Vocabulary*) contain fewer number of items, and hence the reliability is somehow lower.

The items of the test represent various cognitive levels of Bloom's taxonomy (see Bloom et al. 1956; Metfesser, Michael and Kirsner, 1969). The cognitive levels represented by various items of test are *Knowledge, Comprehension, Application* and *Higher Ability* items (see, table 2.17). Higher skills are measured mainly by open-ended, productive type of questions.

Table 2.17 Characteristics of English tests of Grade 5 in various cognitive domains

Topic	Score[1]			Percentage			Reliability			
	E_1	E_2	E_3	E_1	E_2	E_3	E_1	E_2	E_3	Weighted Mean Reliability
Knowledge	16	17	9	26.23	26.15	14.75	0.9	0.83	0.884	0.868
Comprehension	24	26	27	39.34	40.00	44.26	0.79	0.872	0.862	0.843
Application	11	12	15	18.03	18.46	24.59	0.786	0.795	0.81	0.798
Higher ability	10	10	10	16.39	15.38	16.39	0.836	0.858	0.825	0.840
Total	**61**	**65**	**61**	**100.00**	**100.00**	**100.00**	**0.922**	**0.931**	**0.928**	**0.927**

Preparation of Marking Schemes

A consulting company was hired for the tasks of data collection, marking, and data entry. ERO accomplished remaining work such as item writing, pre-testing, and marking of the pre-test papers. ERO with the participation of teachers and subject experts prepared marking scheme for each version of test items, and provided to the consulting company for scoring of the students' answer papers. The marking scheme helped for making scoring more reliable. A marking scheme was prepared by adding strict examples of what kinds of correct answers should be allowed for the marks and what kind of answers should not be credited by marks.

2.5 Test Administration, Marking, and Data Entry

Test Administration

The test was administered on the same day, that is on June 12, 2015, in all sample schools of 23 sample Districts. The administration was outsourced to a consultancy firm, but the officers from the MOE/ERO monitored the process. Monitoring was done also by DEO and RED personnel. Test administration was smooth in almost all the cases. The total number of schools participated in assessment was 1543.

Marking and Data Entry

The marking of the papers as well as data entry were outsourced. The objective questions carried one mark for each question. Subjective items, including short-answer types, carried up to three marks in Mathematics tests. In Nepali and English tests, subjective items were somehow long answer type, carrying up to five marks.

Altogether 73879 students' papers, 2425 teachers' questionnaires and 1543 head teachers' questionnaires were collected into a "marking centre" in Kathmandu. The outsourced firm marked the answer papers in OMR sheet and prepared dataset using OMR machine. The student datasets were screened, verified and cleaned at ERO after receiving the dataset from the consulting company.

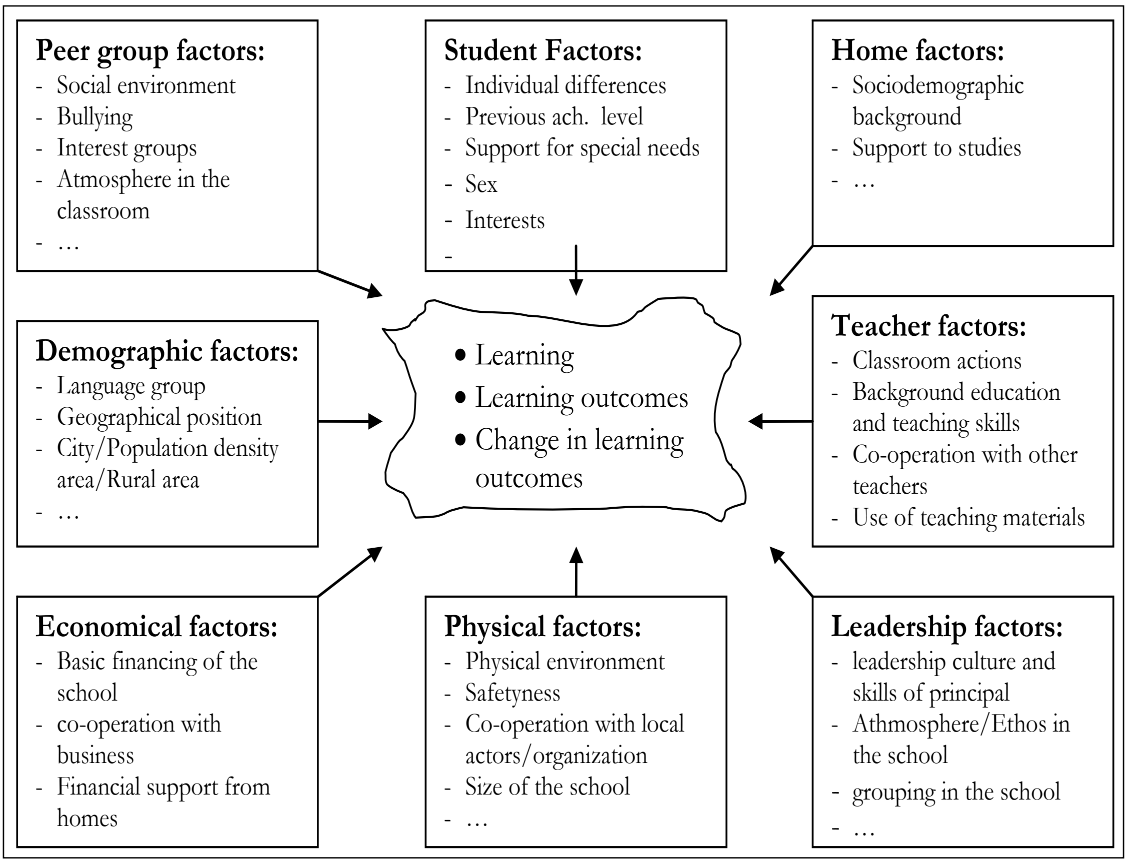
2.6 Background Questionnaires

A set of background questionnaires for students, teachers and head teachers was prepared and administered to find out the factors explaining the differences between the achievements of students as well as schools. Background questionnaires mostly include the contextual information related to the students, their family as well as the school and teaching-learning practices. This section describes the bases and process of selecting background questionnaires.

A Model for Developing Background Questionnaires

A conceptual model for learning presented in Metsämuuronen (2009) and used in previous NASAs (ERO, 2013, 2015a, 2015b) with some contextual modification was taken as a basis for preparing background questionnaires for this assessment. This model suggests that student, family, peer group and school related factors influence students' learning. S***tudent related factors*** include student's individual characteristics such as motivation, attitude, working habits, and so on. F***amily related factors*** include SES, support to the study, literacy in the family and so on. ***Peer group factors*** include a set of related factors such as social support for studies, bullying, atmosphere in the classroom, and so on. The information related to these three factors was included in students' background questionnaires. ***Teacher related factors*** include classroom activities, teaching skills, use of teaching materials, and so on. For the information about teacher related factors, teachers' questionnaires were prepared. ***School related factors*** are divided into two groups: managerial factors and physical factors. To receive the information about the school related factors such as atmosphere in the school, the condition of the school premises, safety, and so on; head teachers' background questionnaire was prepared.

Figure 2.2 Conceptual framework for the background information



(Source: Adopted and contextualized from Metsämuuronen, 2009 as well as ERO, 2013, 2015a, 2015b)

The information about the economic factors was received partly from school head teacher and partly from national educational statistics. Demographic factors were partly included during the sampling from national statistics; and partly it is included in student background questionnaires. Figure 2.2 depicts this model.

Selection of Background Questionnaire

The variables for the background were selected in several phases. As these background questionnaires were adapted with some modifications in language and structure, some modifications were done based on the previous NASA results and responses. While preparing students' questionnaires, some unrelated questions were removed and some low responded ones were modified. Similarly, some questions were modified based on the subject specific information.

Other processes of developing and adapting the background questionnaires were not repeated in this assessment as the background questionnaires have been used since NASA 2011. Therefore, detailed adaptation process from the PISA background questionnaires such as familiarization with the background questionnaire of TIMSS and PISA and the study of regression questionnaire models prepared in PISA to decide the variables were not included at this stage. Without repeating these as in previous NASA (see ERO, 2013; 2015a, chap 2), further processes of revision and modification were done to make questionnaires more relevant. However, each questionnaire was reviewed based on previous responses. In the process of revisiting and finalising the background questionnaires, a number of teachers and experts were involved.

Though the questions were known to be somehow demanding for Grade 3 students, the background questionnaire was the same for Grade 3 and Grade 5 students. The teacher was informed to support the Grade 3 students when needed, for example, with the items such as parents' education and occupation. It is notable, though, that there are several missing values in the Grade 3 datasets; in some variables, almost 20% students did not answer the background questionnaire in previous NASA. The same situation has repeated in this assessment too as there are large missing cases. Therefore, particularly in Grade 3 results a large number of missing cases may have affected the results based on some of the explaining factors. However, in the case of Grade 5, the missing cases are relatively small and the result explains the influences of the factors more accurately.

2.7 Specific Variables Used in the Analysis of Results

Three sets of variables were used carefully while analysing the results: the concept of equated scores, Fennema-Sherman attitude scale, and the indicator for socio-economic status (SES). Three different versions of test items were used in testing student achievement; therefore, the scores were not automatically comparable. For the comparability of different versions of test items, the scores have to be equated before the analysis. The Fennema-Sherman test is a widely used test (e.g. in PISA; TIMSS; NASA, 2011, 2012, 2013) to measure attitude towards school subjects. Similarly, socio-economic variables were used to explain the achievement results and explain the differences between students' achievement.

Equated Scores and IRT Modelling

In final tests, three versions of test items were constructed using certain number of identical items, representing different content areas, which worked as the linking items. Using IRT modelling (Lord, 1980; Béguin, 2000) test scores are equated and finally the comparable latent ability of each student is acquired over the different versions. IRT modelling is the tool for equating test scores in the well-known international comparisons of PISA and TIMSS studies. The testing procedure used in NASA 2011, 2012, 2013 typically involves this kind of complex process facilitated by IRT modelling. IRT modelling is the only credible way to assess how the achievement level was changed from 2012 to 2015, and to compare the results in Nepal with the national as well as international standards such as PISA and TIMSS results.

We can justify the need for equating by giving four distinct reasons: First, to widen the number of items so that the range of topics and sub-topics will be covered in testing. For this purpose, it is preferable to use several versions of tests in testing. In NASA 2015, three versions of tests were used. Second, in order to compare the results of 2012 and 2015, IRT modelling is the most accurate method when equal (or parallel) tests were not used. Third, to compare the national and international level results, IRT modelling is necessary because the only knowledge available of the items was the IRT difficulty parameter β. Additionally, using IRT modelling it is possible to use tests of unequal lengths and make tests free from leaking of the items as well as unfair behaviour among the students. When the tests were of unequal length and somehow different difficulty level, IRT modelling is actually the only sensible method of making the final scores comparable.

Equating the test scores with IRT modelling was applied with the following principles and practices. The scores are transformed into the same scale on the basis of characteristics of IRT models that the latent ability level of a learner (θ) and difficulty level of an item (β) are identical when certain pre-conditions are met (see Wright, 1968). The latent ability level for each student can be determined in the same metric for every test as far as there are the linking items connecting the versions. The estimation was run with OPLM program (Verhelst, Glas and Verstralen, 1995). A brief technical description of the equating process is as follows (see, Béguin, 2000, 17–36):

i. Define the structure of the test so that the linking items are connecting the tests with each other. The values of difficulty parameter of the linking items are the same in each version because the difficulty levels of all other items are calibrated into the same scale as the linking items are.
ii. Use *Conditional Maximum Likelihood* (CML) procedure to estimate the difficulty level (β parameter) for each item.
iii. Use *Marginal Maximum Likelihood* (MML) procedure to estimate the distribution of each student's latent ability (θ parameter) in each version.
iv. Estimate the θ parameter of the scores of each version using means and deviations of distributions of θ and β. This results in a unique latent value, however measured in a common scale, for each observed value of the scores in all versions.

The success of equating depends on three things. First, the linking items should represent a sufficient range of ability level; too easy and too difficult items should, however, be avoided. Second, the linking items should represent a short test inside the test; the items should cover the different content areas as widely as possible. Third, the stable parameters in the equating process are dependent on the sample; the better the sample represents the target population, the better the calibration corresponds with the population parameter. Though the item parameters are to some extent vague, the results are much more accurate than using only the classical matrices (the proportion of correct answers) in comparison.

Normally, in equating the test scores, an average student with average ability would get Theta value zero ($\theta = 0$). The better the student is, the higher is his/her θ above the zero line and parallel, and the weaker the student the lower is θ below the zero line. However, when borrowing the items from the international item bank, their difficulty level is calibrated to that international level where an "average international student" would get the zero for Theta. All the new items were calibrated into that international scale and hence the mean of the "average Nepalese student" does not get the value of zero but either above or below. This makes it possible to assess the achievement level of Nepalese students compared with the international standard.

At the final phase, Mathematics was calibrated into the TIMSS scale. The original scores were transformed into the equated scores and the equated scores were changed to the percentage of maximum score.

Fennema-Sherman Attitude Scale

A shortened version of Fennema–Sherman Attitude Scales (FSAS) (Fennema and Sherman, 1976) are used in several international comparisons, like in Trends in International Mathematics and Science Study (TIMSS) 2007 (Mullis, Martin, and Foy, 2008) and its previous assessments in 1995, 1999, and 2003 as well as in Programme for International Student Assessment (PISA). Original scales include nine dimensions; but in these international comparisons, only three dimensions with four items in each and two negative items in each of the first two dimensions are used. The names of the factors can be "*Liking Math*", "*Self-Efficacy in Math*", and "*Experiencing utility in Math*" (see Metsämuuronen 2012a, 2012b; compare naming in, e.g., Kadijevich, 2006; 2008). Because of students' inconsistent manner in answering the attitude-scale in NASA 2015, only the dimension of "*Experiencing utility in Maths/Nepali/English*" was taken into the measurement instrument for Grade 5 students.

As an example, factor structure analysis as presented in table 2.18 explains whether positive attitude of the students towards the subject results into positive effects on student achievement in Mathematics. There were five questions in the background questionnaire of this assessment to understand the attitude of students in Mathematics in Grade 5. The component matrix shows that the positive attitude towards the subject is positively correlated with student achievement.

Table 2.18 Factor structure analysis for Grade 5 in Mathematics: Component Matrix

Attitude question	Component	
	1	2
BQ_19a Math subject helps me in the daily life.	0.681	-
BQ_19b I need to do better in Math to study the other subjects.	0.662	-
BQ_19c I need to do better in Math to study the subject in the upper Grades.	0.755	-
BQ_19d I like to solve mathematical operations (like division, multiplication, fraction, geometrical shapes and size, percentage etc.).	0.699	-
BQ_19e I need to do better in Math to get a desirable job.	0.734	-

(Extraction Method: Principal Component Analysis)

Socio-economic Status (SES) Eelated Variables

According to Bradley and Corwyn (2002), socio-economic status (SES) is one of the most studied constructs in student learning. There is the general perception that high SES families provide an array of services, goods, parental actions, and social connections for their children that potentially rebound to the benefit of them and a concern that many low SES children lack access to the same resources and experiences, thus putting them at risk for developmental problems (see Brooks-Gunn and Duncan, 1997). Specifically, SES matters because it has been related to health and life outcomes for as long as social groups have existed (Oaks, 2011); and it has been shown to have a strong connection to cognitive and academic attainment (see a convincing literature in, for example, Bradley and Corwyn, 2002; APA, 2007).

In the literature, social status is commonly conceptualized in terms of socio-economic standing, taking into account the various combinations of income, education, and occupation (APA, 2007, p 5). The challenge in measuring SES is that there has not been a complete consensus on precisely what represents economic position or social status (Liberatos, Link and Kelsey, 1988; McLoyd, 1997), economic position or social status and hence there is not a single measure for SES (Bradley and Corwyn, 2002; APA, 2007, p 5). Bardley and Corwyn (2002, p 373) put it as follows: There is consensus that income, education, and occupation together represent SES better than any of these alone (White 1982). However, there is no consensus on (a) how best to composite the set of indicators; (b) whether it works the best to examine relations between SES and child outcomes using a composite, a statistical procedure that includes each indicator, or each indicator singly; or (c) how best to measure each component (Krieger *et al*., 1997).

As in the previous NASA, the following seven indicators of SES were selected into the final SES indicators for this assessment: Father's education, Mother's education, Father's occupation, Mother's occupation, Home possessions, Home accessories and Attending to private school.

As an example, the SES indicators and their cut-offs for Mathematics Grade 5 are presented in table 2.19.

Table 2.19 Indicators of SES in NASA 2015 (Mathematics in Grade 5)

Variable	Cut-off[1]	Effect on total score[2]
Father's education	Less than Grade 10-passed = 0, other = 1	+5% points, $\eta^2 = 0.011$
Mother's education	Less than Grade 10-passed = 0, other = 1	+4% points, $\eta^2 = 0.008$
Father's occupation	Agriculture = 0, other =1	+4% points, $\eta^2 = 0.006$
Mother's occupation	Agriculture = 0, other =1	+3% points, $\eta^2 = 0.004$
Home possessions	5 or less out of 11 possessions = 0, 6 or more = 1	+9% points, $\eta^2 = 0.038$
Home Accessories	1 or 0 accessory out of 4 = 0, 2 or more =1	+10% points, $\eta^2 = 0.056$
Attending to private school	No = 0, Yes = 1	+16% points, $\eta^2 = .106$
Total SES		**+13% points, η^2 =0.085**

Because the variables were of different scales (from nominal to ordinal scales) and because of incomparable scores (from 0–1 to 0–11), all the variables were rescaled first to fit with each other. At the first phase, the variables were analysed with respect to educational outcomes. Decision Tree Analysis (DTA), the data-mining tool in SPSS software, and ANOVA - the basic tool for analysing the differences between the group means - were used to find the best classification of each variable with regard to the statistical differences in learning outcomes. At the second phase, 11 variables comprising the home possessions and three variables comprising the home accessories were summed up and dichotomized based on DTA and ANOVA. At the third phase, all 7 variables for SES were dichotomized based on DTA and ANOVA. Hence, all the variables – regardless of their original scale – were scaled as 0 or 1, where 1 indicates the higher SES (and maximization of learning outcomes). This makes all the individual indicators equal weighted. At the final phase, seven indicators were summed up as the final SES indicators.

The final SES indicators (1) are strictly geared towards educational outcomes (and not health, for example), (2) are balanced with education and occupation though with somehow over-representation of the economic dimension (3 indicators), (3) have moderately high reliability (0.65) – indicating that it can separate, at least, the extremes quite nicely, and (4) can be changed when the society changes.

2.8 Statistical Methods Used in Analysis

While analysing the results, some statistical tools and concepts have been used which are mainly the mean, standard deviation, percentage, frequency and so on. Similarly, statistically significance, effect size and explanatory power are the concepts. Each of them is described here follow.

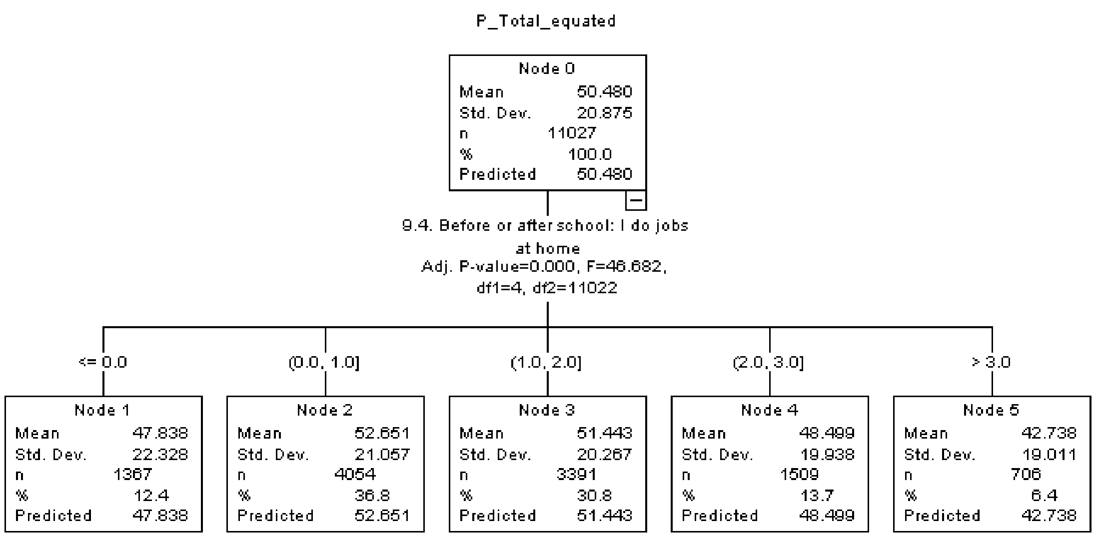
Analytical Tools Used for the Statistical Analysis

The basic tools of statistical description (means, standard deviations, percentages, and frequencies), correlations (Pearson's product moment correlation coefficient), and comparison of two means (t-test) as well as statistical inference (p-values, effect sizes) are used when appropriate in the analysis and reporting. These methods are described in all standard textbooks of statistical description and inference (e.g. Metsämuuronen *et al*., 2013). Analyses are done using SPSS programme. The Analysis of Variance (ANOVA) is used in the General Linear Modelling (GLM) when several means are compared. All the p-values are corrected by using Multi-level modelling (or Hierarchical Linear modelling) by using SPSS Linear Mixed models module when they are not obvious (as $p < 0.001$ is obviously significant with or without correction).

Somehow more exotic method or a set of methods called Decision Tree Analysis (DTA) are used in some cases when willing to find the best predictors of the achievement out of hundreds of possible meaningful variables. DTA is one of the methods used in data mining which is very effective to find statistically the best groupings of the independent variables. DTA produces a chart such as the one below:

DTA (fig. 2.3) shows that when the children spend more than 2 hours per day in household chores, the results in Mathematics are poor. However, when the amount of time spent in household chores is up to 2 hours per day, the achievement level is higher than the average.

Figure 2.3 An example of DTA modelling on paid work and achievement in Mathematics Grade 5



Some Statistical Concepts Used for the Analysis of Results

Within the text, three important concepts have been used: Statistical Significance, Effect Size and Explanatory Power of the variable. ***Statistical significance*** is the p-value, which refers to the possibilities to generalize the result to the population. Behind the p-value (from "probability") is the fact that there is always a measurement error while measuring human mental processes. This means that the result of each individual student as well as each mean score carries error. Especially, when the population is examined by using a sample, all the means carry both measurement error as well as sampling error. In the sample, there can be a small difference between the boys and girls, for example. The p-value tells us how probable the same result could be in the population as a whole. If the probability is $p < 0.001$, this means that the difference would be found at risk of less than 0.1% – only less than one sample out of 1000 samples from the same population, the results differ from those obtained. If the p-value is $p = 0.002$, the risk for a faulty decision (or difference) is 0.2%.

When the sample size is huge – the sample of 17074 students, for instance – the p-value easily gives a signal that the difference between the groups is real in population. In this context, p-value does not tell whether the difference is small or big. For this purpose, there is another statistical concept, which is ***Effect Size (ES).*** Effect size indicates how far the lowest and highest groups are from each other. The especially used indicators of ES are Cohen's d for two means and Cohen's f for several means (Cohen, 1988). Cohen has given boundaries for small, medium and large effect sizes. During the text, these boundaries are used as a "measurement stick" to indicate whether the difference is small, medium or large. The rough boundaries of the small, medium, and large effect sizes are collected in table 2.20.

Table 2.20 Rough boundaries of Effect Sizes

Size	Cohen's d	Cohen's f
Small	< 0.2	< 0.1
Medium	Round 0.4	Round 0.2
High	> 0.8	> 0.4

Technically, ES also gives a preliminary indication as to how well the grouping factors, such as the gender, explain the results. Hence, in the text one may read as "the difference between boys and girls is statistically significant at $p < 0.001$, but the effect size is small." This means that, first, the difference between the boys and girls is real, but second, the difference is very small in reality and third, gender as a grouping variable does not effectively explain the variation in the data.

The third related concept is the ***Explanatory Power of the Variable***, especially when using the Analysis of Variance (ANOVA) as an analytic tool, the output allows the possibility to show how well the factor explains the variation in the data. The usual indicator for this is Eta squared (η^2) which actually is a correlation coefficient between a grouping variable and continuous variable. When the Eta squared equals $\eta^2 = 0.30$, this means that the grouping factor (such as the geographical region) explains 30% of the variation in the dataset. Cohen's f strictly uses this information:

$$f = \sqrt{\frac{\eta^2}{1-\eta^2}}$$

Hence, if $\eta^2 = 0.30$, then the effect size f equals $f = \sqrt{(0.3/0.7)} = \sqrt{0.43} = 0.65$, showing high effect size (see Cohen, 1988, 284).

While comparing two groups, an effect size is measured by Cohen's d which uses the information

$d = \frac{t \cdot (n_1 + n_2)}{\sqrt{df \times n_1 \times n_2}}$, where n_1 and n_2 are number of students in first and second group and df is the degree of freedom.

2.9 Data Analysis and Report Preparation

The report preparation started when the data were received from the consultancy firm. Once the data were received, the first task was the data cleaning process in each subject. During this process, the background variables and items were recoded and many indicators were also created for the purpose of analysis. Data were also verified even from the OMR sheets, mainly when errors were found or data were missing.

The second step for the analysis was the preparation of database for the OPLM design and analysis. During this phase, the item characteristic curves of all the items were analysed and the items having negative correlation were removed from the analysis. At the end of this phase, the version, cognitive level, content area, item types based students' ability scores (theta) were generated.

In the third phase, the three versions of students' score and latent ability (theta) generated from OPLM were used to equate all the three versions of each subject to maintain the difficulty level. Further equating was done based on cognitive level, content areas and item types. All the equating processes were completed manually in Excel application.

During the data analysis, a template of the list of tables and figures was prepared. The template covered list of tables and figures on overall basic results as well as disaggregated results on various strata, results based on various diversity factors, which show equity based on achievement, and factors influencing students' achievement as presented in background questionnaires. Based on the template, tables and figures were generated with appropriate numbering. In most of the table mean scores, standard deviation and coefficient of variance were calculated. While analysing the factors influencing students' achievement using analytic tool, ANOVA variances were calculated and effect sizes were identified using Cohen's *f* and *d*. Several rounds of workshops were organized in order to analyse data. Data analysis was completed in three months.

After analysing the data, draft reports of each subject and Grade were prepared, and then methodological chapter was drafted. Draft report was reviewed and then introductory chapter and concluding chapters were drafted. When the first draft was completed, it was sent to experts for the review of contents. Finally, a language editor was assigned the task of making correction in the report for fine-tuning it and ensuring linguistic correctness. Draft report preparation, reviewing the draft, content editing and language correction of the report took about three months. Several rounds of workshops were organized and experts were consulted in order to analyse data and prepare draft report.

2.10 Professional Standard and Ethical Consideration

This study has been accomplished following and maintaining the professional and ethical norms during the entire process. The study followed the standard norms and procedures of National Assessment during random sampling, preparation of items and test administration. Test items were based on National Curriculum rather than following a particular textbook or reference book. All the data and information presented in this report are based on student achievement test, students' responses on background questionnaires, teachers' and head teachers' responses from sample schools. In order to make assessment technically sound, and analysis and interpretation unbiased, several experts were involved during the process. The task of test administration, marking answer sheets, data entry using OMR, and data cleaning were accomplished through the consultant firm. Furthermore, during the report writing, a team of experts was assigned for editing the contents and finalizing the report by correcting the language. Such external involvement helped to make assessment process and result more credible.

In addition to the above-mentioned professional considerations, basic ethical standards were taken into account during the process of this assessment. This study maintained the confidentiality and anonymity of respondents. Participation of the students and teachers were made voluntary, and no forced and paid recruitment of students and teacher was done. This study has not exposed the information related to any particular student, teacher and school unless self-request was made. ERO can make the report available only to the concerned schools. Partial database can be provided to researchers on request provided that no any individual information will be disclosed. Some of the background information questionnaires and international linking items were adopted from banks of items released by TIMSS and PIRLS accepting their condition of non-profit use.

Chapter 3: Analysis of Assessment Results for Grade 3 in Mathematics

3.1 Introduction

Mathematics as one of the subjects of school curriculum has been accorded 8 credit hours per week. Overall objectives of tecahing the subject are to: develop among the youngsters the basic knowledge and understanding in the use of number and in mathematical operations, foster basic mathematical and numeracy skills required for solving daily life problems, and lay foundations for studying Mathematics in upper Grades.

Along with other subjects, Mathematics has been the one assessed frequently in the national assessment of student achievement in Nepal. Prior to this, Mathematics proficiency of 3rd Graders has been assessed twice, in 1995 (BPEP, 1995) and in 2012(ERO, 2015b).The frequent assessment in the subject is stimulated by the fact that the value for mathematical skills is demanding highly in modern society because of the increased use of information in numeric form, in tables, graphs and plots, charts and so on requiring the ability to calculate in percentage and value of money in every day context. Modern society expects that each citizen is equipped with required mathematical and numeracy skills to handle such information to survive in the digital world.

This result is based on the National Assessment of Student Achievement for Grade 3 in Mathematics conducted by ERO in 2015 in 771 sample schools and 17074 students from 23 Districts of Nepal. The assessment was carried out to assess the curricular competencies as set in the approved national curriculum of Nepal for Grade 3 in Mathematics. This chapter presents and analyses the results of the national assessment of student achievement.

The chapter begins with the introduction, and it presents the overall distribution of achievement of Grade 3 students in Mathematics. While doing so, it describes the achievements in various content areas of the curriculum, levels of cognitive domain and types of items. Considering the temporal perspective, the results of this assessment are also compared with the result of the first cycle of NASA i.e., NASA 2012.

The Third section of this chapter provides results associated with various diversity factors, including District, Ecological Zone, Development Region, school type, school location, language at home, and caste/ethnicity.

Section 4 deals with the selected explanatory factors about the student achievement in Mathematics of Grade 3. The factors that explain achievement in Mathematics included in this section are parents' education and occupation, home possessions and accessories, socio-economic status (SES) of parents, work beyond school hours, age of the student, support provided for the study, availability of textbook, homework given and feedback provided, and negative and positive activities at school. The final section of this chapter briefly summarizes the findings of the assessment results in Mathematics.

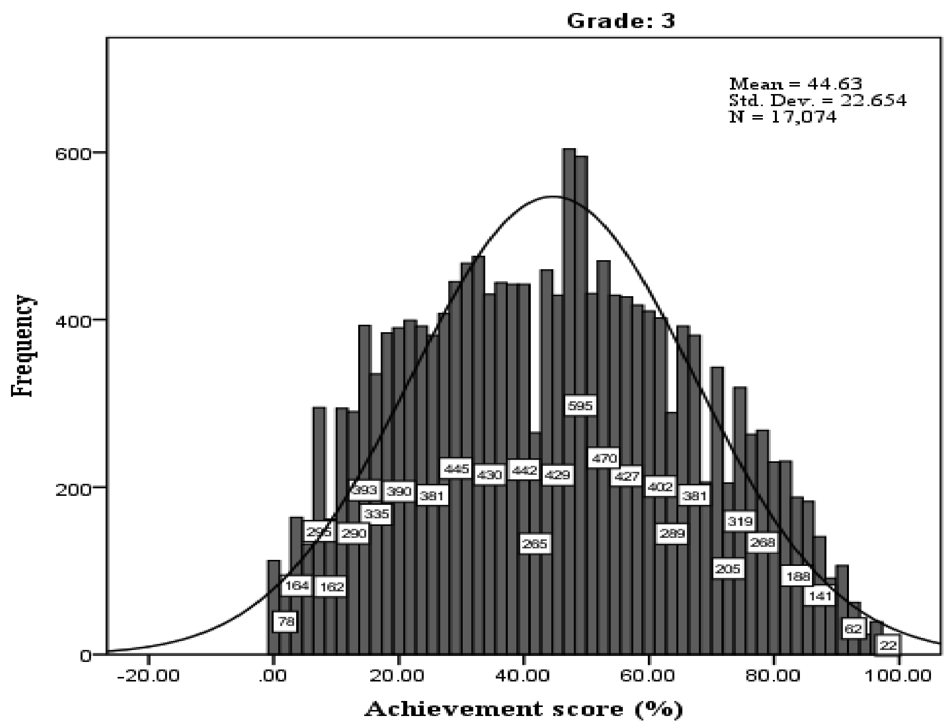
3.2 Basic Achievement Results in Mathematics

This section presents and discusses the overall distribution of assessment scores, achievement by various content areas and various levels of cognitive domain in Mathematics. Similarly, it compares the achievement results of this assessment with the results of NASA 2012 in Mathematics for Grade 3.

Overall Distribution of Achievement Scores

The population is usually distributed normally in a large sample in the study of student achievement. In this assessment of Mathematics, the sample was big enough to form a normal distribution, as there were 17,074 students in the sample. Figure 3.1 shows that the achievement score in Mathematics is normally distributed.

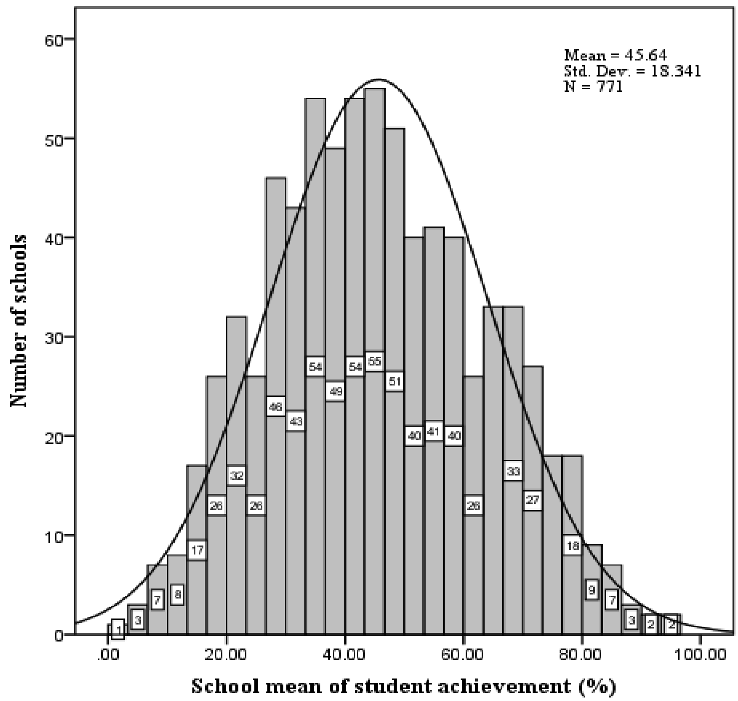
Figure 3.1 Distribution of overall results in Mathematics



The dataset shows that the 3rd Graders' population in Mathematics is normally distributed.

As illustrated in figure 3.2 given below, the schools can be divided into two distinct categories: the high and low performing. The population on the left-hand side has achieved the average of 25.5% of the achievement score, and on the right-hand side, the average achievement is around 63%. The difference in achievement between the populations is remarkable.

Figure 3.2 School mean achievement score in Mathematics



Achievement by Various Content Areas of Mathematics

As per the curricular provisions, the Mathematics test includes five content areas - namely, 1) *Arithmetic, 2) Geometry, 3) Algebra 4) Numeracy.* The number of items and the weightage they carry correspond to the weight allotted in the curriculum.

For comparability, achievement scores in all the content areas were converted into the percentage of maximum score of each content area. Table 3.1 shows the students' achievement in Mathematics by each of the five content areas.

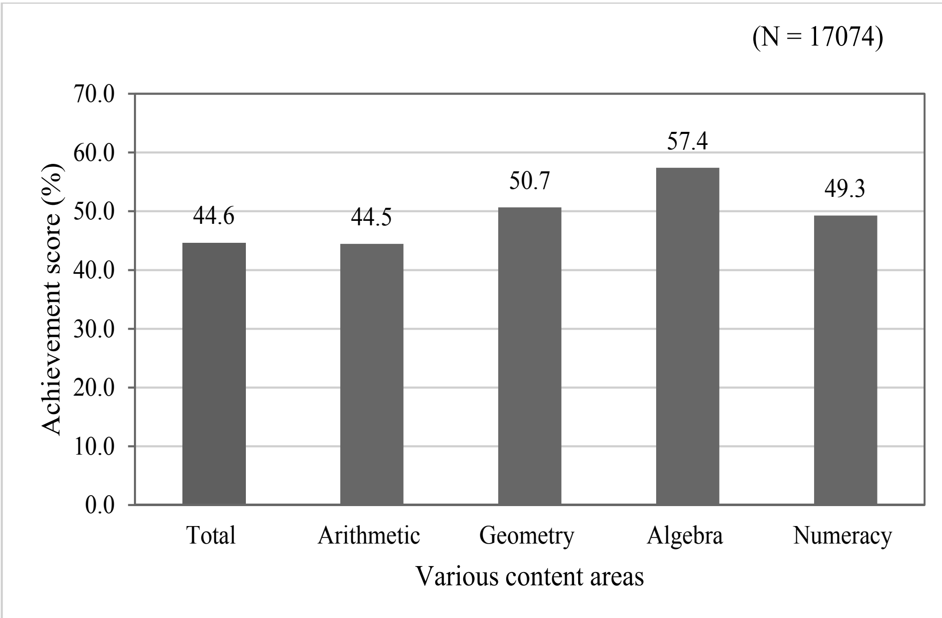
Table 3.1 Achievement score in various content areas in Mathematics

Objectives	Mean	SD	Min	Max
Arithmetic	44.5	22.52	0	100
Geometry	50.7	26.54	0	100
Algebra	57.4	31.93	0	100
Numeracy	49.3	27.27	0	100
Total	**44.6**	**22.65**	**0**	**100**

Table 3.1 shows the variations of achievements in various content areas in Mathematics. The achievement score ranges from 44.5 percent in *Arithmetic* to 57.4 percent in *Algebra* with the variation of 12.8 percent. To compare the maximum and minimum scores, the situation is the

same in all content areas as the maximum score is 100 and the minimum is 0 in each of them. Figure 3.3 compares the variations of achievements in various content areas of Mathematics.

Figure 3.3 Comparison of achievement scores in various content areas in Mathematics

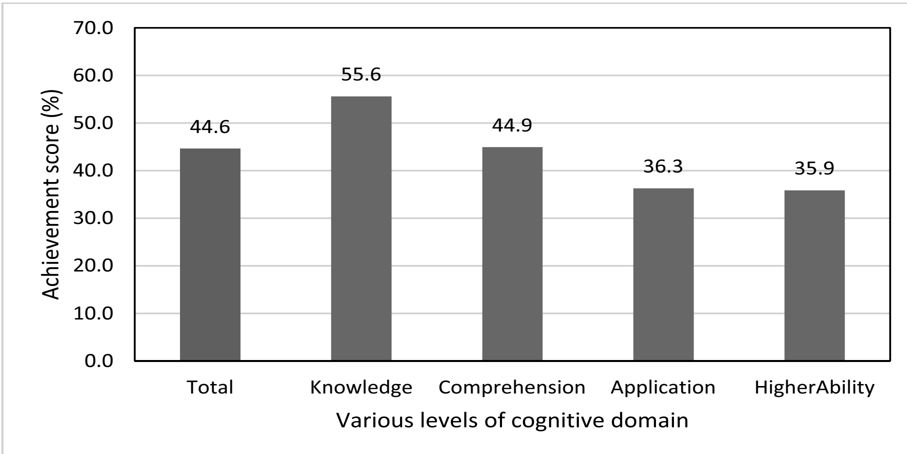


The overall national average achievement score for Mathematics is 44.6%. When the achievement was analysed categorising into four content areas, students were found to be performing poorly in *Arithmetic* (44.5%), which is below the national mean (44.6%), however, the difference is not significant. Similarly, achievement score in other content areas are above the national average – with 50.7% in *Geometry*, 57.4% in *Algebra*, and 49.3% in *Numeracy*. The dataset indicates that the students are somehow weak in *Arithmetic*, but they are performing highest in *Algebra*.

Achievement in Various Levels of Cognitive Domain

The Mathematics test comprised of items that can be grouped under various levels of cognitive domain according to Bloom's taxonomy (Bloom et al. 1956; Metfesser, Michael & Kirsner 1969). The four levels of the domains have been considered while developing items were *Knowledge, Comprehension, Application* and *Higher ability*. Reasoning and problem solving type of items were incorporated under 'H*igher ability*' category because this category covers thinking related to analysis, synthesis and evaluation. The achievement of the students in each level of cognitive domain is shown in figure 3.4.

Figure 3.4 Comparison of achievement score in various levels of cognitive domain in Mathematics



The above figure indicates that, in general, the students' ability to solve complex problems (higher ability) is low as they achieved only 35.9% of the maximum score. However, students are much better in *Knowledge* level questions, mostly recalling type of questions (55.6%).

Analysing the dataset further, performance of a large number of students was found relatively low in higher cognitive abilities in Mathematics. For example, about 20% of the students in Mathematics could not solve any of the tasks requiring higher level skills.

Student Achievement by Type of Items

There were two types of items in the test: *Objective* and *Subjective*. Objective items covered a wide range of content areas, which were very specific to judge because there was only one correct answer; or, one explicit piece of information is needed for correct answer. There were some *Subjective* items in each test version, which required a longer procedure to get full marks. Both the *Objective* and *Subjective* types of items were developed considering various levels of cognitive domains (K*nowledge, Comprehension, Application* and *Higher ability*) and the various difficulty levels, though the subjective items tend to be more demanding. Table 3.2 comprises the basic statistics of achievement in different types of items. It is clear that the students scored poorer marks in *Subjective* items (34.7%) than in *Objective* items (54.3%).

Table 3.2 Students' achievement by types of items in Mathematics

Type of items	N	Mean	SD	SE of Mean
Objective	17074	54.3	22.06	0.169
Subjective	17074	34.7	26.01	0.199

Dataset indicates that students are good at recognizing the correct answer and in very fundamental contents like recognizing numbers, shapes, choosing correct answer from the options, very basic operations such as the basic manipulation of data and numbers, and calculations using a single or few steps. They are much weaker in calculation having more than one-step, questions involved reasoning, problem solving, plotting, proving the theory or formula, and constructing the shapes and figures. In many cases, the students did not even start doing the open-ended questions, and hence, they got low score.

Comparison of Achievement Results between NASA 2015 and 2012

Comparing raw average scores of the achievement between NASA 2015 and NASA 2012, NASA 2012 score (60%) is remarkably higher than that of NASA 2015 (44.6). However, comparing scores of different test may not be preferable. In this case, item response theory (IRT) suggests to compare results of two assessments based on the latent ability (θ). Latent ability (θ) is a mathematical model of estimating the ability shown by the students based on the item difficulty parameters. Difficulty parameters of some linking items taken from NASA 2012 were fixed in the test of NASA 2015 to calibrate the test items and to equate scores of NASA 2015 so that both tests became comparable. While comparing latent ability (θ) for both the tests there is no such differences seen as latent ability of NASA 2015 (0.236) is higher that of NASA 2012 (0.111). Value of theta (θ) ranges from negative to positive value from -3 to +3 with the mean value zero (0) as a reference line.

Table 3.3 compares the raw average score in various content areas of Mathematics in Grade 3 between NASA 2012 and NASA 2015.

Table 3.3 Comparison of achievements between NASA 2012 and 2015 in various content areas

Content area	Percentage of mean score in NASA 2012 (N = 19252)	Percentage of mean score in NASA 2015 (N = 17074)
Total	60	44.6
Arithmetic	61	44.5
Geometry	60	50.7
Algebra	40	57.4
Numeracy	54	49.3

Table 3.3 shows that there is not the same trends of average scores in various content areas in Mathematics such as in NASA 2012 students' score is highest in *Arithmetic*, which is the lowest scoring area in NASA 2015; in NASA 2015 students' average score is highest in *Algebra*, which is the lowest achieving area in NASA 2012..

3.3 Achievement Score by Diversity Factors

Though diversity is a relative and contextual term, in the context of NASA 2015, six diversity factors have been considered – namely geographical/ecological, regional, language, gender, ethnic/caste and economic diversity. NASA 2015 background information questionnaire included the questions related to these six factors of diversity. However, this assessment also considered three additional comparisons. They are by Districts, by school type (community/institutional) and by school location (rural/urban). These comparisons are done particularly to assess the equity status of students based on achievement scores.

Student Achievement by Districts

The achievement of each of the 23 sample Districts is presented in table 3.4. The table shows mean achievement of sample Districts starting from Eastern Development Region. The figure 3.5 presents the mean achievement scores of sample Districts according to the sequence of highest to lowest order of achievement score.

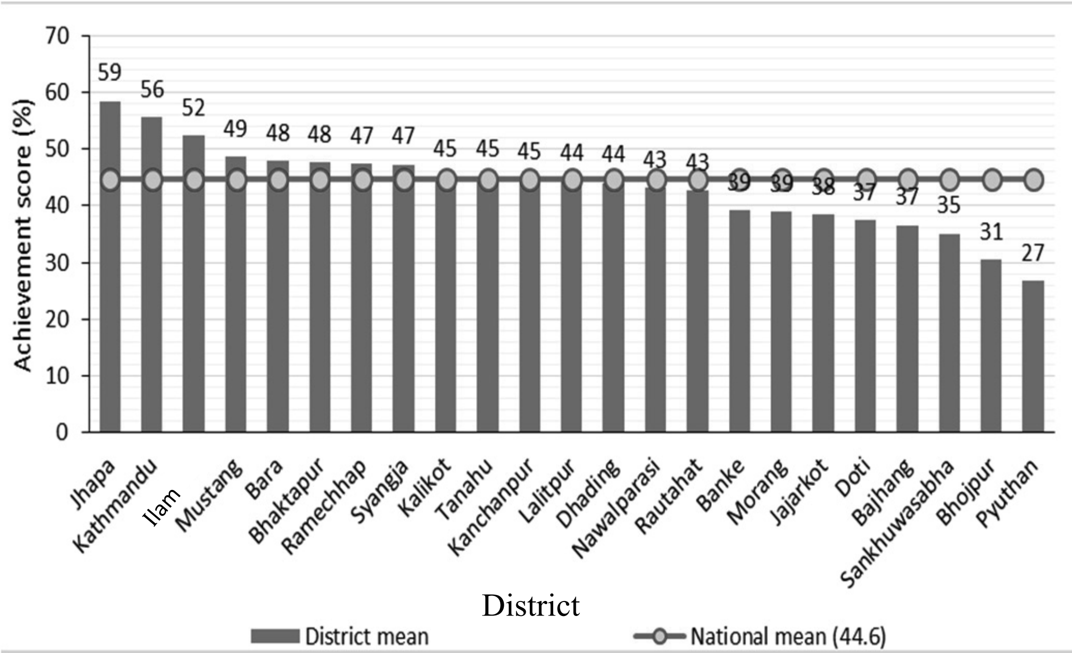
Table 3.4 shows that the Jhapa, Kathmandu and Ilam are the three highest achieving Districts with 58.5, 55.7 and 52.4 percent average score respectively whereas Pyuthan, Bhojpur and Sankhuwasabha are the three least achieving Districts achieving 26.8, 30.5 and 35.1 percent.

Table 3.4 Average achievement scores of sample Districts in Mathematics

District	N	Mean	SD	District	N	Mean	SD
Illam	543	52.4	23.20	Tanahu	965	45.1	19.94
Jhapa	1040	58.5	21.23	Syangja	738	47.3	21.78
Morang	1064	39.0	20.93	Mustang	31	48.7	14.75
Sankhuwasabha	476	35.1	17.94	Nawalparasi	939	43.1	22.99
Bhojpur	300	30.5	22.97	Pyuthan	696	26.8	16.31
Ramechhap	487	47.3	20.93	Banke	847	39.2	20.07
Lalitpur	964	44.5	19.22	Jajarkot	703	38.4	24.19
Bhaktapur	496	47.7	20.76	Kalikot	481	45.1	20.32
Kathmandu	1864	55.7	20.95	Bajhang	636	36.6	20.35
Dhading	892	44.0	22.29	Doti	549	37.5	20.71
Rautahat	661	42.8	21.83	Kanchanpur	569	44.6	23.90
Bara	1133	48.0	24.89	**Total**	**17074**	**44.6**	**22.65**

The figure 3.5 compares the mean achievement of sample Districts against the national mean score in Mathematics. It shows wide variations among the Districts, for example the difference between the high performing District Jhapa (58.5%) and the low performing District Pyuthan (26.8%) is 31.7 percent.

Figure 3.5 Average achievement score of sample Districts in Mathematics



Out of the selected 23 sample Districts, Jhapa is on the top (59%) and Pyuthan on the bottom (27%). Among them, mean score for the 11 Districts is above the national average whereas for 12 Districts, it is below the national average. The data-mining tool of SPSS software – Decision Tree Analysis (DTA)–points out that the lowest performing schools are mostly from Mid-Western and and Far-Western region.

The difference in achievement scores among the Districts is statistically significant ($p < 0.001$). The variation among the Districts explained in achievement is $\eta^2 = 0.111$, which explains 11% of the variation in achievement. Effect size is $f = 0.35$, which indicates that the difference between the lowest achieving and highest achieving Districts is moderately high.

Student Achievement in Various Ecological Zones

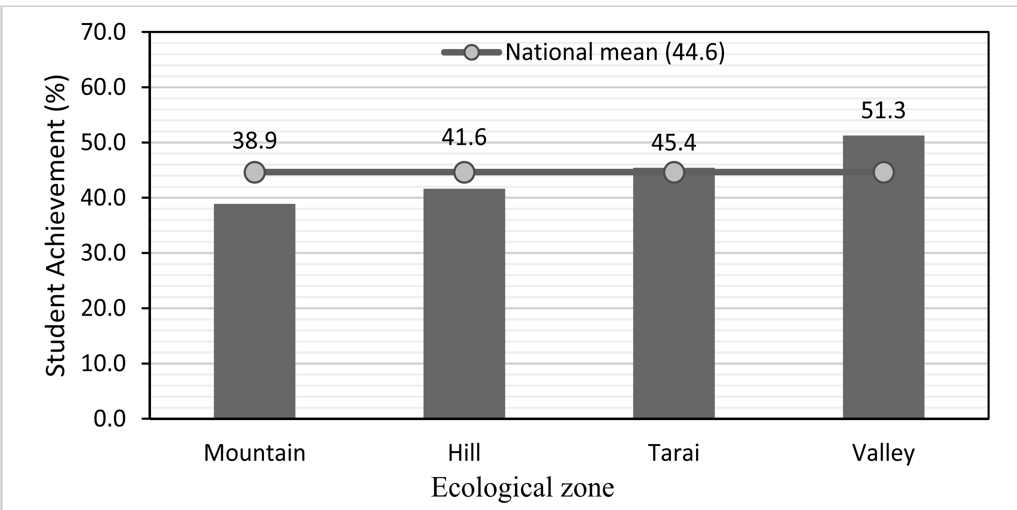
The Mountain, Hill and Tarai are three Ecological Zones of Nepal though the Kathmandu Valley has been considered as a special geographical area because it is the most densely populated area in the country with more opportunities than the others. Not only the population but also the mixed ethnicities, weather condition, economic activities, urbanization, as well as the dense human capacity have been taken into account while considering the Kathmandu Valley as a unique fourth geographical area for the purpose of analysis. The variation in the Ecological Zones in NASA 2015 is presented in table 3.5. The table also includes the number of students, standard deviation, standard error and minimum and maximum scores in each of the categories.

Table 3.5. Achievement scores in Mathematics in various Ecological Zones

Eco Zone	N	Mean	SD	SE of Mean	Minimum	Maximum
Mountain	1624	38.9	20.05	0.498	0	98
Hill	5873	41.6	22.58	0.295	0	98
Tarai	6253	45.4	23.26	0.294	0	96
Valley	3324	51.3	21.07	0.365	0	98
Total	**17074**	**44.6**	**22.65**	**0.173**	**0**	**98**

Figure 3.6 compares the students' achievement among the Ecological Zones as well as with the national mean score in Mathematics.

Figure 3.6 Comparison of achievement scores by Ecological Zones in Mathematics



The data shows that, on average, the students from the Kathmandu Valley have outperformed (51.3%) those from all other Ecological Zones. The students from the Mountain areas performed the lowest (38.9%).

The achievement in the regions differs significantly (p < 0.001). In terms of Tukey's *post hoc* test, all the Zones deviate from each other in a statistically significant manner (at p = 0.01 level). The effect size is f = 0.17, which shows a moderate difference between the highest and lowest performing Ecological Zones. Ecological Zone explains 3% of the variation in the data ($\eta^2 = 0.029$). Excluding the Kathmandu Valley in ANOVA analysis, the Ecological Zone explains 1% of the variation in achievement where the effect size is small (f = 0.10).

Dataset indicates that there is a wide difference between the student performances in four Ecological Zones. Students in the Kathmandu Valley have outperformed (51.3%) the other students. The achievement is the lowest in the Mountain area (39.9%). Moreover, average achievement of Tarai region (45.4%) is higher than that of Hill region (41.6%).

Student Achievement by Development Regions

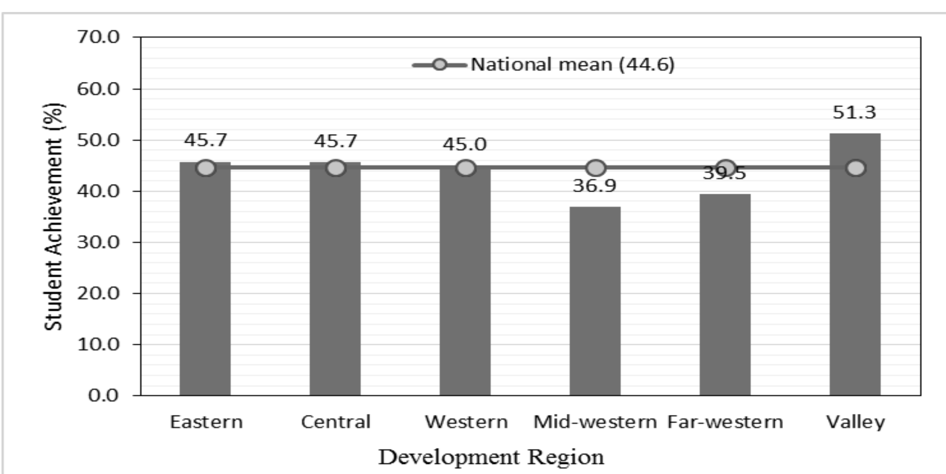
Student achievement varies according to the Development Regions, which are divided into five: 1) Eastern, 2) Central, 3) Western, 4) Mid-Western, and 5) Far-Western. Additionally, the Kathmandu Valley is taken as the additional Region though administratively it falls under the Central Development Region. The mean achievements of the Development Regions are given in table 3.6.

Table 3.6 Student achievement scores in Mathematics in various Development Regions

Region	N	Mean	SD	SE	Minimum	Maximum
Eastern	3423	45.7	23.58	0.403	0	98
Central	3173	45.7	23.06	0.409	0	96
Western	2673	45.0	21.57	0.417	0	98
Mid-Western	2727	36.9	21.37	0.409	0	96
Far-Western	1754	39.5	21.95	0.524	0	98
Valley	3324	51.3	21.07	0.365	0	98
Total	**17074**	**44.7**	**22.65**	**0.173**	**0**	**98**

The highest performance is in the Kathmandu Valley (51.3%). Excluding the Valley, Eastern (45.7%) and Central (45.7%) Development Regions are performing equally while Western Development Region (45%) is performing equal to the national mean. Performance is lowest in two Regions: Mid-western (36.9%) and Far-western (39.5%), which is below the national average. The difference between the Regions is statistically significant ($p < 0.001$) in which Tukey's *Post hoc* test shows that all the Development Regions differ from each other significantly. The effect size is found moderately high (f = 0.21), and it explains 4% ($\eta^2 = 0.04$) of the variation in student achievement. However, if the Kathmandu Valley is excluded from the analysis, the effect size will be medium (f = 0.16) and the Development Regions explain only 2% of the variation in achievement.

Figure 3.7 Comparison of achievement scores by Development regions in Mathematics

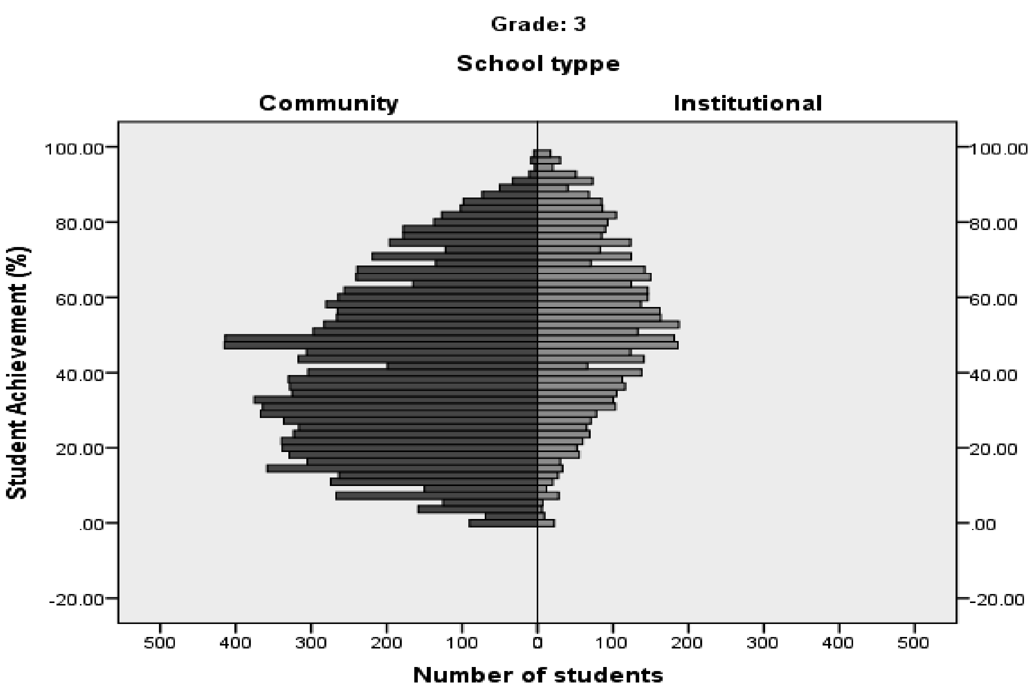


The dataset reveals that there is wide inequality among the Development Regions for children's opportunities to reach an adequate level in Mathematics. The difference between the lowest performing (Mid-Western, 36.9%) and the highest performing Region (the Kathmandu Valley, 51.3%) is remarkable, that is more than 24 percent. If the Kathmandu Valley is excluded from the analysis, the difference will be 8.7 percent. Performance of students in the Mid-western and Far-western Rregions is poor compared to the situation in other Regions.

Student Achievement by School Type

Altogether 572 community schools and 199 institutional schools participated in the assessment of Grade 3 in Mathematics. While plotting community and institutional schools and their mean achievements in a graph as in figure 3.8, community school students on the left hand side are distinctly distributed at least in three kinds of populations (groups): low, middle and high performing. There is a remarkably large number of the students achieving below 35% scores. The right hand side distribution shows the student population of institutional schools, which is shifted slightly upward showing better performance. However, in institutional schools, there are also some students obtaining very low marks in Mathematics. This indicates that the students in institutional schools can be categorized as low to high performing, though majority of them are high performing.

Figure 3.8 Distribution of achievement scores of students in community and institutional schools in Mathematics



All the schools are categorized into 'community' and 'institutional' (usually called 'private') schools. The difference in achievement between the two types of schools is presented in table 3.7.

Table 3.7 Student achievement according to school types in Mathematics

School type	N	Mean	SD	SE of Mean	Min	Max
Community	12316	41.0	22.2	0.20	0.00	98.18
Institutional	4748	54.1	21.0	0.30	0.00	98.18
Objectives	**17064**	**44.6**	**22.7**	**0.17**	**0.00**	**98.18**

As the dataset indicates that the average performance of the students from institutional schools is 54% whereas it is 41% in community schools; that is, the difference is 13 percent, which is remarkable. This shows wide difference in the level of students' achievement between community and institutional schools, whereby the difference is statistically significant ($p < 0.001$) and the effect size is high ($d = 0.65$), which reveals that the gap in achievement between these two types of schools is big. It happened so because most of the institutional schools from the sample are performing high but students in community schools are found in the lowest as well in highest levels in achievement.

The dataset indicates that average performance of the students from institutional schools is 54% whereas it is 41% in community schools, that is, the difference is 13 percent, which is a remarkable difference.

Student Achievement in Rural and Urban Schools

One of the strata of sampling in NASA 2015 was school location. The schools were categorized into rural and urban groups. The disaggregated achievements of the students from rural and urban schools are presented in table 3.8.

Table 3.8 Student achievement in rural and urban schools

School Location	N	Mean	SD	SE of Mean
Urban	4757	50.9	22.17	0.32
Rural	10345	43.5	22.21	0.22
Total	**15102**	**45.8**	**22.46**	**0.18**

The achievement level of the students in the urban schools (50.9%) is higher than that in rural schools (44.45%). The difference in average score is significant ($p < 0.01$); however, the effect size is medium ($d = 0.33$) indicating a small difference between the schools at rural areas and urban areas. Student performance significantly varies based on the location of the schools in Mathematics. Urban schools outperformed (50.9%) the rural schools (43.5%) by 7.4 percent score.

Student Achievement in Various Language Groups

In the context of Nepal, student achievement has been found somehow affected by the language spoken at their homes i.e., the mother tongue. In many cases, mother tongue reflects the ethnic background, and hence, it may be taken as a possible reason for achievement differences in Mathematics.

Based on the total data, about 30.8% of the third Graders speak a language other than Nepali in their home. These “other” languages are quite fragmented; the largest groups in the student dataset are Magar (2.6%), Tharu (2.2%), Tamang (3.4%), Newar (2.7%), and other languages like Abadhi-Bhojpuri, Gurung, Sherpa, Limbu. Beside these language speakers, '*Else'* category of students (as they reported) is 20.5%. Except for Nepali speakers, all other language groups including the language groups stated above (30.8% in total) are categorised as the group “Non-Nepali”. The achievement results are presented in tables 3.9 and 3.10.

Table 3.9 Achievement of Nepali and non-Nepali language groups in Mathematics

Language Group	N	Mean	SD	SE of Mean
Nepali	10242	47.1	22.23	0.22
Non-Nepali	5256	43.7	22.28	0.31
Total	15498	45.9	22.30	0.18

The average achievement of students having their home language Nepali is 47.1% and that of Non-Nepali students is 43.7%; thus, the difference in achievement is 3.4 percent. Table 3.9 shows the achievement of student groups who speak different languages.

Table 3.10 Student achievement by linguistic background in Mathematics

Language at Home	N	Mean	SD	SE of Mean
Nepali	10242	47.1	22.23	0.220
Magar	398	44.0	22.44	1.125
Tharu	345	43.4	22.77	1.226
Tamang	525	45.6	21.07	0.920
Newari	411	49.8	19.98	0.986
Urdu (Muslim)	16	36.5	15.70	3.926
Rai	126	37.9	22.22	1.980
Gurung	94	53.9	19.67	2.029
Sherpa	40	41.6	20.76	3.282

(The language group with less than 10 students has not been considered for this analysis.)

Without further grouping these individual language groups of students, Gurung speaking students outperformed (53.9%) the other language groups. Of the groups, Urdu speaking students have performed the lowest (36.5%) in Mathematics. However, number of students in Nepali language speaking group is high whereas other language speakers are small in the sample.

The difference between the language groups is statistically significant ($p < 0.001$). However, the effect size is low (f=0.10); the language group explains 1% of the variation in the data ($\eta^2 = 0.01$).

Dataset reveals that Nepali speakers outperform non-Nepali speakers by 3.4%. Among all language speakers, Gurung speakers achieved the highest score (53.9%) and Urdu speakers achieved the lowest score (36.5%). Variation between the highest and lowest achieving groups is low.

Student Achievement by Various Ethnic/Caste Groups

Education in Nepal has been influenced in several ways by the legacy of traditional caste system, which is still rooted in the mind-set of people in Nepalese society. Historically, the Brahmans and Chhetris have higher level of educational attainments, but Dalits, for example, have remained outside the educational system. Hence, modern society has made lots of efforts to make education accessible for all children regardless of their caste and ethnicity. The recent National Population Census 2011 shows that the enrolment of Hill Dalits has increased remarkably at the primary level of schooling but their number at the secondary level and higher education is still very small (CBS, 2012). The achievement results of the students from various ethnic/caste groups are presented in table 3.11.

Table 3.11 Student achievement by ethnicity/caste background

Ethnicity/caste	N	Mean	SD	SE	Minimum	Maximum
Brahman	2040	50.0	22.14	0.490	0	98
Chhetri	3488	45.9	21.81	0.369	0	98
Janjati	5690	47.4	21.98	0.291	0	98
Dalit	2090	40.7	21.44	0.469	0	96
Madhesi	1335	44.6	23.55	0.644	0	96
Minority	101	42.1	20.96	2.086	0	91
Other	474	46.2	22.77	1.046	0	98

Average achievement score of the students from Dalit castes is 40.7% whereas the achievement of Brahman caste/ethnicity is the highest (50%). The difference is 9.3 percent. The overall difference among the groups is statistically significant ($p < 0.001$), but the effect size is moderately low ($f = 0.12$). Categorization of students according to their ethnic/caste background explains 1% of the variation in data ($\eta^2 = 0.014$). Shaded cells in table 3.12 represent the lowest achievement, and the cells without shading represent the highest achievement.

Table 3.12 Lowest and highest achievement in various castes/ethnicities by Ecological Zones

Eco-Zone	Mountain	Hill	Tarai	Valley
Brahman	38.0*			*53.3♣*
Chhetri		42.7*		*50.9♣*
Dalit		38.1*		*48.0♣*
Janjati	36.7*			*52.6♣*
Madhesi		31.7*		*50.9♣*
Minority			*46.9♣*	24.5*
Other	34.2*			*53.6♣*

** Lowest achievement ♣ Highest achievement*

Except for Minority groups, other castes/ethnicities perform the highest in the Valley. Besides minority, lowest performance of Brahmans is found in Mountains (38%); Chhetris and Dalits in Hillside (42.7% and 38.1% respectively); Janjati in Mountain (36.7%) and Madhesi in Hill (31.7%).

Overall, dataset shows that achievement of Brahman students is the highest and that of Dalit students is the lowest. Caste/ethnicity wise achievement score varies in most of the Ecologicalregions. All of the castes/ethnicities performed the highest in the Valley; however, minority caste students perform best in Tarai region (46.7%). Brahman, Janjati and other caste categories are performing lowest in the Mountain; Chhetris, Dalits and Madhesi students performed lowest in Hill and minorities performed lowest in Valley.

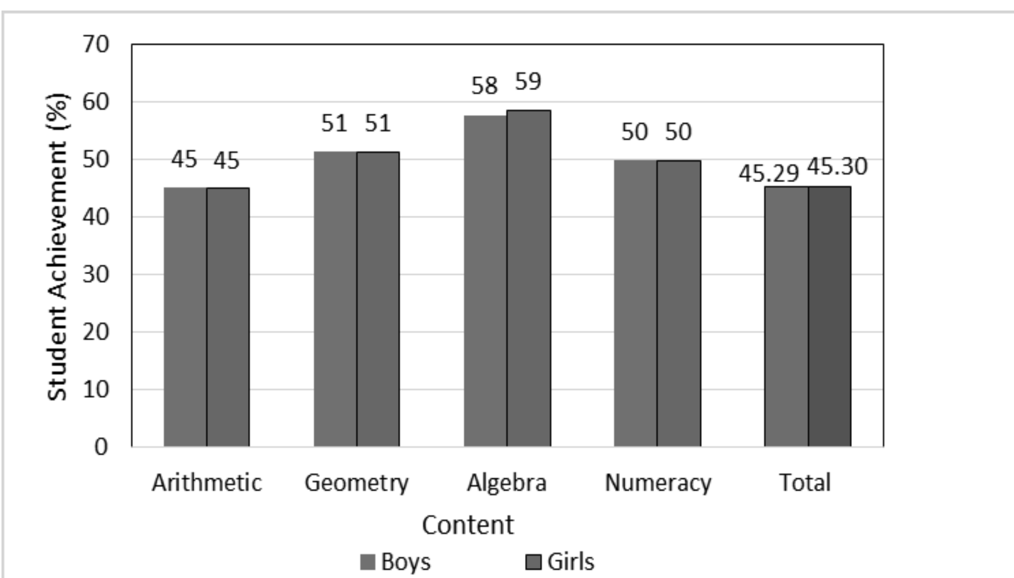
Gender and Student Achievement

In previous assessments, there were significant differences in achievement between boys and girls (see ERO, 2013; EDSC, 2008). Situation has changed radically as the difference in achievement score between boys and girls has reduced remarkably. Basic achievement results of boys and girls are presented in table 3.13. Figure 3.9 further compares the achievement of girls and boys in various content areas. The figure shows that in each content area boys have outperformed girls.

Table 3.13 Student achievement of boys and girls in Mathematics

Gender	N	Mean	SD	SE	Minimum	Maximum
Male	8280	45.3	22.25	0.245	0	98
Female	8047	45.3	22.62	0.252	0	98
Total	**16327**	**45.3**	**22.44**	**0.176**	**0**	**98**

Figure 3.9 Comparison of achievement between boys and girls in different content areas

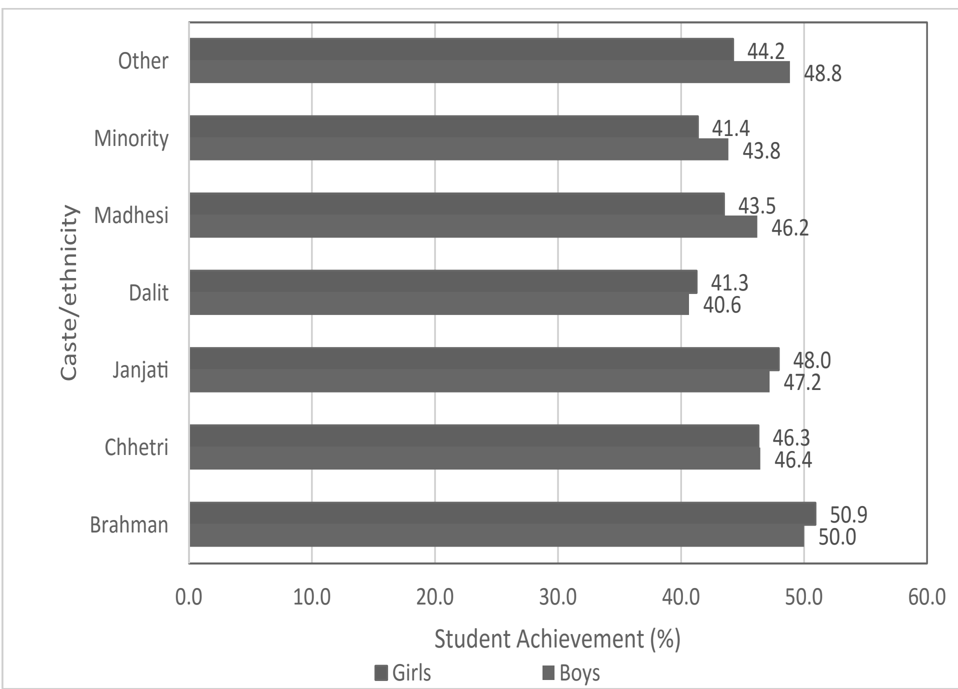


The average achievement of boys (45.29%) and girls (45.30%) is almost equal in Mathematics. However, there is slight difference in the performance of boys (58%) and girls (59%) in Algebra as girls are performing better by 1%. Therefore, there is no significant difference in achievement of boys (45.29%) and girls (45.30%) in Mathematics, which is a positive sign from equity point of view.

Gender and ethnicity/caste

There are gender differences in the achievement of various ethnic/caste groups in Mathematics. For example, the difference in achievement between boys and girls is the highest for Madhesi, Minority and Other communities by 3%, 2% and 5% respectively; that is, boys' performance is higher than the girls' and for minority. However, for other castes/ethnicities, namely Brahman, Janjatis and Dalits, girls have outperformed boys by very small gap (less than 1 percent). Nevertheless, difference is not found in total. Gender gap is prominent in 'Madhesi', 'Minority' and 'others' categories. This comparison is shown in figure 3.10.

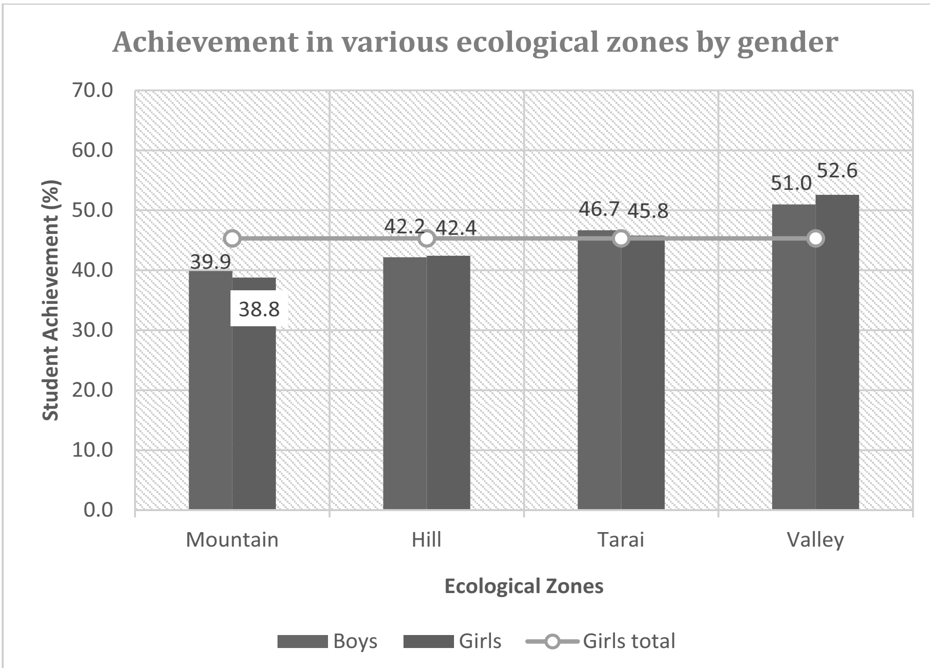
Figure 3.10 Comparison of achievement between boys and girls in various ethnic/caste groups



Gender and Ecological Zone

Achievement of boys and girls over the Ecological Regions is presented in figure 3.11.

Figure 3.11 Comparison of achievement between boys and girls in various Ecological Zones

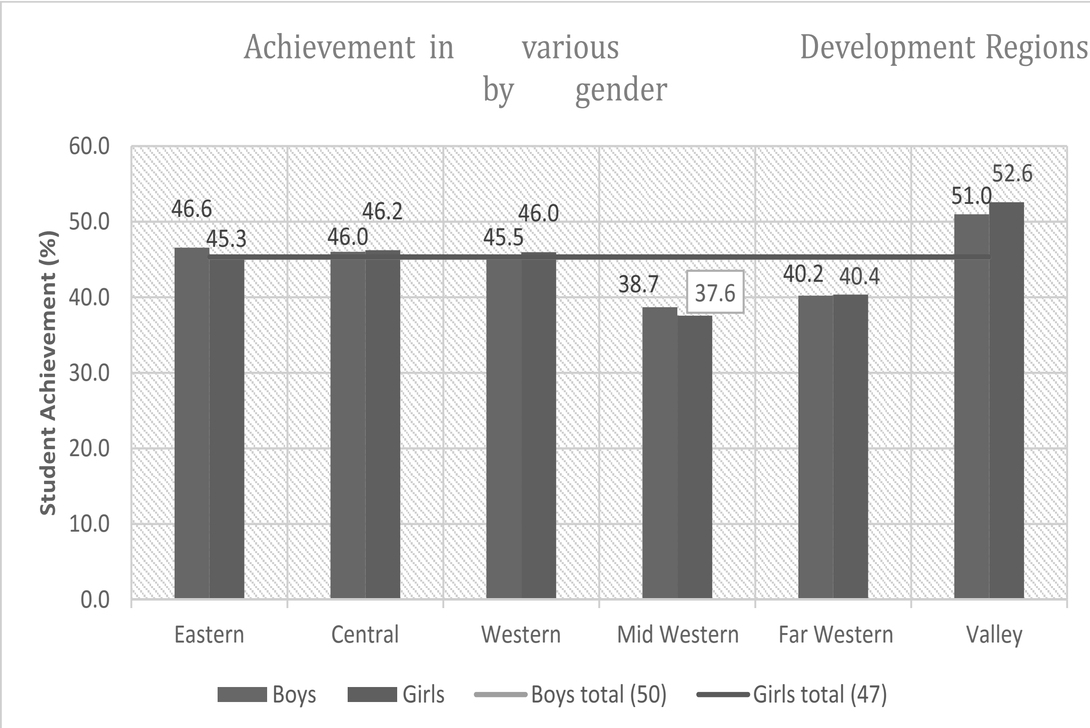


The achievement of girls and boys in Mathematics differs significantly across Ecological belts ($p < 0.05$). In the Kathmandu Valley, girls outperformed the boys; but in other Regions, there is no statistically significant difference between boys and girls, though boys are somehow better than girls (see figure 3.11).

Gender and Development Region

Development region-wise achievement differs between boys and girls; however, the difference is very small except in valley as boys outperform girls by 1 percent point in Eastern and Mid-western region whereas girls outperform boys in the Valley by 1.6 percent – which is presented in comparative form in the figure 3.12.

Figure 3.12 Comparison of achievement between boys and girls in various Development Regions



Dataset shows that achievements of boys and girls are the same in Central and Western Regions. Girls have shown better performance significantly in the Kathmandu Valley ($p < 0.01$). However, there is no significant difference between the achievement of boys and girls in total. It shows that there is no gender disparity in the achievement in Mathematics in Grade 3.

In Grade 3 Mathematics, the performance of boys is found 1 percent higher in Eastern and Mid-western Regions (not significant statistically); girls' performance is 2 percent points higher than boys in the Kathmandu Valley (which is statistically significant); however, the difference between boys and girls is not significant in total. This reveals that there is no gender inequality in the achievement in Mathematics.

3.4 Selected Explanatory Factors and Achievement

Several factors have been handled in the previous section such as geographical factors including Districts, Ecological Zone and Development Region; school related factors including school type and school location. Some individual related factors are also handled, such as home language, caste/ethnicity and gender.

This section deals with socio-economic status (SES) of the students' families, paid work beyond the school time, students' attitude towards Mathematics as a school subject, age of the student, and support provided for studies as the main family and individual related factors. Similarly, this section presents school and teacher-related factors, the availability of textbooks, homework given by the teacher and selected activities in school.

Achievement of Students from Various Socio-economic Status (SES)

The variables indicating the socio-economic status were grouped into seven categories. The seven categories of SES are mother's education, father's education, mother's occupation, father's occupation, home possessions, home accessories, and type of school attended. Finally, SES was estimated based on seven indicators related to the economic, educational, and occupational background of the family. In this section, the education of the parents is further elaborated, and the achievement in Mathematics is analysed in relation the literacy of parents.

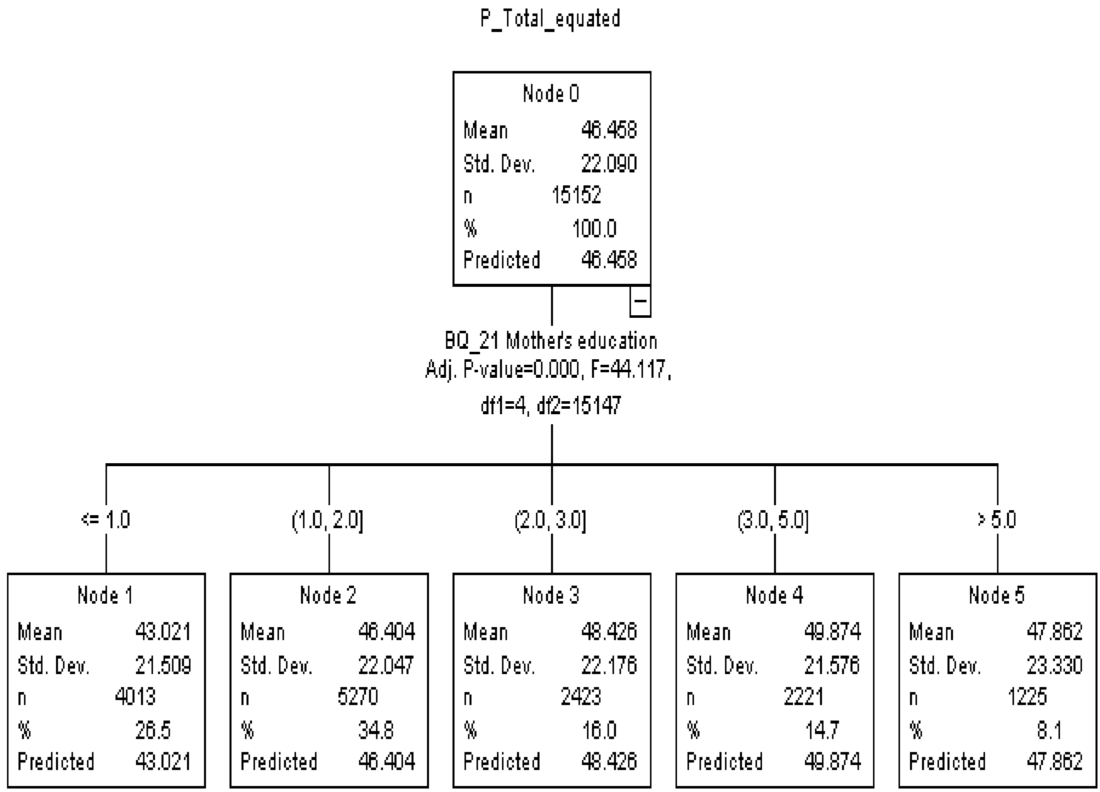
Several SES-related variables were analysed by using a data-mining tool of SPSS – the Decision Tree Analysis (DTA). The method was very effective in finding the cut-offs of the predicting variable, such as mother’s education, and classifying the factor into several groups which differ statistically in the most significant way from each other in relation to student achievement.

Parents' education

In the background questionnaire of NASA 2015, parents’ education is divided into the following seven categories: 1) Illiterate, 2) Literate, 3) Grade 10, 4) SLC, 5) Certificate level, 6) Bachelor’s level, 7) Master’s level and above.

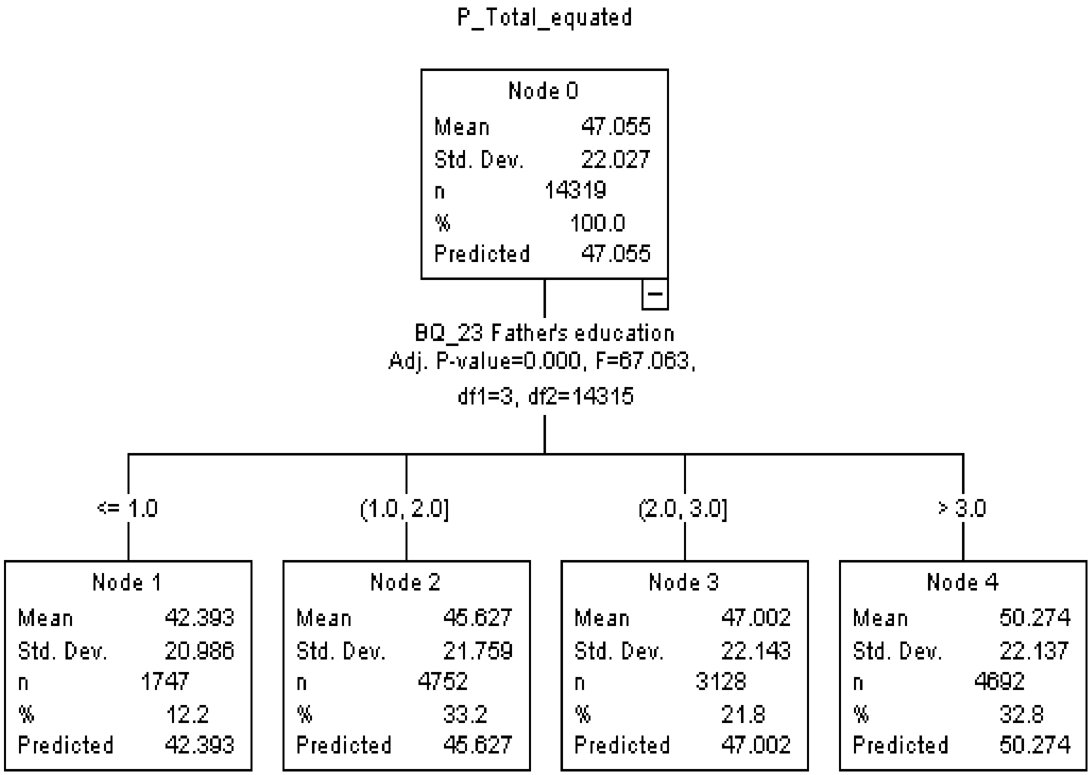
DTA classifies mother’s education into five groups with statistically significant differences in students’ achievement levels. The data shows that students having illiterate mothers obtained only 43% of the maximum score on average, which is the least score on average compared to the students having other categories of mothers in terms of their education. The achievement of students has increased along with the increase in mothers' educational level from literate to Master’s Degree. The figures 3.13 and 3.14 below explain these results. In each group, the number of illiterate and just literate mothers is high enough to make a credible prediction. The difference between each group is statistically significant ($p < 0.001$). Mother’s education explains 1% of variation in achievement ($\eta^2 = 0.01$), indicating a small effect size ($f = 0.11$). In the similar manner, student achievement is also influenced by father's educational level. Father’s education explains 1.5% of variation in achievement ($\eta^2 = 0.015$), indicating a small effect size ($f = 0.12$). As the father is illiterate, student achievement is 42%. When father is just literate, average achievement goes up to 46%.

Figure 3.13 DTA of mother's education and students' achievement in Mathematics



In similar vein, the DTA categorizes father's education into seven groups. Combining both categories, a comparative illustration of relation between father's education and student average achievement percentage is presented in figure 3.14. A positive relationship with parent's education and student achievement is seen.

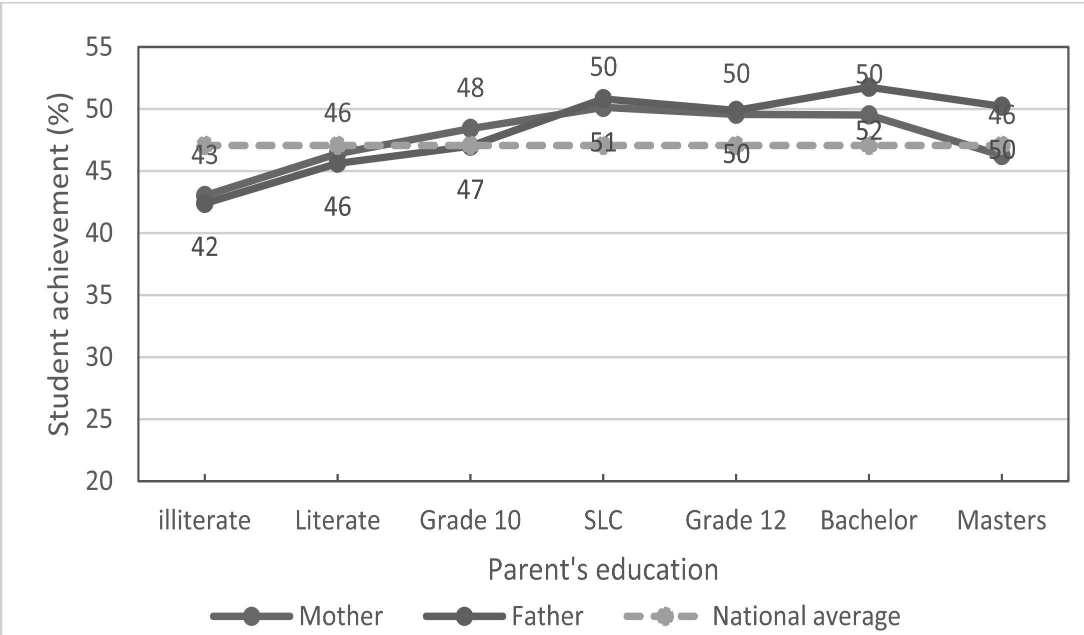
Figure 3.14 DTA of father's education and students' achievement in Mathematics



When parents (father or mother) are illiterate, student achievement is below the national average (42-43 percent). When mother is Grade 10 pass or father is SLC pass, student achievement is above the national average. The difference in achievement due to parent's education is statistically significant ($P < 0.001$). The effect of mother's education is 1 percent and the variation due to mother's education is low (Cohens' $f = 0.13$). Effect of father's education is about 2% and the variation due to father's education is also low as in case of mother.

From the data related both father and mother, a comparative line graph of mother's and father's education level with their corresponding average achievement is shown in figure 3.15.

Figure 3.15 Relation between mother's education and students' achievement in Mathematics



The figure 3.15 shows that the effect of mother's education is more productive than father's educational level as parents are educated up to SLC (about 1% score difference in each level). Above the SLC, father's education seems more effective (up to 4% score difference).

Result explains that the educational level of parents predicts the children's achievement level in Mathematics. When both parents are illiterate, achievement of their children become low compared to the children of other parents with SLC or higher level of education.

Parents' occupation and student achievement

The occupations of mothers were grouped into the following eight categories: 1) agriculture and household chore, 2) only household chore, 3) work abroad, 4) teaching 5) other job/service, 6) business, 7) daily wage and 8) work in other's home (maid). The DTA was used to find statistically the most deviating occupation groups related to student achievement (see fig. 3.16). The differences between the groups are statistically significant ($p < 0.001$). Student achievement is the lowest (41 percent) when his/her mother works in other's house or involves only in household chores, and the student achievement is the highest when his/her mother's job is business (54 percent). It shows that either economic or intellectual ability or both at home gives positive impact on children's achievement. Mothers' occupation shows 3% variation in data ($\eta^2 = 0.03$) with a medium effect size ($f = 0.18$).

Figure 3.16 DTA of mother's occupation and students' achievement in Mathematics

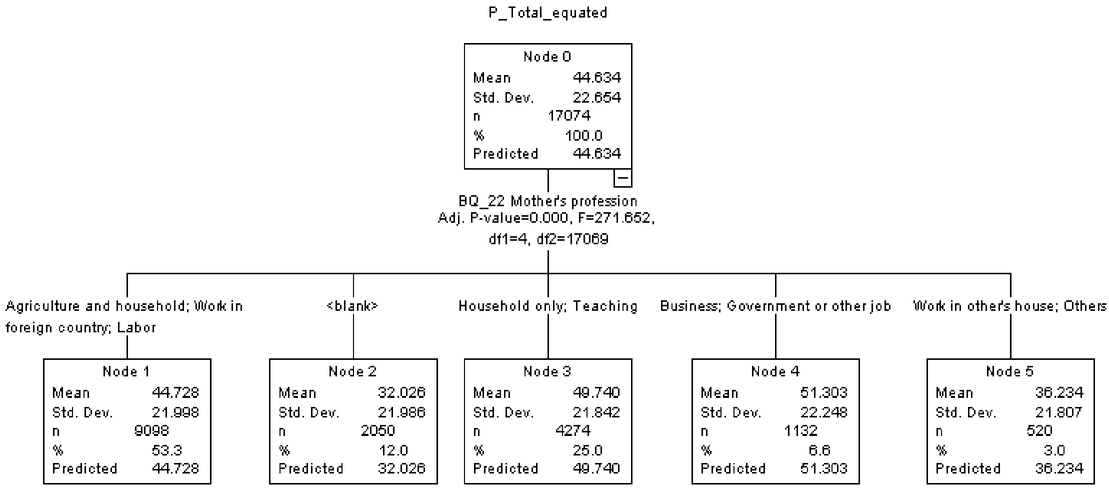
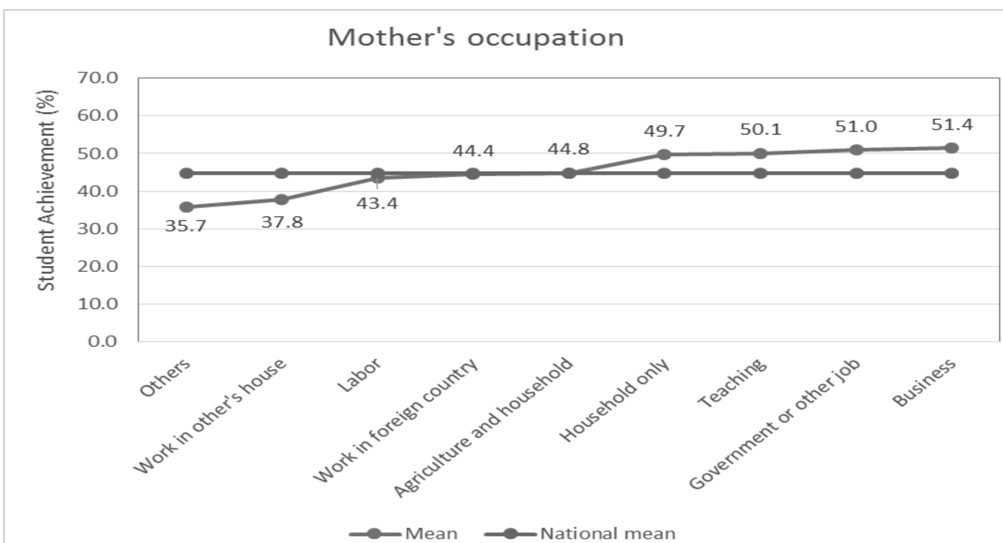


Figure 3.17 Comparison of mothers' occupation and students' achievement in Mathematics



Mothers' occupation can be categorised into three major groups based on the impact on their children's achievements in Mathematics. They are mothers of the students obtaining the average of (1) 43.5% or lower scores who were involved either in labour, work in other's home or in undefined work categorised as 'others'; (2) 44 to 45% who were working in foreign country or involved in agriculture, and (3) 49 to 51.4% who were involved in teaching, household chores, government job or business.Variation between lowest and highest achievement groups due to mother's education is statistically significant ($p < 0.001$); and the variation due to mother's education is medium (Cohen's $f = 0.25$), which explains 6% variation in achievement ($\eta^2 = 0.06$).

Similarly, DTA categorizes fathers' occupation into six groups, as given in figure 3.18. The ANOVA shows the difference between the achievement of students whose father is involved in agriculture or household works (44.3%) and in business, government job or teaching job (53.1%) is statistically significant *($p < 0.001$)*.

Figure 3.18 Comparison of fathers' occupation and students' achievement in Mathematics

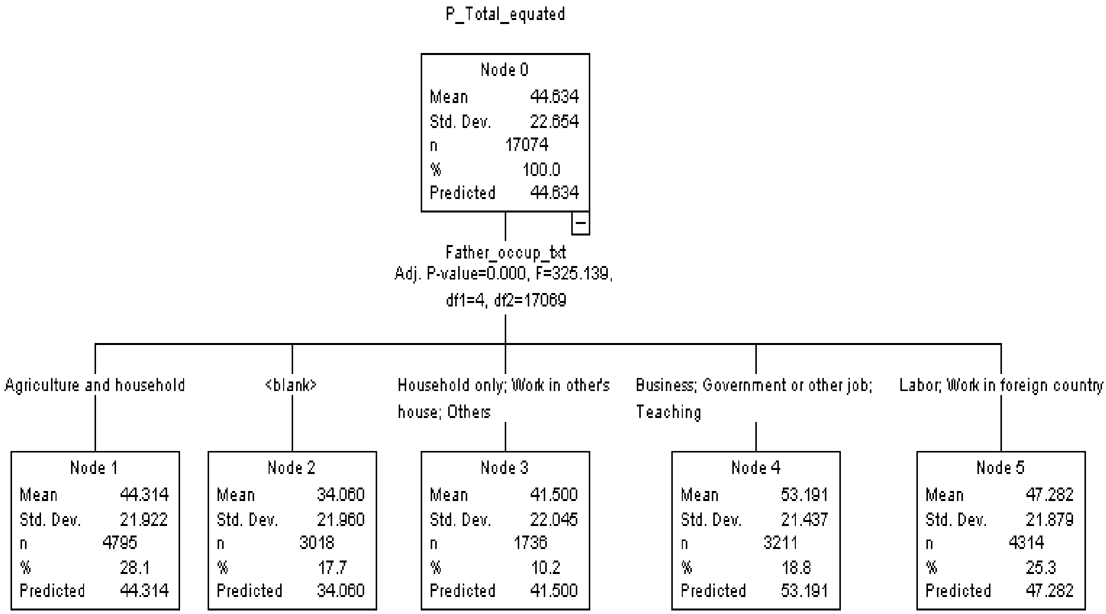
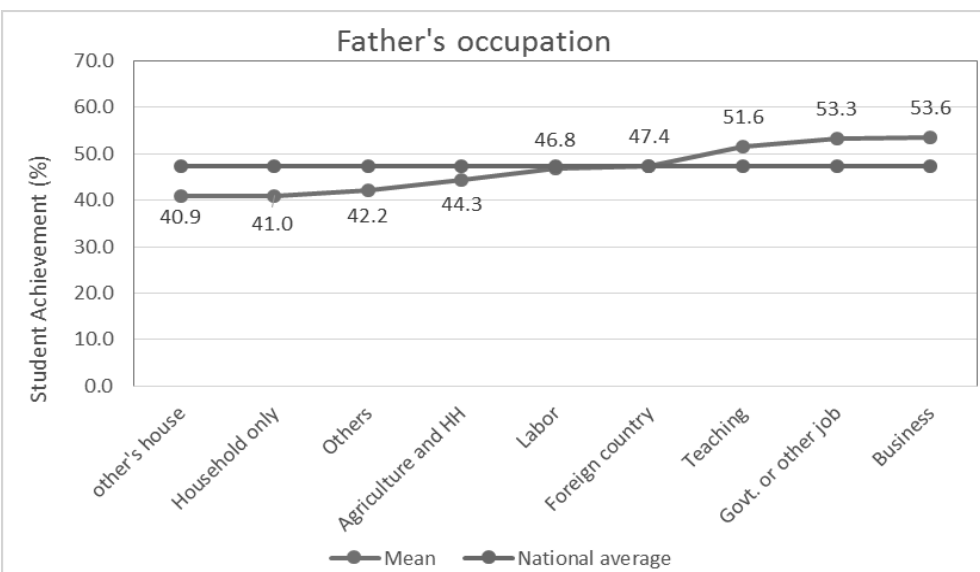


Figure 3.19 Comparison of fathers' occupation and students' achievement in Mathematics



Father's occupation can be categorised into three major groups based on the impact on their children's achievement in Mathematics. They are fathers of the students obtaining average of (1) 44% or lower scores who were involved in household related works; (2) 46 to 47% who were labourers or working in foreign countries, and (3) 52 to 54% who were involved in teaching, government job or business.

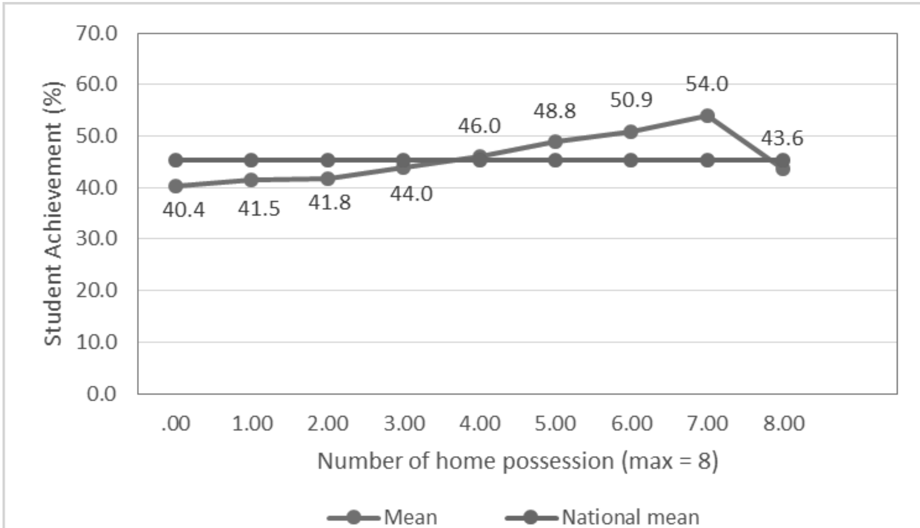
Dataset shows that when parents are involved in labour, household work, agriculture and foreign country, students' achievement is significantly lower than the students whose parents are involved in teaching, government job and business. Particularly, students whose mothers are involved in labour or works at other's home are achieving the lowest score.

Home possessions and accessories

There were two kinds of home possessions defined in the background information questionnaire of students. One is related to the facilities that are supportive for the students' study at home. These facilities are, for example, a table for study, a separate room for student, a peaceful place for study, a computer for school work, software for computer assisted learning, internet facilities, literary magazines, access to classical literature, poetry books, or artistic things like pictures, dictionary and other books that support children for their study. Another type of home possession includes the different types of normal home accessories, such as the number of mobile phones, televisions, computers, and radios.

There were 12 questions in the background questionnaire filled by the students related to home possessions. However, the factors 'a separate room for study', 'computer for school work', 'software for learning', and 'own calculator' have shown negative correlation to the achievement. Besides these negative items, all other 8 positive indicators are recorded, each scored 1 if the student had access to the possession. Adding these items up, the maximum score is 8, indicating that the students have the access to all of the possessions. The result shows that higher access of home possessions provides high score to the students; and hence, the lower the score the fewer are the possessions at home. Figure 3.20 shows relationship between home possessions and achievement level, whereby the achievement level of the student goes up when there is access to home possessions. Pearson product moment correlation coefficient between the achievement level and the factor ($r = 0.15$) is statistically significant ($p < 0.001$) though it is not very high, and the effect size is moderate ($f = 0.19$).

Figure 3.20 Relation between home possessions and achievement in Mathematics



The figure shows the relation between the home possessions and achievement in Mathematics. It is evident from the figure that when the number of accessories increases, the achievement also increases. When home possession is none (zero), the achievement is very low (40.4), when 7 out of 8 are present, the achievement is highest (54.0). However, the result is anomalous (achievement decreases) when all the home possessions are available.

In the background questionnaire, the question asked to the students was: *"How many of the following accessories do you have in your family?"* The options were 0, 1, 2 and 3 or more. After dichotomizing the items individually by using meaningful cut-offs found with ANOVA and DTA, (see, table 3.14), only mobile and television were added. The maximum score was 2, indicating that the students possessed at home a set of all the accessories.

Table 3.14 Dichotomizing the indicators for home accessories

Accessory	Cut-off for 1	Cut-off for 0
Mobile phone	1	0
Television	1	0
Computer	1	0 (negative correlation, so discarded)
Vehicle	1	0 (negative correlation, so discarded)

The relation between the availability of home accessories and achievement is presented in the following figure.

Figure 3.21 Relation between home accessories and achievement in Mathematics

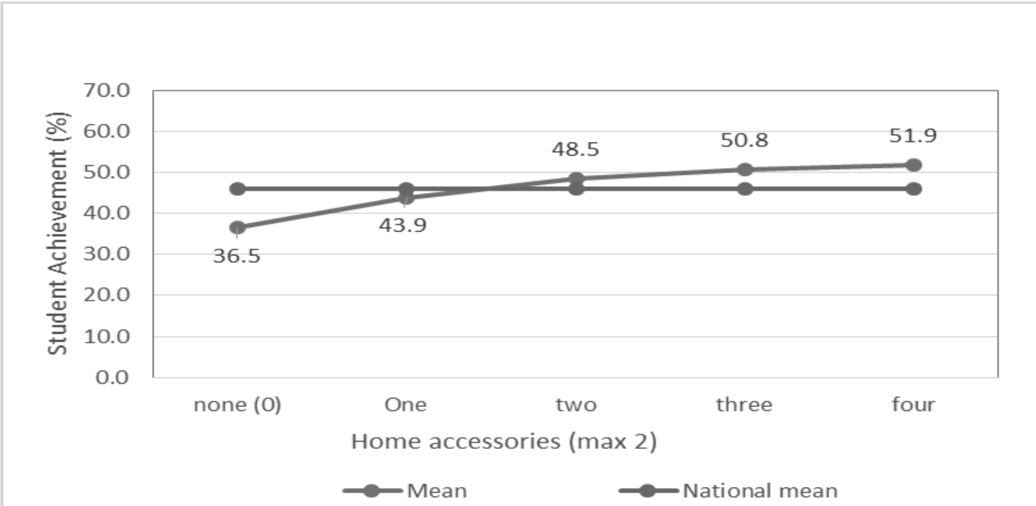


Figure 3.21 clarifies how the number of home accessories affects student achievement. The achievement is 36.5% if none of the accessories is available, and it is 51.9% if all are available. Availability of all the stated facilities indicates a high SES of the family and therefore it can be linked with SES of the family. Correlation between home accessories and achievement is $r = 0.20$ ($p < 0.001$) which is positive, but not very high. The effect size of $f = 0.21$ indicates a high size of difference between the groups. The difference between the lowest group with no mobile phones and television and the highest group with both of accessories is remarkable (15.4%).

Data shows that when the children have 0 to 3 home possessions out of the 8, the achievement level is lower than the national average, and the achievement is higher than the national average for those having more than 3 home possessions. The average achievement score of students from the families with 7 possessions is the highest (i.e., 54%). The same pattern shows with home, as the results are very poor (36.5%) if there is no any one of the four accessories; and when there are all the accessories, the results are remarkably high (51.9%). A negative correlation is found when families have computer, internet, computer assisted learning software and more than one car with the student achievement.

SES and achievement

The socio-economic status was analysed based on seven indicators, which were dichotomized first. The seven variables indicating socio-economic status are: mother's education, father's education, mother's occupation, father's occupation, home possessions (8 items), home accessories (4 items), and the type of school attended by students (private or community school). Scores in each of the variables were added and the total score of each variable was changed into percentage (P-SES). P-SES represents the percentage of SES the student possesses, 100 percent denotes that the student has the highest SES measured with these variables; that is, all the seven indicators of SES are positive, and 0 refers to the lowest possible SES, which means all the seven indicators of SES are negative. Analysis of the PSES by using Univariate GLM (that is the Regression modelling) shows that there is a strong relation between SES and achievement. Figure 3.22 below presents the relationship between SES of the students and the achievement.

Figure 3.22 Relation between the SES and achievement in Mathematics

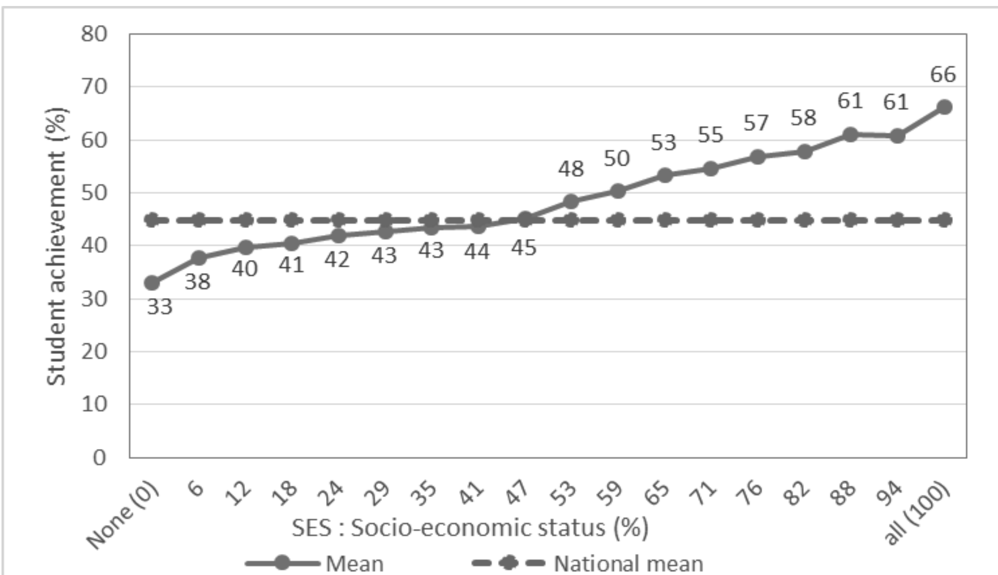
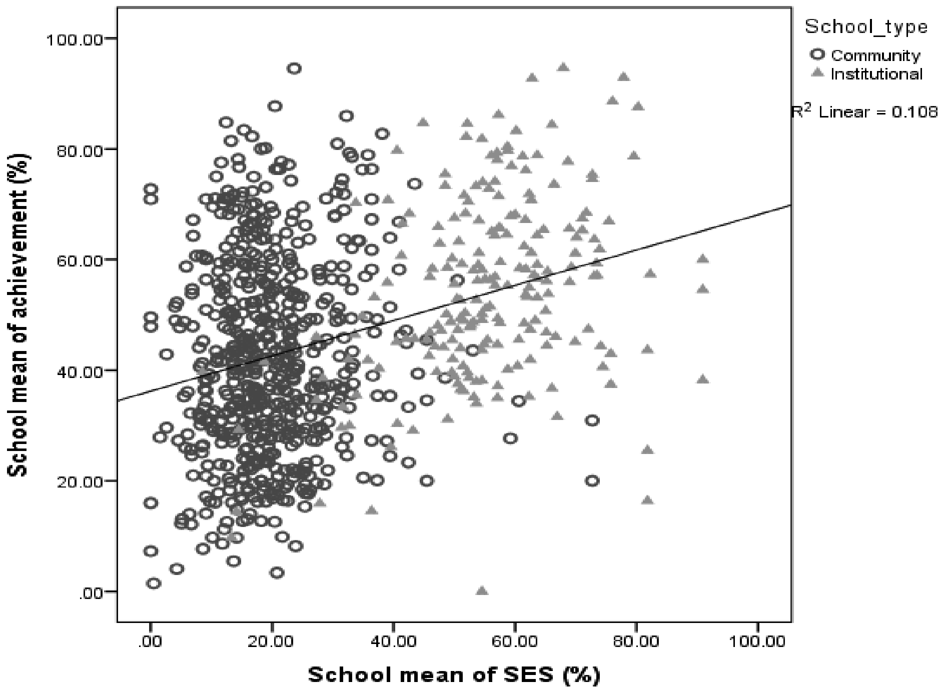


Figure 3.22 shows a positive relationship between SES and the achievement; the correlation between the variables is $r = 0.26$ which is a significant association ($p < 0.001$). The differences between the various SES groups are statistically significant ($p < 0.001$) and the effect size is high ($f = 0.29$); that is, the highest and lowest groups differ from each other remarkably. SES explains 8% of the student variation in data ($\eta^2 = 0.08$).

The dataset suggests that the socio-economic status plays a vital role in student achievement. The difference in achievement between the lowest and highest SES groups is remarkable.

Analysing further with the scatter diagram based on SES and achievement, figure 3.23 shows that two types of schools (community schools in circle and institutional schools in triangle) fall into two groups. The institutional schools with relatively higher SES are concentrated more on relatively high performing group, whereas community schools with relatively lower SES fall from very high performing to very low performing group concentrating more towards low performing group.

Figure 3.23 Distribution of achievement by socio-economic status and type of schools in Mathematics

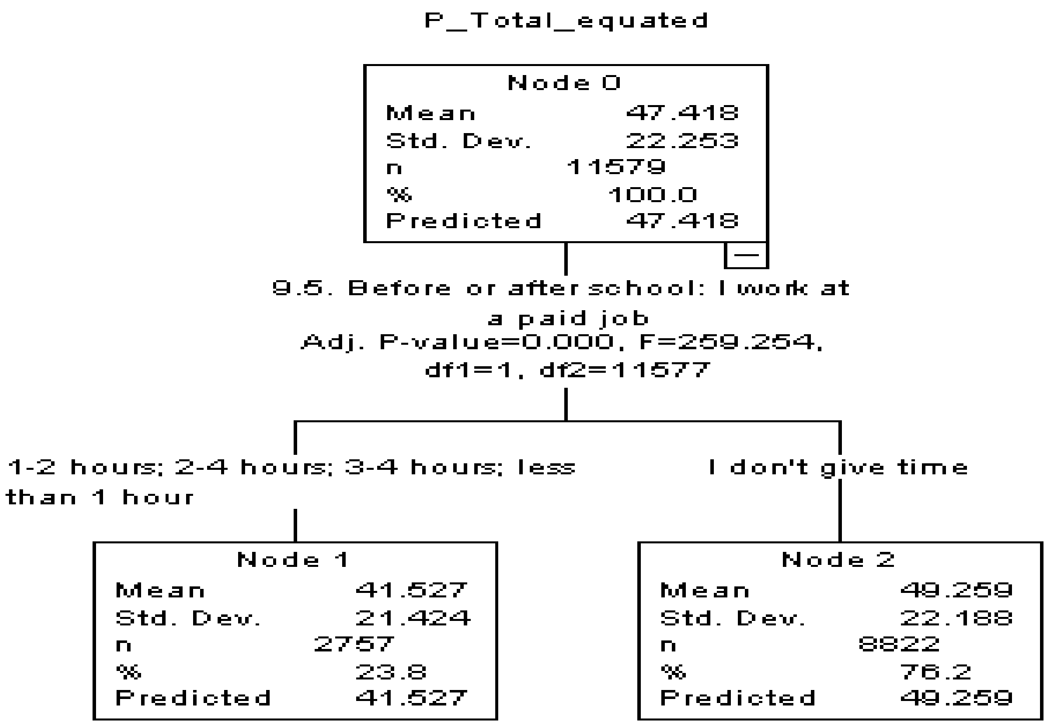


The dataset reveals that the students from institutional schools with relatively high SES are concentrated more towards high performing group, and the students in the community schools with relatively low SES are concentrated more towards low performing group. Overall, SES indicator explains 10.8 percent of the school mean of student achievement.

Working Beyond School Hour and Achievement

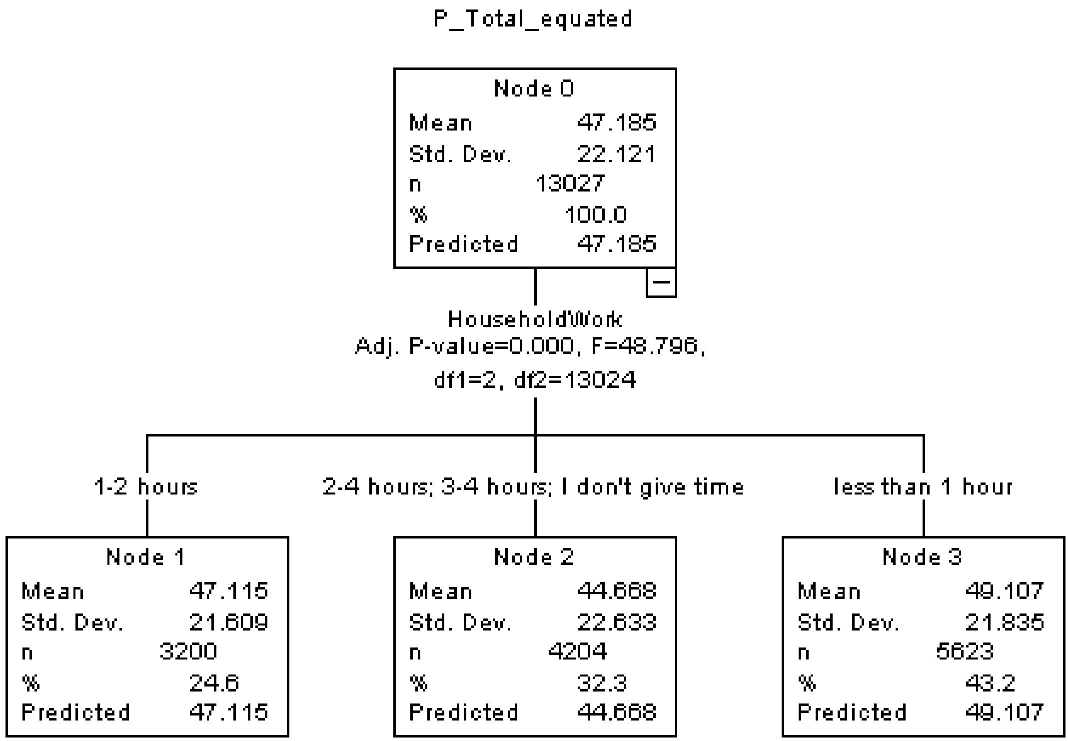
There were questions about student background in the questionnaire related to their activities beyond the school hours. The questions were related to working before/after the school hours in a paid job and involvement in household chores. The values of each variable are divided into five categories and marked on a five-point scale. The scale had 0 point for not working at all, 1 point for the work of less than 1 hour per day, 2 points for the work of 1 to 2 hours per day, 3 points for the work of 2 to 4 hours per day, and 4 points for the work of more than 4 hours per day. Data shows that when the children are not engaged in paid work at all, their results are notably above the national average (49.3%) (see, figure 3.24.). If the students work for paid job even less than an hour to four hours, the results are lower (41.52%) than the national average. Further, if they were working more than 4 hours, the results were remarkably low (23%). The differences are statistically significant ($p < 0.001$) though the effect size is moderate (f = 0.21), as most children do not need to work in a paid job. Working beyond the school hours for earning money indicates that the family is poor and extra earning is needed. When the student needs to work for more than 4 hours per day there is no time or energy to be involved in studies or homework.

Figure 3.24 DTA of paid work and achievement in Mathematics



Next, the DTA (fig. 3.25) shows that when the children spend more than 1 hour per day in household chores, the results are poor. However, when the amount of time spent in household chores is less than one hour per day, the achievement level is higher than the average (49.1%). Differences are significant ($p < 0.001$) though the effect size is very low ($f = 0.008$) as around 38 percent children participate in household work for more than 1 hour per day. However, very low effect size suggests very low effect on achievement.

Figure 3.25 DTA of household work and achievement in Mathematics



The dataset shows that paid work of the students beyond the school hours reduces the achievement. Getting engaged in unpaid household work up to one hour gives positive effect in student achievement, whereas when the children need to work for more than 2 hours per day (either paid or unpaid), the achievement level drops down remarkably.

Students' Attitude towards the Subject of Study and Achievement

The attitude about the subject tells us what the students think about Mathematics and its usefulness in their daily life and future utility. The correlation between Mathematics achievement and attitude toward Mathematics has widely been studied, and it has been found that there is a relationship between the attitude of the students and achievement though the connection is not always clear (see, for example Metsämuuronen 2012a; 2012b; House & Telese, 2008; Shen & Tam, 2008; Kadijevich, 2006; 2008).

In NASA 2015, the same shortened version of Fennema-Sherman Attitude Scales (FSAS) (Fennema & Sherman, 1976), as used in several international comparisons like in TIMSS and PISA studies, was used to identify the relation between attitude towards the subject and their achievement. The original scale included nine dimensions, but in these international comparisons, only three were used with four items on each and two negative items on each of the first two dimensions. The names of the factors can be “Liking Math”, “Self-Efficacy in Math”, and “Experiencing utility in Math” (compare naming in, e.g., Kadijevich, 2006; 2008). Factor analysis was used to identify the factors of the responses in FSAS and the negative items were reversed to make the whole test unidirectional. As in several countries of Asia, the expected factor structure cannot be found in Nepal (for a deconstruction of the test scales, see Metsämuuronen, 2012a; 2012b). Hence, only the total score is used to show the relationship between attitude and achievement. The relation between the attitude, which is divided into deciles of somehow equal number of students, and achievement score, is shown in figure 3.26.

Figure 3.26 Relation between students' attitude and achievement in Mathematics

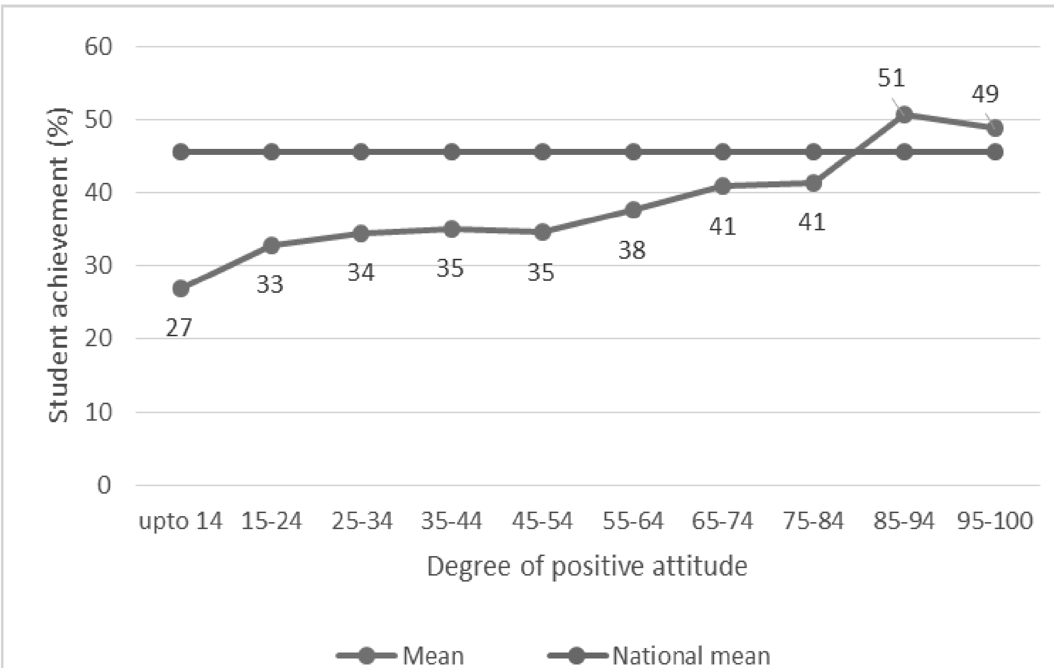


Figure 3.26 shows that the difference between the lowest attitude group (27%) and highest attitude group (54%) is remarkable. However, the real difference in achievement is found only in the two highest attitude groups, that is 49% and 51%. The correlation between positive attitude towards Mathematics and achievement is $r = 0.23$ ($p < 0.001$); the effect size is moderate ($f = 0.26$).

As the data reveals, the more positive the attitude towards Mathematics, the higher is the achievement. The data also supports the fact that positive attitude influences the achievement positively.

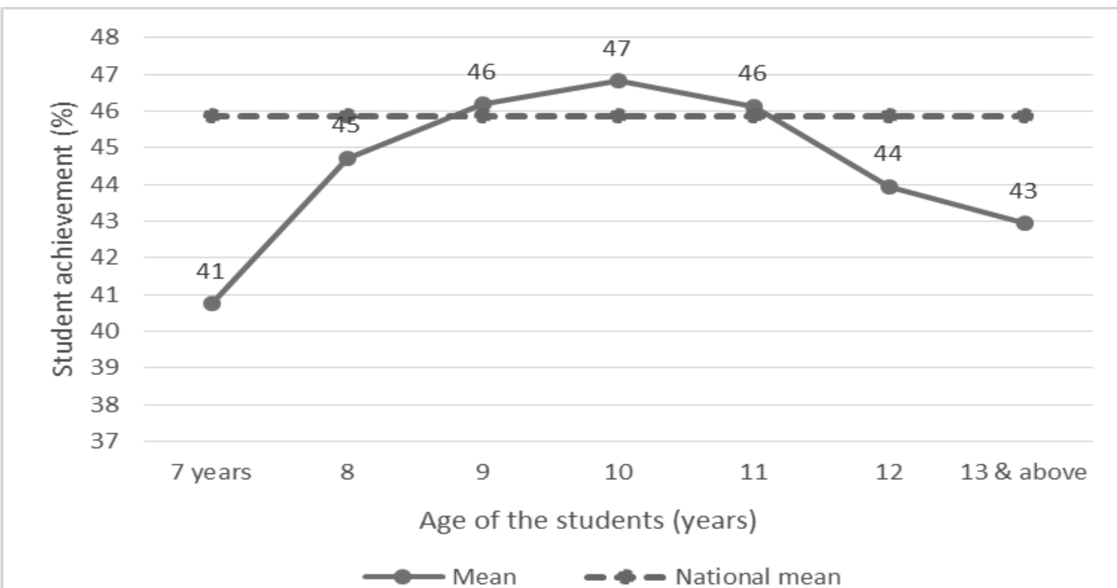
Age of the Student and Achievement

Although official age for Grade three students is 9 years in Nepal, the age of the students attending Grade three varies widely. Some students have mentioned their age below 7 years and some above 13. All the students below 7 were encoded as ‘missing’, and all students above 13 were encoded as '13 years and above'. The descriptive statistics of the mean in each year is presented in table 3.15 and further depicted in figure 3.27.

Table 3.15 Descriptive statistics of the students' achievement in different age groups

Age of the students	N	Mean	SD
7 years	146	40.7	23.20
8 years	1291	44.7	23.30
9 years	3725	46.2	22.01
10 years	5337	46.8	22.17
11 years	2573	46.1	22.28
12 years	1469	43.9	22.45
13 years & above	626	42.9	22.49
Total	**15167**	**45.9**	**22.33**

Figure 3.27 Relation between the age of student and achievement in Mathematics



Data explains that the best achievers are those students, who are at proper age for Grade three, that is, the students of 9 to 11 years. The achievement level is remarkably low when the students are above the age of 13 year. It clearly shows that students with higher age have weaker results. Students of higher age are those who either started school much later than they should have, or repeated the classes. Correlation between the variables is r = -0.015. However, the correlation is not significant.

Dataset indicates that the highest performance is found in the students studying at their correct age group, that is, at the age of 9 to 11 years. Otherwise, the achievement decreases as the age of the student increases, which results in negative correlation.

Support for Study to the Student and Achievement

The relation between the support provided for study and achievement was analysed based on the responses to the question: *"Who supports you when you do not understand what you have studied or felt difficult?"* In the question, students were asked to select only one option. This assessment did not consider the case where there were more than one supporters. The descriptive statistics of the supporters are given in table 3.16.

Table 3.16 Achievement score in Mathematics in relation to the support provided to the students

Support Providers	N	Mean	SD
No one	551	47.2	22.72
Teacher	850	48.9	22.99
Brother/sister	7182	47.1	21.72
Mother	1931	46.2	21.90
Father	2660	45.6	22.85
Tuition	1566	44.0	21.68
Any two or more	2333	32.9	22.71

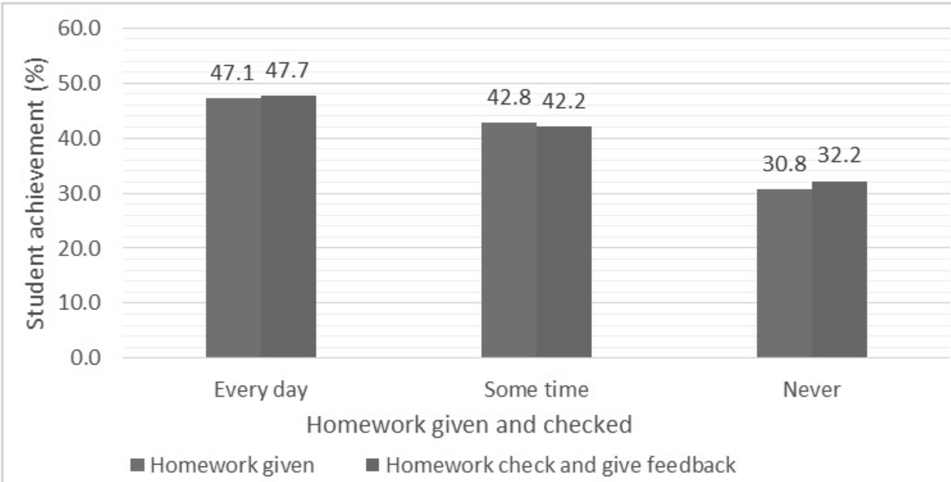
Support in Mathematics is necessary for students to achieve better score. There is about 2 percent difference between the students who did not receive any kind of support and those who received support from the teacher. Interestingly, data shows that those students obtained more score who studied independently (47.2%) than those who received support from their family members (45.6 to 47.1%). Students who obtained lowest score among the groups are those who are dependent on tuition (44%). However, there are some students, who reported that they get help from more than one supporter. They are the lowest achievers (32.9%).

The result shows positive correlation between supports from teacher even after the school hour and student achievement, as the achievement of the students having support from teacher is higher than that of the students receiving support from their family members. However, students studying independently without taking any support from family members and teachers obtained better score than those who had received support from them. Students who need to seek tuition or help from many people are the lowest achievers.

Homework and Achievement

Homework is considered as one of the ways to enhance learning. This can be drill, exercise, or evaluation tool. When homework is regularly checked and feedback is provided, it is likely to boost the level of achievement. Statistics related to homework given and feedback provided is presented in figure 3.28.

Figure 3.28 Effect of homework in students' achievement in Mathematics



If teacher assigns homework and provides feedback to the students regularly, then the students' achievement is higher (47.7%). If homework is not assigned regularly and feedback is not provided, the achievement score is below the national average (32.2%). When home assignment and feedback are never given, then the average achievement score is remarkably low (16.4%). The differences are statistically significant ($p < 0.001$). Those groups with no homework assigned are, however, very small and hence, the effect size is small ($\eta^2 = 0.02$ and 0.02).

Dataset reveals that if the teacher assigns homework and provides feedback to the students every day, the achievement raises higher than without assigning homework or not providing feedback.

Positive and Negative Activities in the School and Student Achievement

The activities of students and teachers determine the learning environment of a school. Bullying, for example, is one of the hindering incidents for the students in school that may affect learning. In students' background questionnaire, several school-related activities were asked to the students. Here, bullying is considered as one of the negative indicators, while students' impressions of school and teacher activities are the examples of positive indicators.

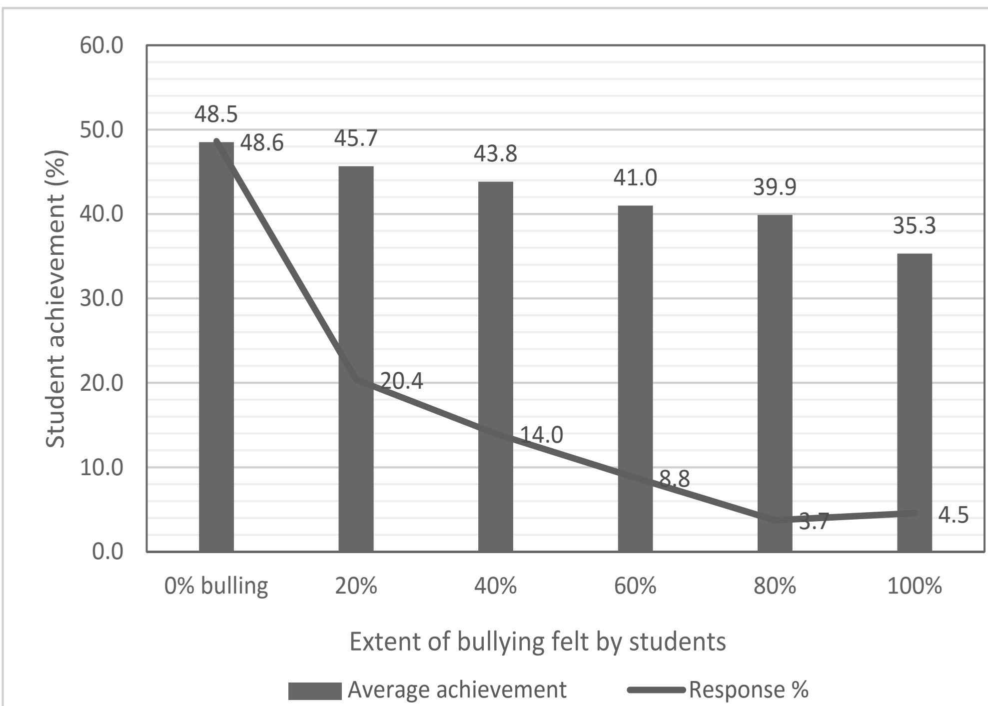
Bullying at school

Bullying is one of the problems in schools as it worsens the learning environment for the students. International Studies like TIMSS and PISA have emphasized to identify such indicators. In NASA 2015 student background questionnaire, five questions indicate the varieties of bullying that tend to happen in schools. All the incidences were stemmed by the phrase "*Which of the following activities happened in your school last month?*" The students' responses are presented in table 3.17 and further depicted in figure 3.29. 'No (%)' indicates the percentage of the students who responded that no such activity had happened in the school; and 'Yes (%)' indicates the percentage of the students who experienced some particular type of bullying within the last month. Alone, 31.3% of the students mentioned that, during the last month, something was stolen, which is an alarming sign of the system.

Table 3.17 Bullying and the achievement

Type of Bullying	No (%)	Yes (%)
BQ_30_1 Incidents: Something of mine was stolen	68.7	31.3
BQ_30_2 Incidents: I was hit or hurt by other student(s)	73.3	26.7
BQ_30_3 Incidents: I was made to do things I didn't want to do by other students	80.8	19.2
BQ_30_4 Incidents: I was made fun of or called nick names	76	24
BQ_30_5 Incidents: fellow students kept me outside without involving me in activities	81.9	18.1

Figure 3.29 Effect of bullying in the achievement of Mathematics



The sum of all five items (that is, the total 100 percent) is considered the indicator of highest bullying. Figure 3.29 shows the extent of bullying with the percentage of bullying and achievement of the students in each category of bullying. The line graph shows that the percentage of students encountered various levels of bullying. If only one activity of bullying is reported, it is categorized as 20% bullying, whereas if all five activities are reported, it is categorized as 100% bullying. Data shows that 48.6% of the students did not encounter any kind of bullying. One can infer from this that the remaining 51.4% have encountered at least one type of bullying, which is a remarkable number of students. About 8% of the students are experiencing a severe kind of bullying, which is the sum of the students who encountered 80% and 100% bullying. It is found, however, that learning outcomes are remarkably low only with 4.5% of the students who have encountered extreme bullying including all five types of harassments (35.3%). Students who did not feel bullying and those who encountered bullying of all five kinds (extreme form of bullying) have almost 13.2 percent achievement gap, although a very small number of students reported all kinds of bullying (n=718). However, the difference is statistically significant ($p< 0.001$) though the effect size is medium ($f = 0.16$). Though the extreme case of severe bullying is rare, bullying is found quite common in schools.

The dataset reveals that a large number of the students (51.4%) have encountered bullying in schools within the last month. Though the phenomenon does not have a moderate effect except in the group of extremely bullied students, all possible efforts should be made to root out the phenomenon from schools.

Positive activities at school

The activities that can boost the learning achievement of students are categorized as positive activities. Such positive activities at the school were asked to the students in two sets of questions listed in table 3.18. The table shows the responses of the students in all four categories, which are in the 4-point rating scale, anchored to 'fully agree' and 'fully disagree'.

Table 3.18 Students' response towards teacher and school-related activities in Mathematics

Teacher and Students activities	**Respondents in %**			
	Fully disagree	**Partially disagree**	**Partially agree**	**Fully agree**
Students get along well with most teachers	1.7	1.6	6.3	90.4
Most teachers are interested in student's well-being	1.7	1.6	6.5	90.1
Most of the teachers really listen to what I have to say	2	3	12.6	82.4
If I need extra help, I will receive it from my teacher	2.1	2	7.8	88.1
Most of my teachers treat me fairly	2.6	2.3	7.6	87.4
I enjoy to be in the school	1.5	1	3.8	93.7
Students in my school like me	1.9	1.9	11.3	85
Friends of my school want to their best in the study	1.7	2.1	6.1	90.1
Teachers' expect good results of their students	2.2	1.7	5	91.2
Average	**1.9**	**1.9**	**7.4**	**88.7**

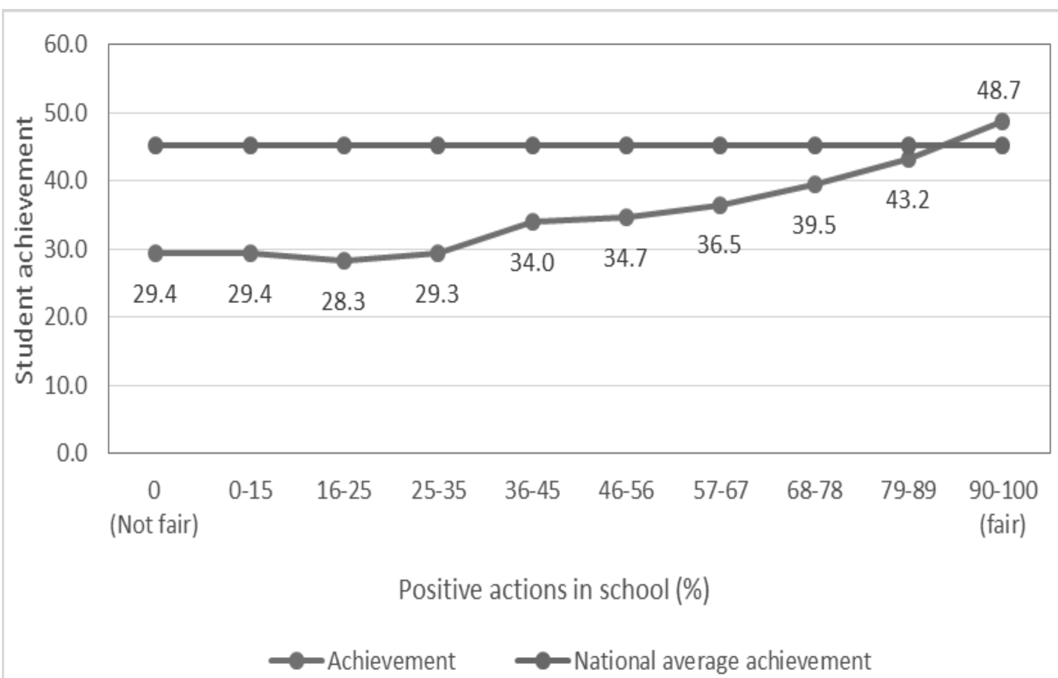
It was further analysed by recording the variables into two categories, that is, 1 for agree and 2 for disagree. Furthermore, the sum of nine indicators was converted into the percentage of maximum score to analyse the level of positive activities and its relation with achievement. The overall result is that the feeling of positive actions in the school relates positively with the student achievement. The correlation between the sum of nine positive activities and achievement is positive ($r = 0.26$, $p < 0.001$).

Univariate ANOVA finds four attitude groups in the indicator. These boundaries and descriptive statistics are presented in table 3.19 and further illustrated in figure 3.30.

Table 3.19 Students' responses towards teacher and school-related activities

% of total score	N	Achievement	SD
Not fair (0%)	116	29.4	19.36
0-15	241	29.4	21.45
16-25	274	28.3	16.92
25-35	272	29.3	19.47
36-45	482	34.0	20.05
46-56	560	34.7	19.76
57-67	529	36.5	20.18
68-78	791	39.5	20.91
79-89	1434	43.2	21.20
90-100	11619	48.7	22.14

Figure 3.30 Relation between positive actions at school and achievement in Mathematics



The data shows that there is a positive connection between the students' feeling about the school activities and achievement. After dividing the indicators into four groups based on DTA, the differences between the groups are statistically significant ($p < 0.001$); however, the effect size is small ($f = 0.27$). The learning achievement is higher than the average only when the students are extremely positive towards school and teachers' behaviour. However, the difference between the most positive and the most negative groups' achievement is notable, that is 19.3 percent score.

Dataset records that when students feel that the actions of the teachers and schools are ultimately good, the results are better than average. At the other extreme, if they feel that such actions are negative, the results are far below the average.

3.5 Summary of Findings

Basic Results

- Achievement score for Grade 3 in Mathematics is normally distributed. However, the schools are divided into two distinct categories: the high and low performing. The low performing population achieve the average of 25.5% of the achievement score and on the high performing population achieve around 63% of the mean. The difference between the populations is remarkable.
- The overall national average achievement score for Mathematics is 44.6%. For *Arithmetic, Geometry, Algebra* and *Numeracy*, the achievement is 44.5%, 50.7%, 57.4% and 49.3% respectively Students are somehow weak in arithmetic, and somehow better in the other content areas.
- Performance of large number of students is relatively low in higher cognitive abilities; about 20% of the students in Mathematics could not solve any one of the tasks requiring the higher skills.
- Students are good at recognizing the correct answer and in very fundamental contents like recognizing numbers, shapes, choosing correct answer from the options, very basic operations such as the basic manipulation of data and numbers, and calculations in single or few steps. They are much weaker in calculating in many steps, reasoning, problem solving, plotting, proving the theory or formula, and constructing the shapes and figures. In many cases, the students did not even start doing the open-ended questions, hence, the low score.
- Comparison of average achievement scores (raw) between NASA 2012 and 2015 shows that achievement in Mathematics in NASA 2015 is significantly lower the than in NASA 2012. While comparing latent ability there is no such variation.

Achievement with Various Diversity Factors

- Two highest performing Districts are Kathmandu from valley and Mustang from Western Development region. Among the low achieving (i.e., below national mean) 12 Districts, the lowest two are Bhojupr and Sankhuwasabha – both from Eastern Development Region. The lowest performing schools are mostly from Mid-Western and Eastern Regions.

- There is a wide difference between the student performances in four Ecological Zones. Students in the Kathmandu Valley have outperformed (51.3%) the other students. The achievement is the lowest in the Mountain area (39.9%). Moreover, average achievement of Tarai region (45.4%) is higher than that of Hill Region (41.6%).
- There is wide inequality among the Development Regions for children's opportunities to reach an adequate level in Mathematics. The difference between the lowest performing Region (Mid-Western, 36.9%) and the highest performing Region (the Kathmandu Valley, 51.3%) is remarkable, which is more than 24 percent. However, if the Kathmandu Valley is excluded from the analysis, the difference will be 8.7 percent. Performance of students in Mid-Western and Far-Western Regions is poor.
- The average performance of the students from institutional schools is 54% whereas it is 41% in community schools; that is, the difference is 13 percent, which is remarkable.
- Student performance varies significantly in terms of the location of the schools in Mathematics. Urban schools outperformed (50.9%) the rural schools (43.5%) by 7.4 percent score. However, the difference is medium.
- Nepali speakers have outperformed Non-Nepali speakers by 3.4%. Among all national language speakers, Gurung speakers achieved the highest score (53.9%) and Rai speakers achieved the lowest score (37.9%). Variation between highest and lowest speaking student groups is small.
- Achievement of students from Brahman caste/ethnicity have performed the highest and Dalit the lowest (42.1%) in total. Caste/ethnicity-wise achievement score varies in most of the Ecologicalregions. All of the castes/ethnicities performed the highest in the Valley; however, minority caste students performed best in Tarai Region (46.7%). Brahman, Janjati and other caste categories are performing lowest in Mountain. Chhetri, Dalit and Madhesi are performing the lowest in Hill; and minorities have performed lowest in the Valley.
- There is no significant difference in achievement of boys (45.29%) and girls (45.30%) in Mathematics. This is a positive sign from equity point of view. However, variation is found across castes/ethnicities, Ecological Zones and Development Regions. Highest difference is in Madhesi communities, that is, 10 percent. Similarly, in the Brahmins/Chhetri and "others" groups, the difference is 5 percent. In the Valley, girls have outperformed boys. Next, girls' performance is higher than that of boys by 2% in the Kathmandu Valley (which is significant).

Selected Explanatory Factors

- Result explains that the educational level of parents predicts the children's achievement level in Mathematics. When both parents are illiterate, achievement of their children is very low compared to the children of other parents with SLC or higher educational level.
- When parents are involved in labour, household work, agriculture and foreign country employment, student achievement is significantly lower than the students whose parents are involved in teaching, government job and business. Particularly, there is the risk of achieving low score when mother is involved in labour or work in other's home.
- When children have 0 to 3 home possessions out of the 8, the achievement level is lower than the national average; and for those having more than 3 home possessions the achievement is higher than the national average. The average achievement score of students from the families with 7 possessions is the highest (i.e. 54%). If there is no any one of the four accessories, the results are very poor (36.5%); and when there are all of the accessories, the results are remarkably high (51.9%). A negative correlation is found between student achievement and home possessions when families have computer, internet, computer assisted learning software, and more than one car.
- Socio-economic status plays a vital role in student achievement. The difference in achievement between the lowest and highest SES groups is remarkable. The students in the institutional schools, with relatively high SES, are concentrated more towards high performing group and the students in the community schools, with relatively low SES, are concentrated more towards low performing group.
- Paid work of the students beyond the school hours reduces their achievement. Getting engaged in unpaid household work up to one hour gives positive effect in student achievement, whereas when the children need to work for more than 2 hours per day (either paid or unpaid), the achievement level drops down remarkably.
- The more positive the attitude is towards Mathematics, the higher is their achievement. Data also supports the fact that positive attitude results in positive achievement.
- The highest performance is found in the students studying at their correct age group, that is, at the age of 9 to 11 years. Beyond this age boundary, the achievement decreases as the age of the student increases.
- Positive correlation is found between the support from teacher even after the school hour and student achievement, as the achievement of the students having support from teacher is higher than that of the students receiving support from their family members. However, students studying independently without taking any support from family members obtained better score than those who received support from them. Students who need to

seek tuition or help from many people are the lowest achievers.

- Dataset is evident that if the teacher assigns homework and provides feedback to the students regularly (every day), the achievement is higher than without assigning homework or engaging the students in homework without giving feedback.
- A large number of the students (51.4%) have encountered bullying in schools. Though the phenomenon does not have a moderate effect except in the group of extremely bullied students, all possible efforts should be put to root out the phenomenon from schools.
- When students feel that the actions of the teachers and schools are ultimately good, the results are better than average. At the other extreme, if they feel that such actions are negative, the results are far below the average.

Chapter 4: Analysis of Assessment Results for Grade 5 in Mathematics

4.1 Introduction

This assessment is based on the curricular competencies as set in the approved national curriculum of Nepal for Grade 5 in Mathematics. As per the approved curriculum, the overall objective of Grade 5 Mathematics is two-fold. One of the objectives is to develop in students the mathematical knowledge and skills needed in solving daily life problems; and the second objective is to provide students with the basic pre-requisites for continuing their study to lower secondary level.

This chapter begins with the introduction and presents the overall distribution of achievement. While doing so, it describes the achievements across various content areas of the curriculum, various levels of cognitive domain and various types of items. Considering temporal perspective, the results of this assessment are also compared with the results of first cycle study i.e., NASA 2012.

The third section compares the level of achievement with international studies- particularly with TIMSS, an international assessment. The fourth section of this chapter provides results associated with various diversity factors - including District, Ecological Zone, Development Rregion, school type, school location, language at home, and caste/ethnicity.

The fifth section of this chapter deals with the selected explanatory factors about the student achievement in Grade 5 Mathematics. The factors that explain achievement in Mathematics included in this section are parents' education and occupation, home possessions and accessories, socio-economic status (SES) of parents, work beyond school hour, age of the student, support provided for the study, availability of textbook, homework given and feedback provided, and negative and positive activities at school. The final section of this chapter briefly summarizes the findings of the assessment results in Mathematics for Grade 5.

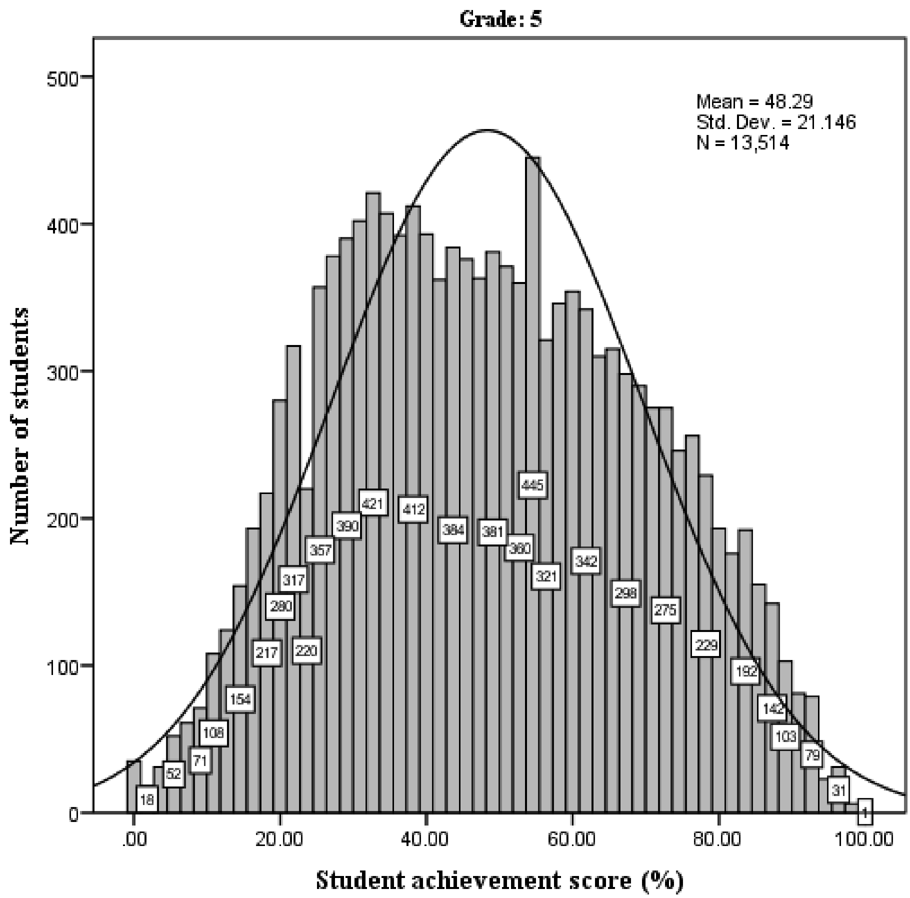
4.2 Basic Achievement Results in Mathematics

This section presents and discusses the overall distribution of assessment scores, achievement by various content areas and levels of cognitive domain in Mathematics. Similarly, it compares the achievement results of this assessment with the results of NASA 2012 in Mathematics for Grade 5.

Overall Distribution of Achievement Scores

The population is usually distributed normally in a large sample in the study of student achievement. In this assessment, the sample was big enough to form a normal distribution, as there were 13514 students in the sample in Mathematics. Figure 4.1 shows that the achievement score in Mathematics is normally distributed (skewness = 0.114). However, based on the distribution of achievement scores, student population can be grouped into two types−namely, low performing, and high performing, although the majority of the students fall in the low-performing group. The low-performing students achieved 15–45% and the high-performing students achieved 45–100% score.

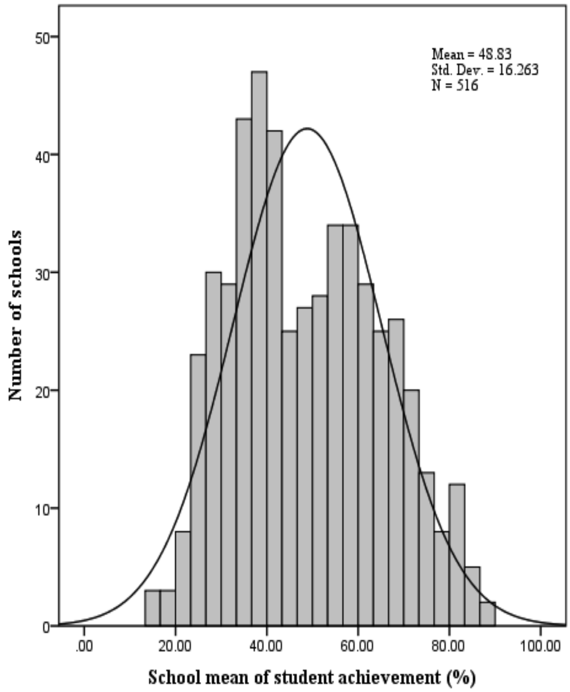
Figure 4.1. Distribution of overall results in Mathematics



The dataset informs that the 5th Graders' main population is low performing, which is evident from the figure 4.1 above where large number of student falls on the left hand side.

As illustrated in figure 4.2 given below, the schools can be divided into two distinct categories: the high and low performing. The population on the left-hand side achieve the average of 36% of the achievement score and on the right-hand side achieve around 66% of the achievement score. The difference between the two groups is remarkable.

Figure 4.2 School mean achievement score in Mathematics



Student Achievement by Various Content Areas in Mathematics

As per the curricular provisions, Mathematics test included four content areas namely: 1) *Arithmetic, 2) Geometry, 3) Algebra, 4) Numeracy.* The number of items and the weightage they carry correspond to the weight allotted in the curriculum.

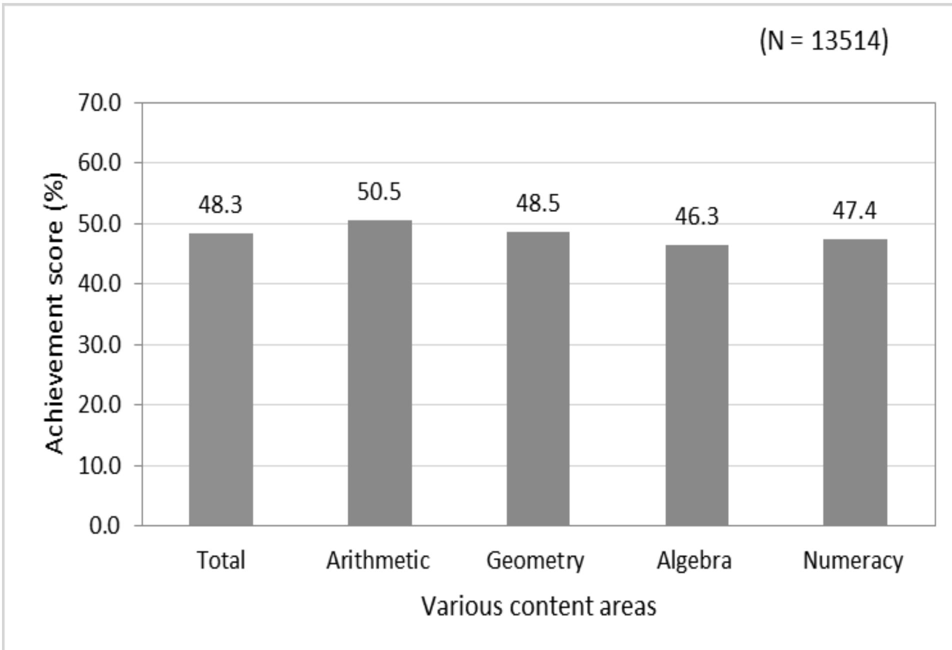
For comparability, achievement scores in all the content areas were converted into a percentage of the maximum score of each content area. Table 4.1 shows students' achievement in Mathematics by each of the four content areas.

Table 4.1 Achievement score in various content areas of Mathematics

Content area	Mean	SD	Min	Max
Arithmetic	50.5	23.08	0	100
Geometry	48.5	24.21	0	100
Algebra	46.3	24.35	0	100
Numeracy	47.4	24.86	0	100
Total	**48.3**	**21.15**	**0**	**100**

The table shows the variations of achievements in various content areas in Mathmatics. The achievement ranges from 46.3 percent in *Algebra* to 50.5 percent in *Arithmetic*, with a variation of 4.2 percent. While comparing the maximum and minimum scores, the situation is the same in all content areas as the maximum score is 100 and the minimum is 0 in each content area. Figure 4.3 compares the variations of achievements in various content areas in Mathematics.

Figure 4.3 Comparison of achievement scores in various content areas in Mathematics

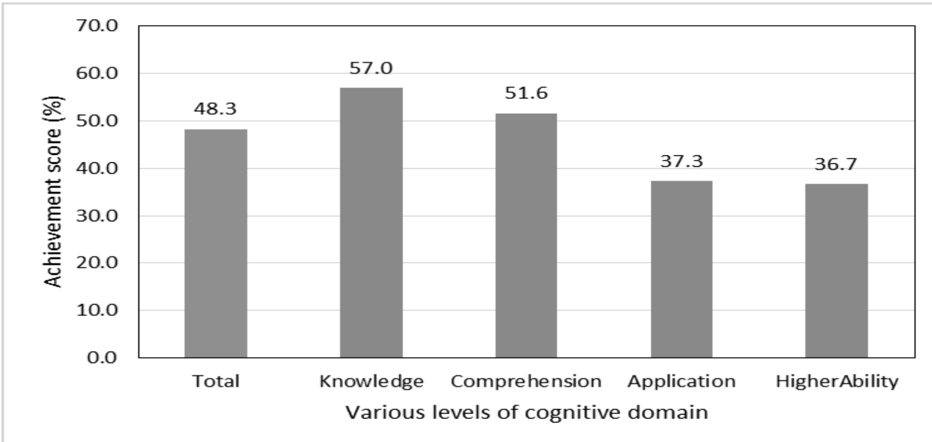


The overall national average achievement score for Mathematics is 48.3%. When the achievement score is analysed categorising it into the four content areas, students are seen performing poorly in *Algebra* (46.3%) and *Numeracy* (47.4%), which is below the national mean (48.3%). Similarly, achievement score in *Arithmetic* (50.5%) and *Geometry* (48.5%) are higher than the national average. The dataset indicates that the students are weak in *Algebra* and *Numeracy* in comparison to comparing other content areas.

Achievement in Various Levels of Cognitive Domain

Mathematics test comprised of items that can be grouped under the various levels of cognitive domain according to Bloom's taxonomy (Bloom *et al.* 1956; Metfesser, Michael & Kirsner 1969). The four domains considered while developing test items were: *Knowledge, Comprehension, Application* and *Higher ability*. Reasoning and problem solving type of items are mostly categorised under higher ability because this category covers analysis, synthesis and evaluation types of thinking. The achievement of the students in each level of cognitive domain is shown in figure 4.4.

Figure 4.4 Comparison of achievement scores in various levels of cognitive domain in Mathematics



The above figure indicates that, in general, the students' ability to solve complex problems (*Higher ability*) is low as they achieved only 36.7% of the maximum score. Students are much better in recalling type of questions as they achieved 57% score.

While analysing the dataset further, performance of a large number of students was found relatively low in higher cognitive abilities in Mathematics. For example, about 25% of the students in Mathematics could not solve any one of the tasks requiring higher skills.

Student Achievement by Type of Items

There were two types of items in the test: *Objective and Subjective*. Objective items covered a wide range of content areas, which were very specific to judge because there was only one correct answer; or one explicit piece of information is needed for correct answer. There were some subjective items in each test version, which required a longer procedure to get full marks. Both the *Objective* and *Subjective* types of items were developed considering various levels of cognitive domain (K*nowledge, Comprehension, Application* and *higher ability*) and the various difficulty levels, though the subjective items tend to be more demanding. Table 4.2 comprises the basic statistics of achievement in different types of items. From the table, it is clear that the students scored low marks in *Subjective* items than in *Objective* items.

Table 4.2 Achievement by type of items in Mathematics

Type of items	N	Mean	SD	SE of Mean
Objective	13514	56.8	21.18	0.182
Subjective	13514	35.1	25.99	0.224

Dataset indicates that students are good at recognizing the correct answer and in very fundamental items like recognizing numbers, shapes, choosing correct answer from the options, very basic operations such as the basic manipulation of data and numbers, and calculations in single or few steps. They are much weaker in calculation having more than one-step, questions involving the process of reasoning, problem solving, plotting, proving the theory or formula, and constructing the shapes and figures. In many cases, the students did not even start doing the open-ended questions; hence, they got low score.

Comparison of Achievement Results of NASA 2015 and 2012

Comparing the mean score of students of Grade 5 in Mathematics, score in NASA 2012 is higher than that of NASA 2015 as the mean scores (raw score) of NASA 2015 and NASA 2012 are 53 and 48.3 percent respectively. However, comparison of achievement in NASA 2015 against NASA 2012 based on the average scores of those assessments may not give accurate picture. Item response theory (IRT) suggests to compare results of two assessments based on the latent ability (θ). Latent ability (θ) is a mathematical model of estimating the ability shown by the students based on the item difficulty parameters. Difficulty parameters of some linking items taken from NASA 2012 were fixed in the test of NASA 2015 to calibrate the test items and equate the scores of NASA 2015 so that both tests became comparable. Although mean scores are presented in the table, it is good to compare latent ability (θ) to compare the results. Values of theta range from negative to positive value (eg. -3 to +3) with mean value zero (0) as a reference line.

Table 4.3 Comparison of achievements between NASA 2012 and 2015 in various content areas (Mean score)

Content Areas	NASA 2012 (N = 13514)	NASA 2015 (N = 17074)
Arithmetic	54	50.5
Geometry	57	48.5
Algebra	49	46.3
Numeracy	44	47.4
Mathematics Total	**53**	**48.3**

While calculating latent ability (θ) it is 0.008 and 0.003 in NASA 2012 and NASA 2015 respectively. This shows that latent ability of both years (NASA 2012 and NASA 2015) is quite close. However, the latent ability of NASA 2015 ($\theta = 0.003$) is slightly lower ($\theta = 0.008$) than that of NASA 2012.

4.3 Comparison with the International Standard

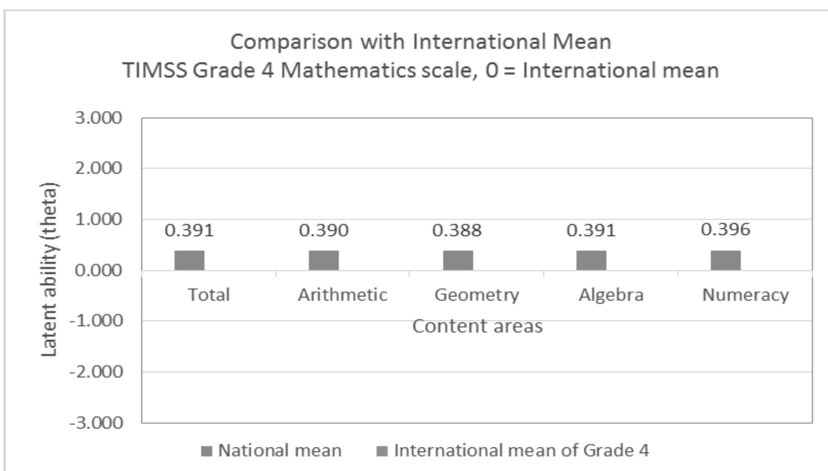
NASA 2015 was also made comparable with the international TIMSS assessment. Some of the items (8 items) of the released TIMSS Grade 4 items were used as linking items. Those items were selected in such a way that they exactly match with Nepal's Mathematics curriculum of primary level. Known difficulty parameters of those international items were fixed in the calibration of the local items. Hence, the international average of $\theta = 0$ was fixed in the Nepalese datasets.

When a student's ability level in NASA 2015 is zero, it corresponds to the average level of the international students.

Figure 4.5 shows the comparison of the students' achievement with the international standards. In the figure, the x-axis shows the content areas of Mathematics and y-axis shows the ability shown by the students. The middle horizontal thick line indicates the international average of Grade 4. As the ability is above the average line, it indicates the ability of Grade 5 students in Nepal is above the international average of Grade 4.

The mean latent ability in Mathematics for Grade 5 students is higher than international average in Grade 4 Mathematics (i.e, TIMSS results in Mathematics). Figure 4.5 shows that the average ability shown by Nepali students in Mathematics as a whole is $\theta = 0.391$, which is above the Grade 4 average ($\theta = 0.00$). This indicates that the students in Mathematics are somewhat better than Grade 4 international average. Similarly, the students show somewhat better ability in *all* the content areas ($\theta > 0.00$).

Figure 4.5 Student achievements in the international TIMSS Mathematics Scale



This result indicates that Nepali students of Grade 5 are performing better than international Grade 4 students. However, it does mean that Grade 5 students are internationally better, but are not performing lower than the average of Grade 4 in TIMSS.

4.4 Achievement Score by Diversity Factors

Though the diversity is a relative and contextual term, in the context of NASA, six diversities have been considered–namely Geographical/Ecological, Regional, language, gender, ethnic/caste and economic diversity. NASA 2015 background information questionnaire included the questions related to these six factors of diversity. However, this assessment also considered three additional comparisons. They are by Districts, by school type (community/institutional) and by school location (rural/urban). These comparisons are made particularly to assess the equity status of students based on achievement scores.

Student Achievement by Districts

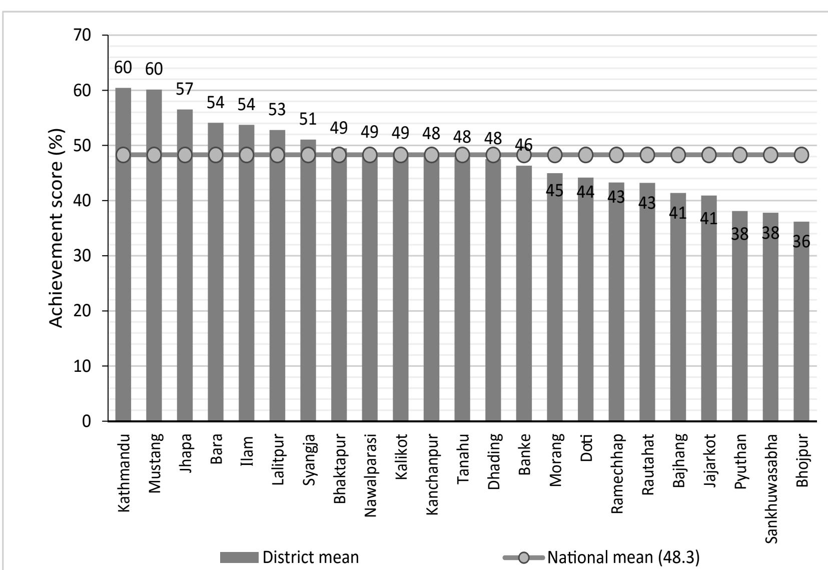
The achievement of each of the 23 sample Districts is presented in table 4.4. The table shows mean achievement of sample Districts in descending order according to the achievement scores.

Table 4.4 Average achievement score of Mathematics in sample Districts

District	N	Mean	SD	District	N	Mean	SD
Kathmandu	1494	60.4	19.35	Dhading	777	47.5	20.57
Mustang	29	60.1	16.33	Banke	576	46.3	20.62
Jhapa	724	56.5	21.25	Morang	788	45.0	21.63
Bara	788	54.1	25.48	Doti	418	44.2	15.69
Illam	429	53.7	22.59	Ramechhap	576	43.3	19.46
Lalitpur	491	52.8	15.90	Rautahat	661	43.2	19.48
Syangja	607	51.1	20.02	Bajhang	522	41.4	18.13
Bhaktapur	245	49.5	20.16	Jajarkot	386	40.9	20.11
Nawalparasi	744	48.7	22.06	Pyuthan	661	38.1	15.29
Kalikot	345	48.5	19.48	Sankhuwasabha	469	37.8	15.79
Kanchanpur	561	48.4	19.66	Bhojpur	432	36.2	20.39
Tanahu	791	47.8	21.02	**Total**	**13514**	**48.3**	**21.15**

Figure 4.6 compares the mean achievement of sample Districts with the national mean score in Mathematics. It shows wide variations among the Districts, for example the difference between the high performing Districts Kathmandu (60.4%) and Mustang (60.1%) and the lowest performing District Bhojpur (36.2%) is 24 percent.

Figure 4.6 Average achievement score of sample Districts in Mathematics



Of the selected Districts in the sample, except for 10 Districts, all Districts' achievement score is below the national average. Out of the 10 higher performing Districts, two highest performing Districts are from the Kathmandu Valley and one Mustang from Western Development Region. Among the low achieving 13 Districts, lowest two Districts are Sankhuwasabha and Bhojupr, which are both from Eastern Development Region. The data-mining tool of the SPSS software – Decision Tree Analysis (DTA) points out that those lowest performing schools are mostly from Mid-Western and Eastern Region.

The difference in achievement scores among the Districts is statistically significant ($p < 0.001$). The variation explained in achievement among the Districts is $\eta^2 = 0.101$, which explains 10% of the variation in achievement. Effect size is $f = 0.33$, which indicates that the difference between the lowest achieving and highest achieving Districts is moderately high.

Student Achievement in Various Ecological Zones

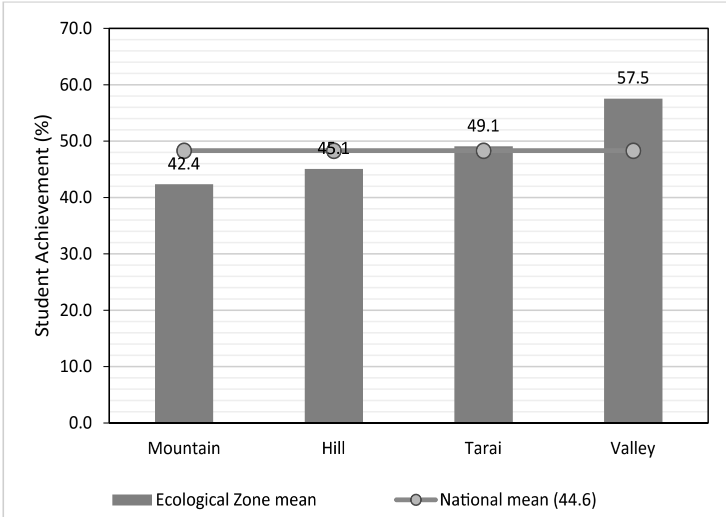
The Mountain, Hill and Tarai are three Ecological Zones in Nepal though the Kathmandu Valley has been considered as a special geographical area because it is the most densely populated area in the country with more opportunities than other areas. Not only from the population point of view but also due to the mixed ethnicities, weather conditions, economic activities, urbanization, as well as the dense human capacity have been taken into account while considering the Kathmandu Valley as a unique fourth geographical area in the analysis. The variation in the Ecological Zones in NASA 2015 is presented in table 4.5. The table has also included the number of students, standard deviation, standard error and minimum and maximum scores in each of the categories.

Table 4.5. Achievement scores in Mathematics in various Ecological Zones

Ecological Zone	N	Mean	SD	SE of Mean	Minimum	Maximum
Mountain	1365	42.4	18.33	0.496	0	93
Hill	5077	45.1	20.28	0.285	0	98
Tarai	4842	49.1	22.17	0.319	0	100
Valley	2230	57.5	19.20	0.407	0	96
Total	**13514**	**48.3**	**21.15**	**0.182**	**0**	**100**

Figure 4.7 compares the students' achievement among various Ecological Zones as well as with the national mean score in Mathematics.

Figure 4.7 Comparison of achievement scores by Ecological Zones in Mathematics



The data shows that, on average, the students from the Kathmandu Valley have outperformed (57.5%) all the other Ecological Zones. The students from the Mountain areas performed the lowest (42.4%). Moreover, average achievement of Tarai Region (49.1%) is higher than Hill (45.1%).

The achievement across the Regions differs significantly ($p < 0.001$). In terms of Tukey's *post hoc* test, all the Zones deviate from each other in a statistically significant manner at $p= 0.01$ level. The effect size is $f = 0.23$, which shows a medium difference between the highest and lowest performing Ecological Zones. Ecological Zone explains 5% of the variation in the data ($\eta^2= 0.049$). Excluding the Kathmandu Valley with the ANOVA analysis, the Ecological Zone explains 1% of the variation in achievement where the effect size is small ($f = 0.11$).

Dataset indicates that there is a wide difference between the student performances in the four Ecological Zones. Students in the Kathmandu Valley have outperformed (57.5%) the students from other Zones. The achievement is the lowest in the Mountain (42.4%). Moreover, average achievement of Tarai Region (45.4%) is higher than Hill Region (41.6%).

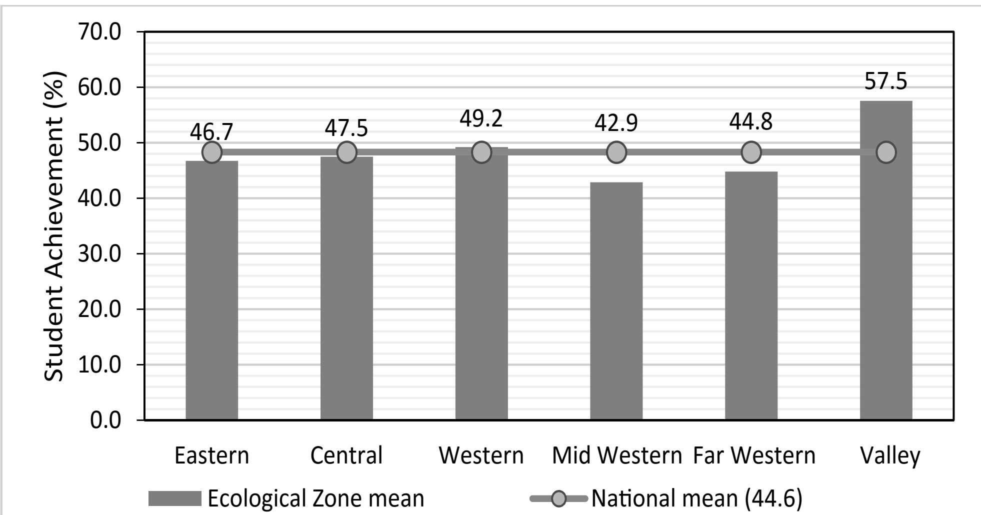
Student Achievement by Development Regions

Student achievement varies according to the Development Regions, which are divided into five: 1) Eastern, 2) Central, 3) Western, 4) Mid-Western, and 5) Far-Western Regions. Additionally, the Kathmandu Valley is taken as the additional Region though administratively it falls under the Central Developmental Region. The mean achievements of the Development Rregions are given in table 4.6.

Table 4.6 Student achievement scores in Mathematics by various Development Regions

Regions	N	Mean	SD	SE of Mean	Minimum	Maximum
Eastern	2842	46.7	22.10	0.415	0	100
Central	2802	47.5	22.07	0.417	0	98
Western	2171	49.2	21.12	0.453	0	98
Mid Western	1968	42.9	19.12	0.431	0	98
Far-western	1501	44.8	18.33	0.473	0	91
Valley	2230	57.5	19.20	0.407	0	96
Total	**13514**	**48.3**	**21.15**	**0.182**	**0**	**100**

Figure 4.8 Comparison of achievement scores by Ecological Zones in Mathematics



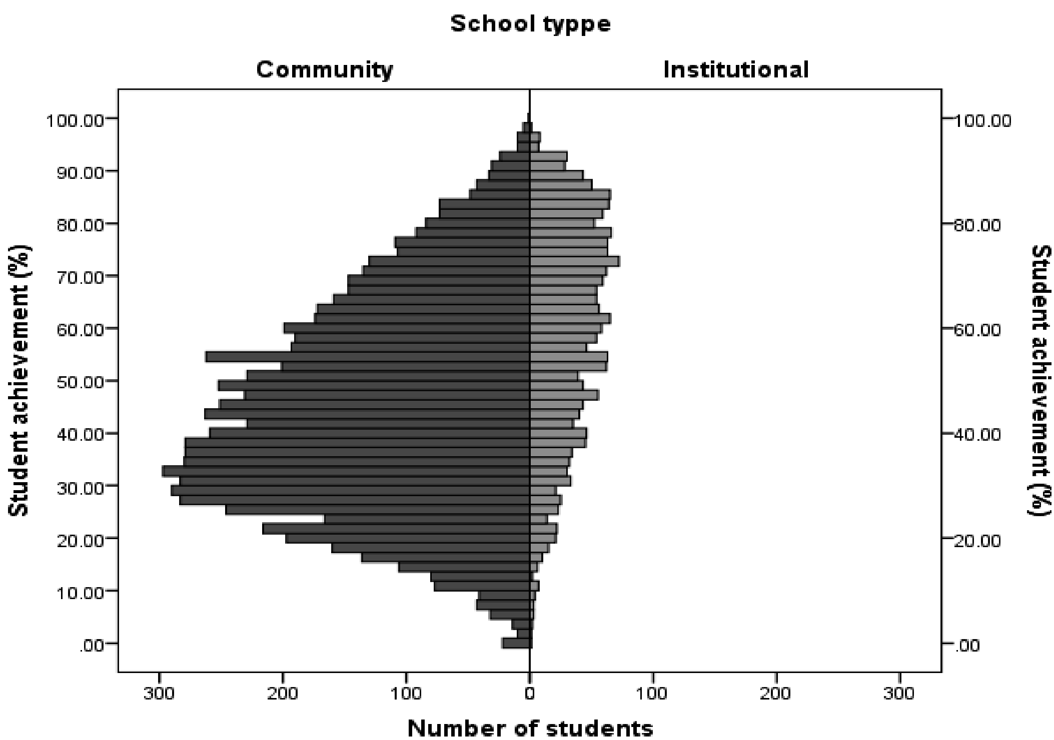
The highest performance is found in the Kathmandu Valley (57.5%). Besides Valley, Western Region is performing high (45%) which is above the national mean; all other Regions are performing below the national average. Performance is lowest in two Regions: Mid-Western (42.9%) and Far-Western (44.8%). The difference between the Regions is statistically significant (p < 0.001), to which Tukey's *Post hoc* test shows that all the Development Regions differ from each other significantly. The effect size is found high (f = 0.23); and it explains 5% (η^2 = 0.049) variation in student achievement. However, if the Kathmandu Valley is excluded from the analysis, the effect size will be medium (f = 0.11), and the Development Region explains only 1% (η^2 = 0.013) variation in achievement.

The dataset reveals that there is wide inequity among the Development Regions for children's opportunities to reach an adequate level in Mathematics. The difference between the lowest performing (Mid-Western, 42.9%) and the highest performing Region (the Kathmandu Valley, 57.5%) is remarkable, that is more than 14.6 percent. However, if the Kathmandu Valley is excluded from the analysis, the difference will be 6 percent. Moreover, except for the Western Region, performance of the other Regions is below the national average.

Student Achievement by School Type

While plotting the community and institutional schools and their mean achievements in a graph as in figure 4.9, it is seen that the community school students on the left hand side are distinctly distributed at least in three kinds of populations (groups). There are a remarkably high number of students achieving below 40% achievement. The right hand side whereby distribution shows the student population of institutional schools, which is shifted upward, majority of students, are showing better performance. However, in institutional schools, there are also some students obtaining very low marks in Mathematics. This indicates that the students in institutional schools are also categorised as low to high performers though majority of them are high performers.

Figure 4.9 Distribution of students' achievement scores in community and institutional schools



All the schools are categorised into community and institutional (usually called private) schools. The difference in achievement between the two types of schools is presented in table 4.7.

Table 4.7 Student achievement by school type in Mathematics

School Type	N	Mean	SD	SE of Mean	Min	Max
Community	10361	44.5	20.29	0.199	0	100
Institutional	3145	60.8	19.00	0.339	0	98
Total	**13506**	**48.3**	**21.14**	**0.182**	**0**	**100**

The average performance of the students from institutional schools is 60.8% whereas it is 44.5% in community schools, that is, the difference is 15 percent, which is remarkable. This shows wide difference in the level of achievement between community and institutional schools, whereby the difference is statistically significant ($p < 0.001$) and the effect size is very high ($d = 0.81$). Within the type of school themselves, variation in student performance is higher in community schools in comparison to institutional schools. It happened so because most of the institutional schools from the sample are performing high but students in community schools are found in the lowest as well in highest level of achievement.

The dataset indicates that the average performance of the students from institutional schools is 60.8% whereas it is 44.5% in community schools, that is, the difference is 15 percent, which is remarkable. Compared to the result of NASA 2015 in Grade 3, variation between community and institutional schools is wider in Grade 5.

Student Achievement in Rural and Urban Schools

One of the strata of sampling in NASA 2015 was school location. The disaggregated achievement scores of the students from rural and urban schools are presented in table 4.8.

Table 4.8 Student achievement in rural and urban schools

School Location	N	Mean	SD	SE of Mean
Urban	3554	54.8	21.06	0.353
Rural	8430	47.4	20.65	0.225
Total	**11984**	**49.6**	**21.05**	**0.192**

The achievement level of the students in the urban schools (54.8%) is remarkably higher than that in rural schools (47.4%)−the difference is 7.7%. The difference in average score is significant ($p < 0.001$) and the effect size is moderately high ($d = 0.36$).

Student performance significantly varies based on the location of the schools. Urban schools outperformed (54.8%) the rural schools (47.4%) by 7.4 percent score in Mathematics. However, the difference is moderate.

Student Achievement in Various Language Groups

In the context of Nepal, student achievement has been found somehow affected by the language spoken at their homes i.e., the mother tongue. In many cases, mother tongue reflects the ethnic background, and hence it may be taken as a possible reason for achievement differences in Mathematics.

Based on the total data, 32.9% of the 5th Graders speak a language other than Nepali in their home at most. These other languages are quite fragmented; the largest groups in the student dataset are Magar (2.7%), Tharu (2.6%), Tamang (3.9%), Newar (2.0%), and other languages like Abadhi-Bhojpuri, Gurung, Sherpa, Limbu speakers are less than 1% and 9.5% students did not respond (missing cases). Besides these language speakers, students of '*Élse*' category are about 20.5%. Except for Nepali speakers, all other language groups including the above stated language groups, total 34.6% are termed as “Non-Nepali”. The results are presented in tables 4.9 and 4.10.

Table 4.9 Achievement of Nepali and non-Nepali home language groups in Mathematics

Language at Home	N	Mean	SD	SE of Mean
Nepali	8404	50.3	21.00	0.23
Non-Nepali	4441	46.4	20.78	0.31
Total	**12845**	**49.0**	**21.00**	**0.19**

The average achievement of students having their home language Nepali is 50.3% and Non-Nepali is 46.4%, the difference in achievement is 3.9 percent.

Table 4.10 shows the achievement of student groups who speak different national languages.

Table 4.10 Student achievement by linguistic background in Mathematics

Language at Home	N	Mean	SD	SE of Mean
Nepali	8404	50.3	21.00	0.229
Magar	359	40.4	16.92	0.893
Tharu	351	50.1	19.37	1.034
Tamang	533	47.6	20.95	0.907
Newar	274	55.3	19.26	1.164
Rai	126	42.1	19.09	1.700
Gurung	61	49.7	20.25	2.593
Sherpa	63	49.8	20.90	2.633
Limbu	37	44.6	19.90	3.272
Else	2630	45.7	21.26	0.415
Total	**12845**	**49.0**	**21.00**	**0.185**

(The language groups with less than 10 students have not been considered for analysis separately.)

Without grouping the achievement of individual language group students, Newari speaking students outperformed (55.3%) the other language groups. Of the groups, Rai speaking students have performed lowest (40.4%) in Mathematics. However, the number of students in Nepali language speakers is very high whereas other language speakers are very few in the sample.

The difference between the language groups is statistically significant ($p < 0.001$). However, the effect size is medium ($f = 0.13$). The language group explains 2% of the variation in the data ($\eta^2 = 0.02$).

Dataset reveals that Nepali speakers outperform non-Nepali speakers by 3.9% points. Among various national language speakers, Newari speaking group achieved the highest score (55.3%) and Magar speaking students achieved the lowest score (40.4%). Variation between highest and lowest speaking student group is medium.

Achievement of the Students of Various Ethnic/Caste Groups

Education in Nepal has been influenced in several ways by the legacy of the traditional caste system which still remains in the mind-set of people in Nepalese society. Historically, the Brahmans and Chhetries have higher level of educational attainments; but Dalits, for example, have remained outside the educational system. Hence, modern society has made lots of efforts to make the education accessible for all children regardless of the caste and ethnicity. The recent National Population Census 2011 shows that the enrolment of Hill Dalits has increased remarkably at the primary level of schooling but their number at the secondary level and higher education is still very small (CBS, 2012). The achievement results of the students from various ethnic/caste groups are presented in table 4.11.

Table 4.11 Student achievement by ethnicity/caste background in Mathematics

Ethnicity/Caste	N	Mean	SD	SE	Minimum	Maximum
Brahman	1610	55.6	21.01	0.524	3.64	98
Chhetri	2962	49.6	20.63	0.379	0	98
Janjati	4864	48.5	20.33	0.291	0	100
Dalit	1654	43.0	19.49	0.479	0	96
Madhesi	1213	50.1	22.51	0.646	0	98
Minority	75	37.5	19.71	2.276	9.09	91
Other	430	50.4	21.46	1.035	0	96

Minority's achievement score in Mathematics (37.5%) is lower among the achievement of all castes/ethnic groups. The difference in achievement scores between the outperformer Brahmans (55.6%) and Minority is 18 percent. The overall difference among the groups is statistically significant ($p < 0.001$), but the effect size is moderately low ($f = 0.16$). Categorisation of students according to their ethnic/caste background explains around 3% variation in data ($\eta^2 = 0.026$).

The following table (table 4.12) shows the status of various castes/ethnicities based on achievement scores in Mathematics in Grade 5.

Table 4.12 Caste/ethnicity-wise minimum and maximum achievement in various Ecological Zones

Eco-Zone	Mountain	Hill	Tarai	Valley
Brahman	41.2*			62.2♣
Chhetri	45.4*			56.8♣
Dalit		41.6*		53.3♣
Janjati	39.5*			56.7♣
Madhesi	35.0*			62.8♣
Minority	31.8*			63.6♣
Other	34.4*			58.0♣

* *Lowest performance;* ♣ *Highest performance*

All the castes/ethnicities perform the highest in the Kathmandu Valley and lowest in Mountain except Dalit as Dalits performed lowest in Hill. Dataset shows that achievement of Brahman students perform the highest (55.6%) and Minorities the lowest (37.5%) in total. Castes/ethnicities wise achievement score varies in most of the Ecological Zones. All the caste/ethnicity performed the highest in the Kathmandu Valley. Madhesi and Minority, although their number is very low in the Kathmandu Valley, performed the highest (62.8% and 63.6% respectively) among various caste/ethnic groups.

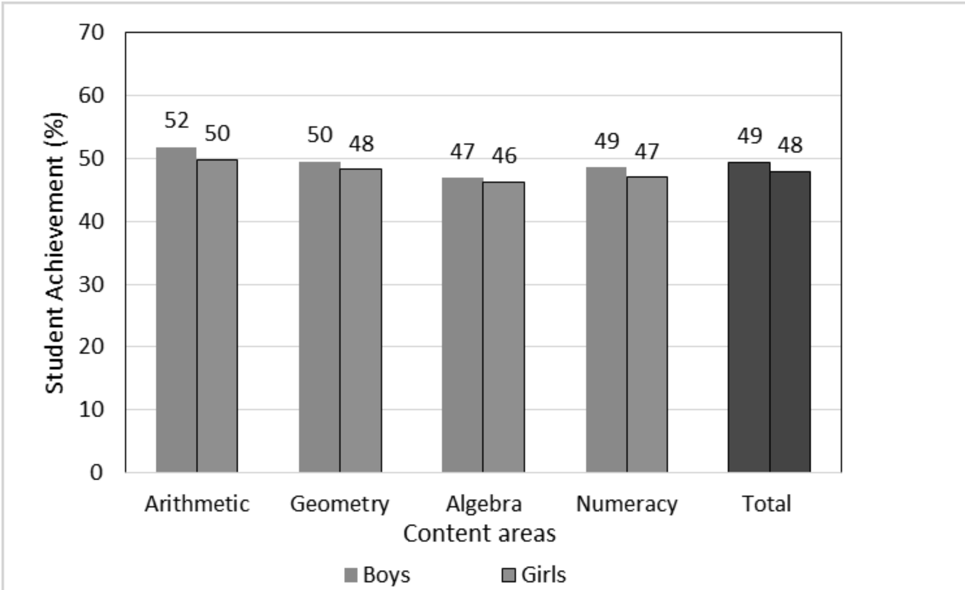
Gender and Student Achievement

In previous assessments, there was a significant difference in achievement between boys and girls (see ERO, 2011; EDSC, 2008). Situation is found improved in Grade 5 Mathematics in this assessment, as there is only 1.5 percent difference in the achievement between boys and girls. Basic achievement results of boys and girls are presented in table 4.13. Figure 4.10 further compares the achievement of girls and boys in various content areas. The figure shows that in each content area boys have outperformed girls.

Table 4.13 Student achievement of boys and girls in Mathematics

Sex	N	Mean	SD	SE	Minimum	Maximum
Male	6501	49.5	20.90	0.259	0	100
Female	6612	47.9	21.11	0.260	0	98
Total	**13113**	**48.7**	**21.02**	**0.184**	**0**	**100**

Figure 4.10 Comparison of achievement between boys and girls in different content areas

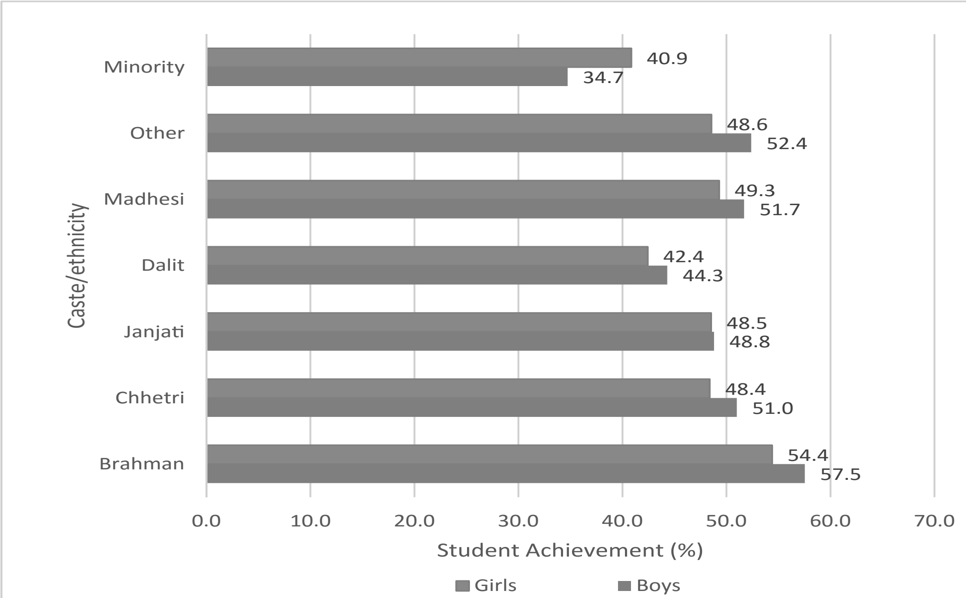


The above table shows that there is statistically significant ($p < 0.001$) difference in the achievement between boys (49%) and girls (48%) in Mathematics. It is noteworthy that in all content areas boys have outperformed girls. However, the effect size is small (f = 0.03), indicating that the difference is very low. Gender does not explain the variation (η^2 = 0.001) in achievement score very much. However, still small difference is seen in achievement between boys (49%) and girls (48%), which is statistically not significant. Variation in achievement is very low, indicating that there is not significant gap between boys and girls in achievement. This is positive sign from the equity point of view.

Gender and Ethnicity/Caste

There are gender differences in the achievement of various ethnic/caste groups in Mathematics. For example, difference in achievements between boys and girls is the highest in Minority communities, that is, girls have outperformed boys by 6 percent. For other castes/ethnicities, the difference ranges from very low (0.27%) in Janajati to high (3.2%) in Brahman. However, boys outperformed girls by 3.82 percent in "other" caste; specific caste/ethnicity cannot be identified (see figure 4.11). The difference is statistically significant at $p < 0.001$. The effect size is medium (f = 0.12), which explains only 1% variation in achievement (η^2 = 0.014).

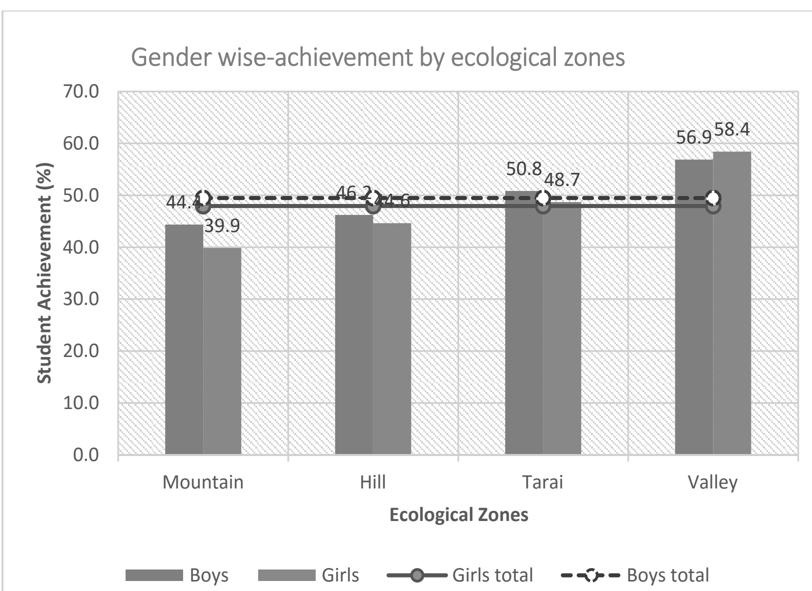
Figure 4.11 Comparison of achievement between boys and girls in various ethnic/caste groups



Gender and Ecological Zone

Achievement of boys and girls across the Ecological Zones is presented comparatively in figure 4.12.

Figure 4.12 Comparison of achievement between boys and girls in various Ecological Zones

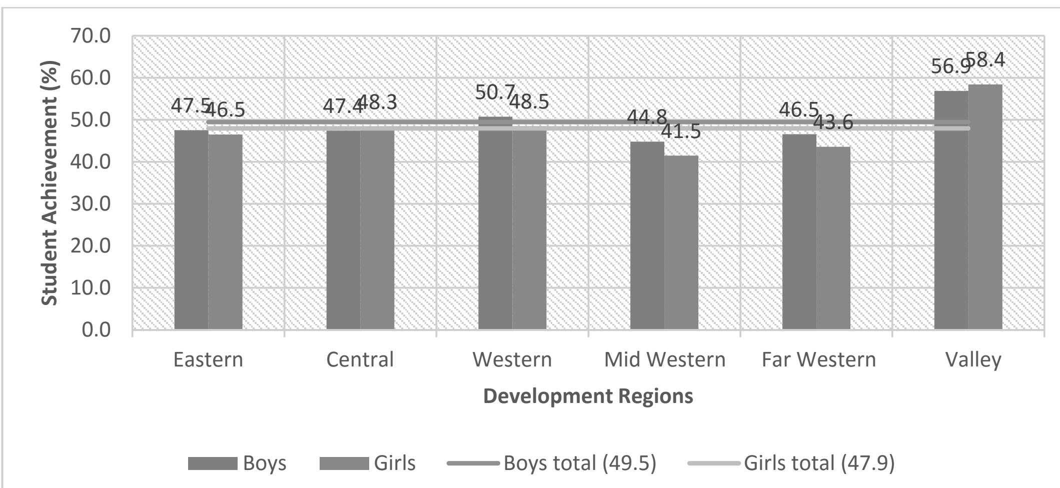


As shown in the figure, achievement of girls and boys in Mathematics differs significantly across Ecological Zones ($p < 0.05$). Except the Kathmandu Valley, in all Ecological Zones boys outperformed girls, but in the Kathmandu Valley girls outperformed boys significantly at ($p = 0.056$), however the difference is very small.

Achievement by Gender and Development Regions

Development Region-wise achievement differs between boys and girls. Boys outperformed girls by around 3 percent in the Mid and Far Western Regions. The difference in achievement between boys and girls is minimum in Eastern Region (1%) whereas girls outperformed boys in the Kathmandu Valley by 2 percent, which is presented in comparative form in figure 4.13.

Figure 4.13 Comparison of achievement between boys and girls in various Development Regions



Dataset shows that girls are performing better than boys in Central Region (by +1%) and the Kathmandu valley (by +2%). Boys are better in Far and Mid Western Regions along with Eastern and Western too. However, the difference is not significant in Eastern Region.

In Grade 5 Mathematics, girls are performing better than boys in Central Region and the Kathmandu Valley. However, boys performed better than the girls in other Regions, which is a significant difference ($p < 0.05$). In total, boys achieved higher (1.5% above) than girls, which is statistically significant, though the difference is small.

4.5 Selected Explanatory Factors and Achievement

Several factors have already been presented in the previous section such as geographical factors including Districts, Ecological Zone and Development Region; school related factors including school type and school location. Some individual related factors were also handled, such as home language, caste/ethnicity and gender. This section deals with socio-economic status (SES) of the students' families, paid work beyond the school time, students' attitude towards Mathematics as a school subject, age of the student, and support provided for the studies as the main family and individual related factors. Similarly, in this section school and teacher related factors, the availability of textbooks, homework given by the teacher and selected activities in school are also considered for analysis.

Achievement Variation of Students in Various Socio-Economic Status (SES)

The variables indicating the socio-economic status were grouped into seven categories. The seven categories of SES are mother's education, father's education, mother's occupation, father's occupation, home possessions, home accessories, and type of school attended. Finally, the SES is estimated based on seven indicators related to the economic, educational, and occupational background of the family. In this section, the education of the parents is further elaborated in that the literacy of the parents is analysed in relation to achievement in Mathematics.

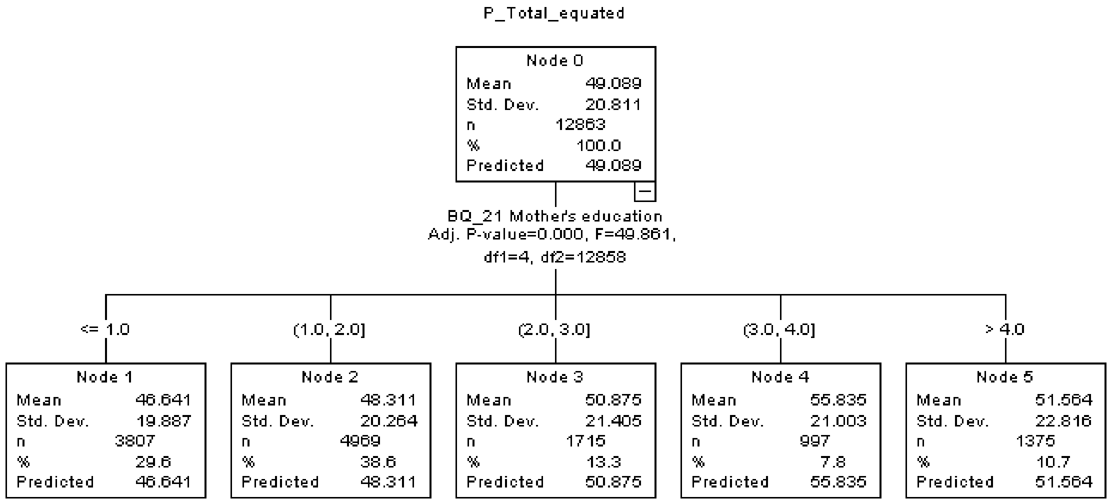
Several SES related variables were analysed by using a data-mining tool of SPSS – the Decision Tree Analysis (DTA). The method was very effective in finding the cut-offs of the predicting variable, such as mother's education, and classifying the factor into several groups which differ statistically in the most significant way from each other in relation to student achievement.

Parents' education

In the background questionnaire of NASA 2013, parents' education is divided into the following seven categories: 1) Illiterate, 2) Literate, 3) Grade 10, 4) SLC, 5) Grade 12/Certificate level, 6) Bachelor's level, 7) Master's level and above.

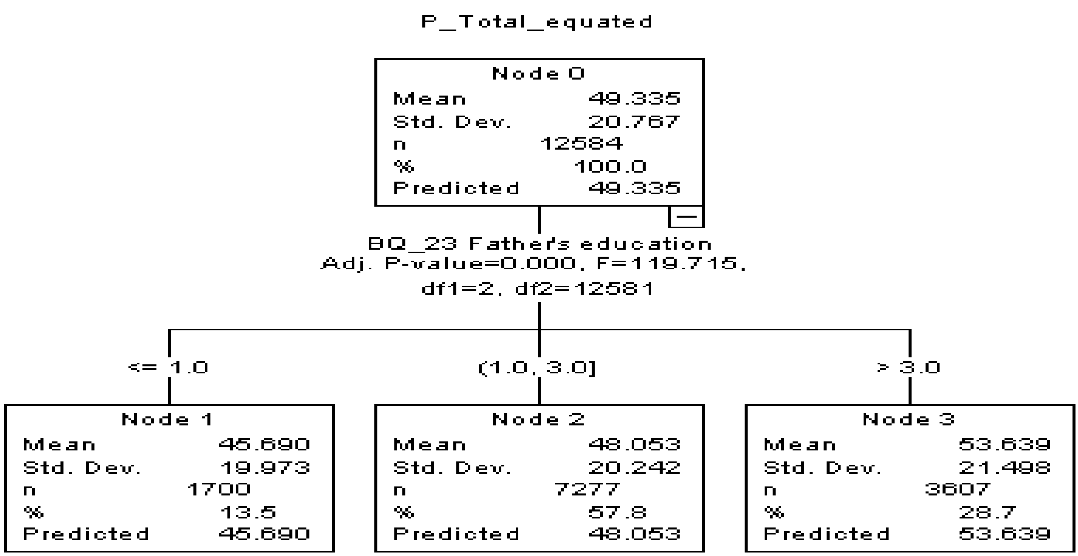
DTA classifies mother's education into seven groups with statistically significant differences in students' achievement levels. The data shows that students having illiterate mothers obtained only 28% of the maximum score on average, which is the least average score in comparison to the students having other categories of mothers in terms of their education. The achievement of students has increased along with the increase in mothers' educational level from literate to Bachelor's degree. However, the achievement of students having mothers with the qualification of Master's level and above has been lower compared to the achievement of those students having mother with Bachelor's degree. For example, when mothers are Bachelor's degree holders, their children obtained above 54 percent scores compared to the illiterate mothers (47%). Only the students having illiterate or just literate mothers are performing below the national average. The figures 4.14(a) and 4.14(b) below explain these results. In each group, the number of illiterate and just literate mothers is high enough to make a credible prediction. The difference between the groups is statistically significant ($p < 0.001$). Mother's education explains 1.6% variation in achievement ($\eta^2 = 0.016$), indicating a medium effect size ($f = 0.14$).

Figure 4.14(a) DTA of mother's education and students' achievement in Mathematics



In a similar vein, DTA categorizes father's education into three groups. Combining both categories, a comparative illustration of relation between parents education and student average achievement percentage is presented in the figure 4.14(b).

Figure 4.14(b) DTA of father's education and students' achievement in Mathematics

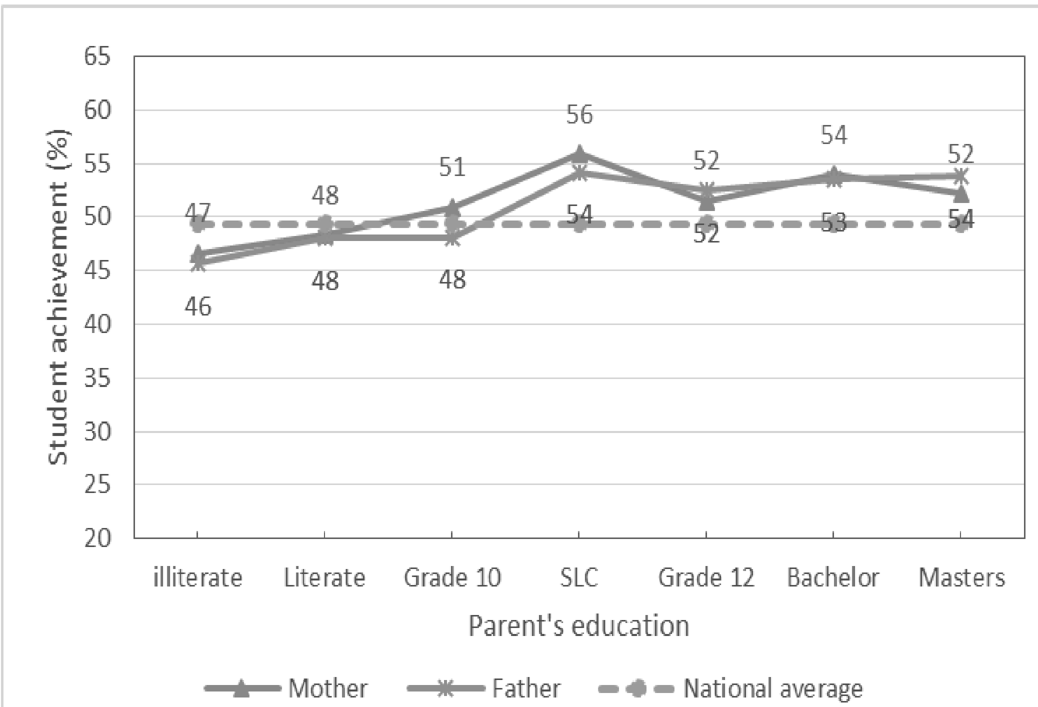


A positive relationship between parent's education and student achievement is found. As one of the parents (either father or mother) is illiterate, student achievement is below the national average (46-47 percent). As mother is Grade 10 pass or father is SLC pass, student achievement is above the national average (mother Grade 10 pass 51%, father

SLC pass 54%). The difference in achievement due to parents' education is statistically significant ($p < 0.001$). The effect of mother's education is about 2% and the variation due to mother education is medium (Cohens' $f = 0.14$). Effect of father's education is about 1% and the variation due to father's education is also medium as mother (Cohens' $f = 0.13$).

From the data of both father's and mother education comparative line graph of mother's and father's education level with their corresponding average achievement is shown in figure 4.15.

Figure 4.15 Relation between parent's education and students' achievement in Mathematics



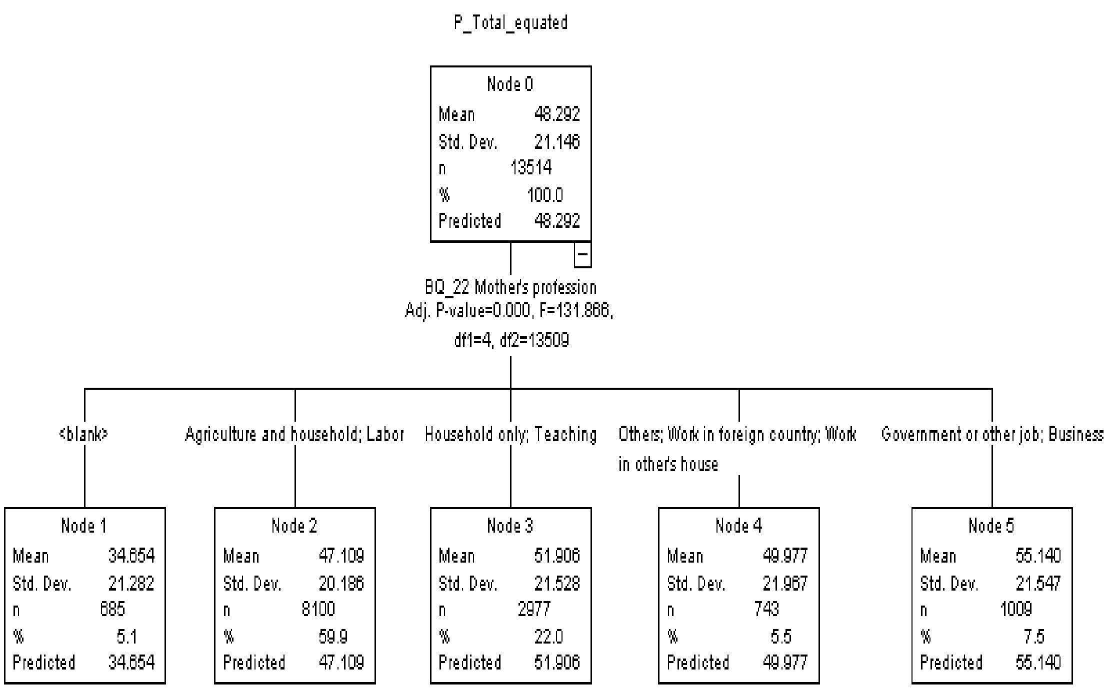
From figure 4.15, it is known that effect of mother's education is more productive than father's educational level as parents are educated up to SLC (about 1% score difference in each level). Above the SLC, father's education seems more effective (upto 2% difference in score at Master's level).

Result explains that the educational level of parents predicts the children's achievement level in Mathematics. When both parents are illiterate, achievement of their children become very low compared to the children of other parents with SLC or higher educational level. Up to SLC, mother's education is more effective than that of parents whereas father's educational level above SLC is found effective.

Parents' occupation and student achievement

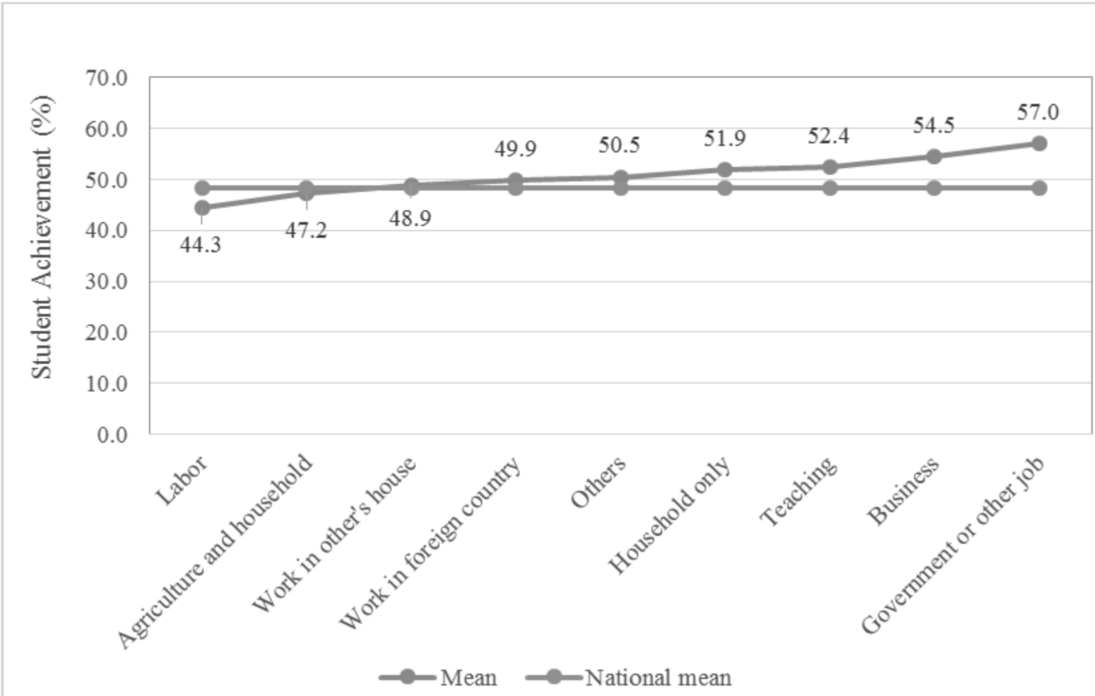
Mother's occupation was grouped into the following eight categories: 1) agriculture and household chore, 2) only household chore, 3) work abroad, 4) teaching 5) other jobs/service, 6) business, 7) daily wage and 8) work in other's home. The DTA was used to find statistically the most deviating occupation groups related to student achievement (see fig. 4.16). The differences between the groups are statistically significant ($p < 0.001$). Student achievement is the lowest i.e., 44.3%. When mother's occupation background is labourer, the achievement is the lowest, whereas the children with the mothers having government service have scored the highest (i.e., 56 percent). It shows that higher economic or intellectual level or both at home give positive impact on children's achievement. Mothers' occupation shows 3.5% variation in data ($\eta^2 = 0.035$) with a medium effect size ($f = 0.19$).

Figure 4.16 DTA of mother's occupation and students' achievement in Mathematics



Besides DTA, a line graph is drawn to show the exact average score of students over their mother's occupation in figure 4.17.

Figure 4.17 Comparison of mothers' occupation and students' achievement in Mathematics



There are three major groups of students according to mother education, (1) mother is either labour or involved in agriculture, they achieved the lowest (44% - 47%), (2) mother is involved in other's home or work in foreign country (around 49%), and (3) whose mother is involved in house hold chore only or involved in any service like teaching, government job or business (52-57%).

Similarly, DTA categorizes fathers' occupation into six groups, which is shown in the tree diagram in figure 4.18. If father is involved in household chore only, student achievement is lowest (42.5%), and the achievement is highest (55.5%) when father is involved in government job, teaching or business. The variation of student achievement due to parents' occupation category is statistically significant ($p < 0.001$). Fathers' occupation shows 4% variation in achievement ($\eta^2 = 0.035$) with a medium effect size ($f = 0.19$).

Figure 4.18 DTA of fathers' occupation and students' achievement in Mathematics

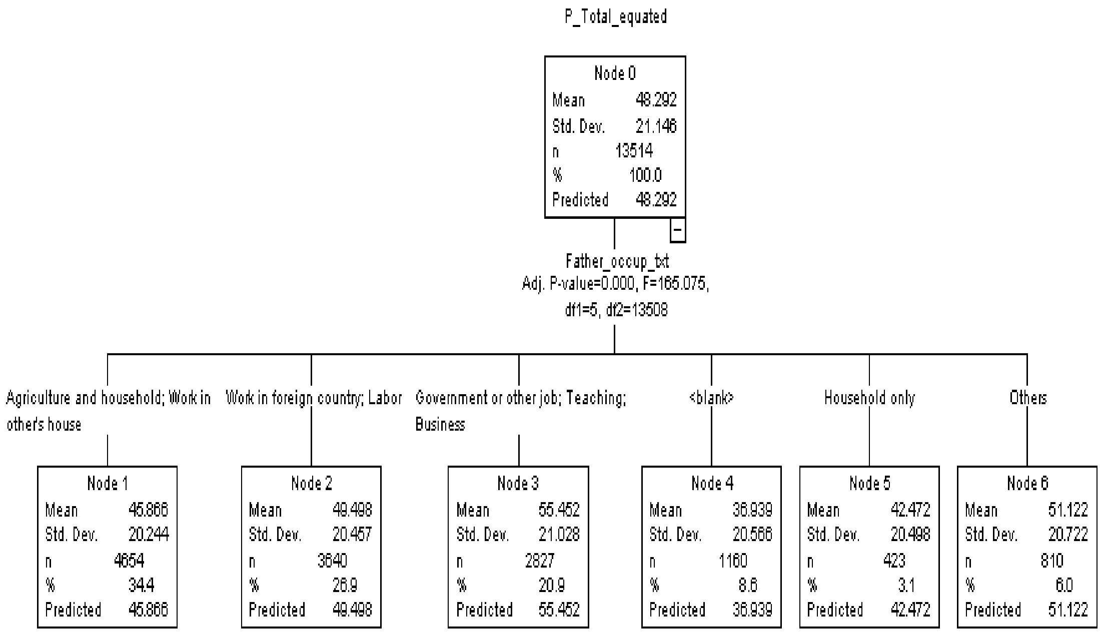
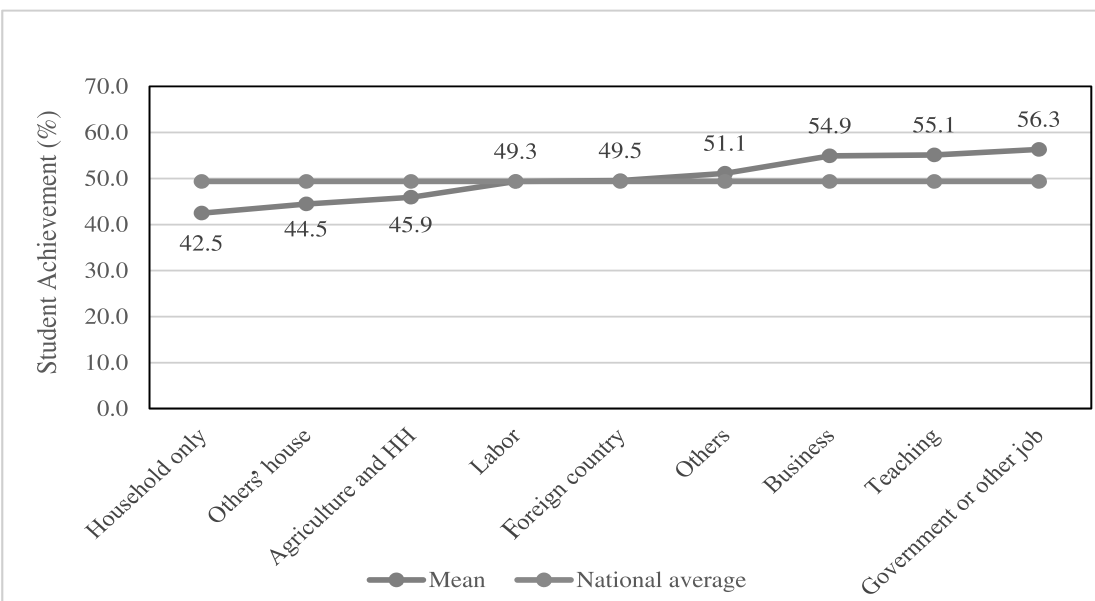


Figure 4.19 Comparison of fathers' occupation and students' achievement in Mathematics



There are three groups in father's occupation (1) lowest achievement (42-46%) when father is involved in the jobs like household works, works in other's home and agriculture, (2) medium achievement (49-51%) when father is involved in labour work or work in foreign country, and (3) highest achievement when father is involved in business, teaching or government job (54-56%).

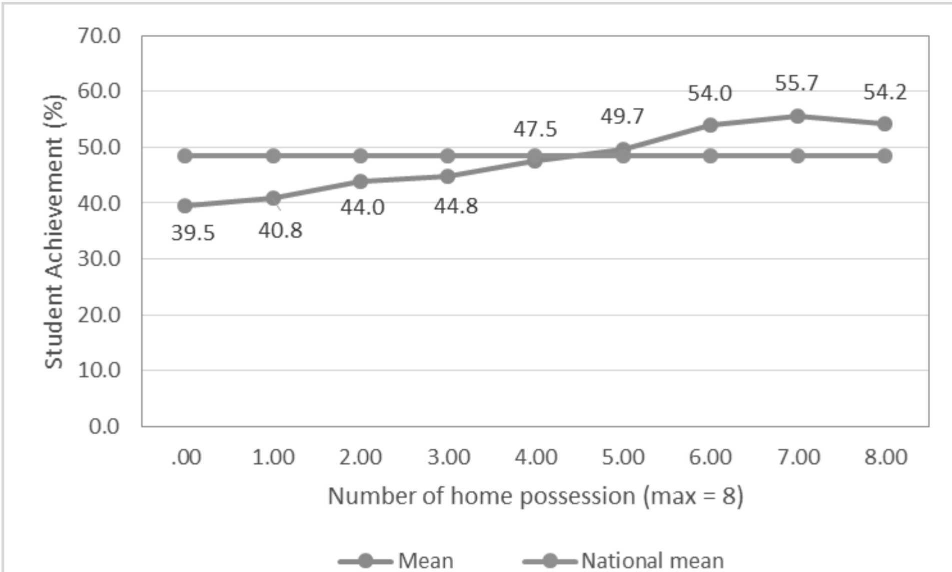
Dataset shows that when parents are involved in labour, household work, agriculture and foreign country, student achievement is significantly lower than the students whose parents are involved in teaching, government job and business. Especially, there is the risk of achieving low score when mother is involved in labour or works in other's home.

Home possessions and accessories

There were two kinds of home possessions defined in the background information questionnaire given to students. One is related to the facilities that are supportive for the students' study at home. These facilities are, for example, a table for study, a separate room for student, a peaceful place for study, a computer for school work, software for computer assisted learning, internet facilities, literary magazines, access to classical literature, poetry books, or artistic things like pictures, dictionary and other books that support children for their study. Another type of home possession includes different types of normal home accessories such as the number of mobile phones, televisions, computers, and radios.

There were 12 questions in the background questionnaire filled by the students related to home possessions. However, a separate room for study, computer for school work, software for assisted learning, and own calculator have shown negative correlation with the achievement, they are grouped in the bunch. Besides these negative items, all other 8 positive indicators are recorded, and each was scored 1 if the student had the access to the possession. Adding these items up, the maximum score is 8, indicating that the students have the access to all of the possessions. The result shows that higher access of home possessions provides high score to the students; hence the lower the score, the fewer are the possessions at home. Figure 4.20 shows relation between home possessions and achievement level where the achievement level of the student goes up when there is access to home possessions. Pearson Product Moment Correlation Coefficient between the achievement level and the factor ($r = 0.23$) is statistically significant ($p < 0.001$) though it is not very high, and the effect size is moderate ($f = 0.20$).

Figure 4.20 Relation between home possessions and achievement in Mathematics



The figure shows the relation between home possessions and achievement in Mathematics. It is evident from the figure that when the number of accessories increases, the achievement also increases. When home possession is none (zero), the achievement is very low (39.5%), when 7 out of 8 are present, the achievement is the highest (55.7%).

In the background questionnaire, the question asked to the students was "how many of the following accessories do you have in your family?" The options were 0, 1, 2 and 3 or more. After dichotomizing the items individually by using meaningful cut-offs found with ANOVA and DTA (see table 4.14), only mobile and television were added as indicators. The maximum score was 2, indicating that the students possessed at home a set of all the accessories.

Table 4.14 Dichotomizing the indicators for home accessories

Accessory	Cut-off for 1	Cut-off for 0
Mobile phone	1	0
Television	1	0
Computer	1	0 (negative correlation, so discarded)
Radio	1	0 (negative correlation, so discarded)

The relation between the availability of home assessories and achievement is presented in the following figure.

Figure 4.21 Relation between home accessories and achievement in Mathematics

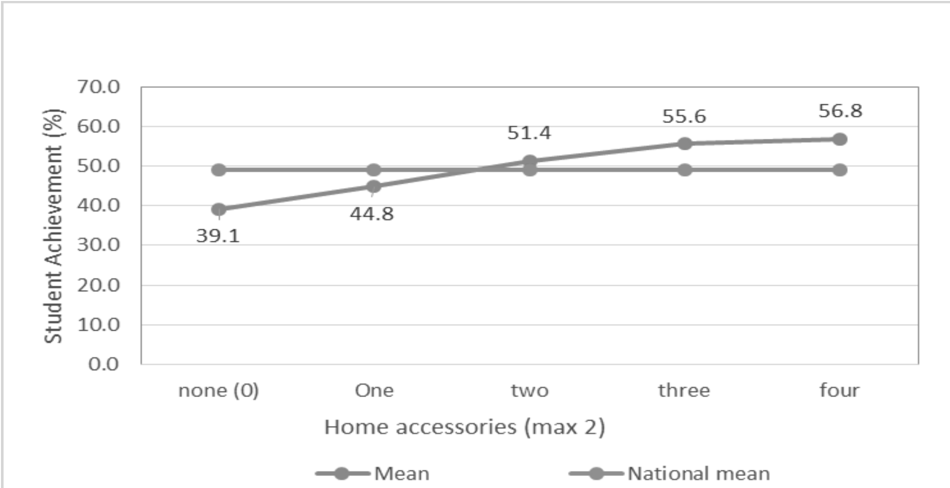


Figure 4.21 clarifies how the number of home accessories affects student achievement. Students' achievement is 39.1% if none of the accessories is available, and it is 56.8% if all of them are available. Availability of all the stated facilities indicates the high SES of the family; and therefore it could be linked with SES of the family. Correlation between home accessories and achievement is $r = 0.25$ ($p < 0.001$) which is positive, but not very high. The effect size of $f = 0.27$ indicates a high size of difference between the groups. The difference between the lowest group with no home accessories and the highest group with all of them accessories is remarkable (-17.7).

Data shows that when the children have 0 to 3 home possessions out of the 8, the achievement level is lower than the national average and those having more than 3 home possessions have the achievement higher than the national average. The average achievement score of students is the highest (55.7%) when the families have 7 possessions. The same pattern shows with home - as if there is no any one of the four accessories, the results are very poor (39.1); and when there are all of the accessories, the results are remarkably high (56.8%). A negative correlation is found between the achievement and possessions when families have computer, internet and computer assisted learning software and more than one car.

SES and achievement

The socio-economic status was analysed based on seven indicators, which were all first dichotomized. The seven variables indicating socio-economic status are: mother's education, father's education, mother's occupation, father's occupation, home possessions, home accessories, and the type of school attended by the students. Score on each of the variables were added and the total score of each variable was changed into percentage (P-SES). P-SES represents the percentage of SES the student possesses, whereby 100 percent denotes the student has the highest

SES measured with these variables, that is, all the seven indicators of SES are positive; and 0 refers to the lowest possible SES, that is, all the seven indicators of SES are negative. The analysis of the PSES by using Univariate GLM (that is the Regression modelling) shows that there is a strong relation between SES and achievement. Figure 4.22 below presents the relationship between SES of the students and the achievement.

Figure 4.22 Relation between the SES and achievement in Mathematics

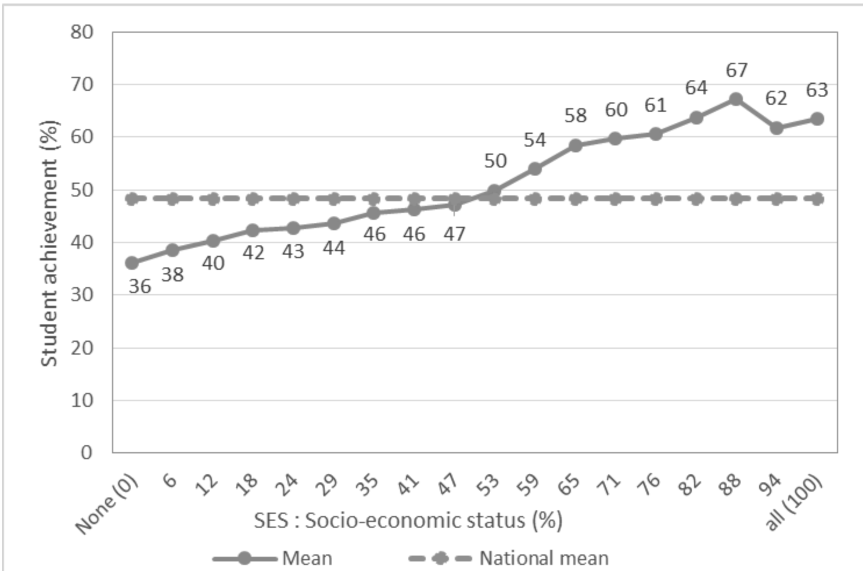
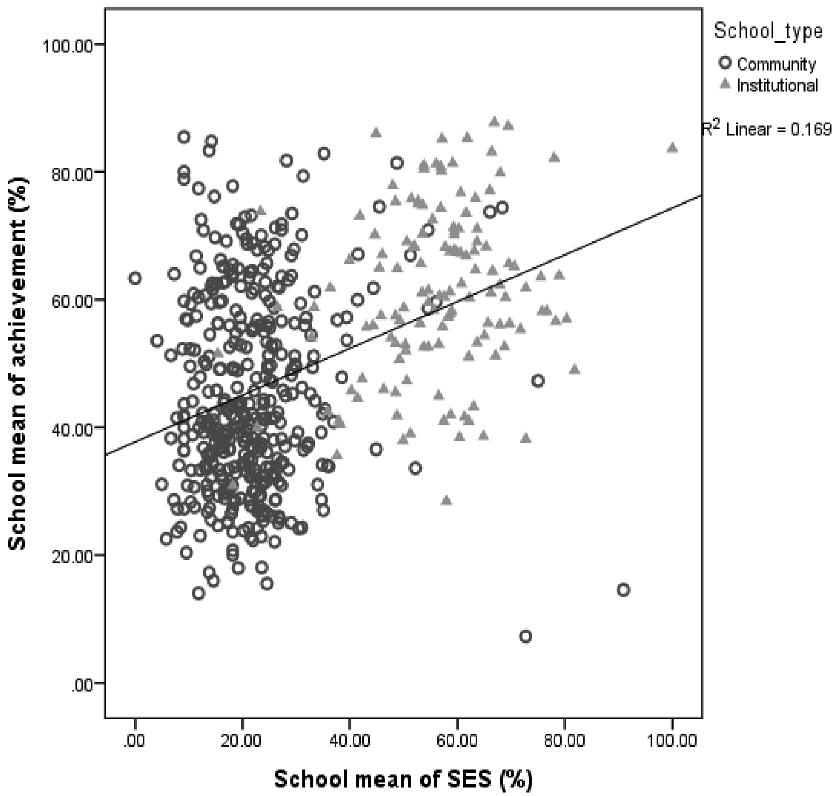


Figure 4.22 shows a positive relationship between SES and achievement; the correlation between the variables is $r = 0.33$ which is significant ($p < 0.001$). The differences between the SES groups are statistically significant ($p < 0.001$) and the effect size is high ($f = 0.36$); that is, the highest and lowest groups differ from each other remarkably. SES explains 12% student variation in data ($\eta^2 = 0.12$).

The dataset suggests that socio-economic status plays a vital role in student achievement. The difference in achievement between the lowest and highest SES groups is remarkable.

Analysing further with the scatter diagram based on SES and achievement, figure 4.23 shows that two types of schools (community schools in circle and institutional schools in triangle) fall into two groups. The institutional schools, with relatively higher SES, are concentrated more on relatively high performing group whereas community schools, with relatively lower SES, fall from very high performing to very low performing group, concentrating more towards low performing group.

Figure 4.23 Distribution of achievement by SES and type of schools in Mathematics

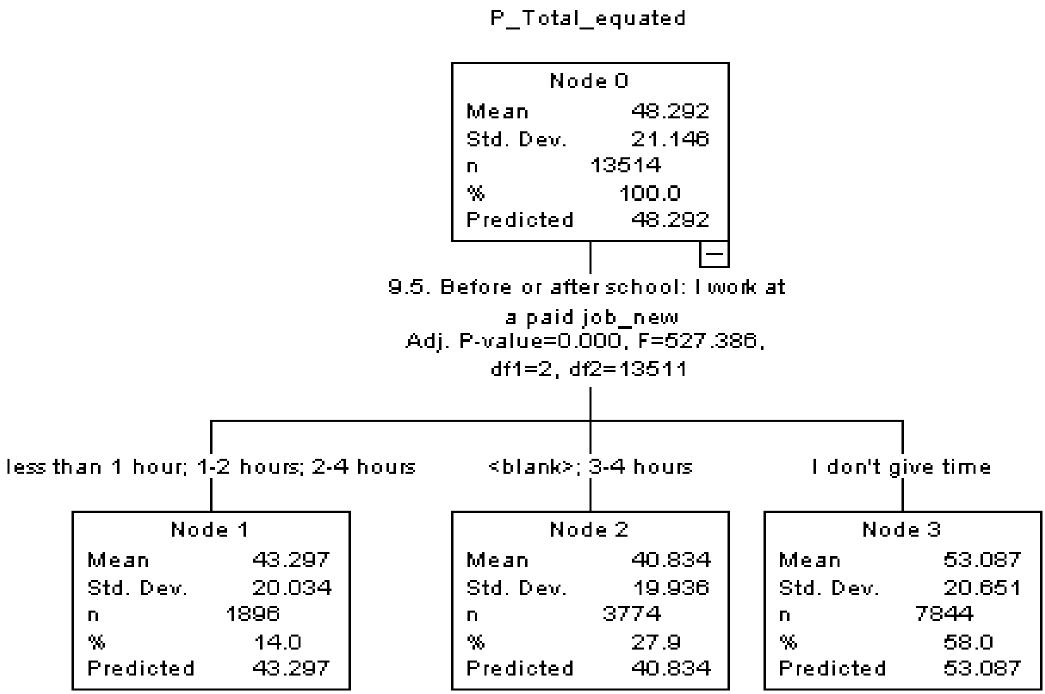


The dataset reveals that the students in the institutional schools with relatively high SES are concentrated more towards high performing group and the students in the community schools, with relatively low SES, are concentrated more towards low performing group.

Working Beyond School Hour and Achievement

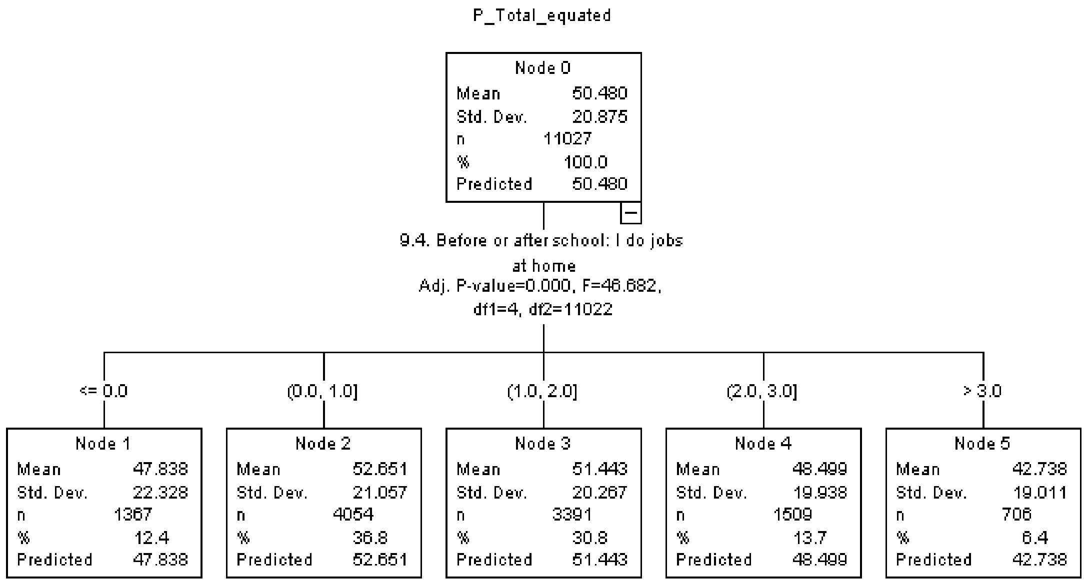
There were questions in the student background questionnaire related to their activities beyond the school hours. The questions were related to working before/after the school hours in a paid job and involvement in household chores. The values of each variable are divided into five categories and marked on a five-point scale. The scale was 0 point for working not at all, 1 point for the work of less than 1 hour per day, 2 points for the work of 1 to 2 hours per day, 3 points for the work of 2 to 4 hours per day, and 4 points for the work of more than 4 hours per day. Data shows that when the children are not engaged in paid work at all, their results are notably above the national average (53.1%) (see figure 4.24). If the students work for paid job, even less than an hour to four hours, the results are lower (41.52%) than the national average. Further, if they were working more than 4 hours, the results were remarkably low (40.8%). The differences are statistically significant ($p < 0.001$) though the effect size is moderate ($f = 0.28$), as most of the children do not need to work in a paid job. Working beyond the school hours for earning money indicates that the family is poor and extra earning is needed. When the student needs to work more than 4 hours per day, there is no time or energy to be involved in school homework.

Figure 4.24 DTA of paid work and achievement in Mathematics



Next, the DTA (fig. 4.25) shows that when the children spend more than 2 hours per day in household chores, the results are poor. However, when the amount of time spent in household chores is up to 2 hours per day, the achievement level is higher than the average (41.4% for upto 2 hours and 52.7% up to 1 hour). Differences are significant ($p < 0.001$) though the effect size is very low ($f = 0.09$) as around 20 percent children participate in household work for more than 2 hours per day. However, very low effect size suggests very low effect on achievement.

Figure 4.25 DTA of household work and achievement in Mathematics



The dataset shows that paid work of the students beyond the school hours reduces the achievement. Getting engaged in unpaid household work up to two hours gives positive effects on student achievement, whereas when the children need to work for more than 2 hours per day either paid or unpaid, the achievement level drops down remarkably.

Student Attitude Towards the Subject of Study and Achievement

The attitude about the subject tells us what the students think about Mathematics and its usefulness in their daily life and future utility. The correlation between Mathematics achievement and attitude toward Mathematics is widely studied, and it has been found that there is a relationship between the attitude of the students and achievement though the connection is not always clear (see, for example Metsämuuronen 2012a; 2012b; House & Telese, 2008; Shen & Tam, 2008; Kadijevich, 2006; 2008).

In NASA 2015, the shortened version of Fennema–Sherman Attitude Scales (FSAS) (Fennema & Sherman, 1976) used in several international comparisons, like in TIMSS and PISA studies, was used to identify the relation between attitude towards the subject and their achievement. The original scale included nine dimensions, but in these international comparisons, only three were used with four items on each and two negative items on each of the first two dimensions. The names of the factors can be "Liking Math", "Self-Efficacy in Math", and "Experiencing utility in Math" (compare naming in, e.g., Kadijevich, 2006; 2008). Factor analysis was used to identify the factors of the responses in FSAS and the negative items were reversed to make the whole test unidirectional. As in several countries of Asia, the expected factor structure cannot be found in

Nepal (for a deconstruction of the test scales, see Metsämuuronen, 2012a; 2012b). Hence, only the total score is used to show the relationship between attitude and achievement. The relation between the attitude, which is divided into deciles of somehow equal number of students, and achievement score, is shown in figure 4.26.

Figure 4.26 Relation between students' attitude and achievement in Mathematics

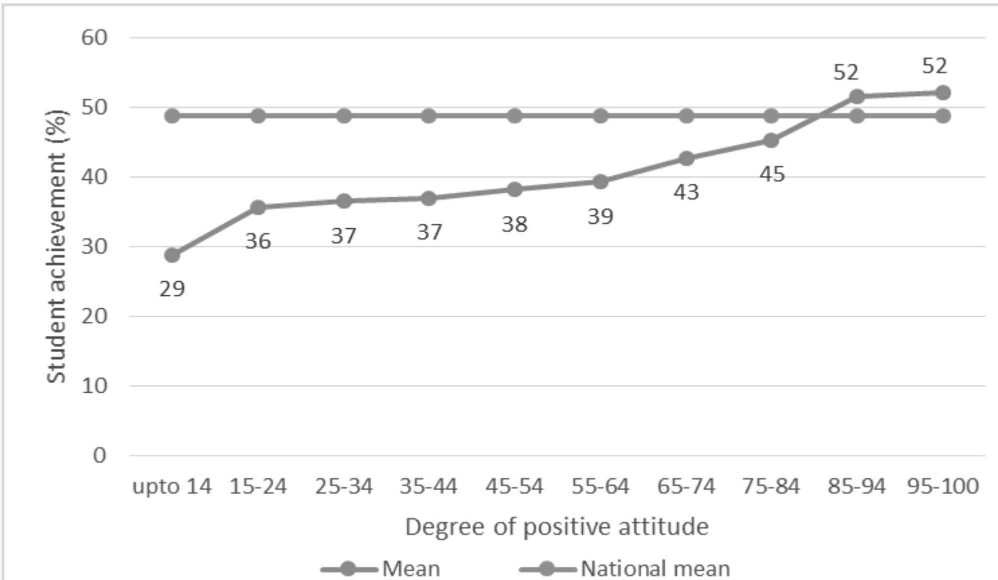


Figure 4.26 shows that the difference between the lowest attitude group (29%) and highest attitude group (52%) is remarkable. However, a big difference in achievement is found only in the two highest attitude groups, that is 52% and 52% and the lowest group. The correlation between positive attitude towards Mathematics and achievement is $r = 0.25$ ($p < 0.001$); and the effect size is moderate ($f = 0.27$).

Data reveals that the more positive the attitude is towards Mathematics, the higher is the achievement. The data also supports the fact that positive attitude influences the achievement positively.

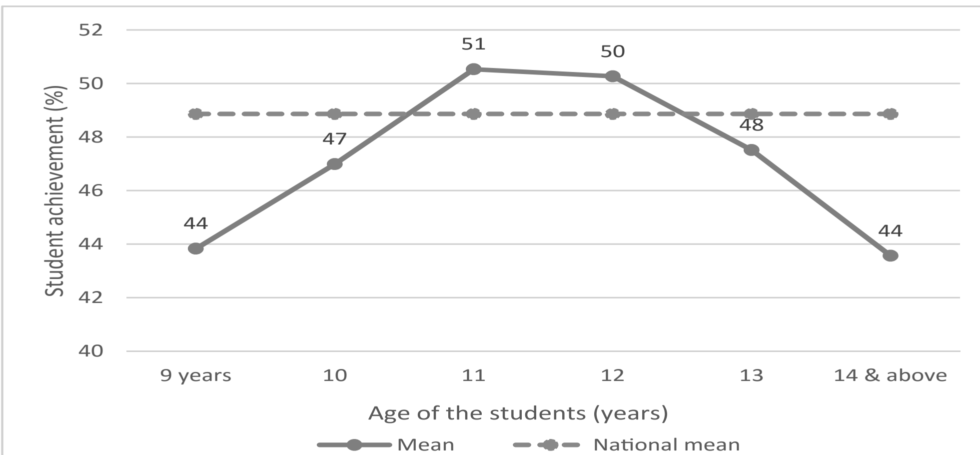
Age of the Student and Achievement

Although official age for Grade 5 students is 11 years in Nepal, the age of the students attending Grade five varies widely. Some students have mentioned their age below nine years and some above 14 years. The age of all the students below 9 years were encoded as '9 years', and all students above 14 were encoded as '14 years and above'. The descriptive statistics of the mean in each year are presented in table 4.15 and further depicted in figure 4.27.

Table 4.15 Descriptive statistics of the students' achievement in different age groups

Age	N	Mean	SD
9 years	151	43.8	20.97
10 years	1211	47.0	21.10
11 years	3376	50.5	20.57
12 years	4612	50.3	21.25
13 years	2069	47.5	20.77
14 years & above	1198	43.6	20.19
Total	**12617**	**48.9**	**20.99**

Figure 4.27 Relation between the age of student and achievement in Mathematics



Data indicates that the best achievers are the students having the proper age for Grade five, that is, the students of 11 to 12 years. The achievement level is remarkably low when the students are above the age of 13 years. It clearly shows that students with higher or lower age have weaker results. Students of higher age are those who either started school much later than they should have, or repeated the classes. Correlation between the variables is r = - 0.015 (p < 0.001) and effect size is found low (f = 0.11). The ANOVA hints that the age explains 1% variation in the achievement level.

Dataset indicates that the highest performance is found in the students studying at their correct age group, that is, at the age of 11 and 12 years. Otherwise, the achievement decreases as the age of the student is lower or higher both, correlation between achievement and age is negative, indicating that higher age, results in lower achievement.

Support for Study to the Student and Achievement

The relation between the support provided for study and achievement was analysed based on the responses to the question: *"Who supports you when you do not understand what you have studied or felt difficult?"* In the question, students were asked to select only one option. This assessment did not consider the case where there were more than one supporters. The descriptive statistics of the supporters are given in table 4.16.

Table 4.16 Achievement score in Mathematics in relation to the support provided to the students

Support Providers	N	Mean	SD
No one	651	52.4	21.40
Teacher	851	53.1	21.96
Mother	922	51.4	21.08
Brother/Sister	6540	50.1	20.33
Father	1729	49.3	21.39
Tuition	1835	42.9	19.71
Any two or more	986	35.2	21.07

Support in Mathematics is necessary for students to achieve better score. There is about 1 percent difference between the students who did not receive any kind of support and those who received support from teacher. Interestingly, the data shows that those students obtained more score who studied independently (52.4%) than those who received support from their family members (49.3 to 51.4%). Students who obtained lowest score among the groups are those who are dependent on tuition (42.9%). However, some students reported that they get help from more than one ways. They are the lowest achievers (35.2%).

The result shows positive correlation between student achievement and supports received from teacher even after the school hour. Students receiving supports from teacher have higher achievement than those students receiving support from their family members. However, students studying independently without taking any support from family members obtained better score than those who had received support from family members. Students who need to seek tuition or help from many people are the lowest achievers.

Availability of Textbook and Student Achievement

The data shows that there were some students who did not have Mathematics textbook even up to the end of academic session. Table 4.17 shows the descriptive statistics of the availability of the textbook of Mathematics and the achievement.

Table 4.17 Availability of textbook of Mathematics and the achievement

Availability of Textbook	N	Mean	SD
Yes	12090	49.6	20.79
No	464	40.5	20.78
Total	**12554**	**49.3**	**20.86**

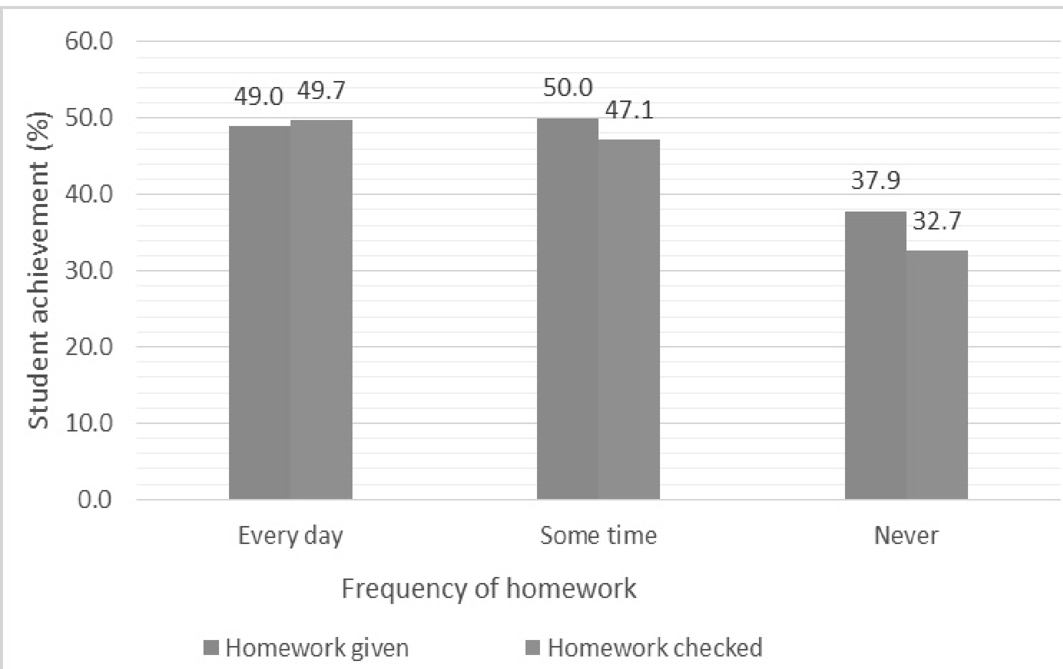
Out of 12,554 students who responded the question, about 4% of the students did not receive Mathematics textbook even up to the end of the academic session. The relation between the availability of textbook and achievement is significant ($p < 0.001$) though the effect size is very small ($d = 0.04$) due to the small group size. The difference in achievement is big, that is 9 percent.

Data shows that 4% students did not receive textbook of Mathematics even up to the end of the academic session. The achievement level of the students without textbook is significantly lower than those who have access to the textbook; however, effect size is small.

Homework and Achievement

Homework is considered as one of the ways to enhance learning, which can be used as drill, exercise and an evaluation tool as well. When homework is regularly checked and feedback is provided, it is likely to boost achievement levels. Statistics related to homework given and feedback provided is presented in figure 4.28.

Figure 4.28 Effect of homework in students' achievement in Mathematics



If teacher assigns homework and provides feedback to the students regularly, then the students' achievement becomes higher than the national average (49.7%). If homework is not assigned regularly and feedback is not provided to them, the achievement score is below the national average (32.7%). When home assignment and feedback are never given, then the average achievement score is remarkably low (11%). The differences are statistically significant ($p < 0.001$). Those groups with no homework assigned are, however, very small and hence, the effect size is small ($\eta^2 = 0.003$ and 0.009).

Dataset is evident that if the teacher assigns homework and provides feedback to the students regularly, the achievement is higher than those students who were not assigned homework or not provided feedback.

Positive and Negative Activities in the School and Student Achievement

The activities of students and teachers determine the learning environment of a school. Bullying, for example, is one of the hindering incidents for the students in school that may affect learning. In the students' background questionnaire, several schools related activities were asked to the students. Here, bullying is considered as one of the negative indicators and students' impressions of school and teacher activities are taken as the examples of positive indicators.

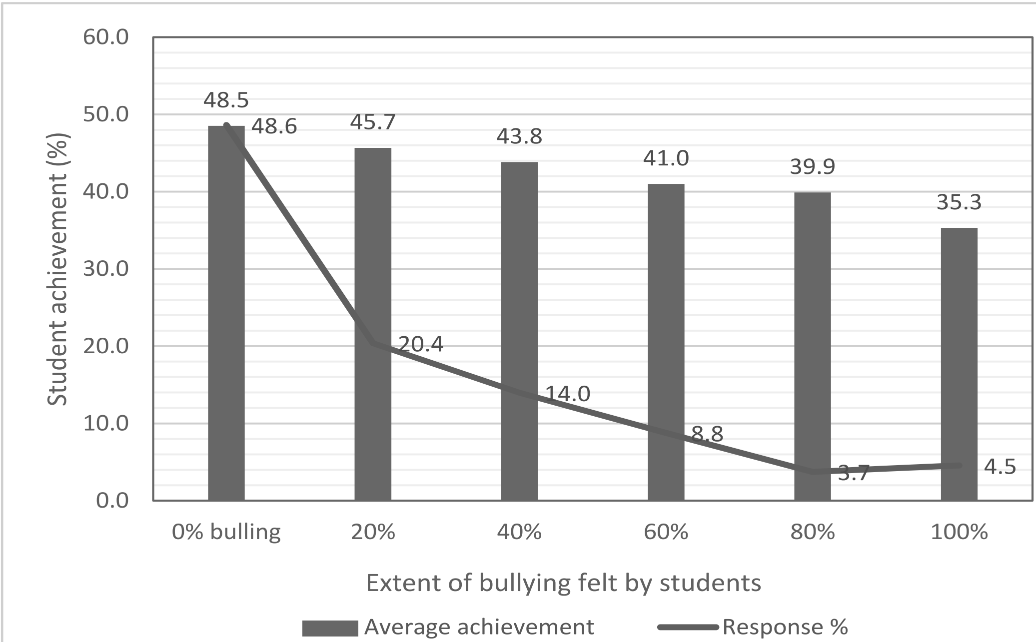
Bullying at school

Bullying is one of the problems in schools as it worsens the learning environment for the students. International Studies like TIMSS and PISA emphasize to identify such indicators. In NASA 2015 student background questionnaire, five questions indicate the varieties of bullying that tend to happen in schools. All the incidences were stemmed by the phrase "*Which of the following activities happened in your school last month?*" The students' responses are presented in 4.18 and further depicted in figure 4.29. 'No (%)' indicates the percentage of the students' response of no such activity happened in the school, and 'Yes (%)' indicates the percentage of the students who experienced the particular type of bullying within the last month. Alone, 32% of the students mentioned that, during the last month, something of their own was stolen, which is an alarming sign of the system.

Table 4.18 Bullying and the achievement

Type of Bullying	No (%)	Yes (%)
BQ_30_1 Incidents: Something of mine was stolen	68	32
BQ_30_2 Incidents: I was hit or hurt by other student(s)	76.2	23.8
BQ_30_3 Incidents: I was made to do things I didn't want to do by other students	84.9	15.1
BQ_30_4 Incidents: I was made fun of or called names	77.2	22.8
BQ_30_5 Incidents: fellow students kept outside without involving me in activities	87	13

Figure 4.29 Effect of bullying in the achievement of Mathematics



The sum of all five items; that is, total of 100 percent, is considered an indicator of highest bullying. Figure 4.29 shows the extent of bullying in percentage and achievement of the students in each category of bullying. The line graph shows that the percentage of students encountered various levels of bullying. If only one activity of bullying is reported, it is categorized as 20% bullying, whereas if all five activities are reported, it is categorized as 100% bullying. Data shows that 48.1% students did not encounter any kind of bullying in a month. One can infer from this that the remaining 51.9% have encountered at least one type of bullying, which is a remarkable number of students. About 5.3% of the students are experiencing a severe kind of bullying, which is the sum of the students who encountered 80% and 100% bullying. It is found, however, that learning outcomes are remarkably low only with 2.0% of the students who have encountered extreme bullying including all five types of harassments (39.5%). Students who did not feel bullying and those who encountered extreme bullying of all five kinds have almost 10.9 percent achievement gap although a very small number of students reported all kinds of bullying (n = 263). However, the difference is statistically significant (p < 0.001) though the effect size is small (f = 0.11). Though the extreme case of severe bullying is rare, bullying is found quite common in schools.

The dataset reveals that a large number of students (51.9%) have encountered bullying in schools. Though the phenomenon does not have a great effect except in the group of extremely bullied students. All possible efforts should be put to root out the phenomenon from schools.

Positive activities at school

The activities that can boost the learning achievement of students are categorized as positive activities. Such positive activities at school were asked to the students in two sets of questions listed in table 4.19. The table shows that the responses of the students in all four categories, which are in the 4-point rating scale, anchored to fully agree and fully disagree.

Table 4.19 Students' response towards teacher and school-related activities in Mathematics at school

Teacher and Students activities	Respondents in %			
	Fully disagree	Partially disagree	Partially agree	Fully agree
Students get along well with most teachers	1.9	1.2	7.8	89.1
Most teachers are interested in student's well-being	2	1.3	6.3	90.4
Most of the teachers really listen to what I have to say	2.3	3	18.6	76.1
If I need extra help, I will receive it from my teacher	2.2	2.5	8.6	86.7
Most of my teachers treat me fairly	2.9	2	9.3	85.8
I enjoy to be in the school	1.6	0.7	3.2	94.4
Students in my school like me	1.9	1.6	10.5	86.1
Friends of my school want to their best in the study	2	1.7	5.6	90.6
Teachers' expect good results of their students	2.2	1.4	3.8	92.5
Average	2.1	1.7	8.2	88.0

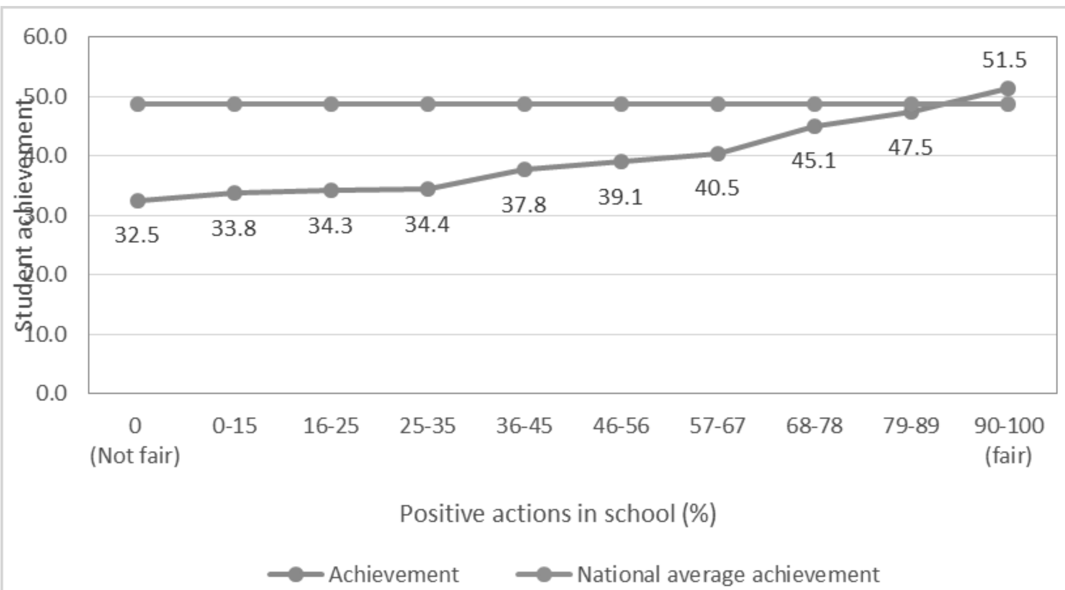
It was further analysed by recoding the variables into two categories, that is, 1 for agree and 2 for disagree. Furthermore, the sum of nine indicators is converted into the percentage of maximum score to analyse the level of positive activities and its relation with achievement. The overall result is that the feeling of the positive actions in the school relates positively with the student achievement. The correlation between the sum of nine positive activities and achievement is positive ($r = 0.24$, $p < 0.001$).

Univariate ANOVA finds four attitude groups in the indicator. These boundaries and descriptive statistics are presented in table 4.20 and further illustrated in figure 4.30.

Table 4.20 Students' response towards teacher and school-related activities in schools

% of Total Score	N	Achievement	SD
Not fair (0%)	85	32.5	18.80
0-15	151	33.8	18.19
16-25	215	34.3	17.66
25-35	254	34.4	16.93
36-45	271	37.8	17.85
46-56	392	39.1	19.87
57-67	448	40.5	18.90
68-78	800	45.1	20.53
79-89	1590	47.5	21.43
90-100	9016	51.5	20.53

Figure 4.30 Relation between positive actions at school and achievement in Mathematics



The data shows that there is a positive correlation between the students' feeling about the school activities and achievement. After dividing the indicator into four groups based on DTA, the differences between the groups are statistically significant ($p < 0.001$); however, the effect size is small ($f = 0.25$). The learning achievement is higher than the average only when the students are extremely positive towards school and teachers' behaviour. However, the difference between the most positive and the most negative groups' achievement is notable, that is 19 percent score.

Dataset records that when students feel that the actions of the teachers and the schools are ultimately good, the results are better than average. At the other extreme, if they feel that such actions are negative, the results are far below the average.

4.6 Summary of Findings

The main findings of the assessment in Mathematics for Grade 5 in NASA 2015 are as follows:

Basic results

- In Grade 5, Mathematics achievement score is distributed normally. However, based on the distribution of achievement scores, the student population can be grouped into two types−namely, low performing, and high performing, although the majority of the students fall in the low-performing group. The low-performing students achieved 15–45% and the high-performing students achieved 45–100%.
- The overall national average achievement score for Mathematics is 48.3%, but the students' achievement varies in the four content areas, for example, score in *Algebra* is 46.3% and *Numeracy* is 47.4%, which are below the nationl mean. Similarly, achievement score in *Arithmetic* (50.5%) and *Geometry* (48.5%) are higher than the national average. Dataset indicates that the students are weak in *Algebra* and *Numeracy* than in other content areas.
- Dataset indicates that students are good at recognizing the correct answer and in very fundamental contents like recognizing numbers, shapes, choosing correct answer from the options, very basic operations such as the basic manipulation of data and numbers, and calculations in single or few steps. They are much weaker in calculatons having more than one step, questions involving the process of reasoning, problem solving, plotting, proving the theory or formula, and constructing the shapes and figures. In many cases, the students did not even start doing the open-ended questions; hence, they got low score.
- Latent ability of both years (NASA 2012 and NASA 2015) is quite close. However, the latent ability of 2015 is slightly lower than that of 2012.

Achievement with Diversity Factors

- The difference in achievement between the Districts is significantly wide. The variation explained in achievement among the Districts is 10%. Achievement of students in Kathmandu (60%) and Mustang (60%) are on the top whereas the students in Bhojpur (36%), Pyuthan (36%) and Sankhuwasabha (36%) are on the bottom as they secured lowest mean score.
- Dataset indicates that there is a wide difference between the student performances in four Ecological Zones. Students in the Kathmandu Valley have outperformed (57.5%) the other students. The achievement is the lowest in the Mountain (42.4%). Moreover, average achievement of Tarai Region (45.4%) is higher than Hill (41.6%).
- The dataset reveals that there is wide inequity among the Development Regions for children's opportunities to reach an adequate level in Mathematics. The difference between

the lowest performing (Mid-Western, 42.9%) and the highest performing Region (the Kathmandu Valley, 57.5%) is remarkable, that is more than 14.6 percent. For the Regions, except the Kathmandu Valley and Western Region, performance is below the national average. However, if the Kathmandu Valley is excluded from the analysis, the difference in achievement between the highest performing (Western) and lowest performing Region (Mid-western) will be reduced to 6.3 percent.

- The dataset indicates that the average performance of the students from institutional schools is 60.8% whereas it is 44.5% in community schools, that is, the difference is 15 percent, which is remarkable. Compared to the result of NASA 2015 in Grade 3, variation between community and institutional schools is wide in Grade 5.
- Student performance significantly varies based on the location of the schools. Urban schools outperformed (54.8%) the rural schools (47.4%) by 7.4 percent score in Mathematics. However, the difference is medium.
- Nepali speakers outperform non-Nepali speakers by 3.9 person points. Among the various language groups, Newari speakers achieved the highest score (55.3%) and Magar speakers achieved the lowest score (40.4%). Variation between highest and lowest speaking student groups is medium.
- Dataset shows that Brahmans students perform the highest (55.6%) and Minorities the lowest (37.5%). Caste/ethnicity-wise achievement score varies in most of the Ecological Zones. All of the castes/ethnicities performed the highest in the Kathmandu Valley. Students from Madhesi and Minority ethnic groups, although their number is very low in the Valley, performed the highest (62.8% and 63.6% respectively) among all castes/ethnicities.
- There is a small difference in achievement between boys (49%) and girls (48%), and this is statistically significant. Variation in achievement between them is very low, indicating that there is not significant gap between boys and girls in achievement. This is positive sign from equity point of view.
- In Grade 5 Mathematics, girls are performing better than boys in Central region and the Kathmandu valley. However, boys performed better than girls in other regions and the difference is significant in all regions ($p < 0.05$) except in the Central region. In total boys achieved higher (1.5% above) than girls, which is statistically significant, though the difference is low.

Selected Explanatory Factors and Achievement

- Result explains that the educational level of the parents predicts the children's achievement level in Mathematics. When both parents are illiterate, achievement of their children becomes very low compared to the children of other parents with SLC or higher educational level. Mother's education upto SLC is more effective than that of father's whereas above SLC, father's educational level is found effective in average achievement score.

- Dataset shows that when parents are involved in labour, household work, agriculture and foreign country, student achievement is significantly lower than the students whose parents are involved in teaching, government job and business. Especially, there is the risk of achieving low score when mother is involved in labour or work in other's home.
- When the children have 0 to 3 home possessions out of the 8, the achievement level is lower than the national average; and those having more than 3 home possessions have the achievement higher than the national average. The average achievement score of students from the families with 7 possessions is the highest (i.e., 55.7%). The same pattern shows with homes where there is not any one of the four accessories. The results are very poor in such cases (39.1%); and when there are all the accessories, the results are remarkably high (56.8%). A negative correlation is found between home possessions and student achievement when families have computer, internet, computer assisted learning software and more than one car with the student achievement.
- The dataset suggests that the socio-economic status plays a vital role in student achievement. The difference in achievement between the lowest and highest SES groups is remarkable.
- The dataset reveals that the students in the institutional schools with relatively high SES are concentrated more towards high performing group and the students in the community schools with relatively low SES are concentrated more towards low performing group. Overall, SES indicator explains 17 percent of the school mean of student achievement.
- The dataset shows that paid work of the students beyond school hours reduces the achievement. Getting engaged in unpaid household work up to two hours gives positive effect in student achievement, when the children need to work for more than 2 hours per day (either paid or unpaid), the achievement level drops down remarkably.
- As data reveals, the more positive the attitude towards Mathematics, the higher is the achievement. Data also supports the fact that positive attitude influences the achievement positively.
- Dataset indicates that highest performance is found in the students studying at their correct age group, that is, at the age of 11 and 12 years. Otherwise, the achievement decreases as the age of the students is lower or higher both. Correlation between achievement and age is negative, so higher age results in lower the achievement.

- The result shows positive correlation between support from teacher (even after the school hour) and achievement. Assessment of students receiving support from teacher is higher than that of the students receiving support from their family members. However, students studying independently without taking any support from family members and teachers obtained better score than those who received support from them. Students who need to seek tuition or help from many people are the lowest achievers.
- Data shows that 4% of the students did not receive textbook in Mathematics even up to the end of the academic session. The achievement level of the students without textbook is significantly lower than those who have access to the textbook; however, effect size is very small.
- Dataset is evident that if the teacher assigns homework and provides feedback to the students regularly, the achievement rises higher than without assigning homework or not providing feedback.
- The dataset reveals that a large number of students (51.9%) have encountered bullying in schools. Though the phenomenon does not have a great effect except in the group of extremely bullied students.
- Dataset records that when students feel that the actions of teachers and schools are ultimately good, the results are better than average. At the other extreme, if they feel that such actions are negative, the results are far below the average.

Chapter 5: Analysis of Assessment Results for Grade 3 in Nepali

5.1 Introduction

Nepali is one of the compulsory subjects in Grade 3. Out of total 34 credit hours per week, eight credit hours are allocated for Nepali subject in the curriculum approved by the Government of Nepal. This assessment is based on the learning outcomes/curricular competencies as set in the approved curriculum for Grade 3 students in Nepali subject. The overall objective of Grade 3 Nepali is to develop the basic language proficiency of Nepali in four language skills – listening, speaking, reading and writing as expected by the curriculum of Nepali (CDC, 2062 BS: p. 6).

Proficiency of 3rd Graders in Nepali language has been assessed frequently in the national assessment since 1995. Before 2011, 3rd Graders' achievements in Nepali language were assessed three times in 1995, 1997 and 2001 (See BPEP, 1995; EDSC, 1997 & 2001). The first large-scale National Assessment of Student Achievement (NASA) 2012 also focused on assessing 3rd Graders' achievement in Nepali. Now, again the 3rd Graders' achievement in the subject is in focus in this assessment. The frequent assessment in this subject is motivated by the fact that reading and writing skills in Nepali are inevitable to live an informed life in modern society, which is surrounded with written information and proficiency in these skills is essential among citizens to be able to acquire and communicate information in written form of various patterns. As Nepali is not only the official language in Nepal but also the lingua franca throughout the country, modern citizens should have an adequate command in this language to be able to handle such information in Nepali and able to tackle the everyday language functions. Most importantly, the basic proficiency in this language since the early Grades is essential for all to succeed further academic life as textbooks, reading materials and medium of instruction are dominantly in Nepali.

Though proficiency in Nepali for 3rd Graders has been assessed since 1995, the results of earlier assessments (1995, 1997 and 2001) are not fully comparable with the later results of 2012 and 2015 because of lack of proper linking procedures between the tests as those previous results did not state the percentage of correct answer by items.

This chapter presents the analysis of the assessment results of Grade 3 in Nepali. First, it describes the overall distribution of achievement scores. Then it presents achievements in various content areas, cognitive domains, and types of items and the comparison of NASA 2012 results in Nepali with NASA 2015. Results related to various diversity factors including District, Ecological Zone, Development Region, school type, school location, language at home, and caste/ethnicity are described in section three of this chapter. Section four of this chapter deals with the selected explanatory factors about students' achievement of Grade 3 students in Nepali. The factors that explain achievement in Nepali subject included in this section are: Parents' education and occupation, home possessions and accessories, SES, working beyond school hour, age of the student, support for study, homework given and feedback provided, activities at school (both negative and positive). The final section summarizes the findings related to the achievement of Grade 3 students in Nepali subject.

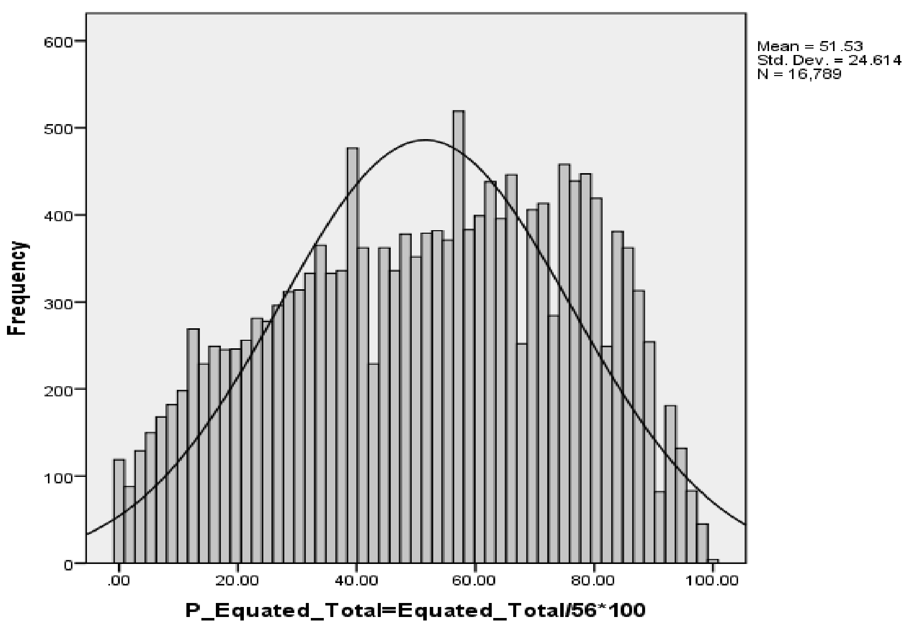
5.2 Basic Results

The basic results of assessment of Grade 3 students in Nepali includes distribution of overall results, results by content areas, levels of cognitive domains and item types and the comparison of results with previous assessments.

Overall Distribution of Achievement Scores

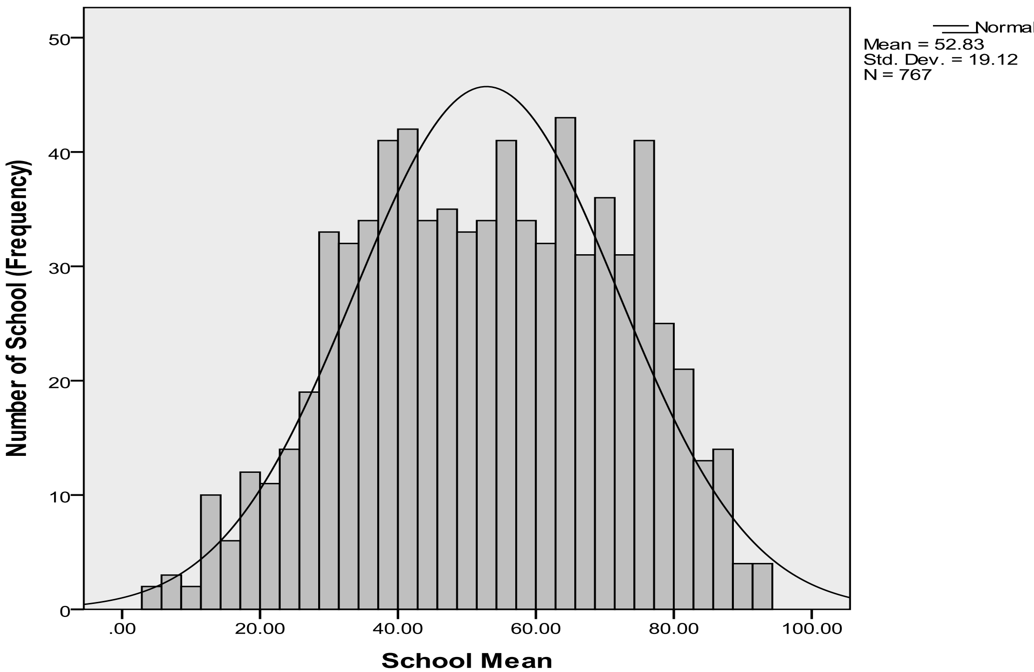
Achievement score is usually distributed normally in a large sample in the students' achievement study when the sample is large enough to assume normal distribution. However, as shown in the figure 5.1, the achievement score in Nepali is not distributed normally as it is slightly right skewed.

Figure 5.1 Distribution of overall achievement scores of Grade 3



The sample size for Nepali subject (16,789 students) was big enough to form a normal distribution. In the final score of Nepali subject, the majority of students are in the middle part of the distribution. Negligible number of students falls in the group of achieving more than 90% whereas significant number of students falls in the group achieving below 20%. Though, the dataset indicates that the Grade 3 population in Nepali subject shows that there is significant population with low score forming a larger plot of low performing students. As illustrated in figure 5.2 given below, the schools' performance is presented based on their students' performance. The figure shows that school performance is distributed normally.

Figure 5.2 Distribution of schools' mean achievement scores of Grade 3 in Nepali



Students' achievement score of the sample schools as presented in figure 5.2 shows that there are two categories of schools, namely high performing and low performing. The maximum average score of schools on the left-hand side is about 30%, whereas it is about 70% for the schools at the right-hand side. It indicates a remarkable difference in achievement score between the high and low performing schools. There more schools in low performing category than high performing category, indicating a wider disparity in learning opportunity among students within the system and persisting distinctly divided two groups of population – the low performing majority and a high performing group that is very small.

Student Achievement in Various Content Areas in Nepali

As per the curriculum, the Nepali language test includes four content areas, namely, 1) R*eading*, 2) W*riting*, 3) G*rammar*, and 4) V*ocabulary*. The weightage in four content areas of the assessment were proportionally equal to the weightage allocated in the curriculum. Even though the assessment cannot cover listening and speaking, grammar and vocabulary are taken as separate content areas.

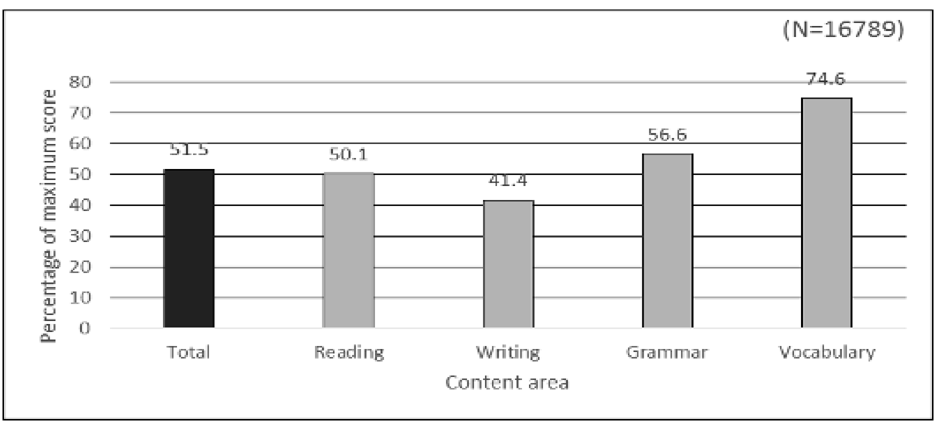
For comparability, achievements in all the content areas were converted into percentage. Table 5.1 shows the students' achievement in Nepali in total and the achievement level in each of the four content areas. Figure 5.3 further illustrates the comparison of achievement scores in various content areas.

Table 5.1 Achievement in various content areas of Grade 3 in Nepali

Content Areas	Mean	SD	Min	Max
Reading	50.1	28.98	0	100
Writing	41.4	28.01	0	100
Grammar	56.6	29.86	0	100
Vocabulary	74.6	26.23	0	100
Nepali Total	**51.5**	**24.61**	**0**	**100**

Table 5.1 shows the variation in achievement in various content areas in Nepali subject. The achievement score is 41 percent in *Writing* and 50 percent in *Reading* with 9 percent variation. Similarly, the variation in G*ramma*r and *Vocabulary* is 18 percent. Comparing the maximum and minimum scores, the situation is the same in all content areas, as the maximum score is 100 and the minimum is 0 in each of the content areas. Figure 5.3 compares the variation of achievement in various content areas.

Figure 5.3 Comparison of achievement scores in various content areas for Grade 3 in Nepali



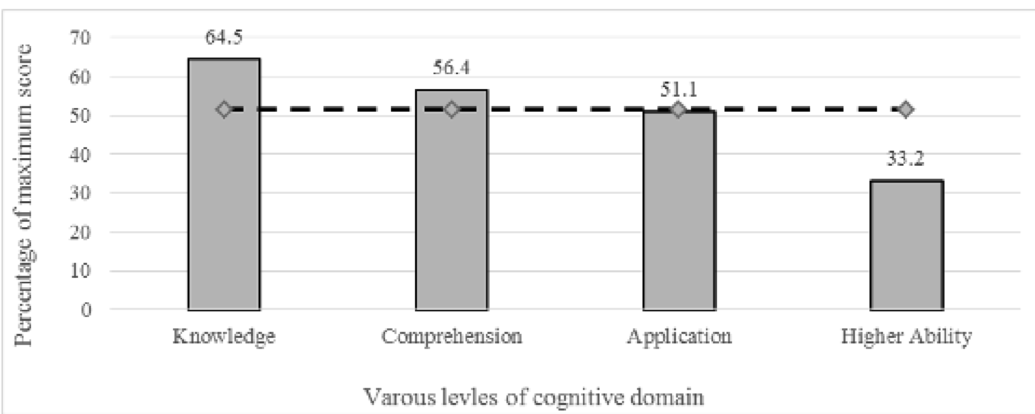
The overall national average achievement score for Nepali is 51.5 percent. Among the various content areas, students are found poorer in W*riting* (41.4%) and R*eading* (50.1%), but the performance is better than the national average in G*rammar* (56.6%) and V*ocabulary* (74.6%). Dataset indicates that the learning achievement of Grade 3 students is the highest in V*ocabulary* and the lowest in W*riting*.

To compare the results by content areas with previous test, the same trends of achievement gaps in learning are noticed. In NASA 2012, Writing was the lowest performed area (54%) compared to others (*Reading* 64.3%, *Grammar* 64.8% and *Vocabulary* 60.2%), and this trend has continued in this test as well. In previous test, the level of difference between *Reading* and *Writing* was more than 10 percent, which has not been narrowed in this test too. The performance in *Vocabulary*, which was 60.2% in the past, has improved in this test by 14 percent (74.6%) points. The overall performance level has lowered down to 51.5% in this test from the previous record of 62.6% (thus reduced by 11 percent in this test), which has further come down by 14 percent in *Reading*, by 13 percent in W*riting* and by 12 percent in G*rammar*. From this scenario, it is evident that the level of learning over the years is seen falling off instead of improving – which is a matter of serious concern for policy and planning.

Achievement in Various Levels of Cognitive Domain

The test items of Grade 3 in Nepali comprise of various levels of cognitive domain according to Bloom's taxonomy (Bloom *et al.,* 1956; Metfesser, Michael & Kirsner, 1969). The categories were K*nowledge, Comprehension, Application and Higher ability* (reasoning/problem solving/ creativity). The achievement of students in each of these cognitive levels is shown in figure 5.4.

Figure 5.4 Comparison of achievement score at various levels of cognitive domain in Nepali



The achievement score is below the national average in higher ability whereas the scores obtained in the knowledge and comprehension domains are above the national average. The performance in application domain is nearer to the national average. The data indicates that students performed better in the lower level cognitive domains than the higher levels. The difference between 'knowledge' and 'higher ability' is found remarkable (32 percent).

Going through the dataset, it is found that a large number of students (25%) were unable to solve any items requiring higher ability. As many as 20 percent students solved less than 10% of the problems requiring the higher cognitive abilities. Similarly, only 7% of them were able to attempt all the higher ability type of items; and they were also not able to secure the full marks allocated in the items.

The dataset shows that the students' ability to solve higher ability items is low (33%) compared to *Comprehension* (56%) and *Knowledge* (64%); that is, students are much better in the recall type of questions than in analytical questions. Similar trend – i.e. comparatively better performance in lower order skills and lower performance in higher order skills – was noticed (72.1% in *Knowledge*, 65.3% in *Comprehension*, 56.2% in *Application* and 37.4% in *Higher* ability) in the earlier test too. Similar to the content-wise results, the achievement level in this test has fallen off in each domain by more than 8 percent in *Knowledge*, 9 percent in *Comprehension*; 4 percent in *Application* and *Higher* ability. From the datasets of both tests, Nepali students are found somehow good at remembering, recalling or reproducing the facts but quite poor in producing and creating new which require additional interventions for improving the situation from curriculum planning and classroom delivery to the assessment practices.

Student Achievement by Type of Items

There were two types of items in the test: O*bjective* and S*ubjective*. *Objective* items covered a wide range of content areas and were very specific to judge because there was only one correct answer or one explicit piece of information required to get the correct answer to a question. On the other hand the S*ubjective* items, contained in each test version, require a longer procedure to get full marks. Both the O*bjective* and *Subjective* types of items were based on various levels of cognitive domain, which include K*nowledge, Comprehension, Application*, and H*igher ability*, covering various difficulty levels. Table 5.2 presents the basic statistics on item types and achievement.

Table 5.2 Achievement by type of items in Nepali

Item type	2015				2012			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Objective	63.0	26.93	0	100	70.5	23.5	0	100
Subjective	48.3	25.54	0	100	51.0	26.5	0	100

Students' mean score in *Subjective* items is much lower (48%) than that in *Objective* items (63%). Most of the *Objective* items were of K*nowledge, Comprehension* and *Application* type whereas most *Subjective* items were higher ability type. While comparing the current results with previous dataset, it is found that students' ability to solve S*ubjective* items has to some extent improved by around 3 percent in this test and the difference in achievement level between S*ubjective* and O*bjective* items was wide in NASA 2012 compared to NASA 2015.

Dataset clearly explains that students are performing well in recognizing the correct answer and in recalling simple facts from the texts, comprehending the basic information from paragraph, table, chart, and a few steps of logical thinking. They are weaker in writing free texts or letters, or preparing synthesis and abstracts from a text. In many cases, the students attempted open-ended tasks like free writing, problem solving and analysis; but the skills were not high enough for obtaining highest marks.

Comparison of NASA 2012 Achievement Results with NASA 2015

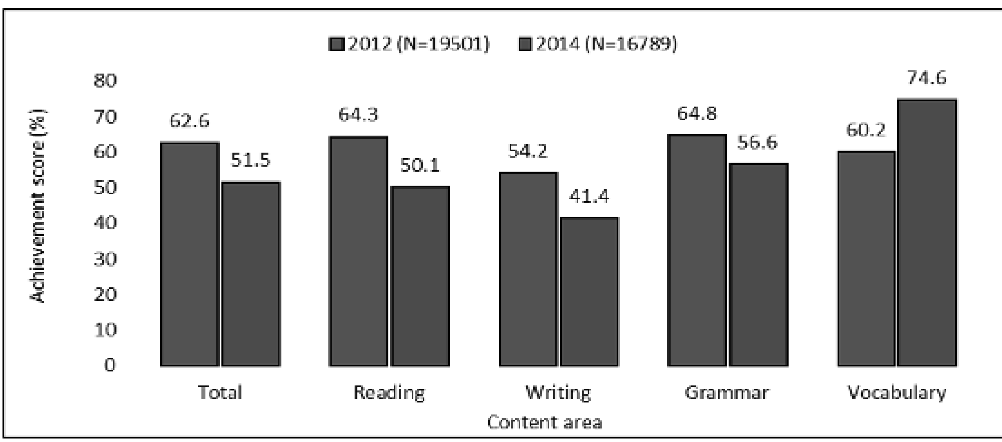
The results of NASA 2014 and 2012 are presented in table 5.3 for the purpose of comparison. However, direct comparison is difficult as the sample students and Districts were not the same in the two studies. The comparison indicates that the mean achievement score of this Grade in Nepali in the year 2015 is even lower than that of 2012.

Table 5.3 Comparison of achievements in 2012 and 2015 in Nepali

Year	N	Mean	SD	SE
2012	*19501*	62.6	24.02	0.17
2015	*16789*	*51.5*	*24.61*	*0.19*

The result in Nepali as a whole and the achievement in various content areas between NASA 2012 and NASA 2015 are presented in figure 5.5.

Figure 5.5 Comparison of results of NASA 2012 and 2015 in various content areas in Nepali



As shown in figure 5.5, the achievement level of 2015 is lower than that of 2012. Except in the case of vocabulary, achievement in other content areas in NASA 2015 is lower than in NASA 2012. The dataset indicates that the achievement score of students in Nepali has not improved over the years.

In both the tests, girls performed better than boys in all the levels of cognitive domain. The level of difference between boys and girls (by 0.3 in K*nowledge*, 1.5 in C*omprehension*, 1.6 in A*pplication* and 1.3 in *Higher ability*) noted earlier is seen to be further widened in this test from 2.3 to 2.7 with greater difference in the case of H*igher ability* (by 2.3 in K*nowledge*, 2.4 in *Comprehension*, 2.3 in *Application* and 2.7 in *Higher ability*).

5.3 Achievement Scores by Diversity Factors

Diversity is a relative and contextual term. As in the previous NASA 2012, this test also considered six diversity factors while analysing the results, which are Ecological, Regional, language, gender, ethnic/caste and socio economic status of the family. NASA 2014 background information questionnaire included questions related to the above six diversities. However, this assessment also considered three additional comparisons. They are by Districts, by school type (community/institutional) and by school location (rural/urban). These comparisons are made to assess the influence of diversity factors in student achievement and equity status of students based on achievement scores.

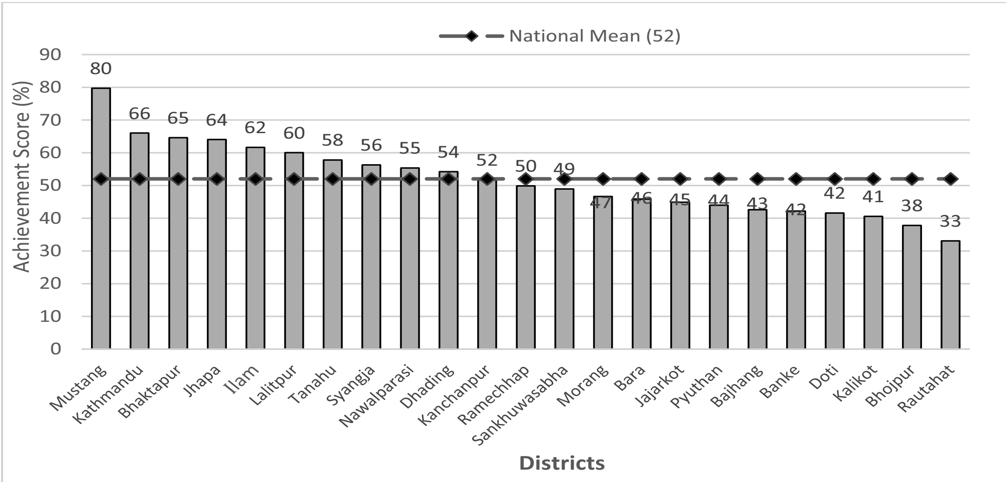
Student Achievement by Districts

Out of 75 Districts, 23 were randomly selected as the sample for this assessment which represent all the three Ecological Zones (Mountain, Hill, Tarai and the Valley) considering the country's geographical features and five Development Regions (Eastern, Central, Western, Mid Western and Far Western) considering administrative division of the country. The achievement of students by District is presented in table 5.4 and figure 5.6. The table below presents District-wise achievement scores which range from 80% (the highest) to 33% (the lowest). The mean represents average percentage of achievement of the particular District. Score is calculated based on the number of students from the Districts that participated in the assessment.

Table 5.4 Average achievement score in sample Districts in Nepali

Districts	N	Mean	SD	Districts	N	Mean	SD
Mustang	29	79.7	15.95	Sankhuwasabha	546	48.9	21.18
Kathmandu	1685	66.0	20.72	Morang	967	46.7	25.44
Bhaktapur	466	64.6	20.11	Bara	1035	45.9	24.78
Jhapa	1083	64.1	21.22	Jajarkot	506	44.9	24.30
Illam	478	61.7	21.82	Pyuthan	733	44.0	22.19
Lalitpur	630	60.0	21.45	Bajhang	692	42.6	21.65
Tanahu	903	57.8	22.30	Banke	816	42.2	24.58
Syangja	761	56.3	22.89	Doti	610	41.6	21.78
Nawalparasi	1019	55.4	23.38	Kalikot	426	40.6	24.10
Dhading	814	54.2	19.98	Bhojpur	490	37.8	22.19
Kanchanpur	614	51.7	25.83	Rautahat	948	33.1	23.47
Ramechhap	538	49.9	22.05	**Total**	**16789**	**24.61**	**51.5**

Figure 5.6 Comparison of average achievement of sample Districts in Nepali



Above data shows that students' achievement is very poor in Rautahat (33%) from Central Development Region, Bhojpur (38%) from Eastern Development Region, Kalikot (41%) from Mid-western Development Region and Doti (42%) from Far-Western Development Region. The best performing District is Mustang (80%) from Western Development Region. Similarly, the second and third best performing Districts are Kathmandu (66%) and Bhaktapur (65%) respectively – both are from the Kathmandu Valley. The difference is very wide (47%) between low performer (Rautahat) to high performer (Mustang). Though Kathmandu and Bhaktapur are still ahead of other Districts after Mustang, they have not been able to retain their position as in the previous tests falling down to 66% and 65% from 81% and 80% respectively. Similarly, Lalitpur was the third outperforming District in the earlier test, which also has lowered down its position to 60% in this test from 78% in the previous one. In the previous test, three top performing Districts were only from the Valley, but now in this test top performing District comes from outside the Valley – which is Mustang from the Western Mountain.

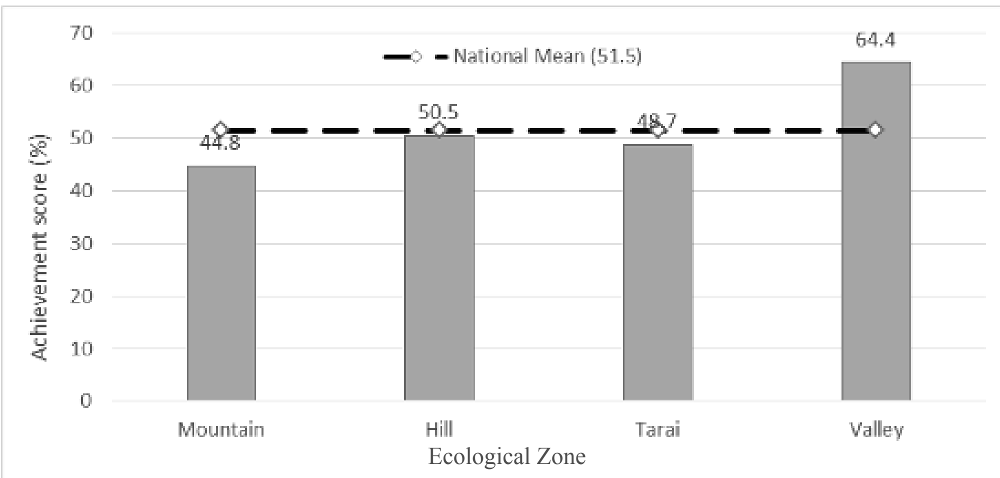
The difference in achievement scores between low and high performing Districts is statistically significant ($p < 0.001$). District explains 15% of the variation in achievement ($\eta^2 = 0.155$). Effect size is high ($f = 0.43$), indicating that the difference between the lowest performing and highest performing District is significant.

The dataset indicates that there is a wide difference in achievement in Nepali among the Districts. Such a wide variation in achievement clearly implies that there is wide inequality among the Districts regarding the issue of providing equal opportunity for children to learn and in ensuring they reach the pre-set goals equally to all. Though the results are bound to 23 sample Districts, there may be other lower performing Districts, which are not included in the sample.

Student Achievement by Ecological Zone

The Mountain, Hill and Tarai are three Ecological Zones of Nepal, though the Kathmandu Valley is taken as a separate geographical area because of being the most densely populated area in the country with more opportunities than other areas. From not only the point of view of population density, but the factors like mixed ethnicities, favourable weather condition, economic activities, urbanization as well as advanced development facilities also make the Kathmandu Valley a unique fourth geographical area in the analysis. The variation in the Ecological Zones in NASA 2015 is presented in figure 5.7.

Figure 5.7 Comparison of achievement across various Ecological Zones in Nepali



The above data shows that, on average, students from the Kathmandu Valley (64%) outperformed the students from the rest Ecological Zones whereas those from the Mountain area performed the lowest (45%). It is noteworthy that some of the lowest performing Districts, such as Kalikot (41%), Bhajang (43%) and Sankhuwasabha (49%) are from Mountain. The difference between the high performing Valley and low performing Mountain is 19 percent.

The achievement across the Regions differs significantly ($p < 0.001$) as the Tukey's *post hoc* test tells us that all the Zones deviate from each other in a statistically significant manner at $p < 0.05$ level. The effect size $f = 0.25$ shows difference between the highest and lowest performing Ecological Zones. Ecological Zone explains 5% of the variation in the data.

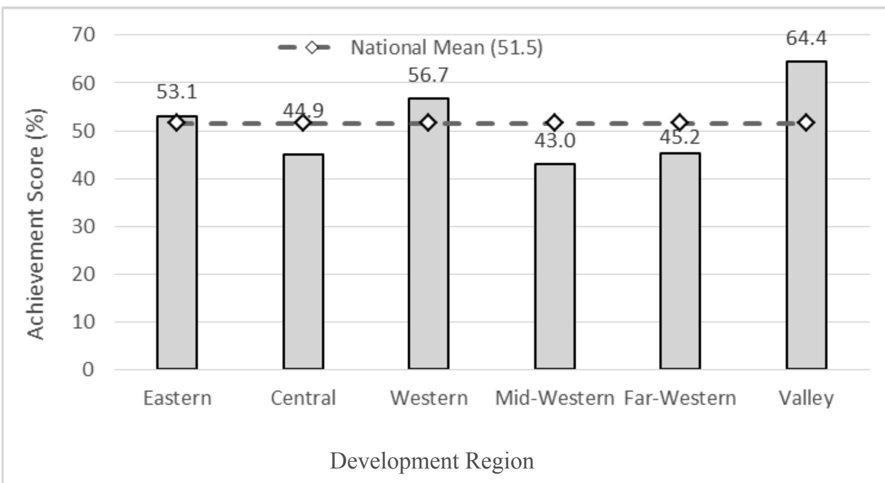
Like in the earlier assessment of 2012, though the Valley still outperforms in this test, it has not maintained its achievement level as in the earlier test (80 percent). In the earlier test, Tarai recorded the lowest performance (57 percent) whereas in this test the Mountain is the lowest achieving Region. In this test, Tarai has raised its position above than Mountain, but it still lags behind the Hill. In comparison to the previous test, all the Ecological belts have achieved lower in this test (see, ERO, 2015b).

Dataset indicates that there is a wide difference in the students' performance among the four Ecological Zones. Students from the Kathmandu Valley outperform the other students. The achievement is the lowest in Mountain Region.

Student Achievement by Development Region

Students' achievement varies according to the Development Regions, which are disaggregated into Eastern, Central, Western, Mid-Western, and Far-Western Regions. Additionally, the Kathmandu Valley has been considered as the additional Region though administratively it falls under the Central Development Region. The mean achievement for each Development Region is presented in figure 5.8.

Figure 5.8 Comparison of the achievement among various Development Regions in Nepali



The above figure shows that the highest achievement is found in the Kathmandu Valley (64%), which is 7 percent higher than the Region having nearest achievement level in the Western Region (57%), and 11 percent higher than Eastern Region (53%). The achievement is the lowest in the Mid-western Region (43%). The Central and Far-Western Regions have more or less equal achievement i.e. 45%. The Valley, Eastern and Western Regions have achievement score higher than national mean. The difference between the Regions is statistically significant ($p < 0.001$), as Tukey's *post hoc* test shows that except for Central and Far Western Regions, other Development Regions differ from each other significantly at least at the significance level of 5%. Development Regions show 9% variation in data, and its effect size is high ($f = 0.32$).

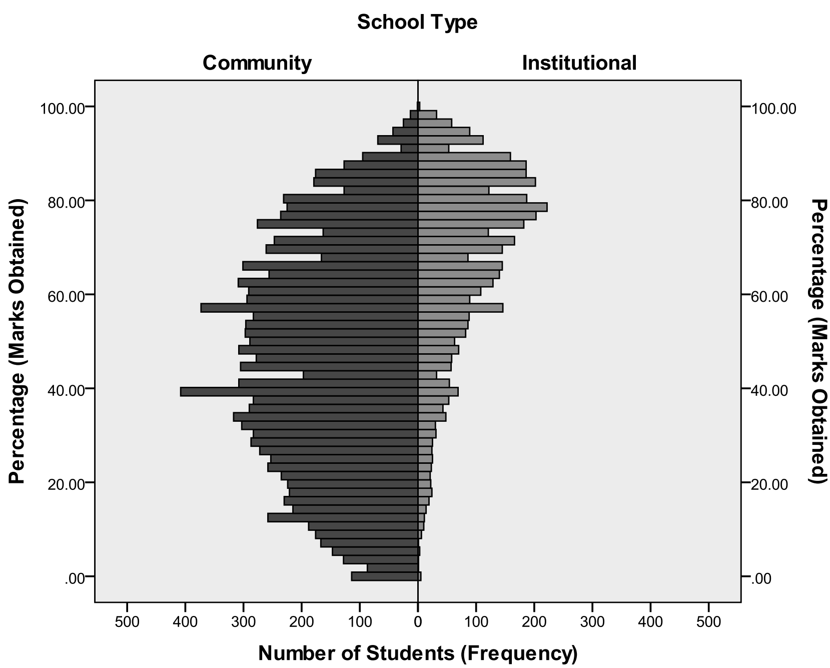
Compared with NASA 2012 data sets, Eastern Region has improved its results from the lowest of all (54%) in earlier test to the third lowest position after Valley and Western Region in this test. Western Region has not maintained its earlier position – the highest achieving as shown in previous test (67%). Similarly, the Central and Far Western Regions have also further lowered down their position being the low achieving Region from the second highest performing Region (59%) in the previous test. Furthermore, all the Regions have achieved less than in earlier test (See, ERO, 2015).

The dataset indicates that the difference between the lowest performing (Mid-Western, 43%) high and the highest performing Regions (the Kathmandu Valley, 66%) is remarkably higher by 21 percent.

Student Achievement by School Type

One way of categorizing schools of Nepal is based on ownership and management. On this basis, there are two types of schools: community and institutional. Community schools, generally funded by the government and managed by the community, are generally known as public schools whereas institutional schools, funded privately, are generally known as private schools. Overall distribution of results in community and institutional schools is presented in figure 5.9.

Figure 5.9 Distribution of the students' mean score in community and institutional schools



The left hand side distribution shows the community school students and the institutional school students on the right hand side. There are some students in the community schools getting equally high marks as in the institutional schools. Figure 5.9 explains that the students in community schools vary from low performers to the high performers, whereas those from institutional schools are concentrated more on relatively high performing group. The distribution of students shows that few institutional school students lie in low performing group while more students from community schools concentrate in this group.

The difference between the achievement of students from community and institutional schools is presented in table 5.5.

Table 5.5 Student achievement by type of school in Nepali

Type of school	N	Mean	SD	SE
Community	12418	45.9	23.7	0.2
Institutional	4370	67.4	19.7	0.3
Total	**16788**	**51.5**	**24.6**	**0.2**

The achievement gap between the community and institutional schools is remarkable. The average achievement score for institutional schools is 67.4% whereas for community schools it is only 45.9%, with a difference of 21.5 percent. The difference is statistically significant ($p < 0.001$) and the effect size is high ($f = 0.42$), showing a wide difference between the community and institutional schools. Category of the students at community and institutional schools explains 15% of the variation in student achievement ($\eta^2 = 0.147$). The variation within the community schools is also remarkable, ranging from 20% to 80%, but the variation is relatively smaller in most private schools.

The dataset indicates that, on an average, the students from institutional schools outperformed those from community schools. Because of the low achieving bigger population from community schools, the entire system seems to be shifted towards low performing.

Student Achievement by Location of Schools

One of the strata of sampling in NASA 2015 was school location. The schools were categorised into rural and urban schools. The achievement of students in rural and urban schools is presented in table 5.6.

Table 5.6 Location of school and student achievement in Nepali

Location of School	N	Mean	SD	SE
Rural	11803	48.9	24.4	0.2
Urban	3742	59.8	23.2	0.4
Total	**15545**	**51.5**	**24.5**	**0.2**

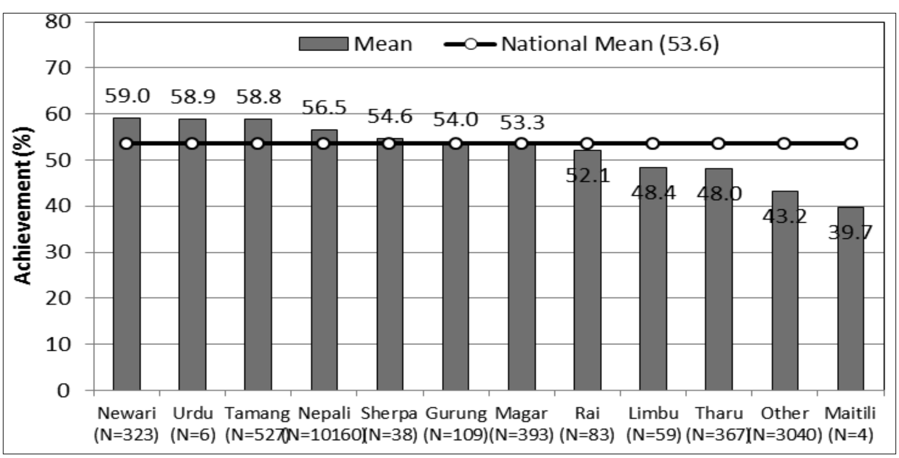
The achievement level of the students from urban schools (59.8%) is remarkably higher than that from rural schools (48.9%). It is notable that institutional schools and schools from the Kathmandu Valley contributed to raise the achievement level of urban schools. Excluding the schools from the Kathmandu Valley, the mean achievement of rural and urban schools is 48% and 55.9% respectively. It indicates that both the community and institutional schools from the Kathmandu Valley contributed to raise the urban schools' achievement. The difference in average score is significant ($p < 0.001$) and the effect size is moderate ($d = 0.45$).

When results of urban and rural schools are compared between NASA 2012 and NASA 2015, the achievement level of urban students was found 16 percent higher in NASA 2012 (75 for urban schools and 59 for rural) whereas it is around 11 percent in 2015. Though the gap between urban and rural was reduced by 5 percent over the years, there is still several interventions to be carried out for rural schools to raise their achievement level.

Language at Home and Student Achievement

Dataset shows that 67.2% of the 3rd Graders, contrary to the 45% in the previous test, speak a language other than Nepali as their first language. These “Other” languages are quite fragmented; the largest groups in the student dataset are Tamang (3.5%), Magar (2.6), Tharu (2.4%), Newari (2.1%) and Gurung (0.7%) and about 20% are not specified in any categories. The language group, like Newari (59%), Tamang (58.8%) and Urdu (58.9, speakers have performed higher than from Nepali speaking group (56.5%). On the other hand, the students from Maithili (39.7%), “Other” (43.2%), Limbu (48.4%), and Tharu (48.0%) speaking background have performed much lower than the national average.

Figure 5.10 Achievement in Nepali by various language groups



Contrary to the previous test, Newari speaking students, achieving relatively low (54% in community schools and 68% in institutional schools) in earlier test, are now at highest level.

Similar to the previous test, Tamang and Magar (68% and 66% in Community Schools, 82% and 84% in Institutional Schools) are performing highest in this test too. After Newari, Urdu speaking students have been able to raise their position from the earlier third (63.6% in community and 81.9% in institutional schools) highest performing language group to the second highest performing one in this test. The difference, which was 4.4% in Community and 3.7% in Institutional Schools between Nepali and Non-Nepali speaking students in the earlier test, have expanded to 8.8% in this test.

For the convenience of statistical analysis, the language groups are broadly categorised into two groups: (1) Nepali speakers and (2) non-Nepali speakers. The results are presented in table 5.7.

Table 5.7 Home language and Student achievement in Nepali

Language Group	N	Mean	SD	SE
Nepali	10160	56.5	23.10	0.23
Non-Nepali	4949	47.6	24.33	0.35

As shown in table 5.11, the difference in achievement between the Nepali and Non-Nepali home language groups is 9 percent. The students who speak Nepali as their mother tongue perform better than others. As in the earlier test, in this test too, the achievement level of students is found dependent on the language spoken in their home – the mother tongue of students. In the context of Nepal, the mother tongue also reflects, to a large extent, the ethnic background and hence the level of difference can be taken as a possible source of inequality in learning.

The difference between the language groups is statistically significant ($p < 0.001$), but the effect size is medium ($f = 0.18$) because the minority groups are small and the division into smaller language groups explains about 3% of the variation in the results ($\eta^2 = 0.031$).

Caste/ Ethnicity and Student Achievement

There are caste/ethnicity based variations in the participation in and achievement of education in Nepal. Historically, the Brahmans and Chhetris have a higher level of educational attainment, but Dalits, for example, had their less participation in or were deprived of education. Hence, the government has made several efforts to make the education possible and accessible for all children regardless of their caste and ethnicity. The recent National Population Census 2011 shows that the enrolment of Hill Dalits has increased remarkably at the lower level of schooling but their number at the secondary and higher education is still very small (CBS, 2012). However, still there are differences in the achievement of various castes and ethnic groups, as presented in table 5.8.

Table 5.8 Student achievement by ethnic/caste background

Caste	N	Mean	SD	SE	Minimum	Maximum
Brahmin	1980	60.6	24.92	0.56	0	100
Chhettri	3328	56.2	23.11	0.40	0	98
Janjati	5517	54.3	23.02	0.31	0	100
Dalit	2160	47.4	23.14	0.50	0	98
Madhesi	1232	45.3	24.60	0.70	0	98
Minorities	84	49.1	23.85	2.60	0	89
Others	435	51.3	25.25	1.21	0	96

The above data shows that the Madhesi students are performing the lowest (45%) in Nepali language and the Bhrahmins have the highest performance (60.6%). When the results between NASA 2012 and NASA 2015 tests are compared, Dalit students have raised their position from the lowest achieving in the earlier year (59%) to higher than Madhesi students in this test, but they still lag behind the others. The Janjati students have also improved their earlier third higher achieving position to the second highest achieving. In both tests, Brahmans and Chhetris have outperformed the rest groups (For details see ERO, 2015b:235). The overall difference between the groups is statistically significant ($p < 0.001$) and the effect size is medium ($f = 0.19$). Caste/ ethnic background explains 3.5% variation in achievement (($\eta^2 = 0.035$). Apart from this, the mean of the various caste/ethnic groups differ from each other statistically at significant at ($p < 0.05$) except for Minorities with all others.

A positive sign from equity point of view is that the Dalit students have performed remarkably better than the national mean (48%) in the Western Mountain (76%) and Western Hill (55%) as well as Western Tarai (55) (see, table 5.9). However, the results are much lower than the average in other regions and very poor in Far western Hill (36%) and Far-western and Eastern Tarai (37%). The number of students in certain strata is small, and hence it is noteworthy not to derive strong implications from the results. However, in the Kathmandu Valley Dalits performed remarkably high (64%).

Table 5.9 Dalit students' achievement in different Ecological and Development Regions

Eco-Zone	Eastern	Central	Western	Mid-Western	Far-Western	Valley
Mountain	45.7		76.2	38.9	37.7	63.6
Hill	41.2	50.0	55.4	46.1	36.0	
Tarai	50.1	37.3	54.9	40.3	45.3	
Valley						

Earlier dataset showed that Dalits have achieved better than national mean (59%) in the Eastern (64%) and Central Mountain (61%), Western Hill (60%) and in the entire Central Region (above 60%) whereas it was lower in the rest part as in the recent test specially in Eastern Hill (50%), Far Western Tarai (53.5%) and Mountain (48.5%).

Figure 5.11 Madhesi students' achievement in various content areas in Nepali

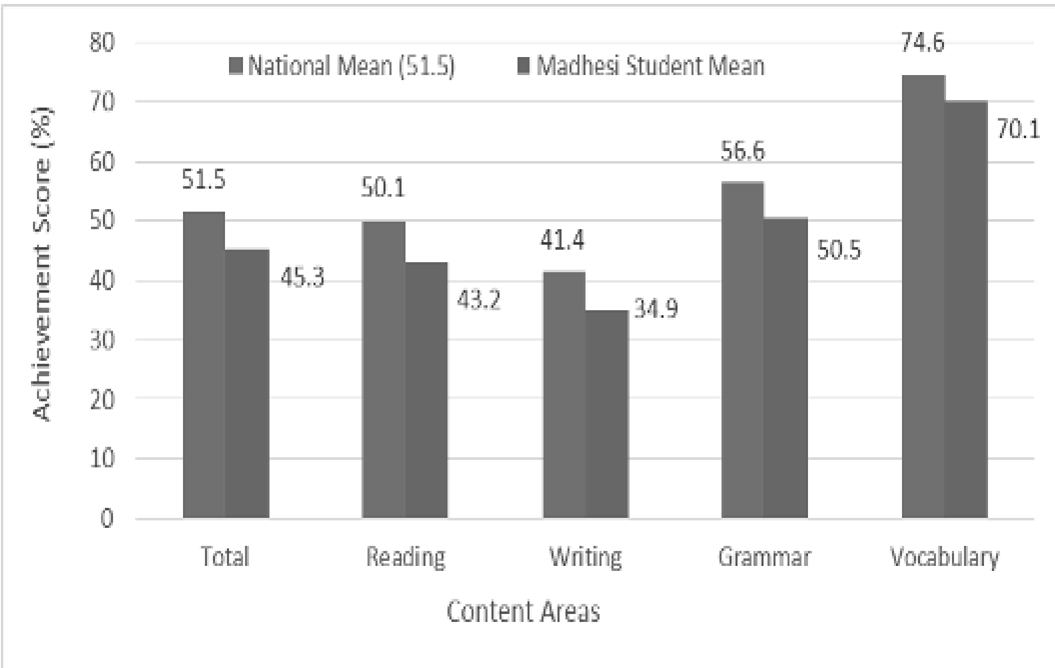
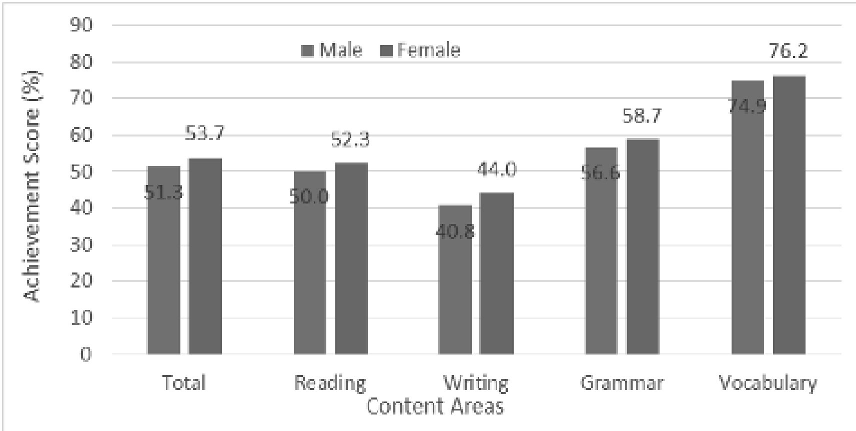


Figure 5.12 shows that the Madhesi students have performed lower than the national mean in all content areas. The difference is highest in Reading, which is about 7 percent. A notable percentage of Madhesi students have not achieved the required level of performance in Reading, Writing, Grammar and Vocabulary.

Achievement by Gender

Basic achievement results of boys and girls are presented in figure 4.13. This figure further compares the achievements of girls and boys in various content areas. The figure shows that remarkable variations are not found between boys and girls in various content areas in Nepali.

Figure 5.12 Comparison of achievement of boys and girls in different content areas of Nepali

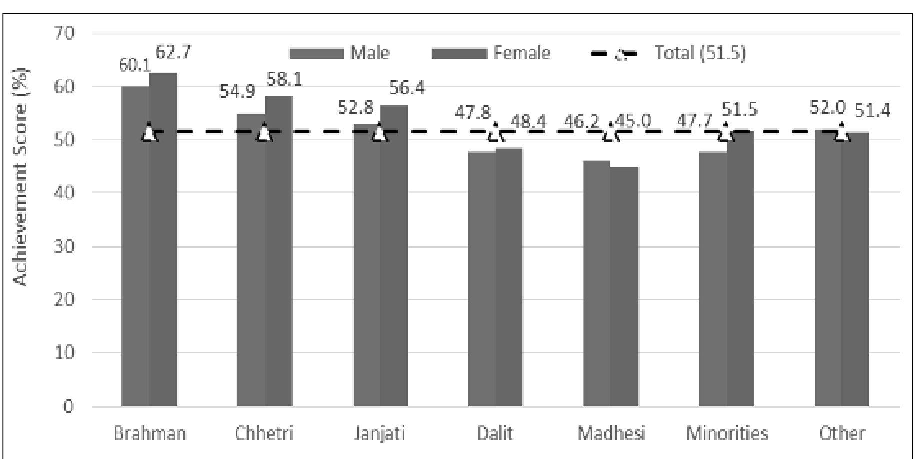


There is statistically significant difference between boys and girls' achievement in all the content areas ($p > 0.001$). However, the effect size and percent of variance explains negligible difference.

In comparison to the earlier test, the achievement level has improved only in vocabulary whereas it has significantly lowered down in other content areas for both the gender. The highest difference noted in reading and writing in the earlier test is persisting in this test too.

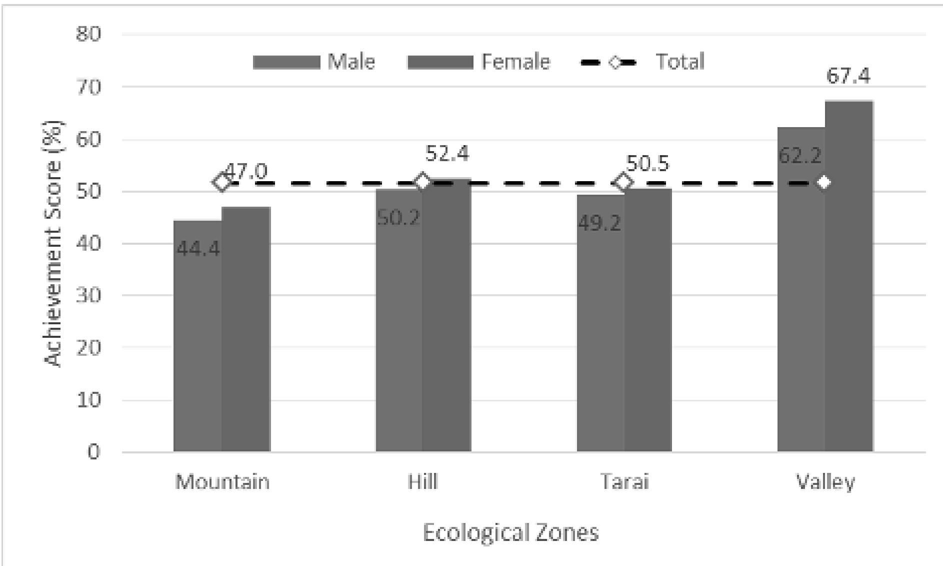
The difference in achievement between boys and girls is the highest among Janjati and Minorities (about 4 percent) where girls have outperformed boys. Madhesi is the only caste/ethnic group where boys performed better than girls though the difference is less (1 percent). Tukey's post hoc test shows that the differences are statistically significant at $p < 0.01$ level. Figure 5.14 presents the performance of boys and girls among different castes/ethnicities.

Figure 5.13 Comparison of achievement between boys and girls in various castes/ethnicities in Nepali



The achievement of girls and boys differs significantly across the Ecological belts. Particularly, in Mountain and Valley girls have outperformed boys. In case of Tarai, there is only 1 percent point difference between boys' and girls' achievement. Similarly, there is no statistically significant difference between boys and girls in Mountain and Tarai whereas it is statistically significant in Hill and the Kathmandu Valley ($p < 0.001$) (see, figure 5.15).

Figure 5.14 Comparison of achievement by gender in various Ecological Zones in Nepali

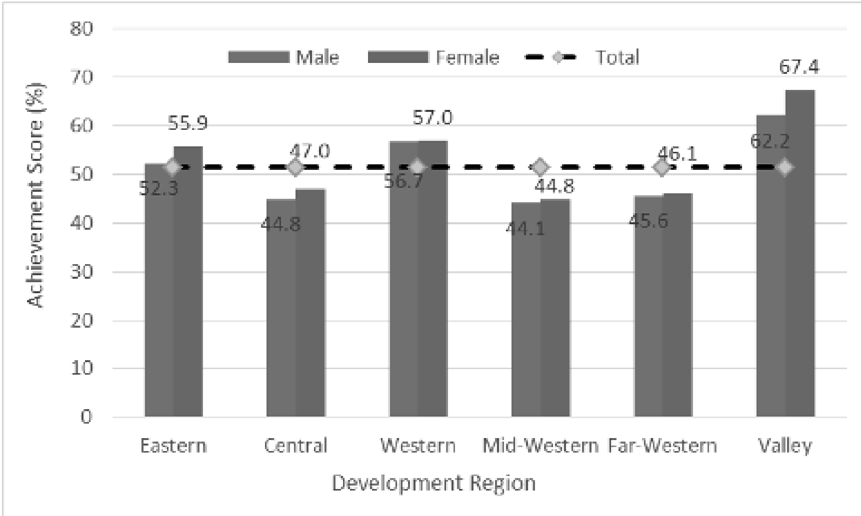


Comparing the achievements of various Ecological Zones, the difference between boys and girls varies and the effect size is also small (male $f = 0.04$ and female $f = 0.06$).

Gender, Development Region and achievement

Mixed results are found among the Development Regions in terms of gender. The difference between boys and girls is found somehow wider in the Kathmandu Valley (5 percent) and Eastern (4 percent).

Figure 5.15 Comparison of achievement of boys and girls in various Development



In general, dataset indicates that boys and girls perform somehow equally. However, there are small variations across different communities and regions. Girls perform better than boys to some extent in all the Development Regions.

Though remarkable difference is not noted between girls and boys in all Development Regions, girls have outperformed boys in all regions. In comparison to the previous test, the achievement gap between boys and girls has widened by more than 3 percent in this test particularly in the Eastern, Central and Valley Regions whereas it has remained constant in the rest parts of the country.

5.4 Selected Explanatory Factors and Achievement

Some of the diversity factors have already been discussed in section 5.3. For example, geographical factors such as Districts, Ecological Zone, and Development Region; school-related factors such as school type and school location; some individual factors related to the students, such as home language, caste/ethnicity and gender. This section deals with socio-economic status (SES) of the students' families, paid work after school, students' attitude towards the Nepali language as a school subject, age of the student, and support provided for studies as the main family and individual related factors. As school and teacher-related factors, assigning homework and providing feedback by the teacher, and selected activities in the school are considered.

Socio Economic Status (SES) of Parents and Student Achievement

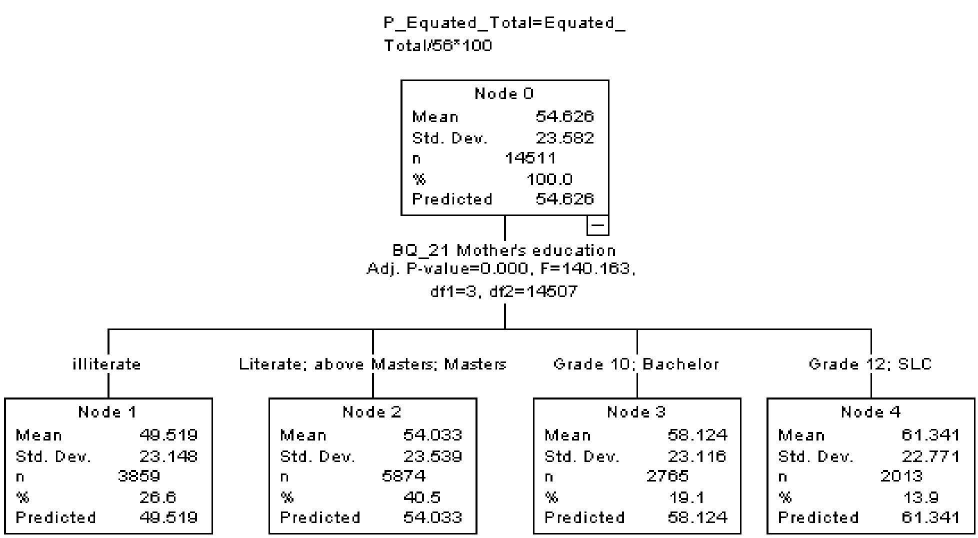
The variables indicating the socio economic status were categorized into parents' education, parents' occupation, home possessions, home accessories, and the type of school attended. Finally, SES is explained based on seven indicators related to economic, educational, and occupational background of the parents. In this section, parents' education is further elaborated and the literacy status of the parents is analysed. Several SES related variables were analysed using a data mining tool of SPSS – the Decision Tree Analysis (DTA). This method is used to find the cut-offs of the predicting variable, i.e., in case of mother's education it classifies the factors into several groups, as it differs statistically in the most significant way from each other in relation to student achievement.

Parents' education

In this assessment, questions related to parents' educational status were included in the background questionnaire. Responses are divided into seven categories as: 1) Illiterate, 2) Literate, 3) Grade 10, 4) SLC, 5) Grade 12 i.e., Certificate level, 6) Bachelor level, 7) Master level and above.

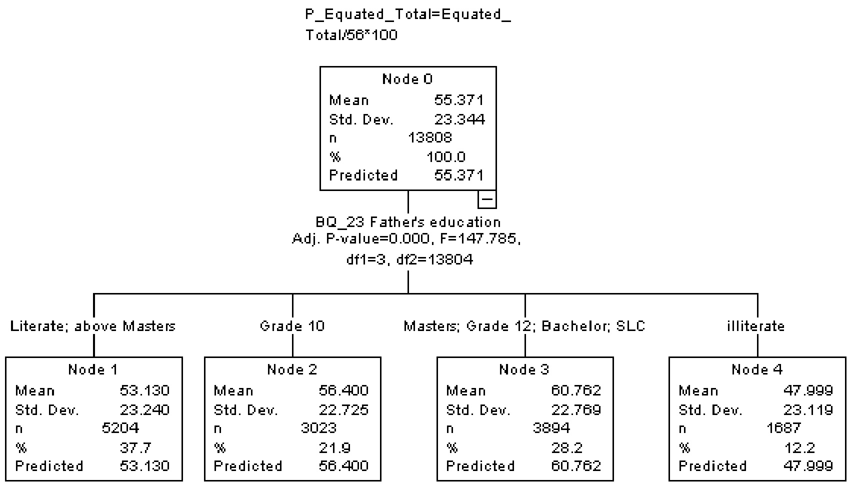
DTA classifies mothers' education into four categories with statistically significant differences in students' achievement levels. Figure 5.17 illustrates these categories with their values as: Illiterate mothers (49%), Literate mothers (54%), SLC and Grade 12 (61%) and mothers' having Grade 10 and Bachelor Degree (58%). In each group, the number of mothers is high enough to make a credible prediction. The difference between each group is statistically significant ($p < 0.001$).

Figure 5.16 DTA of mother's education and students' achievement in Nepali



Similarly, as in figure 5.18, DTA divides father's education into four categories. These categories are Illiterate (47%), Literate (53%), Grade 10 (56%) and SLC, Grade 12, Bachelor, and Master (61%).

Figure 5.17 DTA of father's education and students' achievement in Nepali



From the figure, it can be concluded that children of illiterate father and mother achieved the lowest and the achievement of students increased as fathers' and mothers' education increases up to Bachelor's/Master's degree. The positive relation of parent's education with student achievement is illustrated in figure 5.19. Mother's education explains about 2.9% (η^2=0.029) and father's education 3.1% (η^2=0.031) of the variation. DTA shows that when both parents are illiterate, students achieved very low. On the other hand, they perform better when mother's and father's education is SLC.

Dataset indicates that the educational level of parents affects the children's achievement level. Students' achievement is more encouraging if parents or mother is at least SLC passed, or if father has the qualification of Bachelor's level or higher.

Figure 5.18 Relation between parents' education and student achievement in Nepali

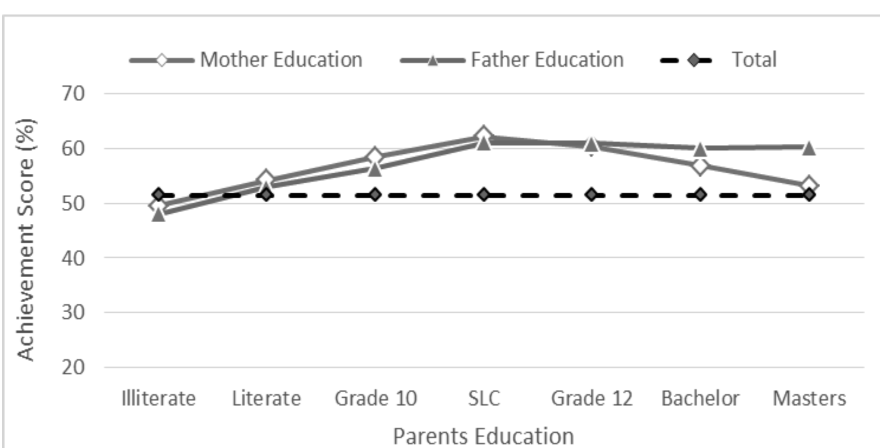
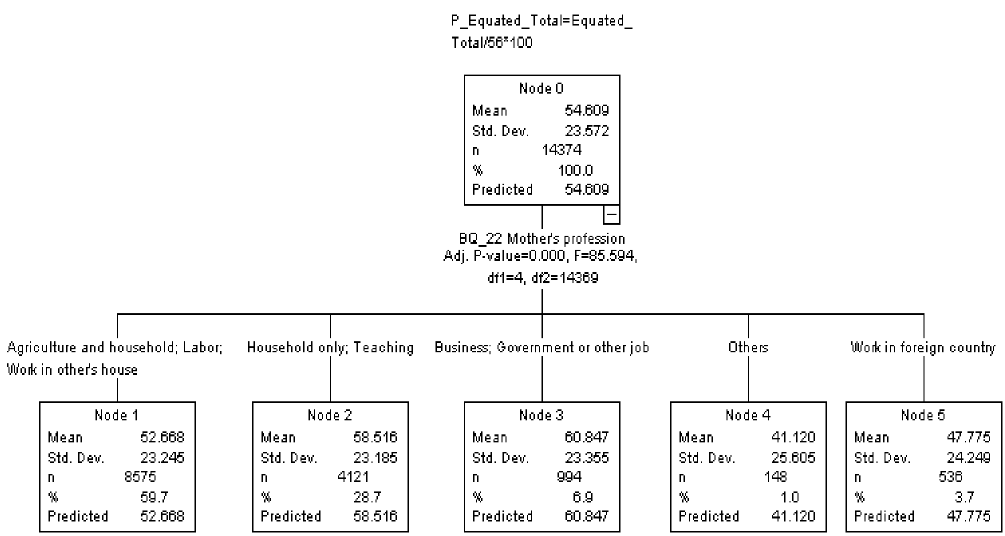


Figure 5.19 shows that father's and mother's level of education has a significant impact on their children's achievement. Students' performance is found below the average national mean if their parents are illiterate. Thus, the fact that the higher the parents' education the better the students' achievement is consistent with the earlier results (see, ERO, 2015b).

Parents' occupation

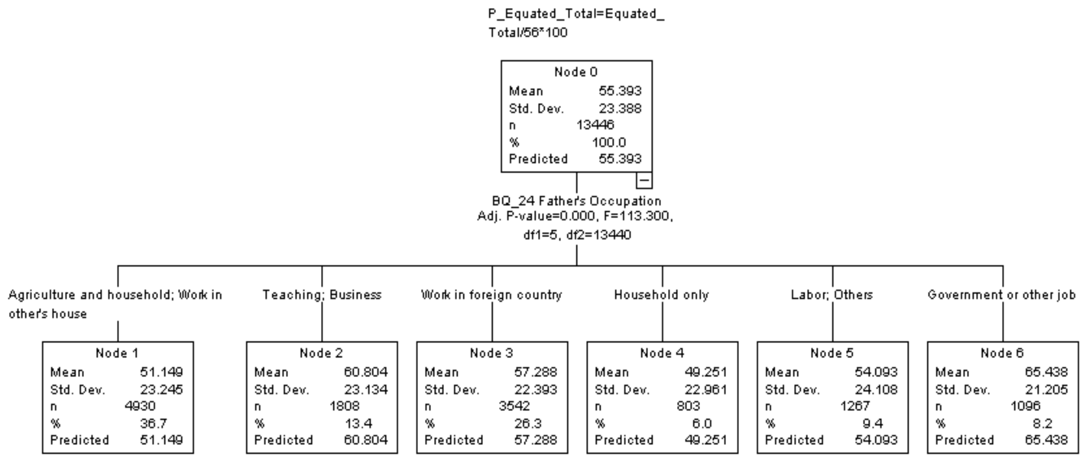
The occupation of parents was categorized into nine groups as: 1) work abroad 2) agriculture and household, 3) household only, 4) teaching, 5) service (government, non-government and private sector job), 6) business, 7) labour, 8) work at others' home, 9) Others. While comparing the mean using ANOVA, student achievement is the lowest when mother's occupational background falls under 'others' (daily wage, driving, not specific) (41%). The result is presented in figure 5.20 as classified by DTA.

Figure 5.19 DTA of mother's occupation and students' achievement in Nepali



In a similar manner, as with the parents' education, DTA was used to find the statistically most deviating groups related to student achievement in terms of parents' occupation. This analysis found four significant nodes. The lowest achievement is in the group where mother comes from non-specific job (others), and agriculture (52%). Significantly, higher achievement is seen when mother is having business and government job (60%) and teaching (58%). However, the result indicates that the mothers who are employed in some kind of job or business are seen beneficial for the achievement of the students. The difference is found statistically significant at ($p < 0.001$) for each group.

Figure 5.20 DTA of father's occupation and students' achievement in Nepali



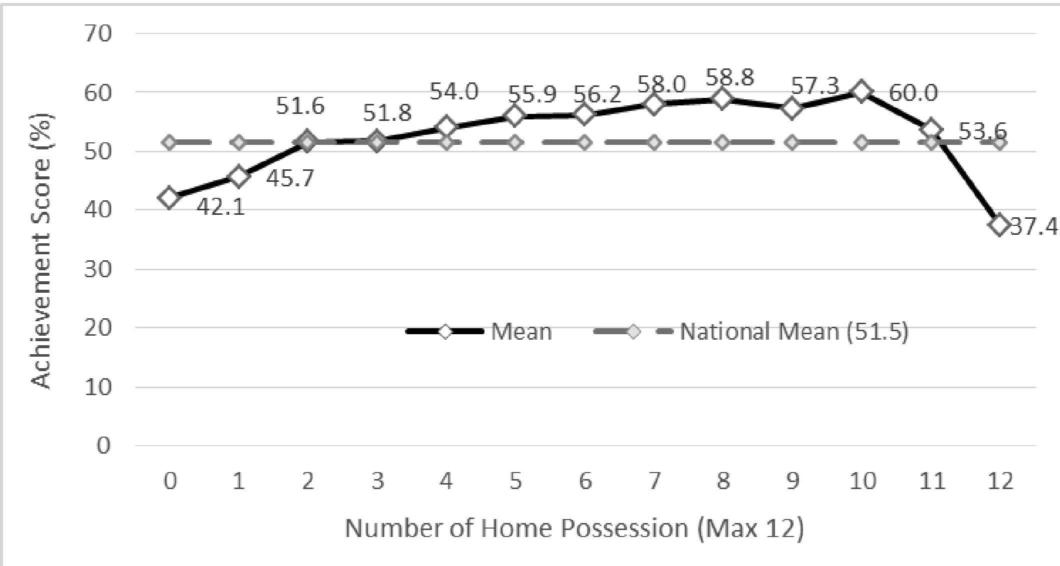
In case of father's occupation, the main division is whether the father works in agriculture or not. If the father is in a teaching profession, the students perform better (52%) than the children of fathers who are working in agriculture. These seven groups are created by DTA, whereby the difference between each group is found statistically significant at $p < 0.001$. The result is consistent with the previous test, as similar trend is noticed in the relation between parental profession and student achievement – the better the profession of parents, the higher the learning achievement.

Home possessions and accessories

This factor takes into account whether the facilities and resources available at home have some effects on the achievement. There were two kinds of home possessions defined in the background information questionnaire distributed to the students. One is related to the facilities that are supportive for the students' study at home. For example, whether they have a table for study, a separate room for them, a peaceful place for study, a computer for homework and additional study, software for the computer assisted learning, internet facilities, literary magazines, access to classical literature, poetry books, or artistic things like pictures, dictionary and other books. Other types of home possession include different types of normal home accessories such as the number of mobile phones, televisions, computers and vehicle.

Related to home possessions, there were 12 questions in the background questionnaire for students. Each question was scored 1 if the student had an access to the possession. Adding these items up, the maximum score was 12 indicating that the student had access to all of the possessions, and hence lower score means fewer possessions at home. Figure 5.22 shows positive correlation between the home possessions and achievement level, except in the case of the highest category (i.e., all the home possessions). Pearson Product Moment correlation coefficient between the achievement level and the factor ($r = 0.11$) is statistically significant ($p < 0.001$) though the value is small.

Figure 5.21 Relation between home possessions and achievement in Nepali



As shown in figure 5.22, the achievement raises with the increase in number of home possessions. However, it is amazing that while students have 11 and 12 possessions the achievement has decreased.

For calculating the SES value regarding home possession, the cut-off for the factors was set on five possessions. If the students possessed 5 items or more as mentioned in background questionnaire, the student was given 1, otherwise 0. The same pattern – the more accessories, the better results – can also be found with home accessories, as shown in figure 5.22. The question in the background questionnaire was set differently compared with home accessories. Regarding the accessories, the question was: "*How many of the following accessories do you have in your family*?" The question was accompanied by the options 0 to 3 (or more). Out of four questions related to accessories, negative correlation was found with two questions related to computer and vehicle. Hence, those two questions were omitted from the analysis of Grade 3 Nepali. The availability of home accessories is dichotomized in two groups. After dichotomizing the items individually by using meaningful cut-offs found with ANOVA and DTA (and maximizing the differences in achievement level), all the three indicators were summed, though the missing one is not included. The maximum score was 2 indicating that the students possessed a set of all the accessories.

Table 5.10 Dichotomizing the indicators of home accessories

Accessory	Cut-off for 1	Cut-off for 0
Mobile phone	2, 3	0 and 1
Television	1–3	0

Figure 5.22 Availability of home accessories and achievement in Nepali

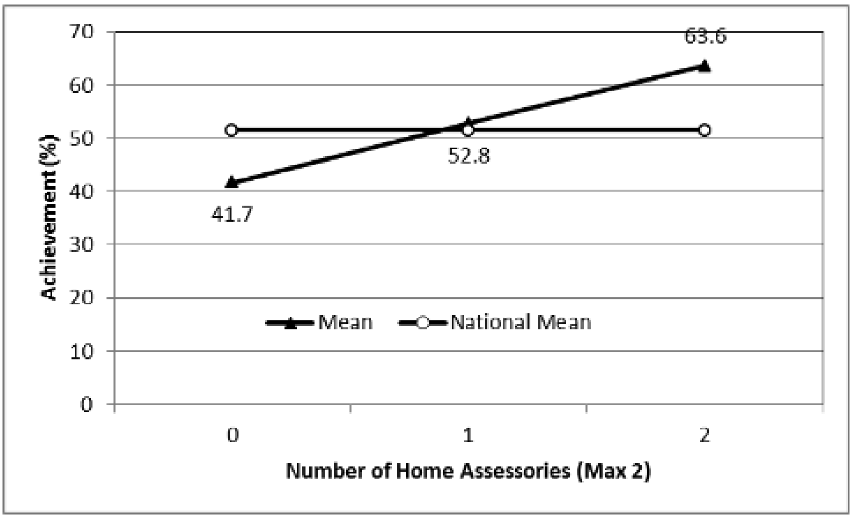


Figure 5.23 clarifies that when the number of home accessories increases, students' achievement also increases ranging from 42% (if none of them are available) to 64% (if any two of them are available). Availability of the stated facilities indicates a higher SES of the family. Therefore, the correlation between home accessories and achievement is found strongly positive (r =.359, $p <$ 0.001) and the effect size is medium.

Data shows that when children have very few home possessions (i.e. 0 to 1 out of the 12), the achievement level is found lower (< 46%), which is statistically significant than the achievement of students of family having more than two home possessions (> 52%). When the family is having seven to ten possessions, the average score is found higher compared with the national average. The same is true for home accessories; that is, when none of the accessory indicators out of two is met, the result is lower than the average (42%) and when there is one met, the result is above the national average (> 52%). If two indicators are met, the result is the highest (64%).

The dataset indicates that either economic or intellectual ability or both at home support children to raise their educational achievement. If the students come from high SES family (having a large number of home possessions and accessories), their performance is found remarkably higher than the performance of other students.

SES and achievement

The value of socio-economic status was calculated based on seven indicators, which were dichotomized first. The variables "mother's education", "father's education", "mother's occupation", "father's occupation", "home possessions", "home accessories", and "type of school" were summed as SES and changed into the percentage of the maximum score (P-SES). Deeper description of the transformations for SES has been given in chapter 2. P-SES represents the percentage of SES of the student's family, whereby 100 means that the student has the highest possible SES where all the seven indicators of SES are positive; and 0 refers to the lowest possible SES where all the seven indicators of SES are negative. The analysis of P-SES by using Univariate GLM (the regression modelling) shows the strong relation between SES and student achievement. Figure 5.24 presents the relationship between SES of the students and their achievement.

Figure 5.23 Relation between the SES and achievement in Nepali

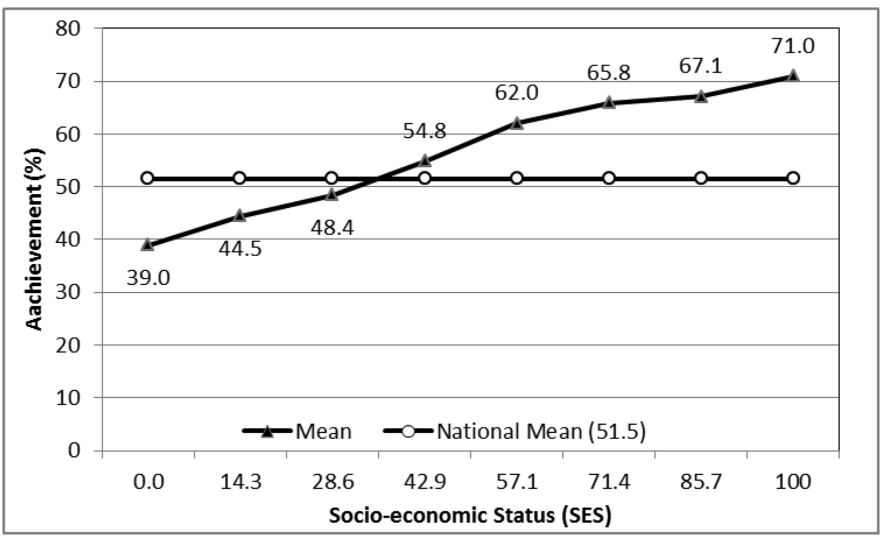
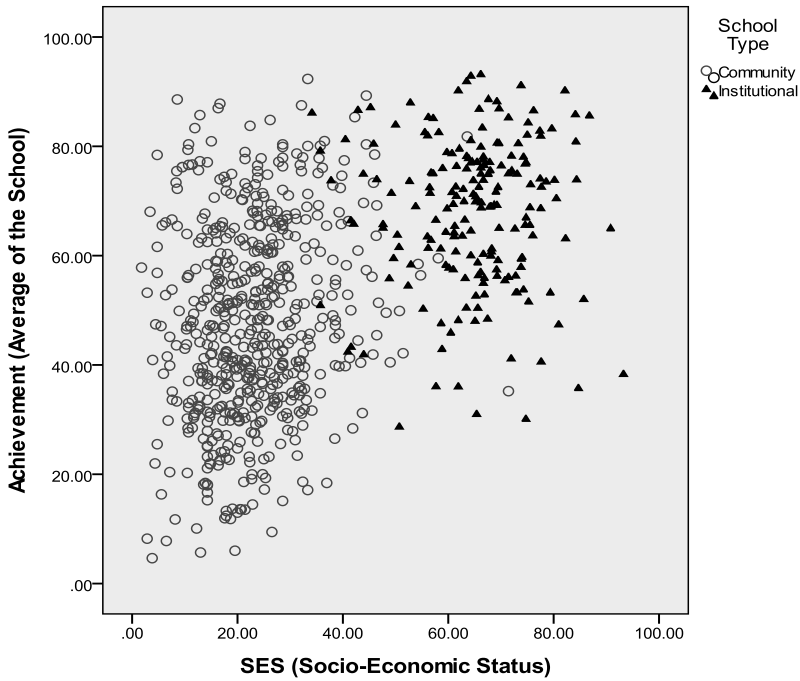


Figure 5.24 shows a positive relationship between SES and achievement as the correlation between the variables is $r = 0.375$, which is statistically significant ($p < 0.001$). The differences between the SES groups are statistically significant ($p < 0.001$) and the effect size is high ($f = 0.41$). SES explains about 14% of the variation in achievement ($\eta^2 = 0.144$).

The dataset indicates that the socio-economic status plays a vital role in student achievement. The difference between the lowest and highest SES groups is remarkable (32 percent). Especially challenging is the situation in the families where the father or both parents are illiterate or both of them work in agriculture.

While plotting the school in terms of mean score and SES in a scatter diagram as in figure 5.25, community and institutional schools (community school in circle and institutional schools in triangle) fall into two distinct groups. Most of the institutional schools, which have high SES value, are among the relatively high performers, whereas the community schools even having low SES value belong to the ranges from the low performing to high performing.

Figure 5.24 Achievement of schools with various levels of SES in Nepali



The dataset indicates that most of the students in the institutional schools have higher SES, and they have achieved relatively higher score in comparison to the students from community schools. Further, the students from the community schools form two kinds of groups: high performing and low performing. Although the variations in achievement among institutional as well as community schools are wide, community schools are concentrated more on the range of 20% to 80% score. Institutional schools are concentrated more on the range of 40% to 90% of score.

Working Beyond School Hour and Achievement

Seven questions were included in the background questionnaire of the students' activities beyond the school time. The values of the variables are divided into five categories: 0 (not at all), 1 (less than 1 hour per day), 2 (1 to 2 hours per day), 3 (2 to 4 hours per day), and 4 (more than 4 hours per day). The GLM Univariate (regression modelling) indicates the cut-off, for example, whether or not the students work in a paid capacity. The relationship is negative when students are engaged in various activities like playing and chatting with friends, playing games, paid job, and entertainment. On the contrary, watching TV 1 to 2 hours a day was found having positive effect on students' achievement, whereas more than two hours spending on watching TV or in activities that do not involve watching TV is seen to be negatively affecting the students' achievement (see figure 5.26).

Figure: 5.25 Relationship between watching TV and achievement in Nepali

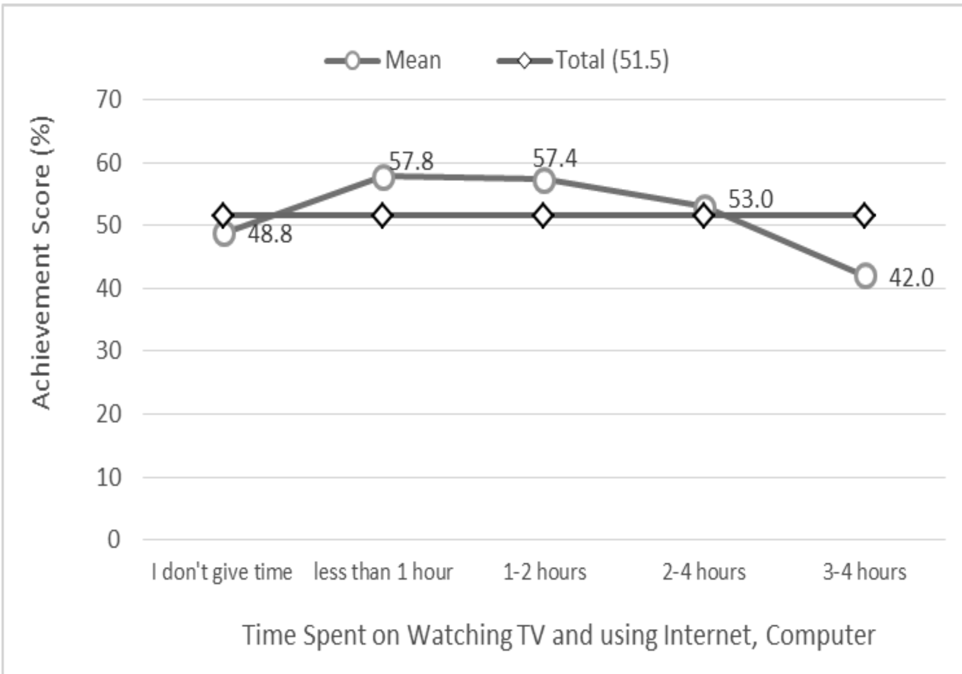
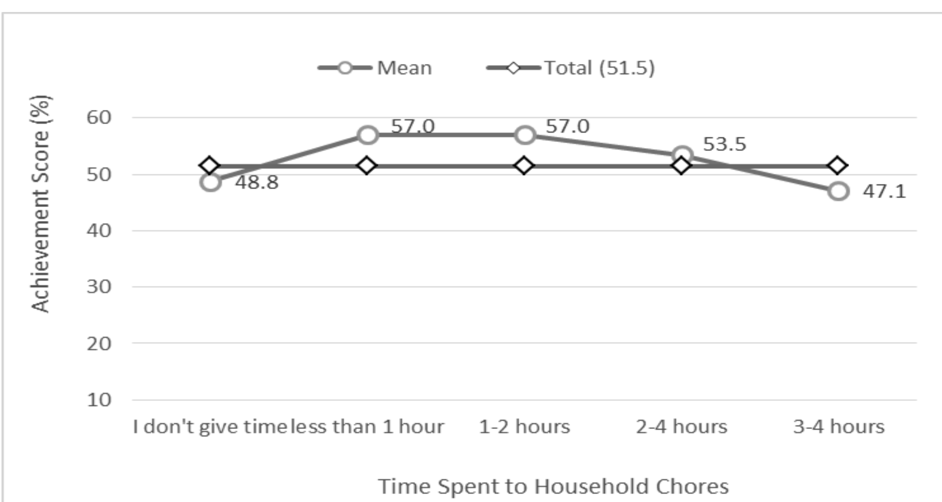
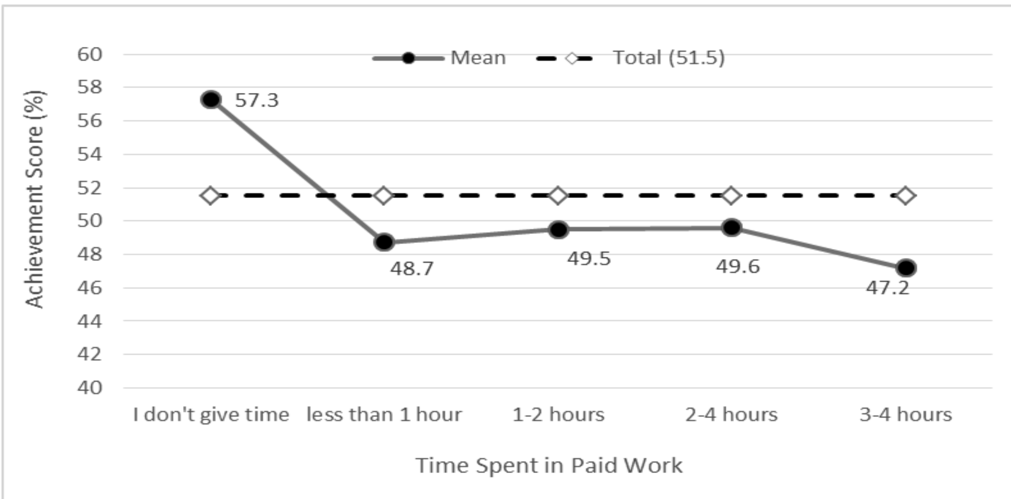


Figure 5.26 Relationship between household chores and achievement in Nepali



The above figure shows that when children do not work at all at home, the results are lower than the national average (49%). If the students are involved in household chores for less than an hour or one to two hours, the results are higher than national average. If students are involved more than three hours in household chores, the results are remarkably lower than the national average. The differences are statistically significant ($p < 0.001$) though the effect size is medium ($f = 0.18$). Out of the total, 21 percent students were reported to have been engaged in household chores for more than one to two hours.

Figure 5.27 Paid work and achievement in Nepali



The DTA shows that when the children have no paid work at all, the results are notably above the national average (57%) (see figure 4.28). If the students are working on a paid job, even if it is for less than an hour, the results are lower than average (48%). The differences are statistically significant ($p < 0.001$) though the effect size is medium (0.16), but a large number of children were not found engaged in paid work. Working before/after school time indicates that the family is poor and the extra earning is needed.

The dataset indicates that working in a paid job or working in unpaid household work for four hours per day noticeably reduces the students' achievement. However, a decent level of household chores that is up to two hours per day does not hamper students' learning.

Student Attitude towards the Subject of Study and Achievement

In the context of Nepali language assessment, the concept of 'attitude' has been explained in terms of what the students think about Nepali subject and its usefulness in their daily life and future. It is assumed that more or less there is some association between the attitude of students and their achievement. The correlation between achievement and attitude towards the subject is widely studied, though the connection is not always clear (see: Metsämuuronen 2012a; 2012b; House & Telese, 2008; Shen & Tam, 2008; Kadijevich, 2006; 2008). In NASA 2011, 2012 and 2013, the same shortened version of Fennema–Sherman Attitude Scales (FSAS) (Fennema & Sherman, 1976), as used in several international comparisons like TIMSS and PISA studies were used. The original scale contains nine dimensions but in these international comparisons only three were used with four items on each dimension and two negative items on each of the first two dimensions (see the detail in chapter 2). The names of the factors can be "Liking Nepali", "Self-Efficacy in Nepali", and "Experiencing Utility in Nepali" (compare naming in, e.g., Kadijevich, 2006; 2008). Factor analysis was used to identify the factors of the responses in FSAS, and the negative items were reversed to make the whole test unidirectional. As in several countries of Asia, the expected factor structure cannot be found in Nepal (for a deconstruction of the test scales, see Metsämuuronen, 2012a; 2012b). Hence, only the total score is used to show the relation between attitude and achievement. The relation between attitudes (as divided into five groups with somehow an equal number of the students, that is, deciles) and the achievement score is shown in figure 5.29.

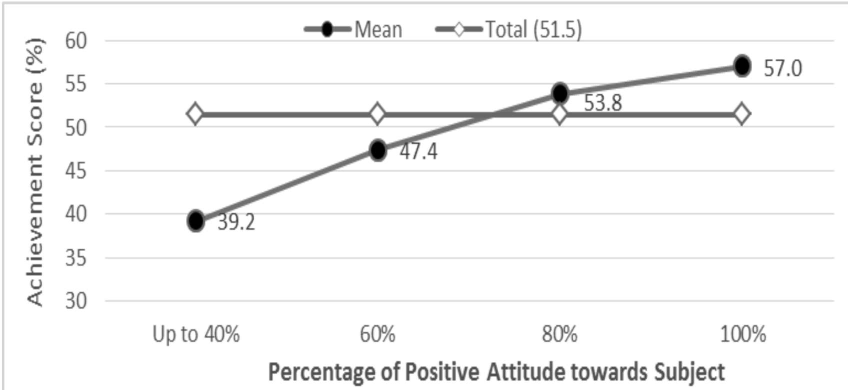
Table 5.11 Students' attitude towards Nepali subject

Students' attitude towards Nepali subject	Respondents in % (valid percentage)			
	Fully agree	Partially agree	Partially disagree	Fully disagree
20a. Nepali subject helps me in the daily life.	84.5	9.9	2.8	2.8
20b. I need to do better in Nepali to learn other subjects.	73.2	18.1	4.3	4.5
20c. I need Nepali for governmental tasks.	82.4	10.1	4.4	3.1
20d. I need Nepali for speaking in all parts of the country.	71.8	17.7	6.5	4.1
20e. I need to do better in Nepali to get a desirable job.	64.5	21.3	8.7	5.5
Average	75.28	15.42	5.34	4.0

Table 5.12 Students' response towards teacher and school related activities in schools

Percentage of positive attitude	N	Mean	SD
Up to 40%	1868	39.2	22.29
60%	1581	47.4	23.73
80%	2559	53.8	22.91
100%	9519	57.0	23.42

Figure 5.28 Relation between attitude and achievement in Nepali



There is a positive correlation between the students' attitude and their achievement ($r = 0.24$) in Nepali. The connection is moderately high ($f = 0.26$); the division of attitude into ten groups explains the achievement level at about 6% (($\eta^2 = 0.06$) variation. The difference between the lowest and highest attitude groups is 18 percent. The connection is clear though it unknown whether the positive attitude is a consequence of high achievement or the other way round. Data indicates that positive attitude towards the Nepali subject is positively correlated with the higher achievement.

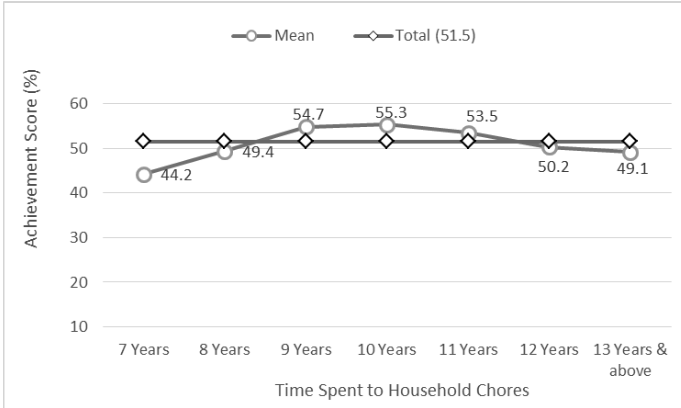
Age of the Students and their Achievement

The age of students attending Grade three varies widely, although the official age for Grade 3 students is 8 in the Nepalese context. Some students have mentioned their age below eight years and some above 13. All the students aged below 8 were encoded as '7 years', and all those above 13 were encoded as '13 years and above'. The descriptive statistics of the mean achievement of each of the age years are presented in table 5.13 and further visualized in figure 5.30.

Table 5.13 Descriptive statistics of the students' achievement in different age groups in Nepali

Age	N	Mean	SD
7 Years	167	44.2	25.78
8 Years	1240	49.4	24.79
9 Years	3545	54.7	23.15
10 Years	5202	55.3	24.10
11 Years	2663	53.5	23.59
12 Years	1434	50.2	24.46
13 Years and above	634	49.1	24.22

Figure 5.29 Relation between the age of students and achievement in Nepali



It is evident that the best achievers are those students who are at the proper age for Grade 3 studies (initially 8 years, while attending assessment they are 9), i.e. 9 years old, scoring 55%. The students have started their study much later, or they have repeated the Grades and therefore studying at higher age. The achievement level is remarkably lower than the average when the students are 13 or higher in age. Correlation between the variables is negative (i.e., - 0.16 at $p <$ 0.054) indicating small effect size ($f = 0.10$). The difference in achievement between proper age and over age students is up to 6 percent. Dataset indicates that the students studying at proper age have obtained the highest achievement score and the achievement lowers down as the age increases.

Support to the Students for Study and their Achievement

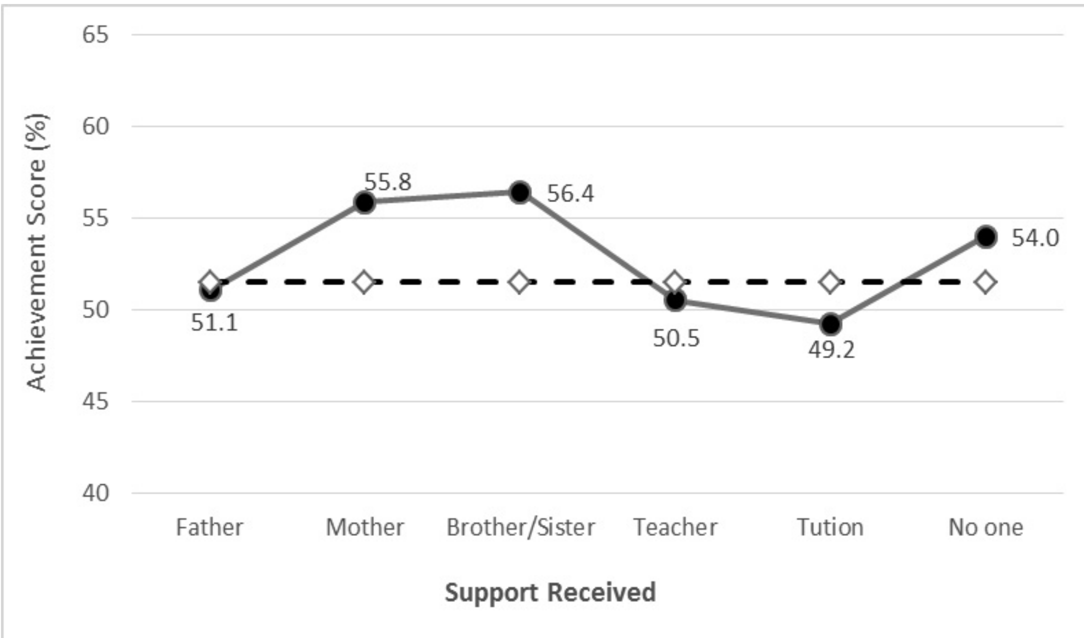
The relation between the support provided to the students for the study and their achievement was analysed based on the information gathered on the question: *"Who supports you when you do not understand what you have read?”*. To the question, there was only one option chosen by the respondent in this study; that is, only one supporter was reported. The descriptive statistics of the supporters are given in table 5.14.

Table 5.14 Descriptive statistics of support to the students and achievement level in Nepali

Support providers	N	Mean	SD
Father	2835	51.1	24.57
Mother	1907	55.8	23.98
Brother/Sister	7091	56.4	22.93
Teacher	811	50.5	25.29
Tuition	1435	49.2	23.18
No one	525	54.0	25.03

Support is seen necessary for the students to gain better than average achievement. Support provided by mothers, brothers and sisters at home raise the achievement level at more or less the same level (around 56%). It is interesting to note that those who received tuition support (49%) and the support from their teacher (50%) obtained lower score against those who did not get any support (54%).

Figure 5.30 Support received and achievement

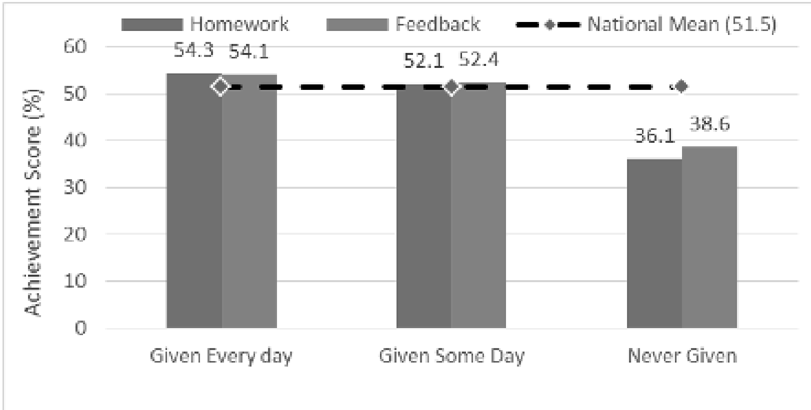


The dataset indicates that the support provided by the mother, brothers and sisters has raised the achievement level more than the support provided by father and teacher.

Homework and Achievement

Homework is considered as one of the ways to enhance learning which can be used as drill, exercise, or an evaluation tool as well. When homework is regularly assigned and feedback is provided to the students, it is likely to boost their achievement level. Statistics related to assigning homework and providing feedback is presented in figure 5.32.

Figure 5.31 Relation between assigning homework and achievement in Nepali



Based on the dataset, it is seen that when the teacher assigns homework and checks/provides feedback regularly to the students, the achievement is higher (54%) than in the case of the students who are not assigned homework (36%). The differences are statistically significant ($p < 0.001$). Those groups without homework are, however, very small, and hence the effect size is small ($f = 0.10$), which explains only 0.9% of the variance in the data (($\eta^2 = 0.009$).

Dataset explains that students' achievement is higher when the teachers assign homework and provide feedback regularly, and it is low when they are taught without assigning homework.

Positive and Negative Activities at School and Student Achievement

The activities of the students and teachers determine the learning environment of the school. Bullying, for example, is one of the hindering incidents in school that affects students' learning. In the background questionnaire for students, several schools related activities were asked - some positive and some negative. Here, bullying is handled as one of the negative indicators and students' impression of school and teacher activities are taken as the examples of positive indicators.

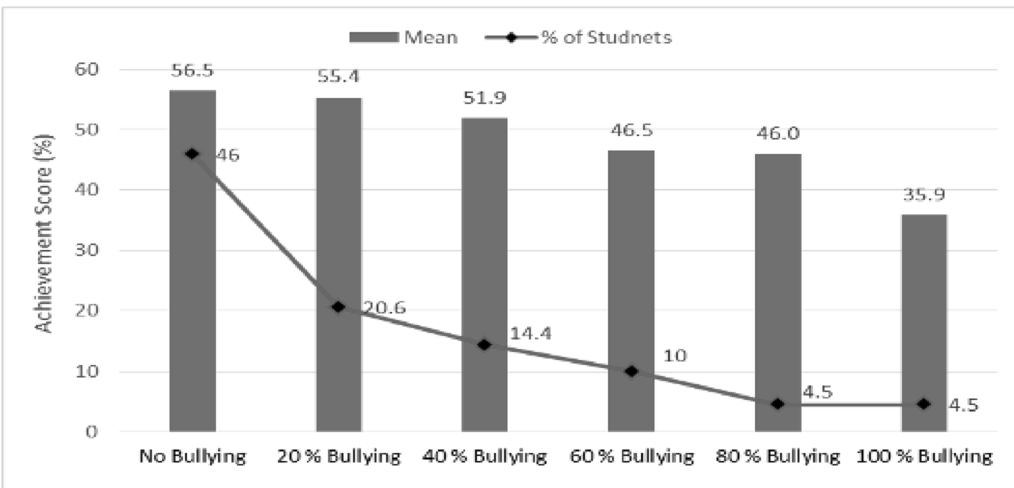
Bullying at school

Bullying is one of the problems in the school that worsens the learning environment. International studies like TIMSS and PISA emphasize to identify such indicators. In NASA 2015 student background questionnaire, five questions indicate the varieties of bullying that tend to happen at schools. All the incidences were stemmed by the phrase "*which of the following activities happened in your school in the last month?*" Students' responses are presented in table 5.15 and further visualized in figure 5.33. 'No (%)' indicates the percentage of the students' response of no such activity happened in the school and 'Yes (%)' indicates the percentage of the students who experienced that particular type of bullying within the last month.

Table 5.15 Bullying and achievement in Nepali

Type of Bullying	No (%)	Yes (%)
Something of mine was stolen	67.9	32.1
I was hit or hurt by other student(s)	70.9	29.1
I was made do the things I didn't want to do by other students	78.3	21.7
I was made fun of or called names	74.4	25.6
Fellow students kept me outside without involving me in activities	81.5	18.5

Figure 5.32 Effect of bullying in student achievement in Nepali



The sum of all five items is taken as an indicator of 100% bullying. Figure 5.33 shows the extent of bullying with the percentage of the students and achievement of the students in each category of bullying. If only one activity of bullying is reported, it is categorized as 20% bullying. If all five activities are reported, it is categorized as 100% bullying. Knowing that 54% of the students did not encounter any bullying during the last month, one can infer that the remaining 46% did encounter at least one type of bullying, which is a remarkable number of students. About 9% of the students had experienced some kind of severe bullying (the sum of 80% and 100% bullying). It is found that learning outcomes are remarkably lower with 14.4% of the students who have encountered more than two different types of bullying (52%). Around 21 percent achievement gap is observed between the students who did not experience bullying and those who encountered extreme bullying of four or five kinds, though there are a few number of students who reported this kind of bullying (n = 2247). The difference is statistically significant (p = 0.001), and the effect size is moderate (f = 0.22). Though extreme cases of severe bullying are rare (n = 702), those students who have faced that kind of extreme bullying have the lowest (36%) achievement score. Overall, bullying is seen quite common incident in schools.

The dataset indicates that a large number of students (46%) have encountered bullying in schools. The phenomenon tends to affect the learning outcomes in almost all the groups of the students who experienced bullying.

Positive activities in school

The activities that can boost the learning achievement of students are termed as positive activities. The students were asked about such positive activities in the school in two sets of questions listed in table 5.16. The table shows the responses of the students in all four categories, which are in the 4-point rating scale, anchored to fully agree and fully disagree.

Table 5.16 Students' responses towards teacher and school related activities in school

Teachers' and students' activities	Respondents in % (valid percentage)			
	Fully agree	Partially agree	Partially disagree	Fully disagree
q27a School: Students get along well with most teachers.	90.1	5.9	1.9	2.1
q27b School: Most teachers are interested in student's well-being.	89.5	6.5	2	2
q27c School: Most of the teachers really listen to what I have to say.	81.4	12.8	2.9	2.9
q27d School: If I need extra help, I will receive it from my teacher.	88.7	6.6	2.2	2.5
q27e School: Most of my teachers treat us fairly.	86.8	7.6	2.5	3.1
q28a School: I like to come and stay in school.	93.4	3.6	1	2
q28b School: Students in my school try to do their best.	84.7	11.1	2	2.2
q28c School: Teacher in the school care about the students.	80.3	5	1.8	2.2
q28d School: Teacher wants the students to do their best.	90.6	4.7	1.8	2.9
Average	87.3	7.1	2.0	2.4

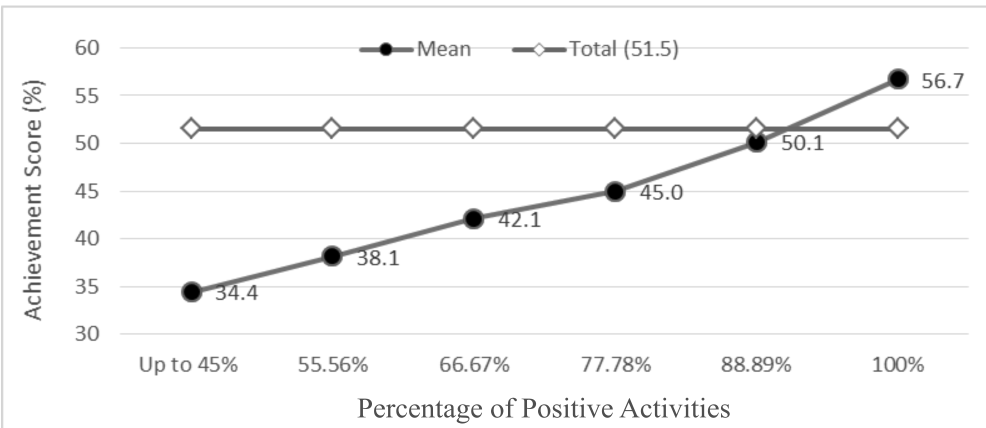
Further analysis was done by recoding the variables which are further grouped into two categories: 1 for agree, and 0 for disagree. Furthermore, the sum of nine indicators is converted into the percentage of maximum score to analyse the level of positive activities and its relation to achievement.

DTA finds four attitude groups in the indicator. These boundaries and descriptive statistics are presented in table 5.17 and illustrated in figure 5.34. The overall result is that feeling the positive actions in the school relates positively with the student achievement. The correlation between the sum of nine positive activities and achievement is found positive ($r = 0.28$, $p < 0.001$).

Table 5.17 Students' response towards teacher and school related activities in schools

Percentage of positive action	N	Mean	SD
up to 45	1242	34.4	21.65
55.56	594	38.1	21.35
66.67	522	42.1	22.71
77.78	698	45.0	22.61
88.89	1410	50.1	22.85
100	11547	56.7	23.35

Figure 5.33 Relation between positive actions in school and achievement in Nepali



The data shows that there is a positive relation between the positive activities of the school and student achievement. The increase in achievement is directly proportional to increase in the intensity of such activities. After dividing the indicator into five groups based on DTA, the differences between the groups are statistically significant ($p < 0.001$), however, the effect size is medium ($f = 0.31$). Learning achievement is much higher than average (57%) only when the students are extremely positive towards school and teachers' behaviour. However, the difference between the most positive group and the most negative group is notable (22%).

Dataset indicates that when students think that the actions of the teachers and the schools are ultimately good, results are higher than the average (57%). At the other extreme of feeling, when such actions are ultimately negative the results are far below the average (34%).

5.5 Summary of Findings

The main findings of the assessment results for Grade 3 in Nepali subject are as follows:

Basic results

- The achievement in Nepali subject is not distributed normally as it is slightly right skewed towards lower performing.

- The average achievement in Nepali subject is 51.5 percent. The achievement is the highest in the content area of V*ocabulary* (75%) and the lowest in R*eading* and *Writing* (41%).
- Students' ability to solve complex problems is quite low. Students are much better in the recalling type of questions, which were geared more towards remembering the things than solving problems.
- The students are performing well in recognizing the correct answer and recalling simple facts from the texts, fundamental thinking, the basic interpretation of paragraph, table, chart and a few steps of logical thinking, and basically in objective items (63%). They are much weaker in producing fluent texts or letters, or preparing synthesis and abstracts from the text, basically in subjective items (48%).

Diversity factors

- There is a wide difference in achievement among the Districts in Nepali subject. The results in Rautahat (33%), Bhojpur (38%), Kalikot (41%), Doti (42%), Banke (42%), Bajhang (43%), Pyuthan (44%) are relatively very low. Out of 23 sample Districts, 12 Districts are below the national average. The gap between low performing District (Rautahat 33%) and high performing District (Mustang 80%) is 47 percent.
- There is a moderate difference in the students' achievement across three Ecological Zones. Students in the Kathmandu Valley (64%) outperform the students from other regions. The achievement is found low in Mountain Region (45%).
- There is a wide gap among Development Regions in students' achievement in the Nepali language. The difference between the lowest performing (Mid-Western Region, 43%) and highest performing Region (Kathmandu Valley, 64%) is remarkable (19 percent).
- On average, the students from institutional schools outperform (67%) the students from community schools (46%). However, some students/schools from community schools perform equally as those from institutional schools.
- The students in the urban schools achieved 11 percent more than those in the rural areas. Most of the high performing institutional schools were from the urban areas. Excluding the Kathmandu Valley, the difference between urban and rural schools has not been that wide.
- Achievement difference is also found among the different language groups. The students from Newari (59%), Tamang (58.8%) and Urdu (58.9, N=6%, N=4) speaking families have performed higher than those from Nepali speaking group (56.5%). On the other hand, the students from Maithili (39.7%, N =4), "Other" (43.2%), Limbu (48.4%), and Tharu (48.0%), achieved lower than national average.
- The performance of Madhesi students is lower than the performance of the other castes/ ethnic groups/communities. Madhesi students' performance was found low in all content areas; moreover, a notable percentage of Madhesi students obtained less score than the national average in reading and writing.
- There is small difference between girls' and boys' performance in the Nepali language. Girls are slightly better performing than boys in all the content areas; and the differences

are found significant. However, there are notable (1.2 percent) differences in achievement between boys and girls in Madhesi students. High difference (4.8) is found in Chhetri students. Differences are also found between boys and girls across the Ecological Zones and Development Regions, but the difference is small in Development Region in comparison to Ecological Zones.

Selected explanatory factors

- Parents' educational level strongly predicts the children's future achievement level in the Nepali language. Especially, achievement level is the lowest when the parents are illiterate.
- Either the high economic status of the family or intellectual capacity or both have positive effect on the children's achievement score. If the father or mother or both are involved in agriculture, labour or unspecified occupation, the students' achievement in Nepali subject is significantly lower than the achievement of the students whose parents are from other occupational groups like business and other job.
- When children have very few home possessions (i.e., 0 to 1 out of the 12), the achievement level is significantly lower (< 46%) than the achievement of those having more than two home possessions (> 52%). With a family having ten possessions, the average score is very high (> 60%) compared with the national average.
- Socio-economic status (SES) plays a strong role in the educational processes in Nepali. The difference between the lowest and highest SES groups is remarkable (32 percent). Especially challenging is the situation in the families where father or both parents are illiterate or both parents work in the agriculture related occupation.
- Household work for three to four hours per day outside the school has remarkably reduced student achievement. However, a decent amount of household work up to two hours per day does not affect learning negatively for the students in Nepali. Rather, it has given positive effect. On the other, the paid work for any time gives negative effect on achievent.
- Positive attitude towards the subject correlates positively with achievement.
- Achievement lowers down as the age increases. The highest performance is among those students who are studying at their appropriate age group.
- Support to the learning provided by brother or sister and mother gives more positive effects on achievement compared to the support provided by father and teacher.
- If the teacher gives homework regularly and provides feedback to the students, achievement is highest; otherwise, it lowers down.
- A large number of students (46%) have encountered at least one kind of bullying in schools within a month. This phenomenon has affected the learning outcomes in almost all groups of students who experienced bullying.
- When students think that the actions of teachers and schools are ultimately good, the results are better than average (57%). At the other extreme, if feeling is ultimately negative, the results are far below the average (34%).

Chapter 6: Analysis of Assessment Results for Grade 5 in Nepali

6.1 Introduction

Nepali language is taught as one of the compulsory subjects in Grade 5. Out of total 39 credit hours per week, eight credit hours are assigned for Nepali subject. This assessment is based on the learning outcomes/ curricular objectives as set in the approved curriculum for Grade 5 in Nepali subject. Overall objective of Grade 5 curriculum in Nepali is to enhance language competencies in four language skills – listening, speaking, reading and writing (CDC, 2062).

As Nepali is one of the core subjects in the school curricula, proficiency of 5th Graders in Nepali language has been assessed frequently in the national assessment since 1995. Before 2011, 5th Graders' achievements in Nepali were assessed three times in 1995, 1997 and 2001 (see, BPEP, 1995; EDSC, 1997 & 2001). The previous large-scale National Assessment of Student Achievement (NASA) 2012 conducted by Education Review Office also focused on assessing 5th Graders' achievement in Nepali. Now, in this assessment too, testin the 5th Graders' achievement in Nepali is again in focus. The frequent assessment in this subject is motivated by the fact that reading and writing skills in Nepali are inevitable to live an informed life in modern society, which is surrounded with written information, and proficiency in these skills among citizens is the main requirement to be able to acquire and communicate information in written forms. As Nepali is not only the official language in Nepal, but also the lingua franca throughout the country. So, citizens should have an adequate command in this language to be able to communicate in Nepali to tackle the everyday language functions. Most importantly, basic proficiency in Nepali in early Grades is essential to succeed in further study as textbooks, reading materials and medium of instructions are dominantly using Nepali language.

Though proficiency in Nepali for 5th Graders have been assessed since 1995, the results of earlier assessments (1995, 1997 and 2001) were not fully comparable with the later results of NASA 2012 and NASA 2015 because of the lack of proper linking procedures between the tests. Previous assessments did not have defined linking procedure as those tests were constructed with utilising traditional testing theory. So, what follows are described comparing only with the results of NASA 2012.

This chapter presents the analysis of assessment results of Grade 5 in Nepali language. The second section of this chapter describes the overall distribution of achievement scores; presents achievements in various content areas, cognitive domains, types of items; and compares the results of NASA 2015 and NASA 2012. The third section presents comparison of results with PIRLS Grade four results. Results with various diversity factors, including District, Ecological Zone, Development region, school type, school location, language at home, and caste/ethnicity are described in section four of this chapter. Further, it deals with the selected explanatory factors about students' achievement in Grade 5 Nepali. The factors that explain achievement in the subject are: parents' education and occupation, home possessions and accessories, SES, working beyond school hour, age of the student, support for study, availability of textbook, homework given and checked/feedback provided, activities at school (both negative and positive). The final section summarizes the findings related to the achievement in Nepali subject.

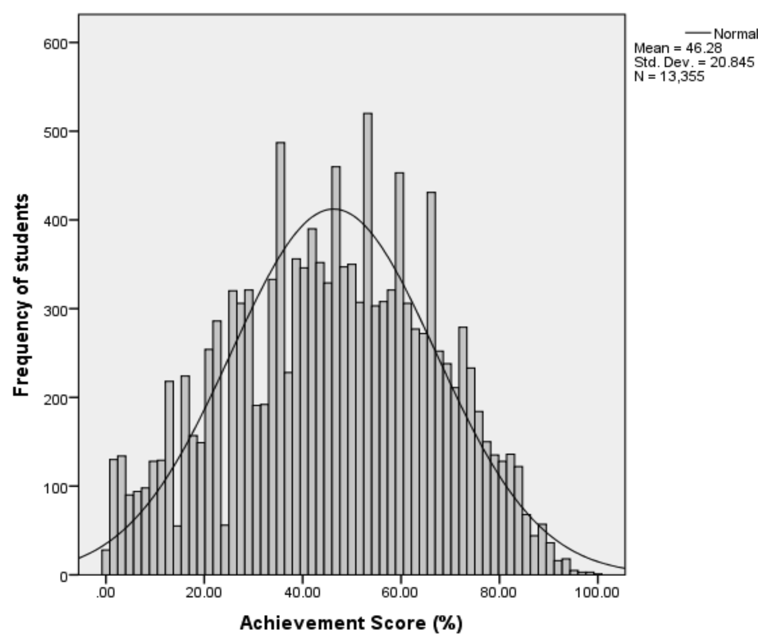
6.2 Basic Achievement Results in Nepali Grade 5

The basic results of assessment of Grade 5 students in Nepali includes distribution of overall results, results by content areas, levels of cognitive domains and item types, and the comparisons of results with previous assessments.

Overall Distribution of Achievement Scores

Achievement score is usually distributed normally in a large sample in the students' achievement study. As shown in figure 6.1, the achievement scores in Nepali subject are distributed normally.

Figure 6.1. Distribution of overall achievement scores in Nepali



The sample size for Nepali subject (13,355 students) was big enough to form a normal distribution. In the final score of Nepali subject, the majority of the students fall into the medium performing section of the population. Negligible number of students falls in the group achieving more than 90% whereas significant number of students falls in the group achieving below 10%. The dataset indicates that Grade 5 population in Nepali subject is distributed normally. As illustrated in figure 6.2 given below, the schools (N = 511) can be divided into two categories – namely, the high performing and low performing.

Figure 6.2 Distribution of schools' mean achievement scores in Nepali

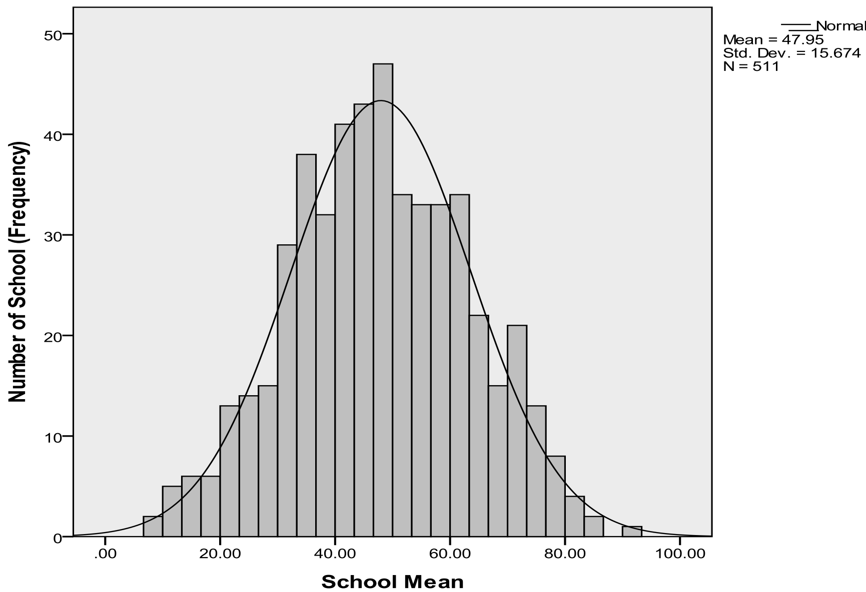


Figure 6.2 shows that, based on students' mean achievement score of each sample school, there are two categories of schools, namely high performing and low performing. The maximum average score of schools on the left-hand side is about 30%, whereas it is about 70% for the schools at the right-hand side. It indicates that there is a remarkable difference in mean achievement score between the high and low performing schools. There are large numbers of schools in low performing category than high performing category indicating a wider disparity in learning opportunity among students within the system, which are distinctly divided into two groups of population – low performing majority and a high performing group that is very small.

As per the curriculum, the Nepali language has five content areas, namely, 1) listening, 2) speaking 3) reading with vocabulary, 4) writing with spelling and punctuation and 5) functional grammar. This type of large-scale learning assessment generally covers only the reading and writing. Therefore, test items were developed based on the four content areas, namely, 1) R*eading* 2) W*riting* 3) G*rammar* and 4) V*ocabulary*. The weightage in the four content areas of assessment were proportionally equal to the weightage allocated in curriculum as far as possible.

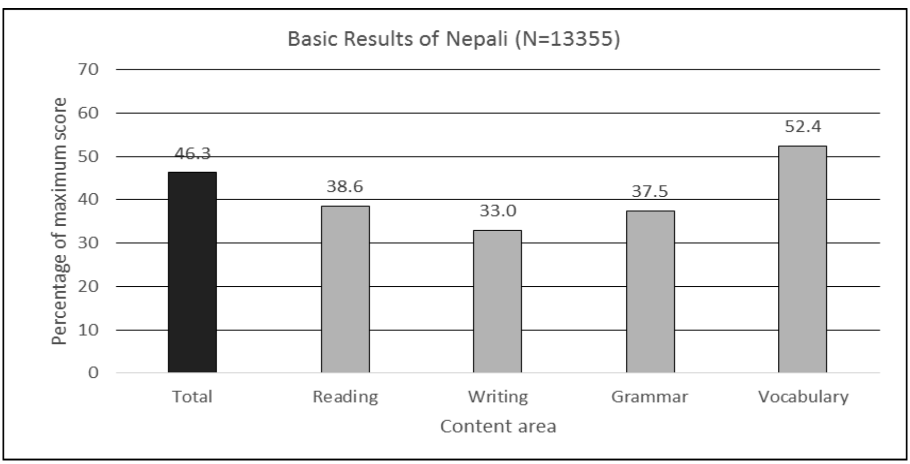
For comparability, achievements in all the content areas were converted into the percentage of the full marks of content areas. Table 6.1 shows the students' achievement in Nepali in total and the achievement level in each of the four content areas. Figure 6.3 further illustrates the comparison of achievement scores in various content areas.

Table 6.1 Achievement in various content areas in Nepali

Content Areas	Mean	SD	Min	Max
Reading	38.6	21.60	0	100
Writing	33.0	25.85	0	100
Grammar	37.5	26.19	0	100
Vocabulary	52.4	27.25	0	100
Nepali Total	**46.3**	**20.84**	**0**	**100**

Table 6.1 shows the variations of achievements in various content areas. The achievement ranges from 33 percent in W*riting* to 52 percent in *Vocabulary* with 19 percent variation. Comparing the maximum and minimum scores, the situation is the same in all content areas, as the maximum score is 100 and minimum is 0 in each of the content areas. Figure 6.3 compares the variation of achievement in various content areas.

Figure 6.3 Comparison of achievement scores in various content areas in Nepali



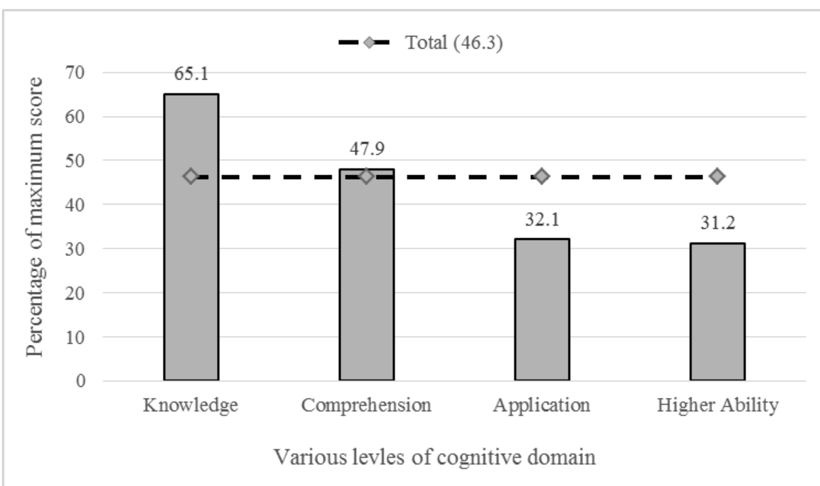
The overall national average achievement score for Nepali is 46% in Grade 5. Of the various content areas, students are weakest in W*riting* (33%), and but their performance is better than the national average in V*ocabulary*. Dataset indicates that the learning achievement in Nepali is the highest in *Vocabulary* and the lowest in *Writing*.

To compare the results by content areas with previous test, a different trend of achievement gaps in learning is noticed. In NASA 2012, V*ocabulary* was the highest performed area (70.1%) in comparison to R*eading* (56.1%), W*riting* (58.1%) and G*rammar* (63.8%). In NASA 2015 test also, *Vocabulary* stands as the highest one. However, in NASA 2012 the *Reading* was the lowest one whereas in NASA 2015 the lowest achieving area is W*riting*. In the previous test, the level of difference between R*eading* and W*riting* was only 2 percent, which is widened in this test by 5.6 percent. The overall performance level has lowered down to 46.9% from 59.7% in the previous test. Content wise wider gap is found between G*rammar* (25.8%) and W*riting* (24.5%) than *Reading* (16.9%) and *Vocabulary* (17%). From this scenario, it is evident that the level of learning over the years is falling off instead of improving, which is a matter of serious thought for policy and planning.

Achievement in Various Levels of Cognitive Domain

The test items in Nepali comprise of various levels of cognitive domain according to Bloom's taxonomy (Bloom *et al.,* 1956; Metfesser, Michael & Kirsner, 1969). The categories were *Knowledge, Comprehension, Application* and *Higher ability* (reasoning/problem solving/ creativity). The achievement of students in each level is shown in figure 6.4.

Figure 6.4 Comparison of achievement score in various levels of cognitive domain in Nepali



The score on the items related to *Higher* a*bility* (31%) and *Application* (32%) are below the national average, whereas the scores in both the *Comprehension* (48%) and *Knowledge* (65%) level items are above the national average. It is apparent that students performed better on the items of lower level of cognitive domain than the higher level.

The dataset shows that a large number (9%) of students were not able to solve even a single item under *Higher ability*. Aound 19 percent of the students could solve just less than 10% of the problems requiring *Higher cognitive abilities*. The dataset shows that the students' achievement in A*pplication* and *Higher ability* items is quite low (32% and 31%) in comparison to C*omprehension*

(48%) and K*nowledge* (65%); that is, students are much better in the recall type of questions than in analytical questions. Similarly, the trend of comparatively better performance in lower order skills and lower performance in higher order skills was noticed (70.5% in *Knowledge*, 47.9% in *Comprehension*, 61.4% in *Application* and 47.0% in *Higher ability*) in the earlier test too. Similar to the content wise results, the achievement level in this test has fallen off in each domain by more than 5 percent in *Knowledge*, 14.8 percent in *Comprehension*; 29.3 percent in *Application* and 15.8 percent in *Higher ability*. From the datasets of both tests, Nepali students are somehow good at remembering, recalling or reproducing the facts but quite poor in producing and creating new which require additional interventions from curriculum planning to classroom delivery and assessment practices for improving the situation.

Student Achievement by Type of Items

There were two types of items in the test: *Objective* and *Subjective*. *Objective* items covered a wide range of content areas and were very specific to judge because there was only one correct answer or one explicit piece of information required to get the correct answer in a question, whereas *Subjective* items require a longer procedure and higher mental ability to get full marks. Both types of items were based on various levels of cognitive domain, which include *Knowledge, Comprehension, Application,* and *Higher ability*– covering various difficulty levels like easy, medium and difficult. Table 6.2 presents the basic statistics of achievement on each type.

Table 6.2 Achievement by types of items in Nepali

Item Types	2015			
	Mean	SD	Min	Max
Objective	55.0	22.6	0	100
Subjective	33.3	22.6	0	100

Students' mean score in *Subjective* items is much lower (33%) than that in Objective items (55%). Most of the *Objective* items were of knowledge, comprehension and application type whereas *Subjective* items were mostly application and *Higher ability* type. When the current results are compared with previous dataset, it is seen that achievement scor on both type of items has further fallen down in this test (by about 22%) and the difference in achievement level between them has further been widened in the later year.

Dataset clearly explains that the students are performing well in recognizing the correct answer and in recalling simple facts from the texts, comprehending the basic information from paragraph, table, chart, and a few steps of logical thinking. They are weaker in writing free texts or letters, or preparing synthesis and abstracts from a text. In many cases, the students attempted the open-ended tasks like free writing, problem solving and analysis; but the skills were not high enough for obtaining high marks.

Comparison of NASA 2015 Results with NASA 2012

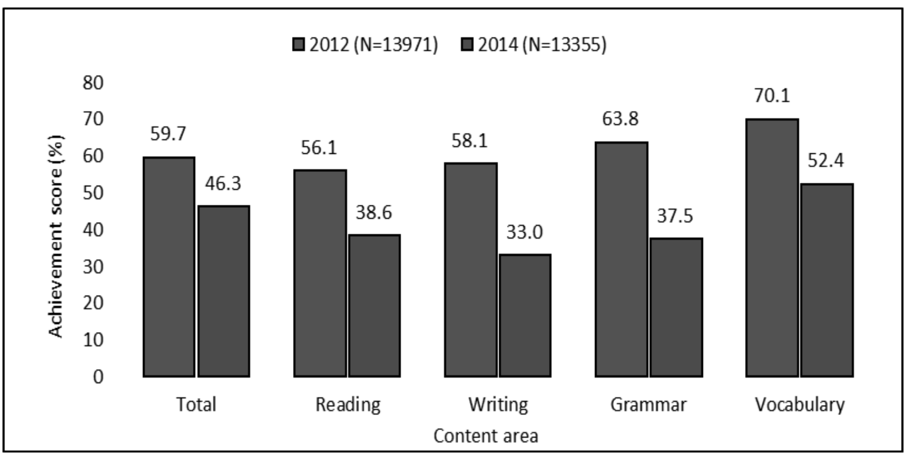
The results of NASA 2015 and NASA 2012 are as presented in table 6.3 for the purpose of comparison. Although the samples were not the same in NASA 2012 and NASA 2015, comparison is done using several linkage items in both the assessments. While comparing the achievement scores of NASA 2015 and NASA 2012, it is found that score in NASA 2015 is much lower than that in NASA 2012.

Table 6.3 Comparison of achievements in 2012 and 2015 in Nepali

Year	N	Mean	SD	SE
2012	13971	59.7	23.10	0.20
2015	13355	46.3	20.84	0.18

The result in Nepali as a whole and the achievement in various content areas of NASA 2015 and NASA 2012 are compared in figure 6.5(a).

Figure 6.5 (a) Comparison of results of NASA 2012 and 2015 in various content areas in Nepali



As shown in figure 6.5 (a), the achievement level in NASA 2015 is remarkably lower than that of NASA 2012. Achievement in all the content areas in NASA 2015 is lower than that of NASA 2012. However, direct comparison is difficult as the sample students and Districts were not the same as previous one. About 20 percent items were used from NASA 2012 test in NASA 2015 assessment to make these two assessments comparable. The dataset indicates that the achievement score of students in Nepali has not improved over the year.

6.3 Comparison of Results with the International Standard of PIRLS

The NASA 2015 was also made comparable with the international PIRLS assessment results. A good number of items (17) of the released PIRLS items were used as linking items. Their known difficulty parameters were fixed in the calibration of the local items. Hence the international average of $\eta^2 = 0$ was fixed in the Nepalese datasets; when a student's ability level in NASA 2015 is zero, it corresponds to the average level of the international students.

Figure 6.5 (b) shows the comparison of the students' achievement with the international standard of PIRLS. In the figure, the x-axis shows the content areas of Nepali and y-axis shows the ability shown by the students. The middle horizontal line indicates the international average. As the ability is below the average (0), the bars are going down whereas when the ability is above the international average, the bars are going upwards.

Figure 6.5 (b) Comparing NASA 2015 results in Nepali with PIRLS reading scale

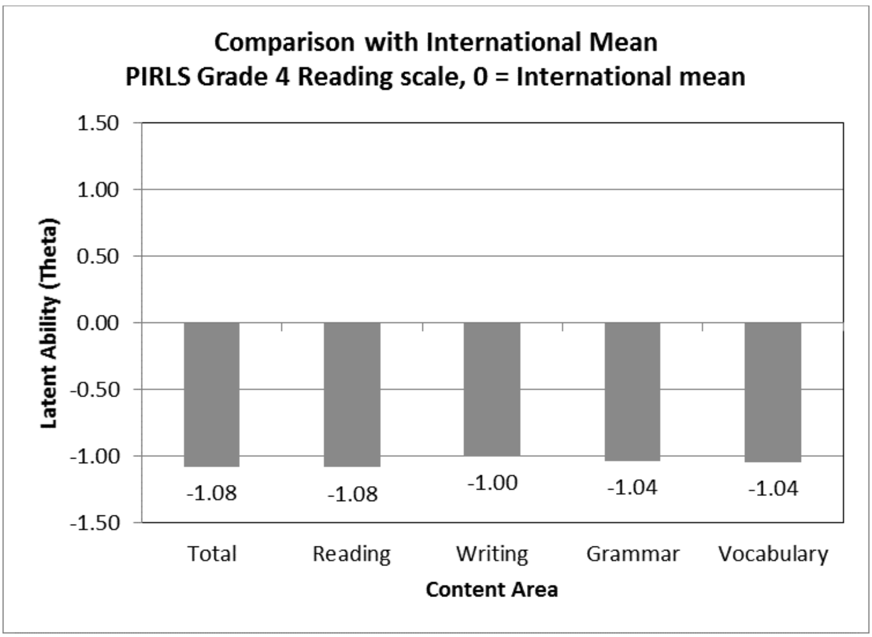


Figure 6.5(b) shows that the average ability shown by Nepali students in Nepali as a whole is below the average line of PIRLS results in reading for Grade 4. This indicates that the students in Nepal are somewhat poor in language showing a lower ability in all the content areas when compared with the international average.

It is good to remember that all the linking items came from the content area of *Reading* and hence there actually is no real equating in the other areas. Especially incomparable are the *Grammar* and *Vocabulary* because, in PIRLS, these areas are not measured at all. However, they are modelled based on proficiency in the reading test.

6.4 Achievement Scores by Diversity Factors

In the context of NASA, six diversity related factors are considered - namely Ecological (Gegraphical), *Regional*, language, gender, ethnicity/caste and economic diversity. NASA 2015 background information questionnaire consisted questions related to the above six diversities. However, this assessment has also considered three additional comparisons. They are by District, by school type (community/institutional) and by school location (rural/urban). These comparisons were made to assess the influence of diverse factors in students' achievement as well as assessing the equity status of students based on achievement scores.

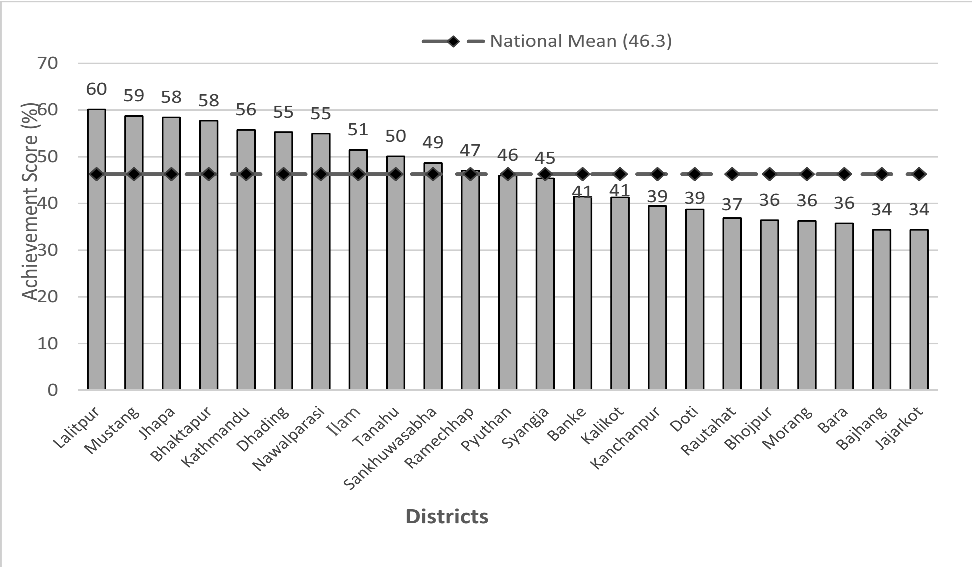
Student Achievement by Districts

Out of 75 Districts, 23 Districts were selected for this assessment – with the representation of all the Ecological Zones (Mountain, Hill, Tarai and the Valley) considering country's geographical features and five Development Regions (Eastern, Central, Western, Mid-Western and Far Western) considering administrative division of the country. The achievement of students by District, which ranged from 34 to 60 mean score, is presented in table 6.4 and in figure 6.6. The table shows the achievement in descending order according to the Districts' mean achievement score. The mean represents average percentage of achievement of the particular District.

Table 6.4 Average achievement score of sample Districts for Grade 5 in Nepali

District	N	Mean	SD	District	N	Mean	SD
Lalitpur	440	60.2	18.37	Syangja	680	45.4	17.99
Mustang	39	58.7	9.64	Banke	639	41.5	20.63
Jhapa	752	58.4	18.38	Kalikot	344	41.3	19.15
Bhaktapur	446	57.7	17.58	Kanchanpur	467	39.5	18.57
Kathmandu	1172	55.7	18.71	Doti	488	38.7	20.51
Dhading	718	55.3	17.52	Rautahat	611	36.9	21.39
Nawalparasi	725	55.0	19.59	Bhojpur	492	36.4	17.42
Illam	444	51.4	20.78	Morang	790	36.2	20.07
Tanahu	802	50.1	20.02	Bara	905	35.7	21.53
Sankhuwasabha	409	48.7	16.51	Bajhang	517	34.4	17.04
Ramechhap	493	47.0	18.79	Jajarkot	391	34.3	18.84
Pyuthan	591	46.0	14.87	**Total**	**13355**	**46.3**	**20.84**

Figure 6.6 Comparison of average achievement of sample Districts for Grade 5 in Nepali



The above figure shows that students' achievement is the lowest in Jajarkot (34%) from Mid-western Development Region and Bajhang (34%) from Far-western Development Region. Likewise, Bara (36%) and Morang (36%) are the lowest performing Districts from Central, and Eastern Development Regions respectively. The best performing five Districts are Lalitpur (60%) from the Kathmandu Valley, Mustang (59%) from Western Region, Jhapa (58%) from Eastern Region, and Bhaktpur (68%) and Kathmandu (56%) from the Kathmandu Valley. The difference between high performing (Lalitpur, 60%) and low performing (Jajarkot, 34%) is very wide (26%). In previous test Bhaktapur (80%), Kathmandu (78%) and Lalitpur (73%) Districts were outperforming. However, in this test Bhaktapur and Kathmandu are lower down to the fourth and fifth positions respectively. Though in this test Lalitpur District is outperforming, compared to previous test it has achieved 13 percent lower. In Nepali for Grade 5, Mustang and Jhapa Districts are performing better in both Grades (three and five) in this year.

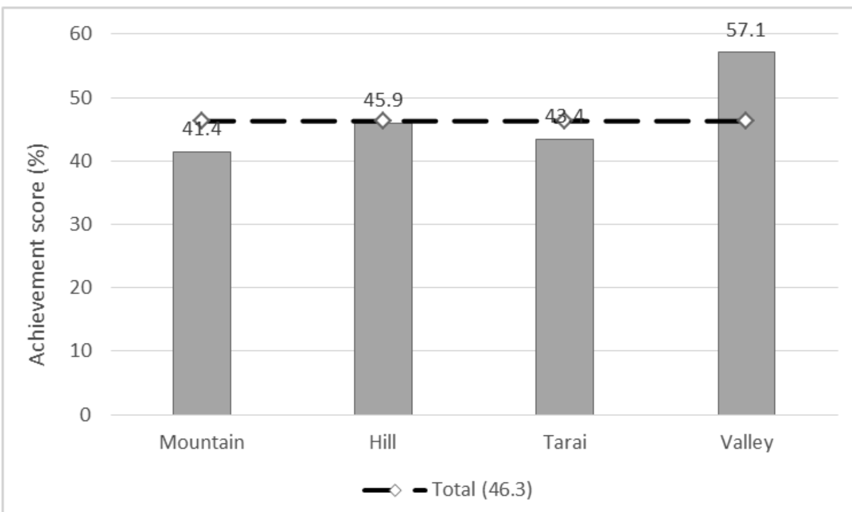
The difference in achievement scores between low and high performing Districts is statistically significant (p < 0.001). District explains 17% of the variation in achievement ((η^2 = 0.170). Effect size is high (f = 0.45), indicating that the difference between the lowest performing and highest performing Districts is significant.

The dataset indicates that there is a wide difference in achievement in Nepali among the 23 Districts. The results in Jajarkot and Bhajhang (34%),Bara, Bhojpur and Morang (36%), Rautahat (37%), Doti and Kanchanpur (39%),Banke and Kalikot (41%) are comparatively low. Such a wide variance in achievement clearly implies that there is wide inequality among the Districts to provide for children an equal opportunity to learn that and ensuring them reach the pre-set goals equally to all. Though the results are bound to 23 sample Districts, there may be other further lower performing Districts, which are not included in the sample.

Student Achievement by Ecological Zone

The Mountain, Hill and Tarai are three Ecological Zones of Nepal, though the Kathmandu Valley can be considered a special geographical area because of being the most densely populated area in the country with more opportunities than other areas. From not only the dense population point of view, but also because of the mixed ethnicities, a favourable climatic condition, economic activities, urbanization as well as advanced development facilities make the Kathmandu Valley as a unique fourth geographical area in the analysis. The variation in the Ecological Zones in NASA 2015 result is presented in figure 6.7.

Figure 6.7 Comparison of achievement of various Ecological Zones in Nepali



The above figure shows that, on average, the students from the Kathmandu Valley (57%) have outperformed the students from the rest Ecological Zones whereas the students from the Mountain area performed the lowest (41%). It is noteworthy that some of the lowest performing Districts, such as Bhajang (34%) and Kalikot (41%) are from Mountain region. Similarly, Districts from the Tarai Region like Bara, Morang, Rautahat and Banke are also low performing.

The achievement across the regions differs significantly ($p < 0.001$) as the Tukey's *post hoc* test tells that except for mountain and Tarai, all the Zones deviate from each other in a statistically significant manner (at $p < 0.001$ level). The effect size ($f = 0.24$) shows great difference between the highest and lowest performing Ecological Zones. Ecological Zone explains 5% of the variation in the data.

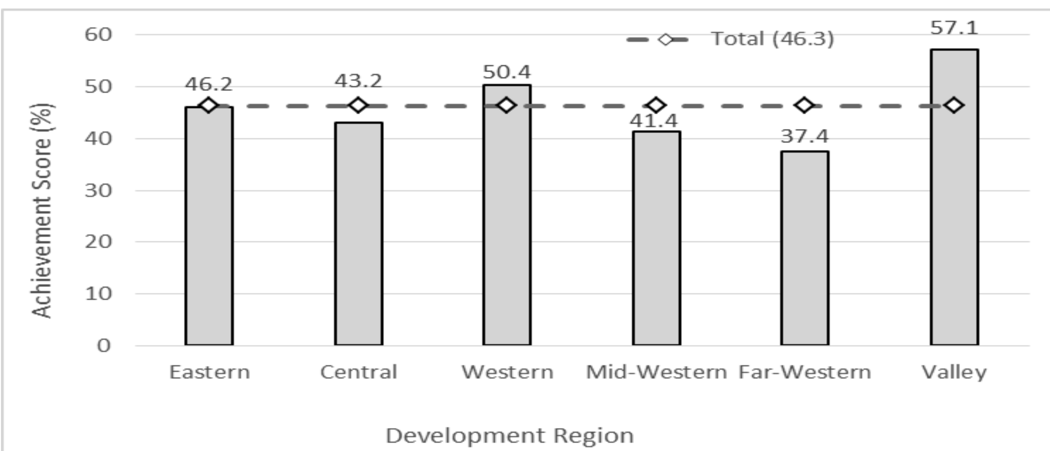
As in the earlier test of 2012, the Valley has outperformed in this test; however, it has not maintained its achievement level as in earlier test (76 percent). In the earlier test the Tarai recorded the lowest performance (52 percent) whereas in this test the Mountain is the lowest achieving of all even behind the Tarai. In this test, Tarai has raised its position above than Mountain. In comparison to the previous test, all the Ecological Zones have achieved lower in this test (see, ERO, 2015b: table 4.1.14 & figure 4.1.11).

Dataset indicates that there is a wide difference in the students' performance among the four Ecological Zones. Students from the Kathmandu Valley outperformed the other students. The achievement is the lowest in Mountain Region.

Student Achievement by Development region

Students' achievement varies according to the Development Regions, which are: Eastern, Central, Western, Mid-Western, and Far-Western. Additionally, the Kathmandu Valley has been considered as the additional Region though administratively it falls under the Central Development region. The mean achievement for each Development region is given in figure 6.8.

Figure 6.8 Comparison of the achievement of various Development Regions in Nepali



The above data shows that the highest achievement is in the Kathmandu Valley (57%), which is 20 percent higher than the Region having lowest achievement level (37%), the Far-Western Region. The achievement is below the national average in Far-western Region (37%), Mid-western region (41%) and Central Region (43%). Western Development Region and the Kathmandu Valley have achieved higher than the national mean. The difference amng the Regions is statistically significant ($p < 0.001$) as Tukey's *post hoc* test shows that excluding Central, Mid-Western Regions, other Development Regions differ from each other significantly with significant level of 0.001%. Development Regions show 8% variation in data and its effect size is medium (f =0.30).

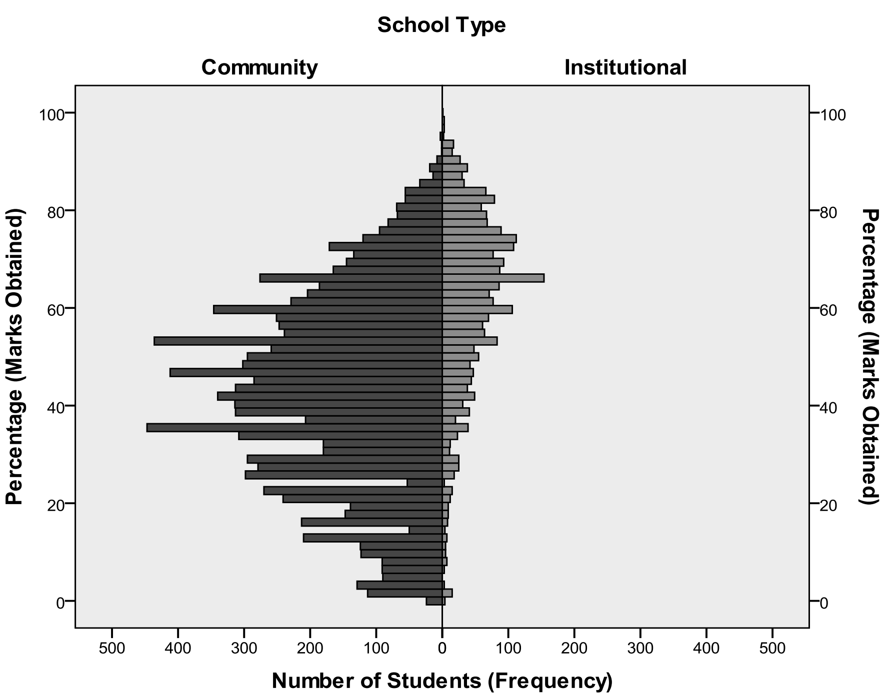
Compared with the 2012 data sets, the Eastern Region has improved its results from the lowest of all (52%) in earlier test to the level of national mean in this test. The Western Region is in the second position after Valley in this test as well. Hence, it has maintained its earlier position, though the score is comparatively low. In previous test, except the Valley and Western Regions, other regions (Eastern 52%, Central 54%, Mid-Western 52% and Far-Western 58%) performed below the national average (60%). However, in this test, Far-Western, Mid-Western and Central Development Regions have performed below the national average. Furthermore, all the Regions have achieved less than their achievement in earlier test (see, ERO, 2015b: table 4.1.15 & figure 4.1.12).

The dataset indicates that the highest performing Region – the Kathmandu Valley (66%) – is remarkably higher (19%) than the lowest performing Region - the Far-Western (37%).

Student Achievement by School Type

One way of categorizing schools in Nepal is based on the ownership and management. On this basis, there are two types of schools covered by this study: community (N = 380) and institutional (N=118). Community schools are funded by the government and managed by the community, and generally known as public schools. Institutional schools are generally known as private schools.

Figure 6.9 Distribution of the students' mean score in community and institutional schools



The left-hand side distribution shows the community school students and the institutional school students on the right-hand side. There are a few numbers of the students in the community schools getting equally high marks as in the institutional schools. Figure 6.9 explains that the students in the community schools vary from low performers to high performers whereas the students from the institutional schools are concentrated more on relatively high performing group.

The differences between the achievement of students from community and institutional schools are presented in table 6.5.

Table 6.5 Student achievement by type of school in Nepali

Type of school	N	Mean	SD	SE
Community	10790	42.8	19.8	0.20
Institutional	2523	61.4	18.4	0.40

The achievement gap between the community and institutional schools is remarkable (19%). The average achievement score for institutional schools is 61% whereas for community schools it is only 43%, which is below the national average. The difference is statistically significant ($p < 0.001$) and the effect size is medium ($f = 0.37$), which shows a wide difference between the community and institutional schools. Category of the students in the community and institutional schools explains 12% of the variation in student achievement ($\eta^2 = 0.12$). The variation within the community schools is also remarkable, ranging from 5% to above 80%; but the variation is relatively smaller in most private schools (see table 6.5). While comparing the results between NASA 2012 and NASA 2015, it is found that the achievement gap which was at 24 percent (average achievement 54 percent for community schools and 78 percent for institutional schools) in the earlier test has come down to 19 percent in this test. The gap seems to be narrowing down by 5 percent though the achievement score in this test is lower than in the previous tests. Despite the fact, there is a lot to do towards reducing the achievement gaps between these two types of schools by raising the learning standards in community schools.

The dataset indicates that, on an average, the students from institutional schools outperformed the students from community schools. Because of the low achieving bigger population from community schools, the entire system is seen to be shifted towards low performing.

Table 6.6 Comparison of results by content areas of two types of schools

Content Areas	NASA 2015		
	Community	Institutional	Total
Total	42.8	61.4	46.3
Reading	35.2	53.0	38.6
Writing	29.3	49.1	33.0
Grammar	33.1	56.6	37.5
Vocabulary	48.6	69.0	52.5

The result of previous test is better in all content areas in both types of schools. Even within the community school, the highest achievement gap is noted in Grammar (26%) and in institutional schools, it is in Writing (27%).

Table 6.7 Comparison of results of two types schools by cognitive domain

Cognitive Domains	NASA 2015		
	Community	Institutional	Total
Total	42.8	61.4	46.3
Knowledge	62.1	78.0	65.1
Comprehension	44.1	64.6	48.0
Application	27.7	50.7	32.1
Higher Ability	27.9	45.7	31.2

Like in content areas, both the students from community and institutional schools performed better in knowledge level. The achievement score is lowest in higher ability in both types of schools.

Student Achievement by Location of Schools

One of the strata for sampling in NASA 2015 was school location. The schools were categorized into rural and urban schools accordingly. The achievement of students in rural and urban schools is presented in table 6.8.

Table 6.8 Location of schools and student achievement in Nepali

School Location	N	Mean	SD	SE
Rural	9857	44.4	20.3	0.2
Urban	2963	53.0	21.4	0.4

The achievement level of students from urban schools (53%) is higher than that from rural schools (44%). It shows that the institutional schools and schools from the Kathmandu Valley contributed to raise the achievement level of urban schools. Excluding the schools from the Kathmandu Valley, the mean achievement of rural and urban schools is found 44% and 49% respectively. The difference in average score is significant ($p < 0.001$) and the effect size is moderate ($f = 0.18$); and the location of schools explains 3% of variation in student achievement ($\eta^2 = 0.03$).

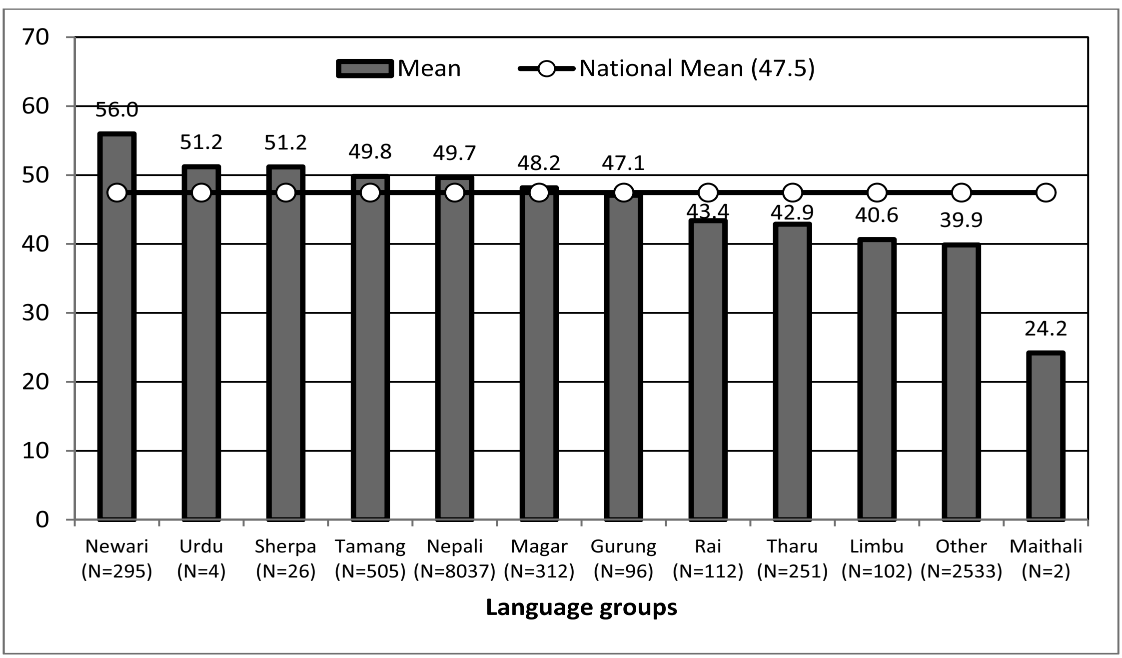
The above data indicates that the students in urban schools obtained 9 percent more score than the students in rural areas. The mean score of rural schools does not differ whether the Kathmandu Valley is included or not from the analysis.

While the results of urban and rural schools are compared between NASA 2012 and NASA 2015, the achievement level of urban students is found 18 percent higher in NASA 2012 (74% for urban schools and 56% for rural) whereas it is only 9 percent in NASA 2015. Though the gap between urban and rural schools is narrowed down over the years, rural schools are weaker achievement level in Nepali.

Language at Home and Student Achievement

Dataset shows that 40% of the 5th Graders speak a language other than Nepali as their first language. These "Other" languages are quite fragmented; the largest groups (19%) in the student dataset are Tamang (3.8%), Magar (2.3%), Newari (2.2%) and Tharu (1.9%). Some other ethnic/ caste groups like Rai, Limbu, and Gurung, are around 1% in the total sample population. There are still some 8% students who are not specified under any language group. The language groups like Newari (56%), Sherpa (51.2%) and Urdu (51.2%, N = 4) speaking have performed higher than the students from Nepali speaking group (49.7%). On the other hand, the students from Maithili (24.2%, N = 2), "Other" (39.9%), Limbu (40.6%), Tharu (42.9%), and Rai (43.4%) speaking background have performed much lower than the national average (fig. 6.10).

Figure 6.10 Language group-wise achievements in Nepali



It is difficult to generalize results of language groups. However, while considering the national average as the reference point, Urdu (57%), Gurung (56%), Newari (54%), Sherpa (52%), Rai (51%), Other (51%) and Tharu (41%) language groups performed below the national average in NASA 2012 test (see, ERO, 2015b). In NASA 2015 test also Rai (43%), Tharu (43%), and Other (40%) groups have performed below the national average. However, Limbu (41%) and Maithali (24%) who performed better in NASA 2012 are far below the national average in NASA 2015 test.

For the convenience of statistical analysis, the language group are further categorized into two groups: (a) Nepali speakers (60%) and (b) "Non-Nepali" speakers (40%). The results are presented in table 6.9.

Table 6.9 Home language and Student achievement in Nepali

Language group	N	Mean	SD	SE
Nepali	8037	49.7	20.05	0.22
Non-Nepali	4238	43.3	20.31	0.31

The difference between the language groups is statistically significant ($p < 0.001$), but the effect size is medium ($d = 0.32$). The students who speak Nepali as their mother tongue performed better than others. Like in the earlier test, in this test too, the achievement level of students is found depending on the language spoken in their home – the mother tongue of students. In the context of Nepal, the mother tongue also reflects the ethnic background largely, and hence the level of difference can be taken as a possible source of inequality in learning.

Ethnicity/Caste and Student Achievement

There are caste/ethnicity based variations in the participation in and achievement of education in Nepal. Historically, the Brahmans and Chhetris have higher level of educational attainment; but Dalits, for example, had less participation in or were deprived of educational opportunities. Hence, the government has made efforts to make education possible and accessible for all children regardless of their caste and ethnicity. The recent National Population Census 2011 shows that the enrolment of Hill Dalits has increased remarkably at the lower level of schooling but their number at the secondary and higher education is still very small (CBS, 2012). However, still there are differences in the achievement of students from various castes and ethnic groups, which are presented in table 6.10.

Table 6.10 Student achievement by ethnic/caste background

Caste/ Ethnicity	N	Mean	SD	SE	Minimum	Maximum
Brahman	1414	54.0	20.31	0.54	0	97
Chhetri	2902	46.6	20.56	0.38	0	100
Janjati	4852	48.7	19.29	0.28	0	98
Dalit	1631	43.5	19.70	0.49	0	98
Madhesi	1159	41.0	21.43	0.63	0	92
Minorities	77	37.0	19.28	2.20	2	82
Other	420	48.1	22.22	1.08	0	94

The above data shows that the Minorities (37%) and Madhesi students (41%) are performing the lowest in the Nepali language. When the results of test in NASA 2012 and NASA 2015 are compared, Madheshi students have performed lower in both tests. In NASA 2015 test, Chhetri students performed even below the Janajatis. In both tests, Brahmans have outperformed the rest groups (see ERO, 2015b: 235). The overall difference between the groups is statistically significant (p < 0.001) and the effect size is small (f = 0.17). Caste/ethnic background explains 2.9% variation in achievement (η^2 = 0.029). Apart from this, except for Chhetri, Janjati and Other students; the mean difference of the various caste/ethnic groups is statistically significant at 95% confidence level.

Although Dalit students performed below national average, a positive sign from equity point of view is that these students have performed remarkably better than the national mean (48%) in the Western Mountain (62%) and Eastern (51%) and Western Tarai (50%) (see table 6.11). However, the results are much lower than the average in other Regions and very poor in Far-Western Tarai (31%) and Mountain (35%) and Central Tarai area (35%). The number of students in particular strata is small, and hence it may be wise not to make too strong implications of the results. However, in the Kathmandu Valley, Dalits have performed higher (51%).

Table 6.11 Dalit students' achievement in Nepali in different ecological and Development Regions

	Eastern	Central	Western	Mid-Western	Far-Western	Valley
Mountain	51.4		62.1	40.3	35.5	
Hill	43.5	48.8	46.6	43.2	42.7	
Tarai	42.4	34.8	49.7	38.6	30.7	
Valley						50.6

Earlier dataset showed that Dalits have achieved better than national mean (60%) in the Eastern (65%), Far-Western (62%) and Western (61%) Mountain, Western Hill (60%) and in the Valley (70%), whereas it was lower in the rest parts especially in Mid-Western Mountain (50%), Eastern (37%) and Western (46%) Tarai.

Figure 6.11 Madhesi students' achievement in various content areas in Nepali

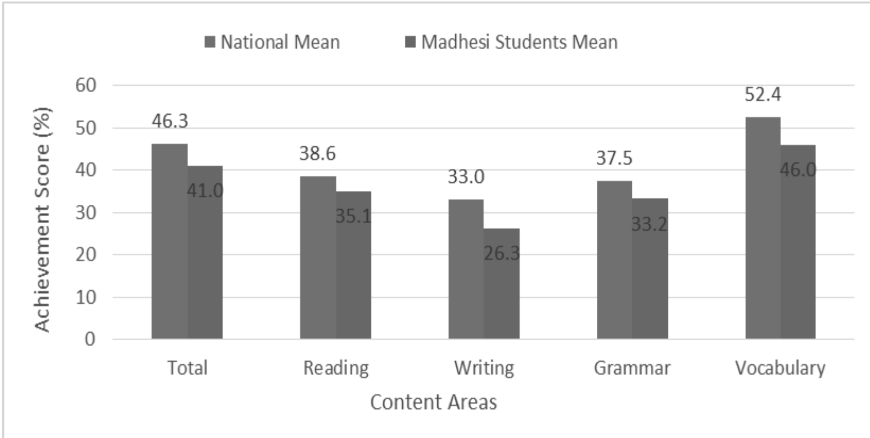
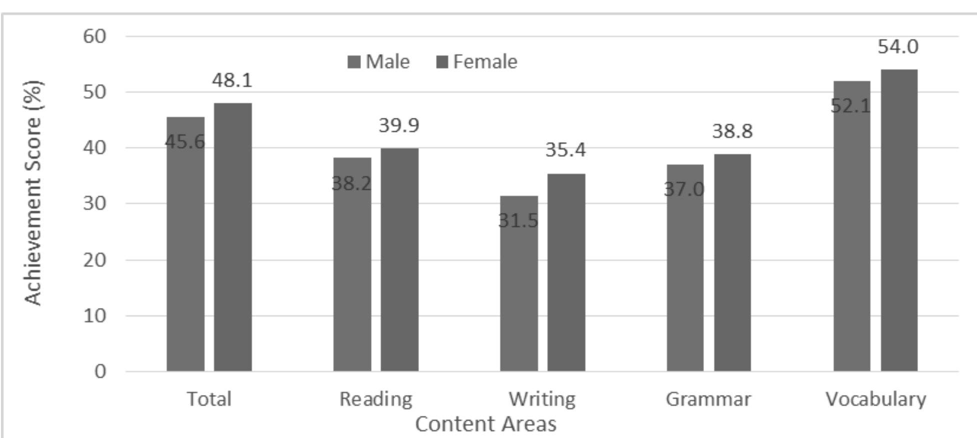


Figure 6.11 shows that the Madhesi students have achieved lower than the national mean in all content areas, particularly performance in W*riting* is found to be poor (26%). A notable percentage of Madhesi students have not achieved the required level of performance in R*eading, Writing* and *Grammar* in Nepali compared with *Vocabulary*. Dataset indicates that Madhesi students' performance in Nepali is lower than the national average except in the case of V*ocabulary*.

Achievement by Gender

Basic achievement results of boys and girls are presented in figure 6.12. This figure further compares the achievements of girls and boys in various content areas. The figure shows remarkable variations between boys and girls in various content areas in Nepali. However, girls performed slightly higher than boys in all content areas.

Figure 6.12 Comparison of achievement of boys and girls in different content areas in Nepali



There is a statistically significant difference between boys and girls ($p > 0.05$) in the total score. Girls are better performers (48%) compared to boys (46%), which is statistically significant ($p = 0.001$). The difference is not significant in *Grammar*, but it is significant in *Vocabulary* ($p < 0.001$), and the effect sizes is negligible ($d = 0.012$).

Table 6.12 Comparison of achievement scores in various contents areas in Nepali

Content Areas	2015			Difference
	Boy	Girl	Total	
Total	**45.6**	**48.1**	**46.9**	**2.5**
Reading	38.2	39.9	39.1	1.7
Writing	31.5	35.4	33.6	3.9
Grammar	37.0	38.8	38.0	1.8
Vocabulary	52.1	54.0	53.1	1.9

The table shows that girls slightly outperformed the boys in each of the content areas in Nepali.

The table 6.13 shows the gender variation in the achievement of students in various levels of cognitive domain for Grade 5 in Nepali. It shows that girls outperformed boys in all content areas in Nepali.

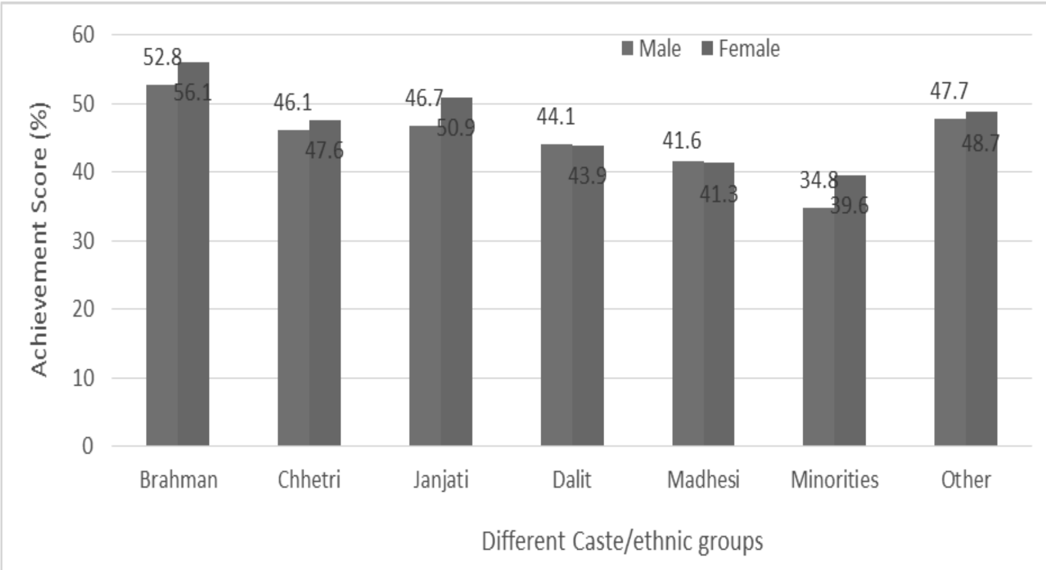
Table 6.13 Variation in achievement by cognitive domains in Nepali

Cognitive Domains	2015			Difference
	Male	Female	Total	
Total	**45.6**	**48.1**	**46.9**	**2.5**
Knowledge	64.9	66.5	65.7	1.6
Comprehension	47.4	49.6	48.6	2.2
Application	30.9	34.1	32.6	3.2
Higher Ability	30.2	33.0	31.7	2.8

Gender, Caste/Ethnicity and Achievement

The difference in achievement between boys and girls is the highest among Minorities (difference is 5 percent) where girls have outperformed boys. Tukey's post hoc test shows that differences are statistically significant at $p < 0.05$ level. Brahman and Janjati girls have performed better than boys, whereas boys and girls perform almost equally in Dalit and Madhesi communities.

Figure 6.13 Comparison of achievement of boys and girls in various castes/ethnicities in Nepali



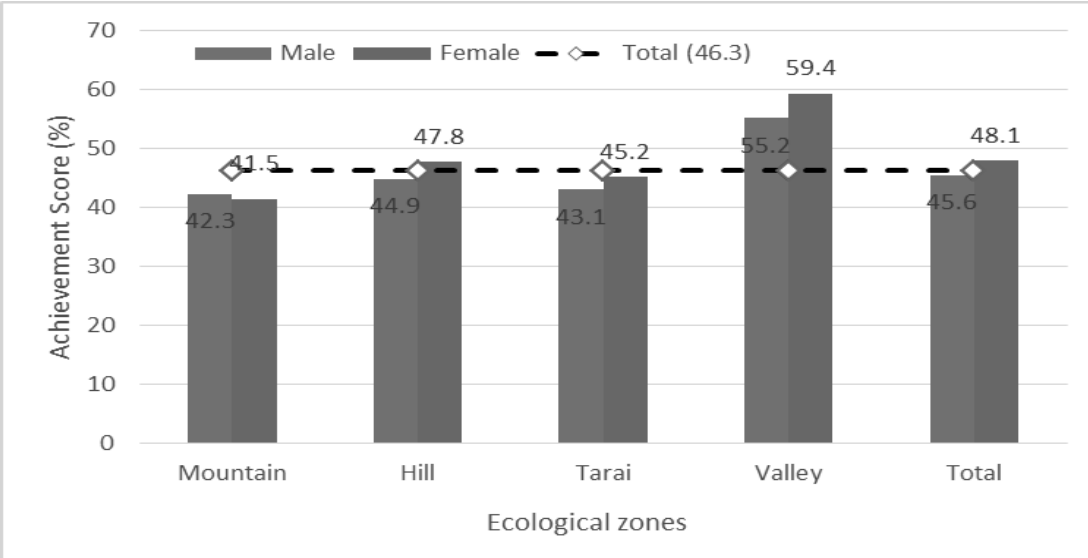
The achievement of girls and boys differs significantly across the Ecological belts. In the Kathmandu Valley, girls have performed better than boys (difference is 4 percent). Similarly, there is no statistically significant difference between boys and girls in Mountain and Hill, whereas it is statistically significant in Tarai and the Kathmandu Valley at $p < 0.001$.

Table 6.14 Comparison of achievement scores by various castes/ethnicities

Caste/Ethnicity	2015			Diff
	Boy	Girl	Total	
Brahman	52.8	56.1	54.5	3.3
Chhetri	46.1	47.6	46.9	1.5
Janati	46.7	50.9	49.0	4.2
Dalit	44.1	43.9	44.0	0.2
Madhesi	41.6	41.3	41.5	0.3
Alpasankhyak	34.8	39.6	37.0	4.8
Other	47.7	48.7	48.2	1.0
Total	**46.4**	**48.9**	**47.7**	**2.5**

The total achievement difference between girls and boys is not more than 2.5 percent in NASA 2015, that is, girls performed better than the boys. While comparing the achievement between girls and boys on the basis of caste/ethnity, the difference is the highest for Minorities (4.8%) and then for Janajati group (4.2%). Achievement gap is found less in Dalit groups in both tests and in Madheshi groups in the later test.

Figure 6.14 Comparison of achievement by gender in various Ecological Zones in Nepali



When it comes to the Ecological Zones, the differences between boys and girls are very small as the effect size is also small ($f = 0.024$).

Table 6.15 Comparison of girls and boys' achievement by Ecological Zones

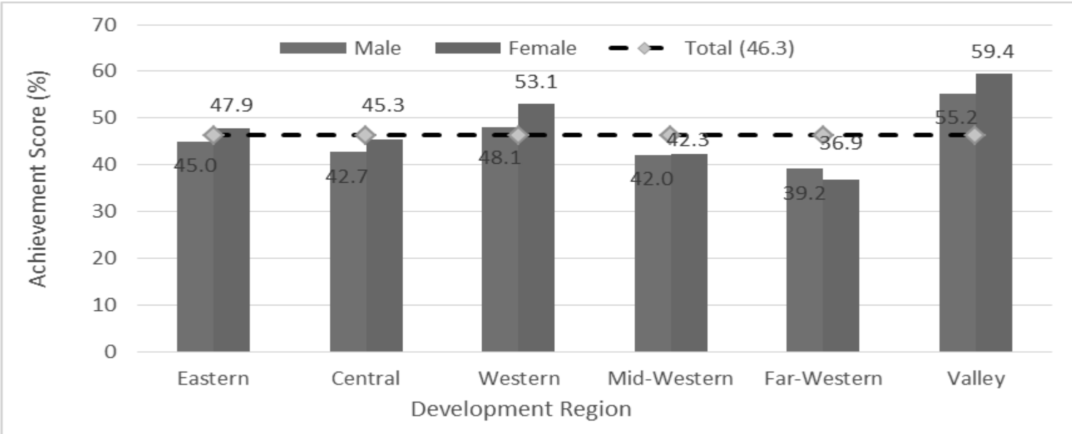
Ecological Zone	2015			Diff
	Boy	Girl	Total	
Mountain	42.3	41.5	41.9	0.8
Hill	44.9	47.8	46.4	2.9
Tarai	43.1	45.2	44.2	2.1
Valley	55.2	59.4	57.4	4.2
Total	**45.6**	**48.1**	**46.9**	**2.5**

The table 6.15 shows that the girls performed slightly lower than boys in Mountain Zone. Although girls' achievement score is slightly lower than that of boys, gender gap is noted less in Mountain (0.8%) than in other Zones. The achievement gap is found high in the Valley.

Gender, Development Region and Achievement

Notable differences are not found among the Development Regions in terms of gender. The difference between boys and girls is somehow wider in the Kathmandu Valley (4 percent) and Western Region (5 percent) (see, figure 6.15). Far-Western Region is the only region where boys have outperformed girls.

Figure 6.15 Comparison of achievement of boys and girls in various Development Regions in Nepali



Dataset indicates that boys and girls have performed almost equally in general. However, there are some variations among different communities and regions.

Table 6.16 Comparison of girls and boys' achievement by Development regions

Dev Region	2015			Diff
	Boy	Girl	Total	
Eastern	45.0	47.9	46.6	2.9
Central	42.7	45.3	44.1	2.5
Western	48.1	53.1	50.7	5.0
Mid-Western	42.0	42.3	42.2	0.2
Far-Western	39.2	36.9	38.0	2.3
Valley	55.2	59.4	57.4	4.2
Total	**45.6**	**48.1**	**46.9**	**2.5**

Except in Far-Western Region in both tests, girls outperformed boys in other Development Regions. The gender gap is found less in Mid-Western Development Region (0.2 %). The gap is found wide in Western (5 percent) and the Valley (4.2 percent points) Regions.

6.5 Selected Explanatory Factors and Achievement

Several factors have already been discussed in section 6.4; for example geographical factors such as Districts, Ecological Zone and Development Region; school-related factors such as school type and school location; and some individual factors related to the students such as home language, caste/ethnicity and gender. This section deals with socio-economic status (SES) of the students' families, paid work after school, students' attitude towards the Nepali language as a school subject, age of the student, and support provided for studies as the main family and individual related factors. As examples of deepening school and teacher-related factors, the availability of textbooks, assigning homework and providing feedback by the teacher, and selected activities in the school are considered.

Socio Economic Status (SES) of Parents and Student Achievement

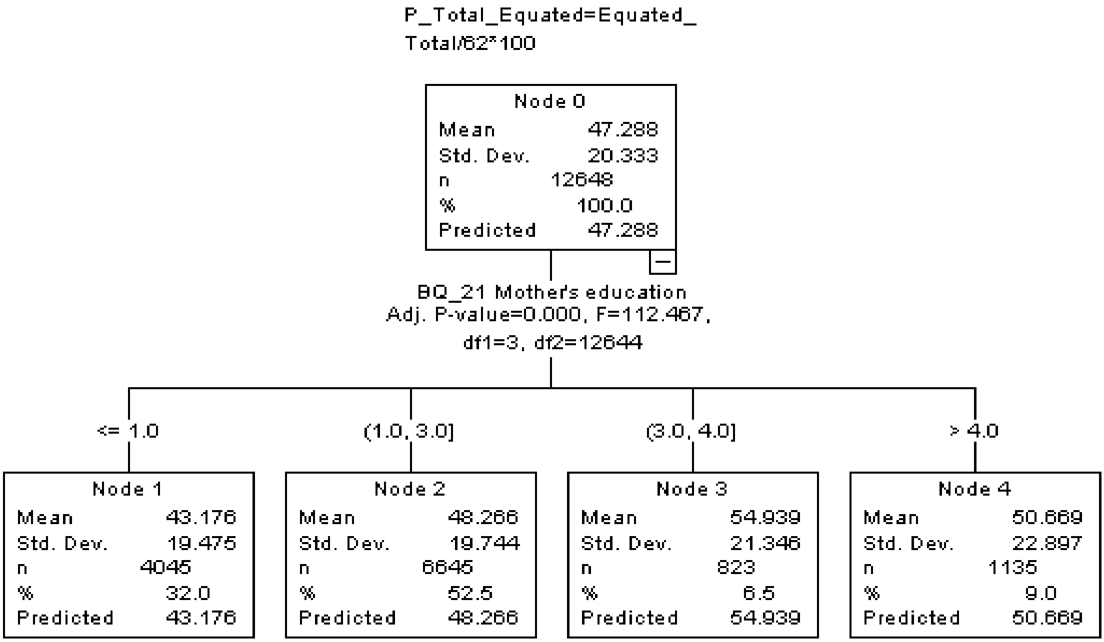
The variables indicating the socio-economic status were categorized into parents' education, parents' occupation, home possessions, home accessories, and type of school attended. Finally, SES is explained based on seven indicators related to economic, educational and occupational background of the parents. In this section, parents' education is further elaborated and the literacy status of the parents is analysed. Several SES related variables were analysed using a data-mining tool of SPSS - the Decision Tree Analysis (DTA). This method is used to find the cut-offs of the predicting variable, i.e., in case of mother's education it classifies the factors into several groups, as it differs statistically in the most significant way from each other in relation to student achievement.

Parents' Education

In this assessment, questions related to parents' educational status were included in the background questionnaire and is divided into seven categories as follows: 1) Illiterate, 2) Literate, 3) Grade 10, 4) SLC, 5) Grade 12 or Certificate level, 6) Bachelor's level, 7) Master's level and above.

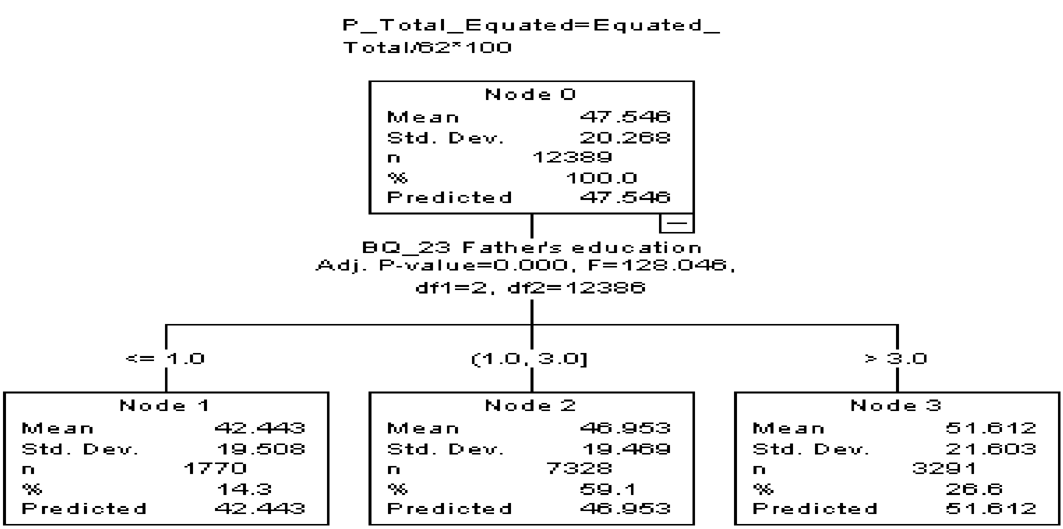
DTA classifies mothers' education into four groups with statistically significant differences in students' achievement levels. Figure 6.16 illustrates these categories with their values as: Illiterate mothers (43%), Literate mothers (48%), SLC passed mothers (55%) and mothers having Grade 12 or Bachelor's or Master's degree (50%). In each group, the number of mothers is high enough to make a credible prediction. The difference across the groups is statistically significant ($p < 0.001$).

Figure 6.16 DTA of mother's education and students' achievement in Nepali



Similarly, as in figure 6.17, DTA divides father's education into three broad categories. These categories are Illiterate (42%), Literate and Grade 10 passed (47%), and SLC and above (52%).

Figure 6.17 DTA of father's education and students' achievement in Nepali



From the figure, it can be concluded that students of illiterate fathers have achieved lowest and the achievement of students increases as fathers' education increases up to Bachelor degree. The positive relation between parent's education and student achievement is illustrated in figure 6.18. A strong relationship was found with mother's education. DTA shows that when both parents are illiterate, students achieved very low (42—43%), on the other hand they achieved highest (52−42%) when mother's education is SLC and father education is Grade 12 or higher level.

Dataset indicates that the educational level of parents affects the children’s achievement level. Achievement of the students is more influencing if either parents or mother is at least SLC passed, or if father has the qualification of Grade 12 or higher level.

Figure 6.18 Relation between parents’ education and student achievement in Nepali

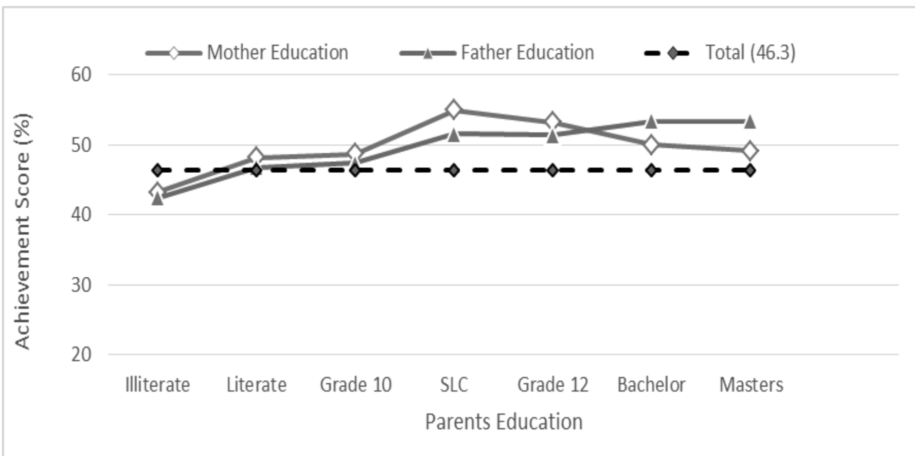
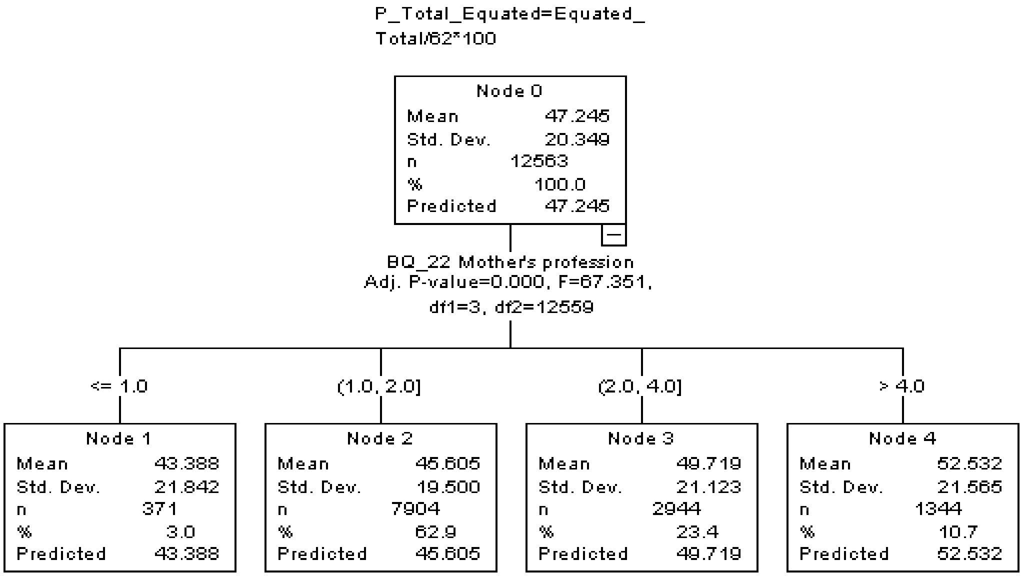


Figure 6.18 shows that father’s and mother’s level of education has a significant impact on their children's achievement. Students' performance is below the national average if their parents are either illiterate or just literate. Thus, the relation is seen as the higher the parents' education, the better is the students' achievement.

Parents' Occupation

The occupation of parents was categorized into eight groups as: 1) work in foreign country, 2) agriculture and household, 3) only household, 4) teaching, 5) service (government, non-government and private sector job), 6) business, 7) labour, 8) work at others' home, 9) others. While comparing the mean using ANOVA, student achievement is the lowest when mother works abroad (43%) or works at other’s home (44%). It is statistically significant and lower while comparing with the mother involved in teaching (50%), business and other job (54%). The result is presented in figure 6.19 as classified by DTA.

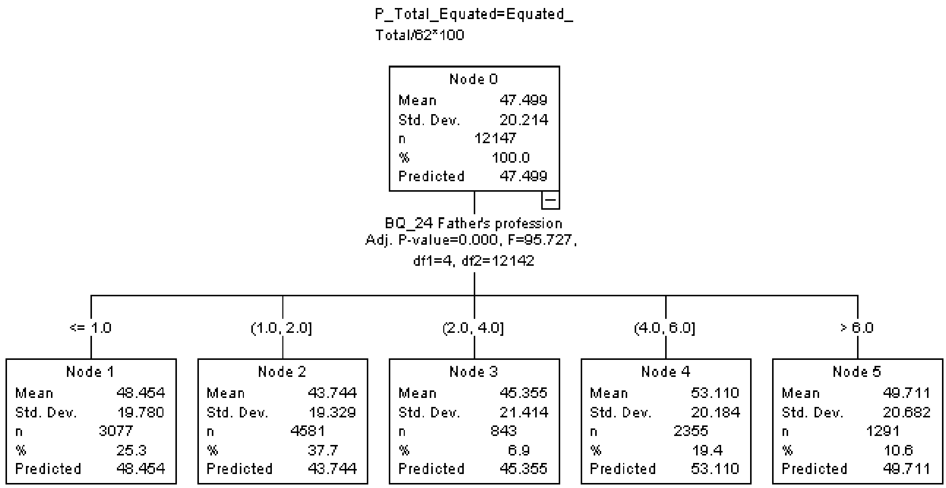
Figure 6.19 DTA of mother's occupation and students' achievement in Nepali



In a similar manner, as with parents' education, DTA was used to find the statistically most deviating groups related to student achievement in terms of parents' occupation. This analysis found four significant nodes. The lowest achievement is in the group where the mother works abroad (43%) and comes from agriculture background (45%) (see, figure 6.19). When a mother is in teaching profession, it can contribute 5 percent points score more than a mother from agriculture background and 7 percent points more than the mother who works in foreign countries. Higher achievement is seen when the mother is involved in service or business (52%).

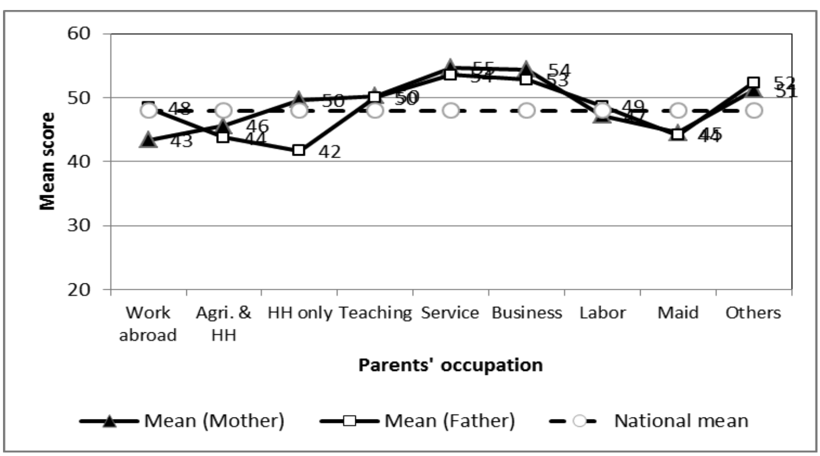
In case of father's occupation, DTA categorized students' achievement into five groups. The lowest achievement score is noted when father works in agriculture. If the father is in a teaching profession or in business, the students have performed better (53%) than the children of fathers who are working in agriculture. Students have achieved score above the national mean even if fathers are from the labour work and involved in other jobs, (see figure 6.20 and 6.21). However, the difference among the groups is statistically significant at $p < 0.001$.

Figure 6.20 DTA of father's occupation and students' achievement in Nepali



A more elaborated picture is depicted in figure 6.21. However, the result indicates that the mothers who are employed in some kind of job or business are seen beneficial for the achievement of students. The difference is found statistically significant (at $p < 0.001$) for each group.

Figure 6.21 Mother's and father's occupation and students' achievement in Nepali

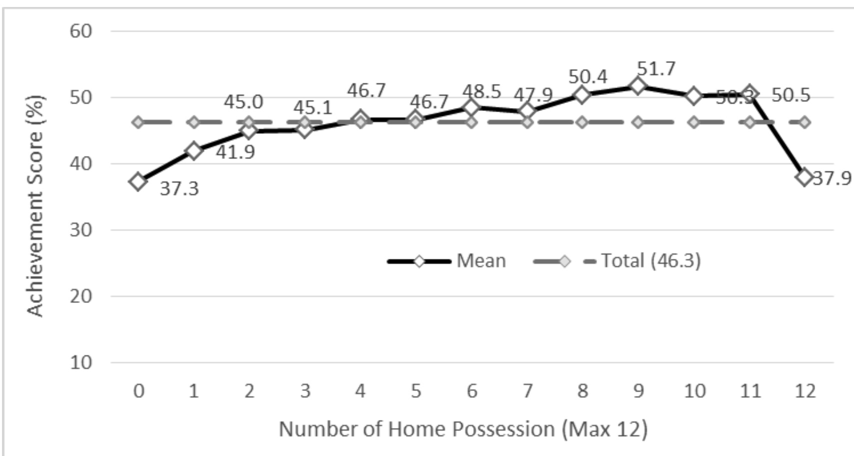


Home Possessions, Accessories and Student Achievement

This factor takes into account whether the facilities and resources available at home have some effects on student achievement. There were two kinds of home possessions defined in the background information questionnaire for the students. One is related to the facilities that are supportive for the students' study at home. For example, whether they have a table for study, a separate room for them, a peaceful place for study, a computer for homework and additional study, software for the computer assisted learning, internet facilities, literary magazines, access to classical literature, poetry books, or artistic things like pictures, dictionary and other books. Other types of home possession include different types of normal home accessories such as the number of mobile phones, televisions, computers and radios.

Related to home possessions, there were 12 questions in the background questionnaire for students. Each question was scored 1 if the student had an access to the possession. Adding these items up, the maximum score was 12 indicating that the student had access to all of the possessions, and hence lower score means fewer possessions at home. Figure 6.22 shows positive correlation between the home possessions and achievement level, except in the case of the highest category (i.e., all the home possessions). Pearson Product Moment correlation coefficient between the achievement level and the factor ($r = 0.11$) is statistically significant ($p < 0.001$) though the value is small.

Figure 6.22 Relation between home possessions and achievement in Nepali



For calculating the SES value regarding home possessions, the cut-off for the factors was set on four possessions. If the students possessed 5 items or more as mentioned in background questionnaire, the student was given 1, otherwise 0. The same pattern – the more accessories, the better results can also be found with home accessories, as in figure 6.23. The question in the background questionnaire was set differently compared with home possessions. Regarding the accessories, the question was: "How many of the following accessories do you have in your family?" The question was accompanied by the options 0 to 3 (or more). The availability of home accessories is dichotomized in three groups. After dichotomizing the items individually by using meaningful cut-offs found with ANOVA and DTA (and maximizing the differences in achievement level, see table 6.17), all the three indicators were summed, though the missing is not included. The maximum score was 4 indicating that the students possessed a set of all the accessories.

Table 6.17 Dichotomizing the indicators for home accessories

Accessory	Cut-off for 1	Cut-off for 0
Mobile phone	2, 3	0 and 1
Television	1–3	0
Computer	1–3	0
Vehicle (motorcycle or car)	1-3	0

Figure 6.23 Availability of home accessories and achievement in Nepali

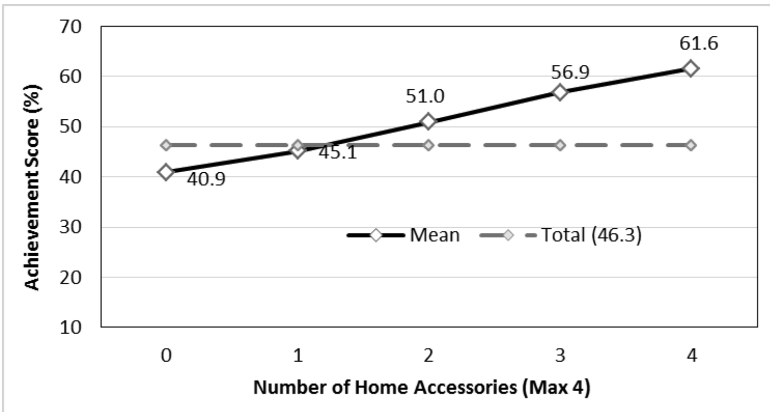


Figure 6.23 clarifies that when the number of home accessories increases, students' achievement also increases ranging from 41% (if none of them are available) to 62% (if any three of them are available). Availability of the stated facilities indicates a higher SES of the family. Therefore, the correlation between home accessories and achievement is found strongly positive (r = 0.27), ($p <$ 0.001) and the effect size is medium.

Data shows (figure 6.22) that when children have very few home possessions (i.e. 0 to 2 out of the 12), the achievement level is found low (< 37%), which is statistically significant than the achievement of students of families having more than two home possessions (> 45%). When the family is having ten to eleven possessions, the average score is higher compared to the national average. The same is the case for home accessories (figure 6.23), that is, when none or only one accessory indicator out of four is met, the results are lower than the average (40–45%) and when there are two or more met, the results are remarkably higher (> 51%). If four indicators are met, the results are the highest (61%).

The dataset indicates that either economic or intellectual ability or both at home support the children to raise their educational achievement. If the students come from high SES family (having a large number of home possessions and accessories), their performance is found remarkably higher than the performance of other students.

SES and Achievement

The value of socio-economic status was calculated based on seven indicators, which were dichotomized first. The variables including mother's education, father's education, mother's occupation, father's occupation, home possessions, home accessories, and the type of school were summed as SES and changed into the percentage of the maximum score (P-SES). Deeper description of the transformations for SES has been given in chapter 2. The P-SES represents the percentage of SES of the student's family; 100 means that the student has the highest possible SES where all the seven indicators of SES are positive and 0 refers to the lowest possible SES where all the seven indicators of SES are negative. The analysis of P-SES by using Univariate GLM (the Regression modelling) shows the strong relation between SES and student achievement. Figure 6.24 presents the relationship between SES of the students and their achievement.

Figure 6.24 Relation between SES and student achievement in Nepali

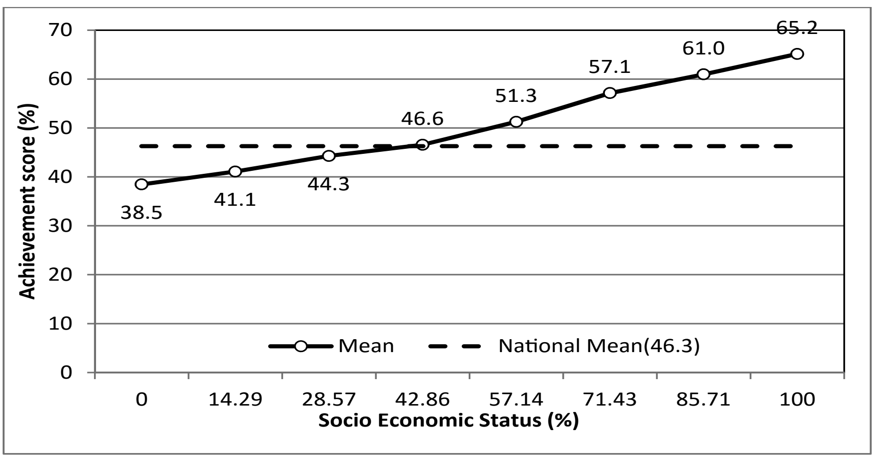
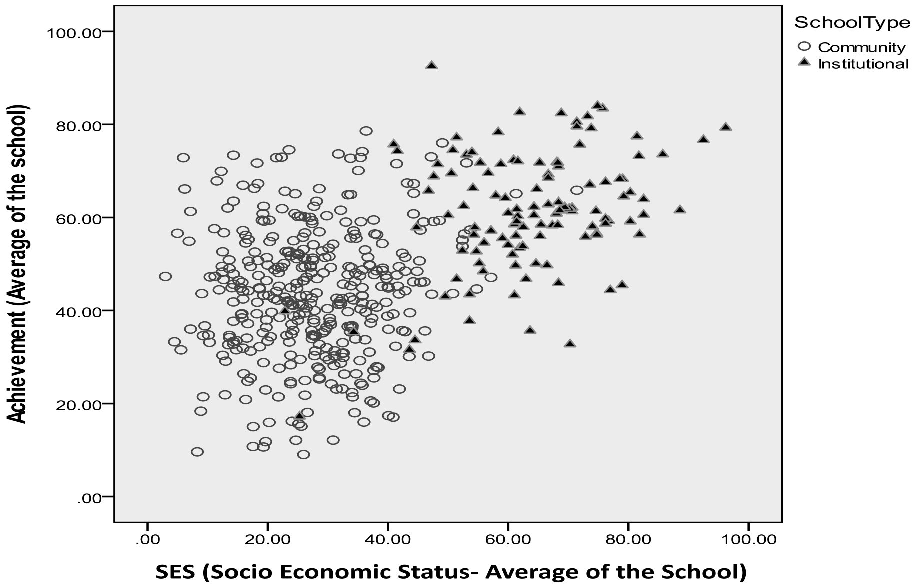


Figure 6.24 shows a positive relationship between SES and achievement as the correlation between the variables is r = 0.302, which is statistically significant (p < 0.001). The differences between the SES groups are statistically significant (p < 0.001) and the effect size is high (f = 0.32). SES explains about 9% of the variation in achievement (η^2 = 0.094).

The dataset indicates that socio-economic status plays a vital role in student achievement. The difference between the lowest and highest SES groups is remarkable (27 percent). Especially challenging is the situation in the families where the father or both parents are illiterate or both of them work in agriculture.

While plotting the school in terms of mean score and SES in a scatter diagram as in figure 6.25, community and institutional schools (community school in a circle and institutional school in a triangle) fall into two distinct groups. Most of the institutional schools, which have high SES value, are among the relatively high performers, whereas the community schools even having low SES value belong to the ranges from the low performing to high performing.

Figure 6.25 Achievement of schools with various levels of SES in Nepali



The dataset indicates that most of the students in the institutional schools have high SES, and they have achieved relatively higher score in comparison to the students from community schools. Further, the students from the community schools form two kinds of groups: high-performing and low-performing schools. Although the variations in achievement among institutional as well as community schools are wide, community schools are concentrated more on the range of 10% to 50% score. Institutional schools are concentrated more on the range of 40 to 90% of score.

Working Beyond School Hour and Achievement

Seven questions were incorporated in the background questionnaire of the students' activities beyond their school time. The values of the variables are divided into five categories: 0 (no at all), 1 (less than 1 hour per day), 2 (1 to 2 hours per day), 3 (2 to 4 hours per day), and 4 (more than 4 hours per day). The GLM Univariate (regression modelling) indicates that cut-off, for example, whether or not the students work in a paid capacity. The relationship is negative when students are engaged in various activities like playing and chatting with friends, playing games, paid job, and entertainment. On the contrary, watching TV for 1 to 2 hours a day has positive effect on students' achievement, whereas watching TV more than two hours or not watching TV negatively affected the students' achievement (see figure 6.26).

Figure: 6.26 Relationship between watching TV and achievement in Nepali

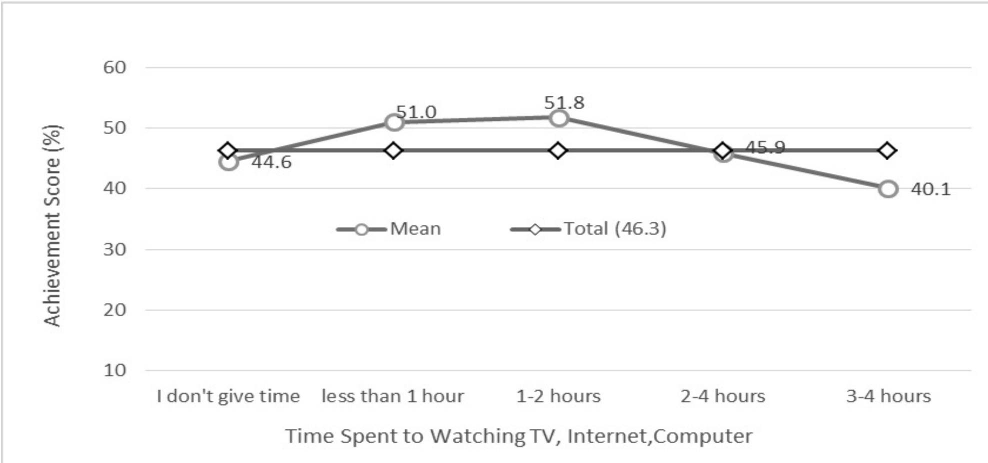
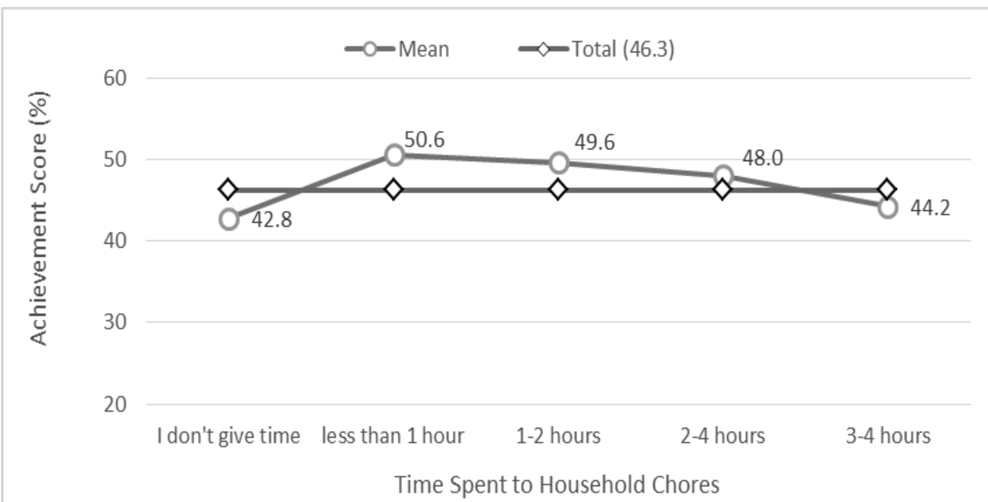
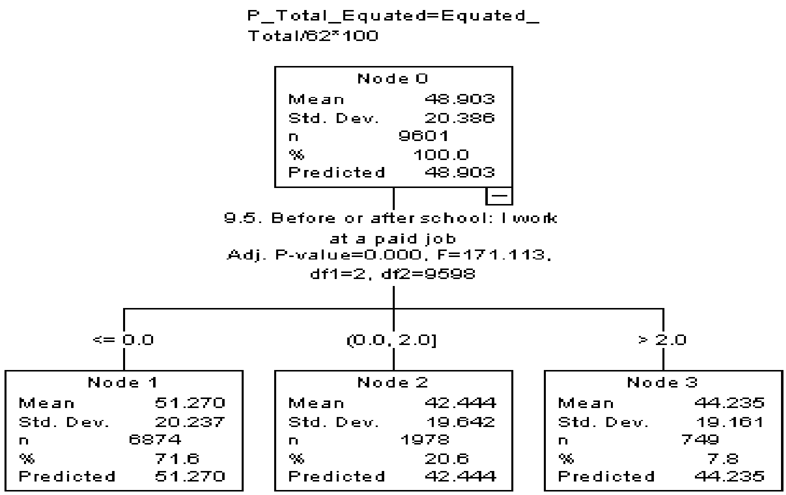


Figure 6.27 Relationship between household chores and achievement in Nepali



The above figure shows that, when children do not work at all at home, the results are lower than the national average (46%). If the students are involved in household chores less than an hour or one to two hours a day, there is an increase in achievement score, but when working more than 2 hours a day the results start declining, and working more than 4 hours or above a day in household chores reduces achievement remarkably. The differences are statistically significant ($p < 0.001$) though the effect size is medium ($f = 0.14$). Out of the total, 18 percent students were reported to have been engaged in household chores for more than two hours a day.

Figure 6.28 DTA of paid work and achievement in Nepali



The DTA shows that when the children have no paid work at all, the results are notably above the national average (51%) (see, figure 6.28). If the students are working on a paid job, even if for less than one hour, the results are remarkably lower than the average (42%). The differences are statistically significant ($p < 0.001$) though the effect size is medium (0.35), and a large number of the children were not engaged in paid work. Working before/after school time for payment indicates that the family is poor and the extra earning is needed.

The dataset indicates that working in a paid job or working in unpaid household work for four hours per day noticeably reduces the students' achievement. However, a decent level of household chores (unpaid) that is up to two hours per day supports the learning for the students.

Students' Attitude towards the Subject and Achievement

In the context of Nepali language assessment, attitude has been explained in terms of what the students think about Nepali subject and its usefulness in their daily life and future. There is more or less some sort of association between the students' attitude and their achievement. The correlation between achievement and attitude towards the subject is widely studied, though the connection is not always clear, (see, for example Metsämuuronen 2012a; 2012b; House & Telese, 2008; Shen & Tam, 2008; Kadijevich, 2006; 2008). In NASA 2015, the same shortened version of Fennema–Sherman Attitude Scales (FSAS, Fennema & Sherman, 1976), as used in several international comparisons like TIMSS and PISA studies are used. The original scale contains nine dimensions; but in the international comparisons, only three were used with four items in each dimension and two negative items in each of the first two dimensions (see the detail in chapter 2). The names of the factors can be "Liking Nepali", "Self-Efficacy in Nepali", and "Experiencing the Utility of Nepali" (compare naming in, e.g., Kadijevich, 2006; 2008). Factor analysis was used to identify the factors of the responses in FSAS and the negative items were reversed to make the whole test unidirectional. As in several countries of Asia, the expected factor structure cannot be found in Nepal (for a deconstruction of the test scales, see Metsämuuronen, 2012a; 2012b). Hence, only the total score is used to show the connection of attitude and achievement. The relation between the attitude, as divided into four groups with somehow an equal number of the students (that is, deciles) and the achievement score is shown in table 6.19 and figure 6.29.

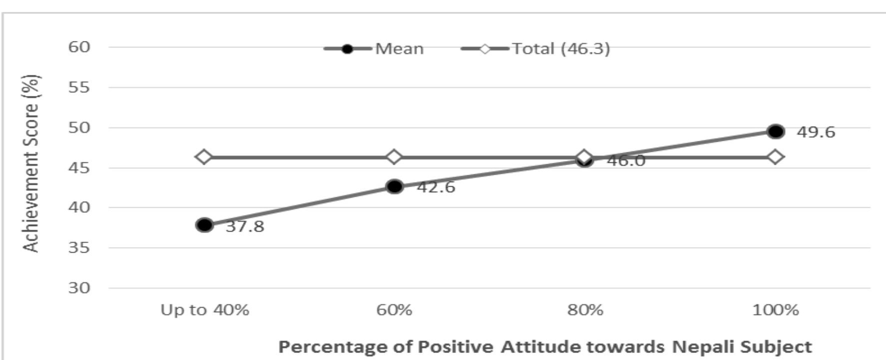
Table 6.18 Students' response towards teacher and school related activities

Students' attitude towards Nepali	**Respondents in % (valid percentage)**			
	Fully agree	**Partially agree**	**Partially disagree**	**Fully disagree**
20a. Nepali subject helps me in the daily life.	86.2	9.8	1.7	2.3
20b. I need to do better in Nepali to learn other subjects.	76.7	16.5	3.1	3.7
20c. I need Nepali language for governmental task.	82.0	11.1	3.8	3.1
20d. I need Nepali for speaking in all part of the country.	70.3	20.2	5.5	3.9
20e. I need to do better in Nepali to get a desirable job.	65.8	21.9	6.7	5.6
Average	86.2	9.8	1.7	2.3

Table 6.19 Attitude towards the subject and student achievement

Percentage of positive attitude	**N**	**Mean**	**SD**
Up to 40%	1375	37.8	20.11
60%	1242	42.6	21.01
80%	2159	46.0	20.56
100%	7970	49.6	19.88

Figure 6.29 Relation between attitude and achievement in Nepali



There is clear positive correlation between the students' attitude and their achievement (r = 0.19). The connection is moderately high (f = 0.19); the division of attitude into ten groups explains the achievement level at about 3% (η^2 = 0.03) variation. The difference between the lowest and highest attitude group is 12 percent. Data indicates that positive attitude towards the Nepali subject is positively correlated with the higher achievement.

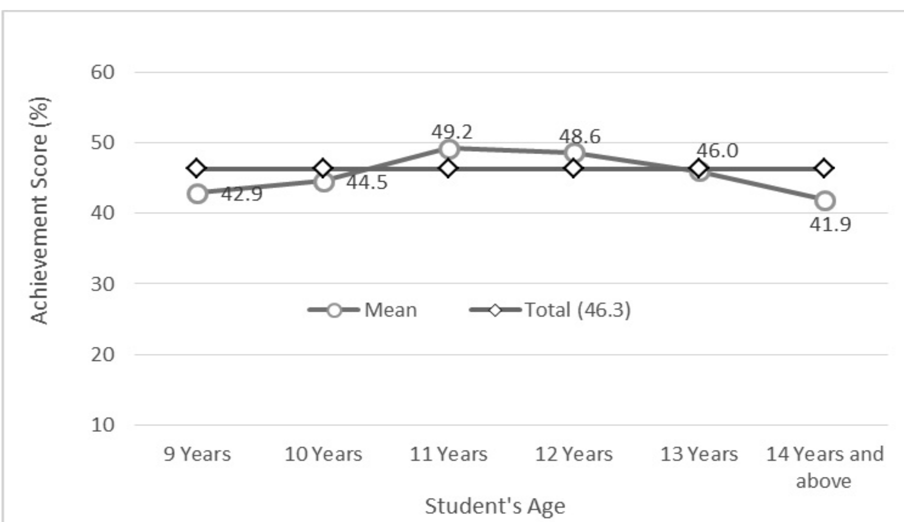
***Age of the* Students and their Achievement**

The age of the students attending Grade 5 varies widely, although the official age for Grade 5 students is 10 years in the Nepalese context. Some students have mentioned their age below ten years and some above 18. All the students aged below 9 were encoded as '9 years', and all those above 14 were encoded as '14 years and above'. The descriptive statistics of the mean by each age year are presented in table 6.21 and figure 6.30.

Table 6.20 Descriptive statistics of the students' achievement in different age groups

Age	N	Mean	SD
9 Years	140	42.9	20.34
10 Years	1202	44.5	20.68
11 Years	3216	49.2	20.22
12 Years	4534	48.6	20.52
13 Years	2035	46.0	20.19
14 Years and above	1244	41.9	19.57

Figure 6.30 Relation between the age of the students and achievement in Nepali



It is evident that the best achievers are those students who are at the proper age for Grade 5 studies, i.e.,11 and 12 years old, scoring about 49% (during test administration time the students were in Grade 6, so the dataset presents proper age 11 year, but not 10 years). The students studying Grade 5 with higher age means either the students have started their study much later, or they have repeated the Grades. The achievement level is remarkably lower than the average when the students are 14 years or above in age. Correlation between the variables is negative (i.e., - 0.48 at $p < 0.001$), indicating moderate effect size ($f = 0.12$). The ANOVA hints that the age explains 1% ($\eta^2 = 0.01$) variation in results. The difference in achievement between proper age and over age students is up to 7 percent. Dataset indicates that the students studying at proper age have obtained the highest achievement score and the achievement decreases as the age increases.

Support to the Students for Study and their Achievement

The relation between the support provided to the students for study and their achievement was analysed based on the information gathered on the question *"Who supports you when you do not understand what you have read?”*. In the question, there was only one option to be chosen by the student, that is, only one supporter could be reported. The descriptive statistics of the supporters are given in table 6.21.

Table 6.21 Descriptive statistics of support to the students and achievement level in Nepali

Support Providers	N	Mean	SD
Sister/Brother	6288	49.2	19.91
No one	691	48.8	21.76
Mother	835	48.8	21.39
Father	1649	45.9	21.34
Teacher	1013	43.9	20.80
Tution	1782	42.7	18.72

Support is necessary for the students to gain better than average achievement. Mother's support at home raises the achievement level (49%) in comparison to the support from father (45%) and teacher (44%). It is interesting to note that those who were supported by their teacher obtained lower score (44%) against those who did not get any support (49%). The highest performing students get support from their sister and brother.

The dataset indicates that the support provided by the sister/brother and mother has raised the achievement level more than the support provided by father, teacher and tuition.

Availability of Textbook and Student Achievement

Some students did not have the Nepali textbook up to the end of the academic session. Table 6.22 shows the descriptive statistics of availability of the Nepali textbook and the achievement.

Table 6.22 Availability of Nepali textbook and achievement

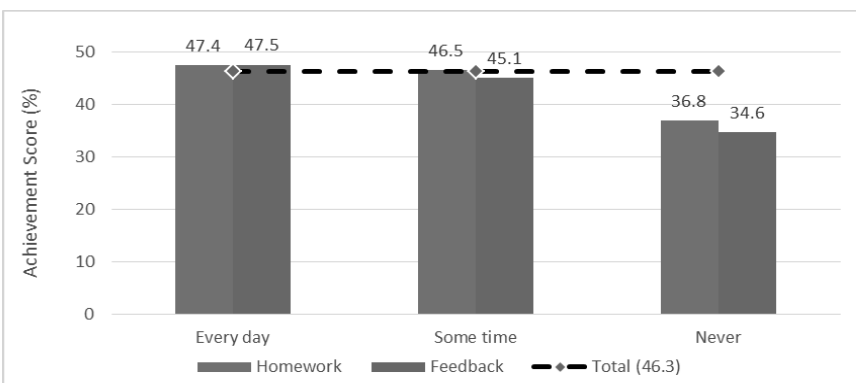
Availability of textbook	N	Mean	SD
Yes	11734	47.7	20.32
No	582	40.4	19.72

Out of 14629 students who responded to the question, 4.4% did not receive textbook even up to the end of the session. The relation between the availability of textbook and achievement is significant ($p < 0.001$) though the effect size is medium ($d = 0.36$) due to the small sample size. The difference in achievement is 7%. Data shows that 4.4% students lack the textbook in Nepali. The achievement level of students without textbooks is significantly lower than the achievement of those who have access to the textbooks.

Homework and Achievement

Homework is considered as one of the ways to enhance learning which can be used as drill, exercise and an evaluation tool as well. When homework is regularly given and feedback is provided to the students, it is likely to boost their achievement level. Statistics related to assigning homework and providing feedback, and its relationship with learning achievement is presented in figure 6.31.

Figure 6.31 Relation between assigning homework and achievement in Nepali



The dataset reveals the fact that, if the teacher assigns homework and checks/provides feedback regularly to the students, achievement is higher (48%) than in the case of those students who are not assigned homework regularly (45%) or not assigned at all (36%). The differences are statistically significant ($p < 0.001$). Those groups without homework are, however, very small, and hence the effect size is small ($f = 0.08$), which explains only 0.2% of the variance in the data ($\eta^2 = 0.02$).

Dataset explains that if the teacher assigns homework and provides feedback regularly to the students, the achievement is higher than in the case of teaching without assigning homework.

Positive and Negative Activities at School and Student Achievement

The activities of students and teachers determine the learning environment of the school. Bullying, for example, is one of the hindering incidents for the students in schools – which may affect their learning. In the background questionnaire for students, several school-related activities were asked - some positive and some negative. Here, bullying is handled as one of the negative indicators and students' impression of school and teacher activities are taken as the example of positive indicators.

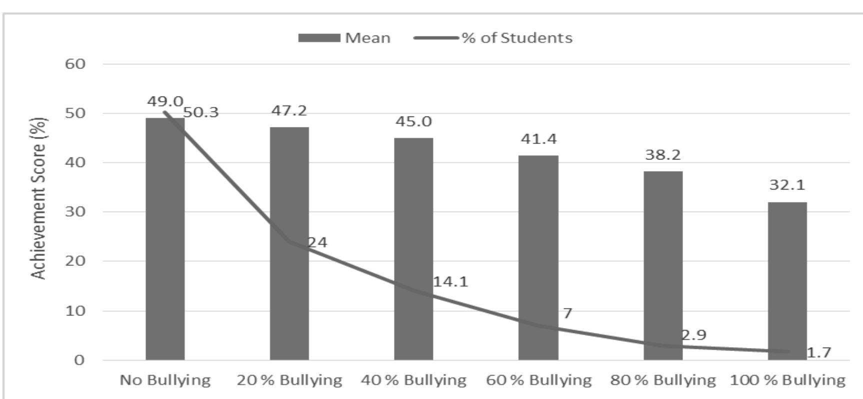
Bullying at School

Bullying is one of the problems in schools that worsens the learning environment. International studies like TIMSS and PISA emphasize to identify such indicators. In NASA 2014 student background questionnaire, five questions indicate the varieties of bullying that tend to happen in schools. All the incidences were stemmed by the phrase "*which of the following activities happened in your school in the last month?*" Students' responses are presented in table 6.23 and depicted in figure 6.32. 'No (%)' indicates the percentage of the students' response of no such activity happened in the school and 'Yes (%)' indicates the percentage of the students who experienced the particular type of bullying within the last month.

Table 6.23 Bullying and achievement in Nepali

Type of Bullying	No (%)	Yes (%)
Something of mine was stolen.	71.9	28.1
I was made fun of or called names.	78.6	21.4
I was hit or hurt by other student(s).	79.4	20.6
I was made to do things I didn't want to do by other students.	84.7	15.3
Fellow students kept outside without involving me in activities.	86.2	13.8

Figure 6.32 Effect of bullying in student achievement in Nepali



The sum of all five items is taken as an indicator of 100% bullying. Figure 6.32 shows the extent of bullying with the percentage of the students and their achievement in each category of bullying. If only one activity of bullying is reported, it is categorized as 20% bullying. If all five activities are reported, it is categorized as 100% bullying. Knowing that 49.7% of the students did not encounter any bullying during the last month, one can infer that the remaining 50.3% did encounter at least one type of bullying, which is a remarkable number of students. About 4.6% of the students had experienced some kinds of severe bullying (the sum of 80% and 100% bullying). It is found that learning outcomes are remarkably lower with 14% of the students who have encountered more than two different types of bullying (45%). Students who did not experience bullying and those who encountered extreme bullying of four or five kinds have 11-17 percent achievement gap; though there are a few students who reported this kind of bullying (n = 582). The difference is statistically significant ($p = 0.001$), but the effect size is small ($f = 0.16$). Though extreme cases of severe bullying are rare (n = 213), bullying seems to be quite common incident in schools.

The dataset indicates that a large number of students (49%) have encountered bullying in schools within a month. The phenomenon tends to affect the learning outcomes in almost all the groups of the students who experienced bullying.

Positive Activities in School

The activities that can boost the learning achievement of students are termed as positive activities. The students were asked about such positive activities in the school in two sets of questions listed in table 6.24. The table shows the responses of the students in all four categories, which are in the 4-point rating scale, anchored to fully agree and fully disagree.

Table 6.24 Students' responses towards teacher and school related activities in school

Teachers' and students' activities	Respondents in % (valid percentage)			
	Strongly agree	Agree	Disagree	Strongly disagree
BQ_28a School: Students get along well with most teachers	88.8	7.7	1.5	1.9
BQ_28b School: Most teachers are interested in student's well-being	90.3	6	1.7	2
BQ_28c School: Most of the teachers really listen to what I have to say	77.7	17.2	2.8	2.2
BQ_28d School: If I need extra help, I will receive it from my teacher	87.1	9	1.8	2.2
BQ_28e School: Most of my teachers treat me fairly	86.1	9	1.9	2.9
BQ_29a School: I enjoy to be in the school	93.8	3.5	0.9	1.8
BQ_29b School: Students in my school like me	84.2	12.2	1.6	2
BQ_29c School: Friends of my school want to their best in the study	90.1	6.1	1.5	2.2
BQ_29_d School: Teachers' expect good results of their students	92.3	4	1.2	2.5
Average	87.8	8.3	1.7	2.2

Further analysis was carried out by recoding the variables that are grouped into two categories:

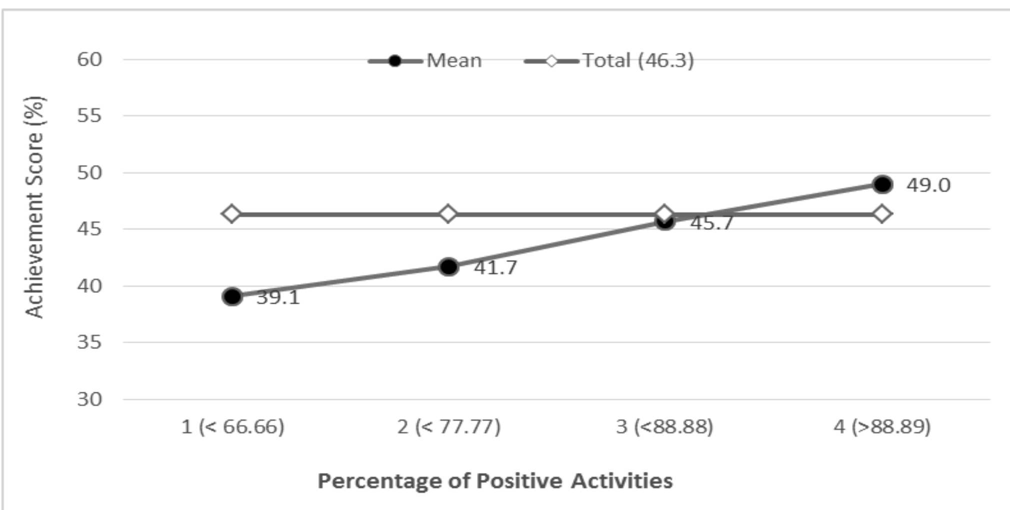
1 for agree, and 0 for disagree. Furthermore, the sum of nine indicators is converted into the percentage of maximum score to analyse the level of positive activities and its relation to the achievement.

DTA finds four attitude groups in the indicator. These boundaries and descriptive statistics are presented in table 6.25 and depicted in figure 6.33. The overall result is that feeling the positive actions in the school relates positively with the student achievement. The correlation between the sum of nine positive activities and achievement is found positive ($r = 0.18$, $p < 0.001$).

Table 6.25 Students' response towards teacher and school related activities

Percentage of positive action	N	Mean	SD
1 (> 66.66)	2015	39.1	21.09
2 (< 77.77)	797	41.7	20.32
3 (< 88.88)	1442	45.7	21.14
4 (≥ 88.88)	8775	49.0	20.00

Figure 6.33 Relation between positive actions in school and achievement in Nepali



The data shows that there is a positive relation between the positive activities of the school and the achievement. The increase in achievement is directly proportional to increase in the intensity of such activities. After dividing the indicator into four groups based on DTA, the differences between the groups are statistically significant ($p < 0.001$), however, the effect size is small ($f = 0.18$). The learning achievement of students are much higher than the average when the students are extremely positive towards school and teachers' behaviour. The difference between the most positive group and the most negative group is notable (10%).

Dataset indicates that when the students think that the actions of the teachers and the schools are ultimately good, the results are higher than the average (49%). At the other extreme of feeling, when such actions are ultimately negative the results are far below the average (39%).

6.6 Summary of Findings

The main findings of Grade 5 assessment in Nepali subject are as follows:

Basic results

- The achievement in Nepali subject is normally distributed as there are two distinct population groups: high and low performing.
- The average achievement in Nepali subject is 46.3 percent. The achievement is the highest in the content area of V*ocabulary* (52.4%) and the lowest in W*riting* (33.0%).
- Students' ability to solve complex problems is quite poor. Students are much better in the recalling type of questions (65%), which are geared more towards remembering the things than solving problems. Students' performance was found low in application than higher ability.
- The students are performing well in recognizing the correct answer and recalling simple facts from the texts, fundamental thinking, the basic interpretation of paragraph, table, chart and a few steps of logical thinking, and basically in objective items (55%). They are much weaker in producing fluent texts, letters, or preparing synthesis and abstracts from the text, basically in subjective items (33%).

Diversity factors

- There is a wide difference in achievement among the Districts in Nepali language. The results in Jajarkot and Bajhang (34%), Bhojpur, Morang and Bara (36%), Rautahat (37%), Doti and Kanchanpur (39%), Kalikot and Banke (41%) are relatively poor. Out of the 23 sample Districts, 11 Districts are below the national average. The gap between the lowest performing District (Jajarkot 34%) and highest performing District (Lalipur 60%) is 26 percent.
- There is a moderate difference in the students' achievement across the three Ecological Zones. Students in the Kathmandu Valley outperform (57%) the students from other regions. The achievement is found low in Mountain area (41%).
- There is a wide gap in students' achievement in the Nepali language among the Development Regions. The difference between the lowest performing Region (Far Western Region, 37%) and the highest performing Region (Kathmandu Valley, 57%) is remarkable, which is 20 percent.
- On an average, the students from institutional schools (61%) have outperformed those from community schools (43%). However, some students from community schools have performed equally as those from institutional schools.
- The students in the urban schools gained 9 percent more than the students in the rural areas. Most of the high performing institutional schools were from the urban areas. Excluding the Kathmandu Valley, the difference between urban and rural schools is 5 percent.

- Achievement difference is also found across the language groups. The students with Nepali mother language have performed better than the students who belonging to other language groups. However, the students from the language groups like Newari (56%), Sherpa (51.2%) and Urdu (51.2%, N = 4) have performed higher than those from Nepali speaking group (49.7%). On the other hand, the students from Maithili (24.2%, N = 2), "Other" (39.9%), Limbu (40.6%), Tharu (42.9%), and Rai (43.4%) speaking students performed much lower than the national average.
- The performance of Minorities (37%) and Madhesi (41%) is lower than the performance of the other castes/ethnic groups/communities. A notable percentage of Madhesi students obtained less score than the national average in all content areas.
- There is a difference between girls' and boys' performance in Nepali. Girls are slightly better performing than boys in all the content areas. Dalit and Madhesi boys are slightly better than girls in Nepali; however, the difference is not more than 1%. Differences are also found between boys and girls across the Ecological Zones and Development Regions.

Selected explanatory factors

- Parents' educational level strongly predicts the children's future achievement level in the Nepali language. Especially, achievement level is the lowest when the parents are illiterate.
- Either the high economic status of the family or better intellectual capacity or both have positive effects on their children's achievement score. If the father or mother or both are working abroad or they are involved in agriculture-related occupation, the students' achievement in Nepali subject is significantly lower than that of the students whose parents are from other occupational groups.
- When the children have very few home possessions i.e., 0 to 2 out of the 12, the achievement level is significantly lower (< 37%) than the achievement of those having more than two home possessions (> 45%). With a family having ten to eleven possessions, the average score is very high (> 51%) compared with the national average.
- Socio economic status (SES) of the family plays a strong role in the students' learning achievement in Nepali. The difference between the lowest and highest SES groups is remarkable (27 percent). Especially challenging is the situation in the families where the father or both parents are illiterate or if both parents work in the agriculture-related occupation.
- Either some paid work or unpaid household work for four hours per day outside the school has remarkably reduced the students' achievement. However, a decent amount of household work up to two hours per day does not affect the students' learning negatively in Nepali rather it has given positive effect.
- Positive attitude towards the subject correlates positively with learning achievement in Nepali.
- Achievement decreases as the age increases. The highest performance is with those students studying within their appropriate age group.
- Support to the learning provided by brother or sister, mother and no one has given more

positive effects on achievement level, compared to the support provided by father, teacher and tuition.

- It is a big challenge for students to learn without textbook, as many as 4.4% of the students did not have textbooks of Nepali subject even by the end of the academic session. The achievement level of these students is significantly lower (40%) than those who have access to the textbook (48%).
- If the teacher gives homework regularly and provides feedback to the students, their achievement is the highest.
- Students in a large number (50%) have encountered at least one kind of bullying in schools within a month. The phenomenon has affected their learning outcomes in almost all the groups of the students who experienced bullying.
- When the students think that the actions of teachers and schools are ultimately good, the results are better than average (49%). At the other extreme, if feeling is ultimately negative, the results are far below the average (39%).

Chapter 7: Analysis of Assessment Results for Grade 5 in English

7.1 Introduction

English is one of the compulsory subjects in the national curriculum of Nepal. It is the second language (L2) for most of the students in Nepal, but most institutional schools as well as some community schools have been using English as the medium of instruction. English is also a 'lingua franca' globally as it is widely used in the areas such as tourism, hotels or restaurants, business, and so on. Therefore, the importance of English language is growing day by day.

Although previous national assessments (see BPEP, 1994; CERID, 1993; 1998; 1999; CERSOD, 2001; Fulbright, 2008) were administered in Grade 5 and they tested English language proficiency, these assessments were not fully comparable with each other due to the lack of linking procedures among them. Therefore, it is difficult to compare the proficiency levels based on one of those studies against others. Proficiency in English was assessed for Grade 5 in NASA 2012 as well as NASA 2015, and made comparable using IRT modelling.

In NASA 2015, achievement of the Grade 5 students in English was assessed through the achievement tests among 13,146 students from 513 schools of 23 selected Districts all over the country. There was a representation from all the Ecological Zones and Development regions in all the strata including the types of schools and school location. Based on these strata and various diversity, basic results are analysed in disaggregated form. Furthermore, using the background information questionnaires, several influencing factors in student achievement are identified and analysed.

This chapter begins with the analysis of basic results of assessment in English for the students of Grade 5, which includes the overall distribution of scores, results in the different content areas and various cognitive levels. It then describes the effects of different diversity factors that explain the differences in achievement. Finally, it summarises the major findings of the assessment for Grade 5 in English.

7.2 Basic Results

The basic results of assessment in English includes distribution of overall results, results by content areas, levels of cognitive domains and item types, and the comparisons of results with previous assessments.

Distribution of Overall Results

The sample size for English subject (13,146 students) was big enough to form a normal distribution. Though the achievement scores in Grade 5 English is not distributed normally as the distribution is slightly skewed towards the left side. The distribution shows three distinct groups of students: high performing, medium performing and low performing (fig 7.1). The majority of the students are in the medium-performing level and negligible number falls in the group of achieving more than 90%, whereas significant number falls in the group achieving below 10%.

Figure 7.1 Distribution of overall achievement scores in English

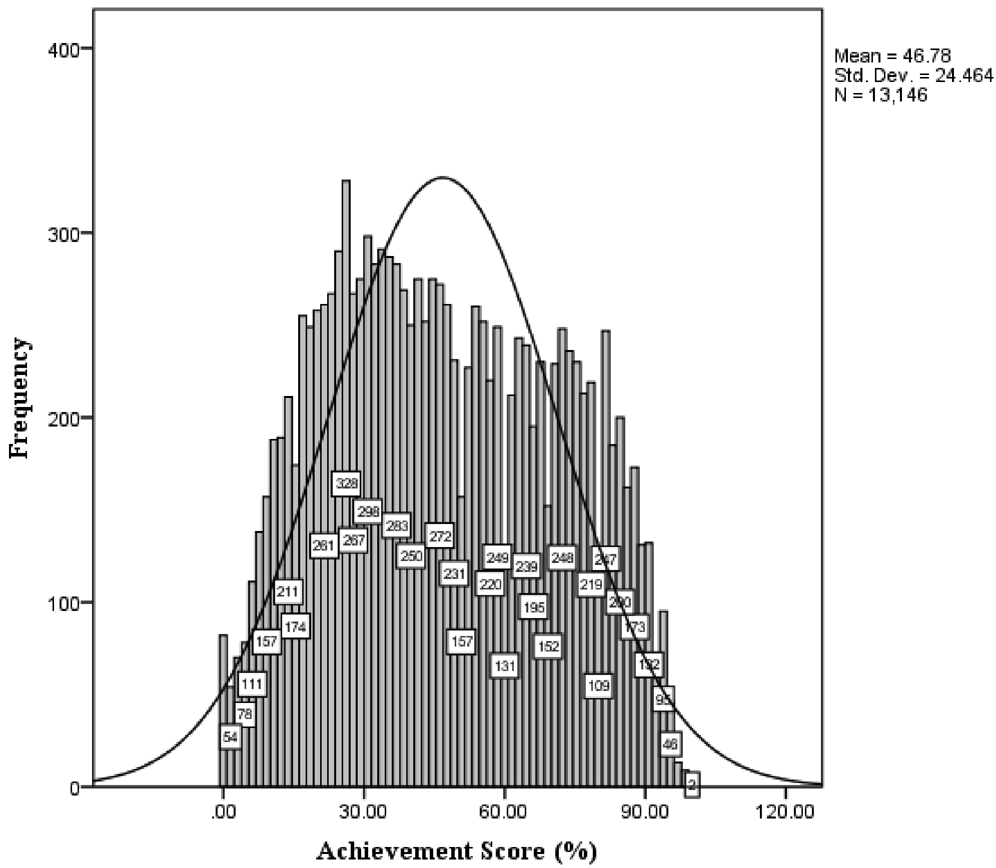
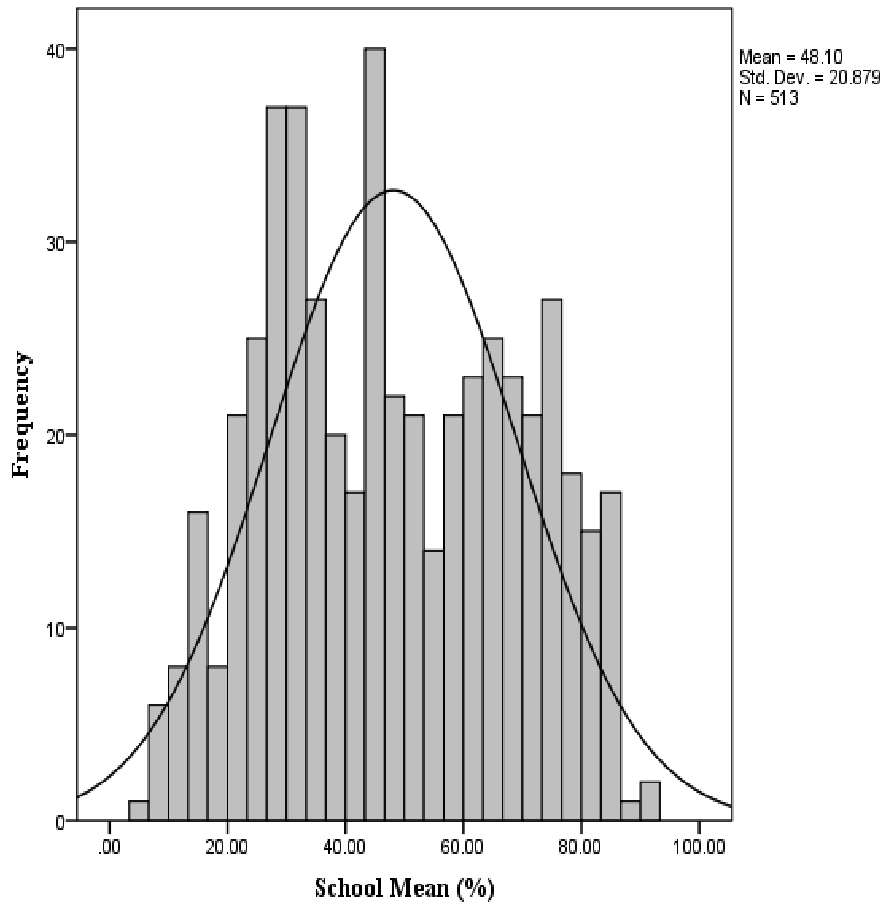


Figure 7.2 shows that the sample schools can be divided into three categories – namely, the high performers, average performers and low performers – based on their mean achievement scores in Grade 5 English.

Figure 7.2 Distribution of schools' mean achievement scores in English



In figure 7.2, the maximum average score of schools on the left-hand side is about 40%, the average score in central part is about 20%, and the minimum average score on the right hand side is about 30%. It indicates a remarkable difference in achievement scores among the different performing levels of schools.

Student Achievement in Various Content Areas

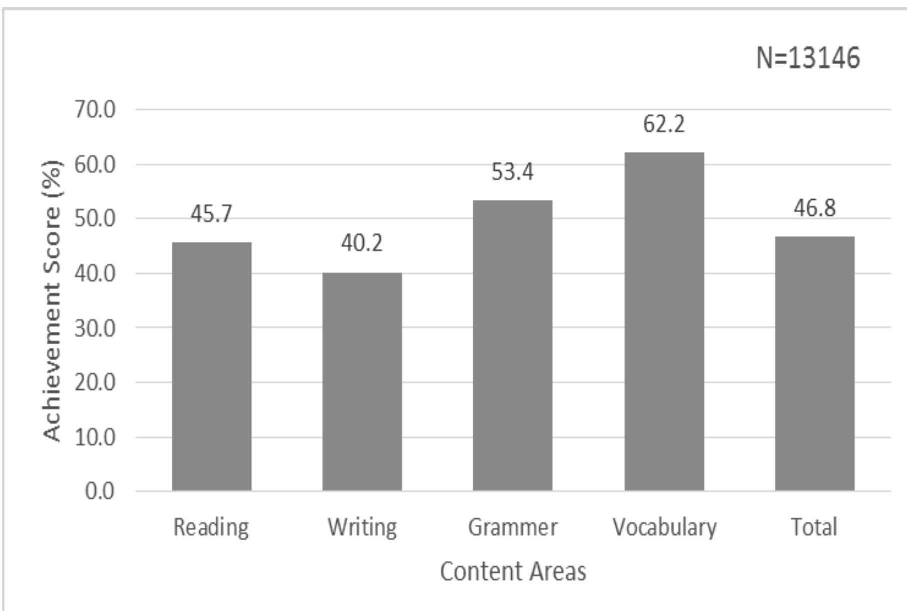
The assessment in English covers four content areas, namely, 1) *Reading*, 2) *Writing*, 3) *Grammar*, and 4) *Vocabulary*. The weightage in these four content areas of assessment were made proportional to the weightage allocated in the curriculum. The achievement scores were converted into percentage for comparability. Table 7.1 shows the students' achievement in each of the four content areas and figure 7.3 further depicts it.

Table 7.1 Achievement in various content areas in English language

Content Areas	Mean	Std.Deviation	Minimum	Maximum
Reading	45.7	24.79	0	100
Writing	40.2	33.32	0	100
Grammar	53.4	26.42	0	100
Vocabulary	62.2	37.76	0	100
Total	**46.8**	**24.46**	**0**	**100**

The table shows high variations of achievements in various content areas of the English language. The achievement ranges from 40.2 percent in W*riting* to 62.2 percent in *Vocabulary* with 22 percent variation. But there is no difference in minimum achievement score and maximum achievement score in any of the four content areas. Figure 7.3 further illustrates the variation of achievement in various content areas.

Figure 7.3 Comparison of achievement scores in various content areas in English

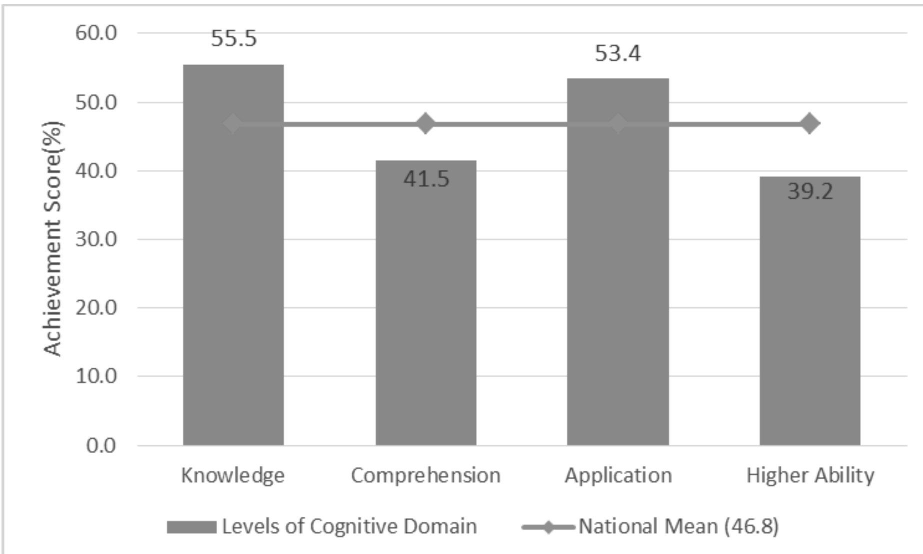


The overall national average achievement score for English language is 46.8%. The students are weakest in W*riting* (40.2%), followed by *Reading* (45.6%) as both the scores are lower than the national average (46.8%). However, the students achieved highest in V*ocabulary* (62.2%) followed by G*rammar* (53.4%), as both the achievements are higher than the national average. Dataset indicates that the students are performing low in R*eading* nd W*riting* in comparison to V*ocabulary* and G*rammar*.

Achievement in Various Levels of Cognitive Domain

The English language test items are categorized into different levels of cognitive domain according to Bloom's taxonomy (Bloom *et al.,* 1956; Metfesser, Michael & Kirsner, 1969). The categories include K*nowledge, Comprehension, Application* and H*igher ability* (reasoning/problem solving/ creativity). The achievement of students in each level is shown in figure 7.4.

Figure 7.4 Comparison of achievement scores in various levels of cognitive domain in English



As illustrated in the figure, the achievement scores in *comprehension* and H*igher ability* levels are below the national average, whereas the scores in K*nowledge* and the application levels are above the national average. It shows that achievement is better in the levels of K*nowledge* and A*pplication* than in C*omprehension* and the higher ability level. The dataset shows that students' ability to comprehend and perform with higher ability is quite low compared to *Knowledge* and A*pplication*; that is, students are much better in the recall type of questions in comparison to the achievement in analytical questions.

Student Achievement by Type of Items

The test items were of mainly two types: objective and subjective. *Objective* items covered a wide range of content areas with weightage of one mark each item, whereby only one correct answer or one explicit piece of information was required to get the correct answer. However, *Subjective* items require a longer description to get full marks. Both types of items were based on various levels of cognitive domain with various difficulty levels. Table 7.2 presents the basic statistics on item type and achievement.

Table 7.2 Achievement by type of items in the English language

Item Type	Mean	Std. Deviation	Minimum	Maximum
Objective	53.7	23.36	0.00	100.00
Subjective	39.7	29.90	0.00	100.00

The mean achievement score in S*ubjective* items is much lower (39.7%) than that in *Objective* items (53.7%). Most of the *Objective* items were of knowledge level and some were of *Application* and C*omprehension* type, whereas *Subjective* items mostly belonged to higher ability level with few of them at C*omprehension* and *Application* levels.

The dataset clarifies that the students are performing well in recognizing the correct answer, recalling simple facts from the texts, and applying the knowledge in grammatical forms. They are weak in comprehending the basic information, writing the texts and in logical reasoning. In many cases, students attempted the open-ended task like free writing, problem solving and analysis; but the skills were not high enough to obtain highest marks.

Comparison of the results of NASA 2015 with NASA 2012

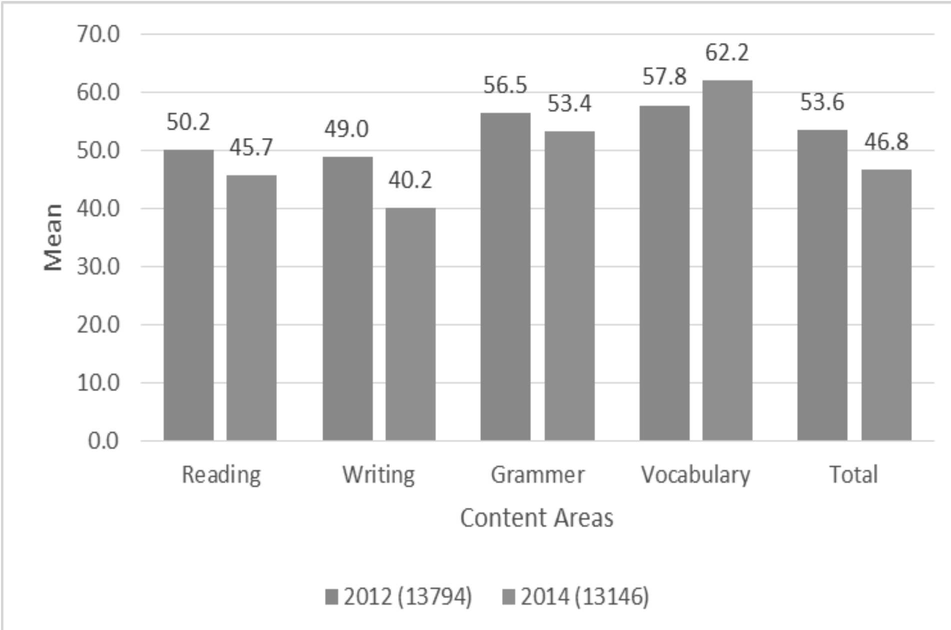
The results of NASA 2015 and NASA 2012 are as presented in table 7.3. These two assessments used different tools and samples, thus might not be fully comparable. However, if we compare average achievement scores of these two years, it is 7 percent lower in the year of 2015 than in the earlier one.

Table 7.3 Comparison of achievements in 2012 and 2015 in English

Year	N	Mean	Std. Error of Mean	Std. Deviation	Minimum	Maximum
2012	13794	53.6	0.21	24.24	0.00	97.47
2015	13146	46.8	0.21	24.46	0.00	100.00

Furthermore, the results in English by content areas of NASA 2015 and NASA 2012 are illustrated in figure 7.5.

Figure 7.5 Comparison of results of NASA 2012 and 2015 in various content areas in English



The achievement level in NASA 2015 is slightly lower than that of NASA 2012. Except in the case of *Vocabulary*, achievement in other content areas in NASA 2015 is lower than that of NASA 2012. The highest gap between these two assessments score is found in *Writing* as the score of NASA 2012 is about 9 percent higher than NASA 2015.

7.3 Achievement Scores by Diversity Factors

Nepal is a much diversified country and therefore diversity is a major factor in achievement variation. NASA 2015 covered six diversity factors - namely Ecological and Development Region, District, language, gender, ethnicity and socio economic status. The background information questionnaire covered the information related to these diversities. Furthermore, three additional comparisons were made by Districts, school type and school location. These comparisons are also considered to assess the equity status.

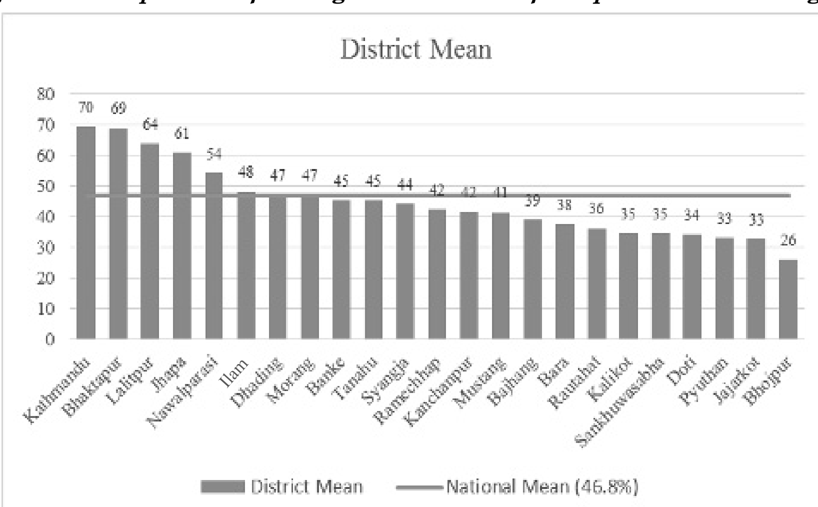
Student Achievement by Districts

The selection process of Districts is described in detail in the methodology chapter. Out of 75 Districts, 23 Districts were selected to represent the Ecological Zones and Development Regions. The achievement of students by District is presented in table 7.4 and further illustrated in figure 7.6. In the figure, the Districts are organized and presented in the descending order by the mean achievement score of the District. The mean represents average percentage of achievement score of the particular District.

Table 7.4 Average achievement scores of sample Districts in English

District Name	N	Mean	Std. Deviation	District Name	N	Mean	Std. Deviation
Illam	435	48.1	27.42	Tanahu	893	45.3	21.94
Jhapa	740	60.8	22.06	Syangja	673	44.1	23.15
Morang	820	47.1	24.08	Mustang	52	41.5	14.02
Sankhuwasabha	381	34.6	18.33	Nawalparasi	734	54.2	22.93
Bhojpur	454	26.3	16.78	Pyuthan	532	33.1	17.62
Ramechhap	447	42.4	21.25	Banke	674	45.3	25.32
Lalitpur	460	63.8	19.86	Jajarkot	461	32.7	20.18
Bhaktapur	273	68.9	17.85	Kalikot	270	34.7	16.90
Kathmandu	1394	69.5	18.52	Bajhang	490	39.0	22.66
Dhading	670	47.3	20.29	Doti	434	34.3	17.80
Rautahat	580	36.3	23.98	Kanchanpur	476	41.6	20.58
Bara	804	37.6	20.85	**Total**	**13146**	**46.8**	**24.46**

Figure 7.6 Comparison of average achievement of sample Districts in English



The figure shows that the achievement score is very low in Bhojpur (26.3%) followed by Jajarkot (32.7%), and Pyuthan (33.1%). These low achieving Districts are from the Hill, whereby Bhojpur is from the Eastern Development region and Jajarkot and Pyuthan are from Mid-Western Development Region. The best performing three Districts are Kathmandu (70%), Bhaktapur (69%) and Lalitpur (64%) from the Kathmandu Valley. The difference is very wide (about 44%) between the low performing (Bhojpur) and high performing (Kathmandu) Districts in English.

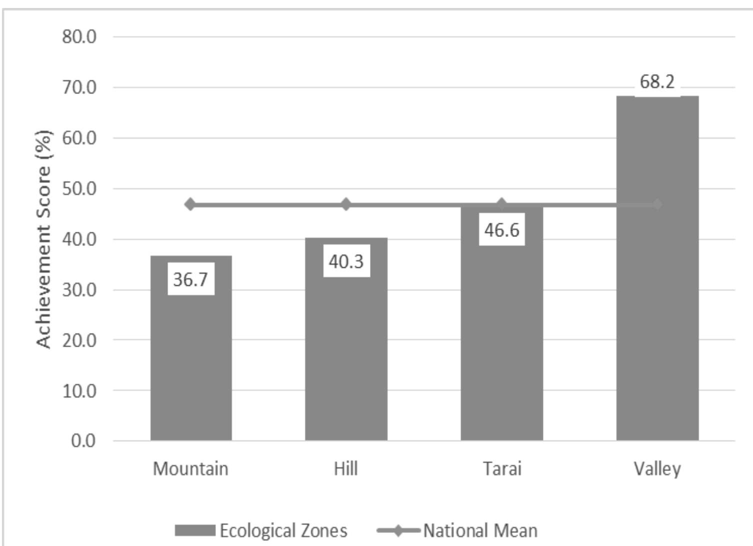
Furthermore, among 23 sample Districts, 15 Districts have achieved scores less than national average.

The difference in achievement scores between low and high performing Districts is statistically significant ($p < 0.001$). District explains 24% of the variation in achievement ($\eta^2 = 0.243$). Effect size is found high ($f = 0.57$). It clearly indicates that the difference between the lowest performing and the highest performing District is remarkable. The dataset indicates that there is a wide difference in achievement in English among the Districts.

Student Achievement by Ecological Zone

In addition to three different Ecological Zones in Nepal−Mountain, Hill and Tarai; Kathmandu Valley is also taken as a special Ecological Zone because of its unique character with most densely populated area, mixed ethnicities and socio-economic status. The variation in achievement across various Ecological Zones is illustrated in figure 7.7.

Figure 7.7 Comparison of achievement of various Ecological Zones in English



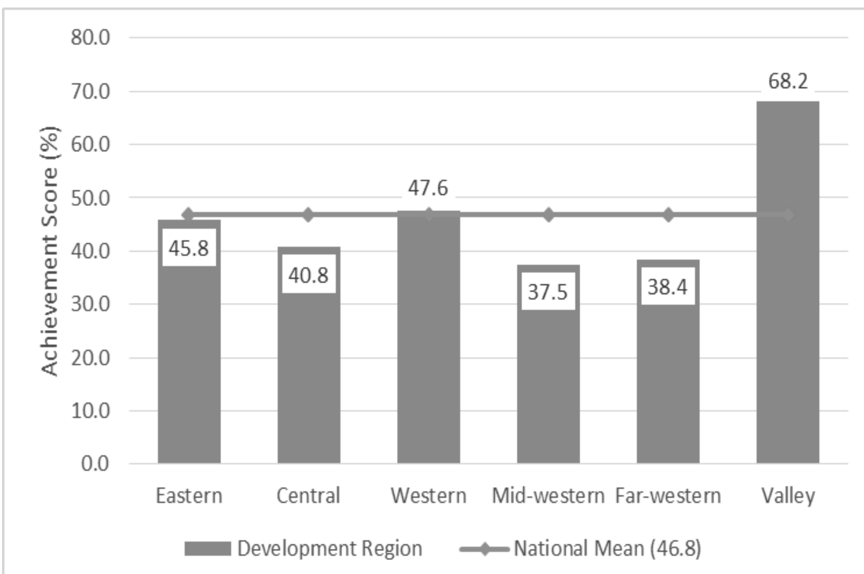
The figure shows that, in general the students from the Kathmandu Valley (68.2%) outperformed the students from the other Ecological Zones, and the students from the Mountain area performed the lowest (36.7%) followed by Hill (40.3%). The achievements in all the Districts from Mountain are lower than the national average, for example, the achievements of Mustang (41.2%), Bajhang (39.0%), Sankhuwasabha (34.6%) and Kalikot (34.7%) are lower than the national average.

The achievement across the Ecological Zones differs significantly ($p < 0.001$) as the Tukey's *post hoc* test explains that all the zones deviate from each other in a statistically significant manner at $p < 0.001$ level. The effect size ($f = 0.45$) shows wide difference between the highest and lowest performing Ecological Zones. Ecological Zone explains 17% of the variation in the data. Dataset indicates moderate differences across the Ecological Zones.

Student Achievement by Development Regions

Five Development Regions have different characteristics. Though administratively the Kathmandu Valley falls under the Central Development region, it is taken as an additional region due to its unique features. Therefore, this analysis is based on six regions: five Development Regions and the Kathmandu Valley. The mean achievement of each Development Region is illustrated in figure 7.8.

Figure 7.8 Comparison of the achievement of various Development Regions in English

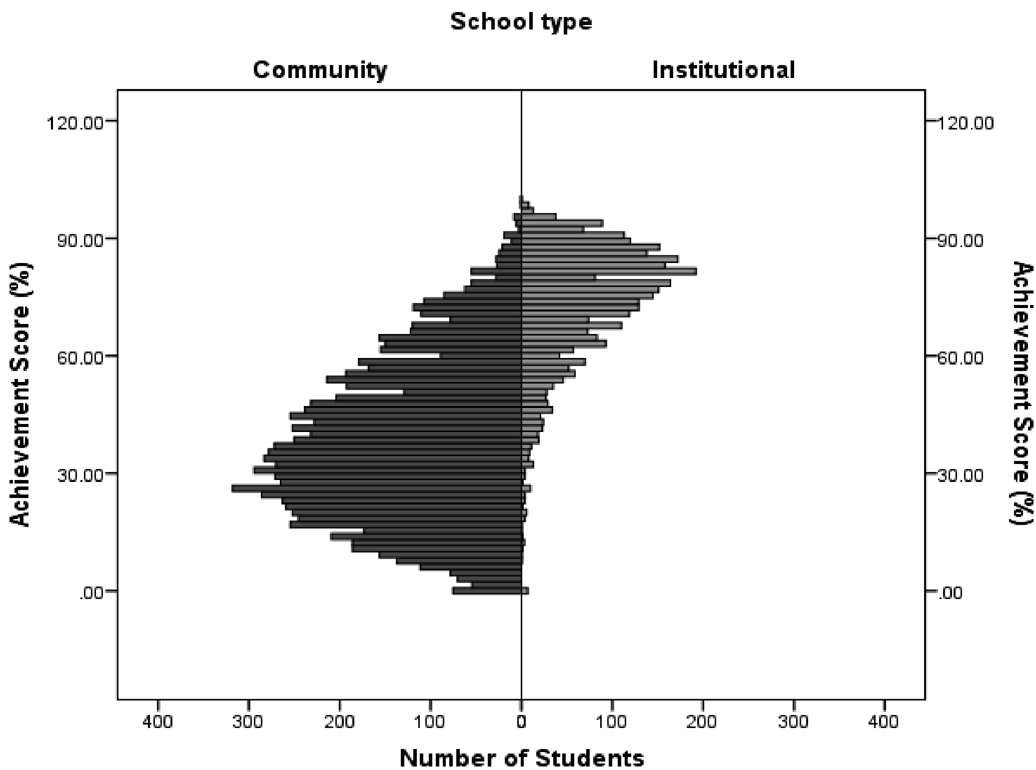


The figure shows that the Kathmandu Valley has the highest level of achievement (68.2%), which is 20.6 percent higher than the region having the nearest lower achievement level, that is, the Western region (47.6%). The achievement is the lowest in the Mid Western Region (37.5%) followed by Far Western (38.4%) and Central (40.8%) Regions. Western and Eastern Regions are close to the national mean. The difference between the Regions is statistically significant ($p < 0.001$) as Tukey's *post hoc* test shows that the difference among Development Regions is significant. Development Regions show 17% variation in data and its effect size is high ($f = 0.45$). The dataset indicates that the variation in performance level among the Development Regions is significant.

Student Achievement by School Type

There are two types of schools in Nepal based on ownership and management. They are: community and institutional. Community schools are generally funded by the government and managed by the community, and they are also known as public schools. On the other hand, institutional schools are generally funded by a group or an individual and managed by a company or trust, and they are known as private schools. Overall distribution of results in community and institutional schools are illustrated in figure 7.9.

Figure 7.9 Distribution of the students' scores in community and institutional schools in English



In the above figure, the distribution of marks obtained by the students of community schools is on left-hand side and by the students of institutional schools on the right-hand side. There are a number of students in community schools getting equally high marks as in institutional schools, but the number of students is small. The figure explains that the students from both types of schools vary from low performers to the high performers. However, the students from community schools are concentrated in low performing groups, while those from the institutional schools are concentrated more on relatively high performing groups.

Furthermore, the difference between the achievement of students from community and institutional schools are presented in table 7.5.

Table 7.5 Student achievement by type of school in the English language

School Type	N	Mean	Std. Deviation	Minimum	Maximum	Std. Error of Mean
Community	9850	38.0	20.17	0.00	100	0.203
Institutional	3296	73.1	15.68	0.00	100	0.273
Total	**13146**	**46.8**	**24.46**	**0.00**	**100**	**0.213**

The achievement gap between the community and institutional schools is remarkable. The average achievement score for institutional schools is 73.1% whereas for community schools it is only 38.0% with a wide difference at 35.1 percent points. The difference is statistically significant ($p < 0.001$) and the effect size is high ($f = 0.80$), showing a wide difference between the community and institutional schools. Category of the students at the community and institutional schools explains 39% of the variation in student achievement ($\eta^2 = 0.39$). The dataset indicates that generally the students in institutional schools have outperformed those in community schools.

Student Achievement by Location of Schools

Based on the location, schools were categorised into rural and urban schools. The achievement of students in rural and urban category of schools is presented in table 7.6.

Table 7.6 Location of school and student achievement in English

School Location	N	Mean	Std.Deviation	Minimum	Maximum	Std. Error of Mean
Urban	3475	66.0	21.52	0	100	.365
Rural	9671	39.9	21.59	0	100	.220
Total	**13146**	**46.8**	**24.46**	**0**	**100**	**.213**

The table shows that the achievement level of the students from urban schools (66.0%) is remarkably higher than that from rural schools (39.9%). There might be different causes of such differences, but someone may explain that the achievement level of urban schools is raised due to the performance level of institutional schools and the schools from the Kathmandu Valley. The difference in average score is significant ($p < 0.001$) and the effect size is high ($f = 0.53$), and the location of schools explains 22% of variation in student achievement ($\eta^2 = 0.22$). The data indicates that the students of urban schools obtained higher scores than the students of rural areas.

Language at Home and Student Achievement

The dataset of NASA 2015 in English shows that 32.7% of the 5th Graders speak a language other than Nepali as their first language. These “other” languages are quite fragmented; the largest groups in the student dataset are Magar (3.9%), followed by Tharu (2, 8%), Tamang (2.8%) and Newari (2.0%). After categorising the languages into 10 groups (excluding Nepali), the remaining 19.8% students were classified into the group “Others”. For the convenience of statistical analysis, all the other language speakers were grouped into “Non-Nepali”. The results are presented in tables 7.7 and 7.8 and further illustrated in figure 7.10.

Table 7.7 Home Language and Student Achievement in English

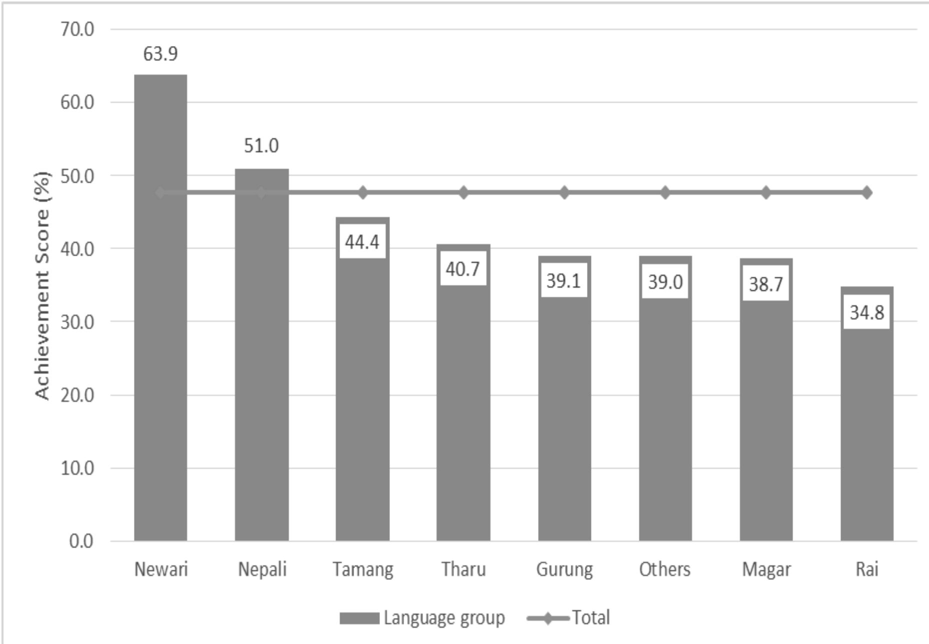
Home Language	N	Mean	Minimum	Maximum	Std. Deviation	Std. Error of Mean
Nepali	8301	51.0	0	100	24.64	0.27
Non-Nepali	4084	41.0	0	97	22.15	0.35
Total	**12385**	**47.7**	**0**	**100**	**24.30**	**0.22**

The table shows that there is a notable difference between the Nepali and other language groups mentioned as “Non-Nepali”. However, when we mentioned all the minor languages as in table 7.8, it is found that the students from Newari (63.9%) speaking have performed higher compared to the students from Nepali speaking (51.0%). On the other hand, the students from Rai (34.8%), Magar (38.7%), “Others” (39.0%), Gurung (39.1%), Tharu (40.7%) and Tamang (44.4%) speaking performed much lower in comparison to the students from Nepali speaking communities.

Table 7.8 Achievement of students from various language groups in English

Home Language	N	Mean	Minimum	Maximum	Std. Devation	Std. Error of Mean
Newari	251	63.9	0.00	100.00	21.72	1.37
Nepali	8301	51.0	0.00	93.85	24.64	0.27
Tamang	342	44.4	0.00	87.69	20.48	1.11
Tharu	346	40.7	0.00	89.23	19.33	1.04
Gurung	120	39.1	0.00	96.92	17.79	1.62
Others	2454	39.0	3.08	84.62	22.03	0.44
Magar	484	38.7	9.23	87.69	20.03	0.91
Rai	87	34.8	0.00	95.38	21.18	2.27
Total	12385	47.7	0.00	100.00	24.30	0.22

Figure 7.10 Comparison of the achievement of students from various language groups in English



The difference in achievements among the language groups is statistically significant ($p < 0.001$); but the effect size is medium ($f = 0.25$). The minority language groups are small and the division into smaller language groups explains about 6% of the variation in the data ($\eta^2 = 0.059$). One of the notable results from the above figure is that in English subject the students from Newari speaking groups outperformed the students from Nepali speaking communities. However, the results are different in other language groups as their performance is lower than that of the students having Nepali as home language.

Ethnicity and Student Achievement

There are ethnicity based variations in participation and achievement of education in Nepal. Historically, the Brahmans and Chhetris have higher level of educational attainment. Because of the various booster initiatives carried out by the government to promote equity, the National Population Census 2011 shows that the enrolment of Hill Dalits has increased remarkably at the lower level of schooling but their number at the secondary and higher education is still very small (CBS, 2012). Although, in this test, Dalit students were not categorised as Hill Dalits and Tarai Dalits, it does not show improvement in their achievements. Table 7.9 shows the lower performance of Dalit students and it shows that there is a gap in the achievement of various ethnic communities. Table 7.9 explores the relation between ethnicity and student achievement.

Table 7.9 Student achievement by ethnicity in English

Caste	N	Mean	Minimum	Maximum	Std. Deviation	Std. Error of Mean
Brahman	1727	58.9	0.0	100.0	24.39	0.59
Others	495	49.7	0.0	96.9	26.65	1.20
Chhetri	2816	48.4	0.0	98.5	23.82	0.45
Janajati	4670	47.1	0.0	100.0	23.64	0.35
Minority	68	42.2	0.0	93.8	25.42	3.08
Madhesi	1007	41.4	0.0	95.4	23.89	0.75
Dalit	1588	39.5	0.0	98.5	20.90	0.52
Total	12371	47.7	0.0	100.0	24.23	0.22

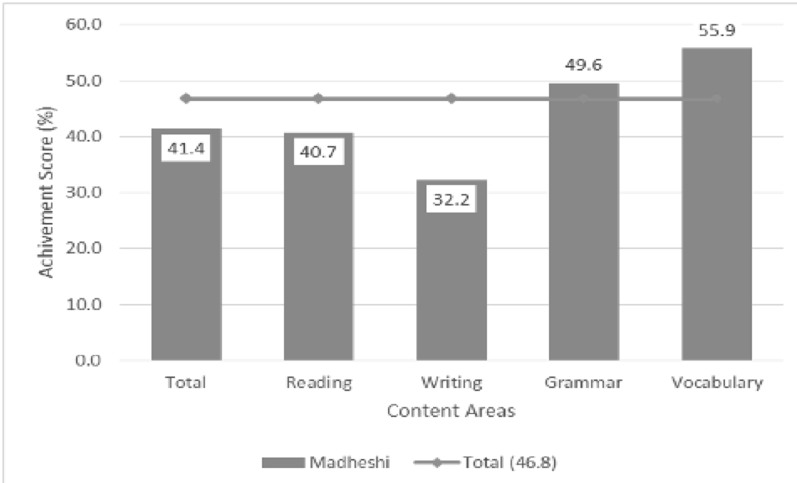
The above data shows that Dalit students are performing the lowest (39.5%) in English. The overall difference among the groups is statistically significant ($p < 0.001$) and the effect size is medium ($f = 0.232$). Caste/ethnic background explains 5% variation in achievement ($\eta^2 = 0.051$). The result shows the lowest achievement status of Dalits (39.5%) followed by Madhesi (41.4%) and Ethnic minority (42.2%), which are below the national average (46.8%).

Table 7.10 shows the status of Dalit students' achievement scores in different Geographical and Ecological Zones. The status of Dalits is better in Kathmandu Valley (57.1%) and poorest in the Eastern Mountain (27.9%).

Table 5.10 Dalit students' achievement in various Ecological and Development Regions

Development Region	Mountain	Hill	Tarai	Valley	Total
Eastern	27.9	31.1	46.2		40.0
Central		44.1	35.1		40.0
Western	36.5	40.1	47.3		41.3
Mid-western	31.4	33.3	42.6		35.2
Far-western	36.8	33.0	37.6		35.3
Valley				57.1	57.1
Total	**33.2**	**37.8**	**41.9**	**57.1**	**39.5**

Figure 7.11 Madhesi students' achievement in various content areas in English

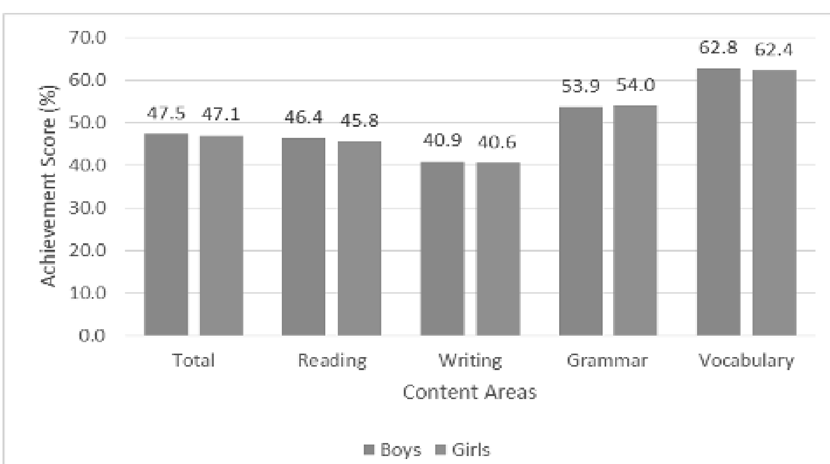


The figure 7.11 shows that the achievement score of Madhesi students is in the same pattern as national achievement, as the lowest score is in W*riting* and the highest in V*ocabulary*. The average scores in G*rammar* and *Vocabulary* are higher than the national mean. The dataset indicates that the Madhesi students' performance in English is lower in R*eading* and W*riting* in comparison to G*rammar* and V*ocabulary*.

Achievement by Gender

Basic achievement results of boys and girls are presented in figure 7.12. This figure further compares also the achievements of girls and boys in various content areas, which shows that there is no remarkable difference between boys and girls in various content areas of the language.

Figure 7.12 Comparison of achievement of boys and girls in different content areas of English

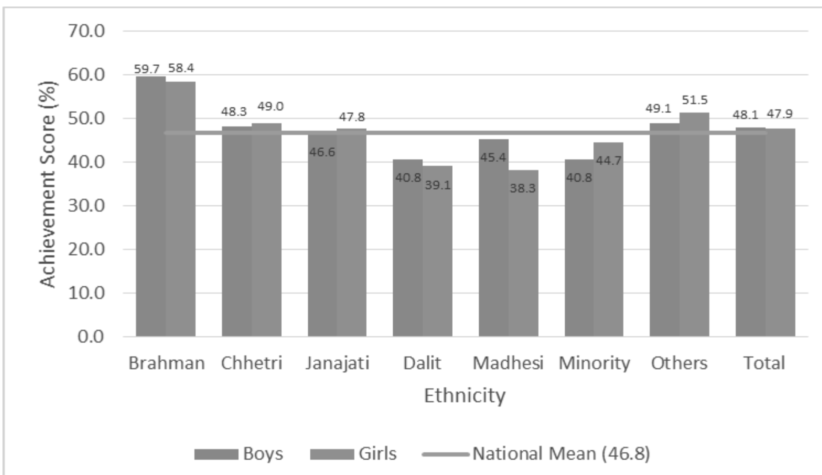


There is no statistically significant difference in the achievement between boys (47.5%) and girls (47.1%) ($p > 0.05$). Boys are performing better in comparison to girls but the difference is not found statistically significant ($p = 0.001$). In Grammar, girls performed slightly better than boys, however, the difference is insignificant ($p < 0.001$).

Gender, Ethnicity and Achievement

The difference in achievement between boys and girls is the highest among Madhesis (difference is about 7 %) where boys have outperformed girls. The differences are statistically significant at $p < 0.001$ and the effect size is medium ($f = 0.227$). Gender in various ethnic background explains 5% variation in achievement ($\eta^2 = 0.049$).

Figure 7.13 Comparison of boys and girls' achievement across ethnicities in English

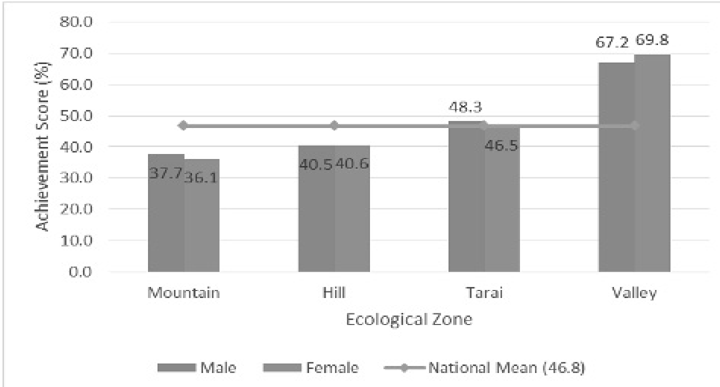


Gender disparity in achievement is widest (6.9%) in Madhesi community. In the Madhesi community, girls' performance is low (38.3%). Girls' performance is better in Chhetri, Janajati, Minority and others communities, whereas boys have performed better in Brahman and Dalit communities.

Gender, Ecological Zone and Achievement

The achievement of girls and boys differs significantly across the Ecological belts. Particularly, in Mountain and Tarai Regions boys have outperformed girls. However, in the Kathmandu Valley girls performed better than boys. In Hill, girls' performance is found slightly better.

Figure 7.14 Comparison of achievement by gender in various Ecological Zones in English

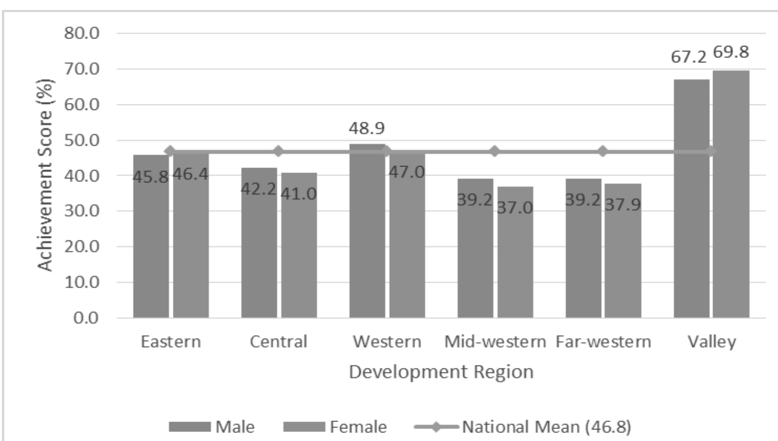


The differences between boys and girls are found statistically significant and the effect size is high (f=0.451). Gender in Ecological Zones explains 17% variation in achievement (η2 = 0.169).

Gender, Development Region and Achievement

There is a noticeable difference among the Development Regions in terms of gender. The girls' performance is found better than that of boys in the Kathmandu Valley and Eastern Region, and in the rest of the regions, boys' performance is better than that of girls.

Figure 7.15 Comparison of achievement of boys and girls in various Development Regions

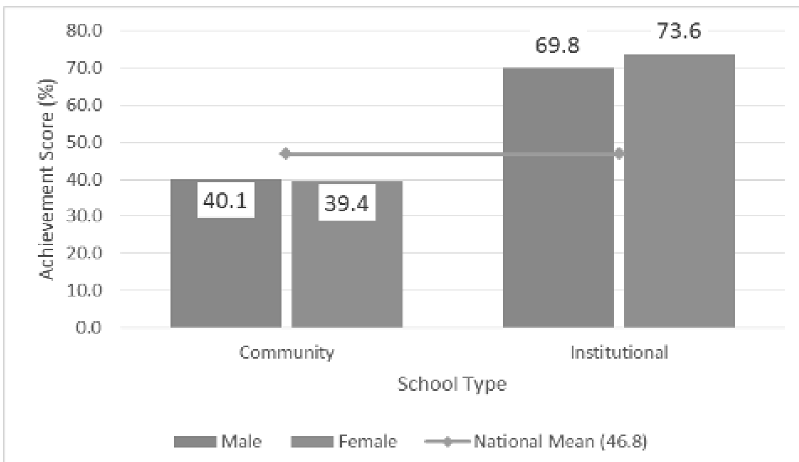


The differences between boys and girls in Ecological Zones is found statistically significant with high effect size (f = 0.451), and gender in Ecological Zones explains 17% variation in achievement (η^2 = 0.169). The dataset indicates that girls' performance is higher in the Kathmandu Valley and Eastern region.

Gender, School Type and Achievement

There is a remarkable difference between the school types in terms of gender. Figure 7.16 illustrates the differences. Girls' performance is better than that of boys in institutional schools (3.8%), whereas boys' performance is slightly higher than that of girls in community schools.There is a significant difference in school type by gender at $p < 0.001$. The effect size is very high ($f = 0.708$) and it explains 33% variations in achievement ($\eta^2 = 0.334$).

Figure 7.16 Comparison of achievement of boys and girls by school types in English



7.4 Selected Explanatory Factors and Achievement

Some basic factors explaining the differences in achievement have already been discussed in section 7.3. This section deals with socio-economic status (SES) of the students' families, paid work after school, students' attitude towards English subject, age of the student, and support provided for the studies as the main family and individual related factors affecting students' achievement in English. As a sample of deepening school and teacher related factors availability of textbooks, assigning homework and providing feedback by the teacher, and selected activities in the school are also considered in this section.

Socio Economic Status (SES) of Parents and Student Achievement

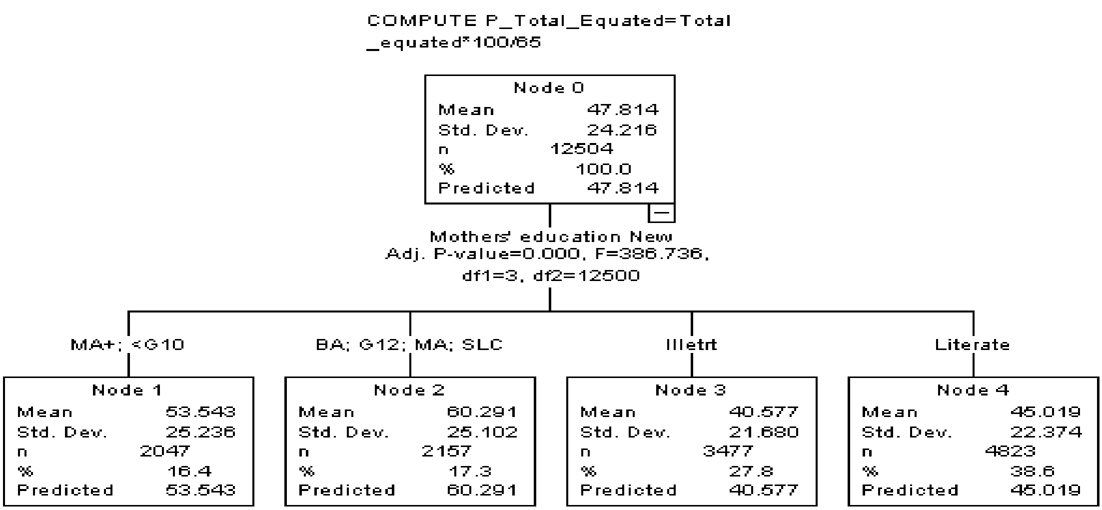
The variables indicating socio economic status (SES) were categorized into parents' education, parents' occupation, home possessions, home accessories, and type of school the student attended. Finally, SES is explained based on seven indicators related to socio-economic, educational, and occupational background of the parents. In this section, parents' education is further elaborated and the literacy status of parents is analysed. Several SES related variables were analysed using a data mining tool of SPSS-the Decision Tree Analysis (DTA). This method is used to find the cut-offs of the predicting variable, i.e., in case of mother's education it classifies the factors into several groups, as it differs statistically in the most significant way from each other in relation to student achievement.

Parents' Education

In this assessment, questions related to parents' educational status were included in the background questionnaire and is divided into eight categories as: 1) Illiterate, 2) Literate, 3) Grade 10, 4) SLC, 5) Grade 12 / Proficiency Certificate Level, 6) Bachelor level, 7) Master level and above.

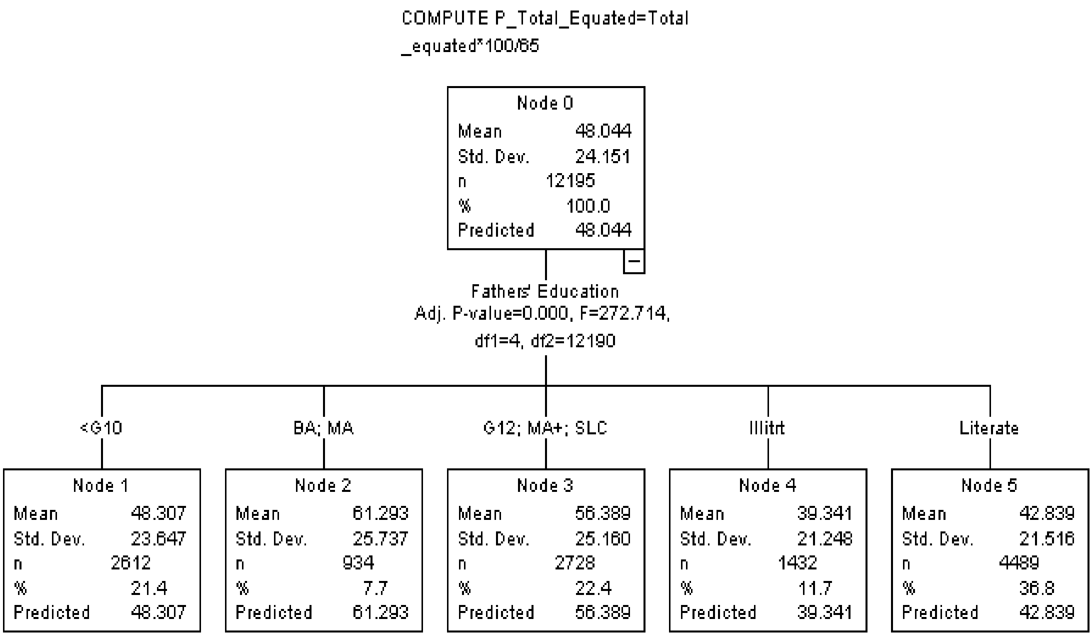
DTA classifies mothers' education into four groups with statistically significant differences in students' achievement levels. Figure 7.17 illustrates these categories with their values as: Illiterate mothers (40.58%), Literate mothers (45.02%), Above Masters and less than Grade 10 mothers (53.54%), and the mothers' having SLC, certificate level/Grade 12, Bachelor's or Master's degree (60.29%). In each group, the number of mothers is high enough to make a credible prediction. The difference between each of the groups is statistically significant ($p < 0.001$).

Figure 7.17 DTA of mother's education and students' achievement in English



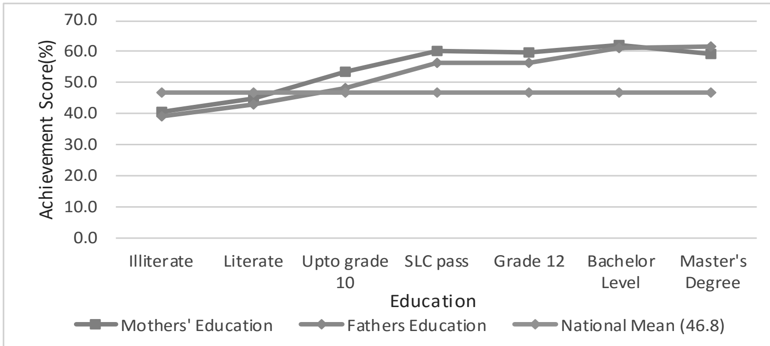
Similarly, as in figure 7.18, DTA divides father's education into similar five categories, as in the case of mother. These categories are: Illiterate (39.34%), Literate (46.84%), up to Grade 10 (48.31%), SLC, Grade 12/PCL (56.38%), Bachelor's and Master's (61.29%). The number of fathers is also high enough to make a credible prediction and the difference between each of the groups is statistically significant ($p < 0.001$). This dataset explores that the students of illiterate father achieved the lowest and the achievement of students increased as fathers' education increased up to Bachelor's degree.

Figure 7.18 DTA of father's education and students' achievement in English



Based on these two DTAs, positive relation is seen between student achievement and parent's education. The following figure (fig 7.19) further illustrates the relationship of student achievement with parental education. Mother's education explains about 11% ($\eta^2 = 0.114$) and father's education explains about 11% ($\eta^2 = 0.108$) of the variation; however, strong relationship is found with the mother's education. The dataset shows that when both parents are illiterate, students achieved very low (about 40%); on the other hand, they have achieved highest (about 62%) when both are having education of Bachelor's degree. The figure below shows that mothers' education can contribute more than fathers' education for improving learning achievement of students. In this way, the dataset indicates that the educational level of parents affects the children's achievement level. Achievement of the students is more influencing if mother is at least SLC passed, or if father has the qualification of Bachelor's level or higher.

Figure 7.19 Relation between parents' education and student achievement in English

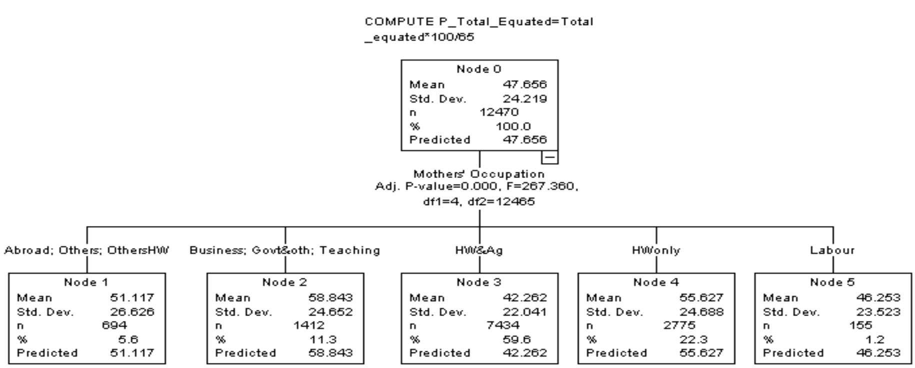


This figure shows that father's and mother's level of education has a significant impact on their children's achievement. Students' performance is below the average national mean if their parents are either illiterate or just literate.

Parents' Occupation

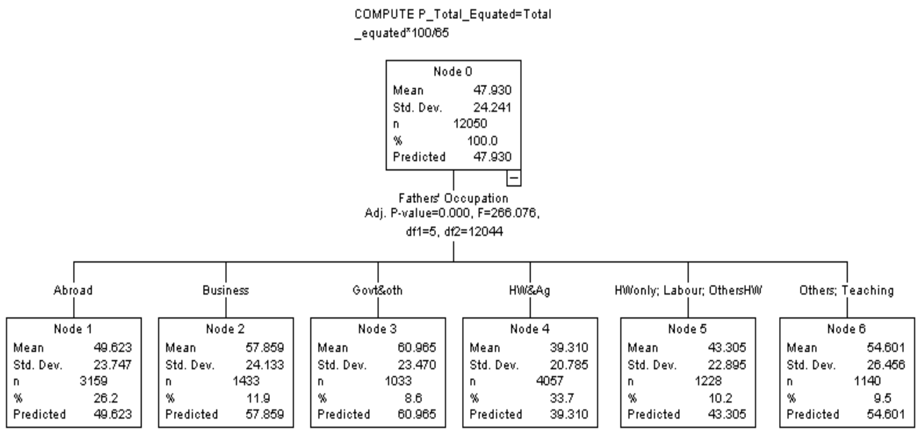
In a similar manner with the parents' education, DTA was used to find out the statistically most deviating groups related to student achievement in terms of parents' occupation. This analysis found four significant nodes. The lowest achievement is in the group where the mother comes from the background of agriculture, daily wages and working abroad. The occupation of parents was categorised into eight groups as: 1) agriculture, 2) teaching, 3) service (government, non-government and private sector job), 4) business, 5) working abroad, 6) work at others' home, 7) daily wages work, and 8) work only at home. While comparing the mean using ANOVA, student achievement is the lowest when the mother's occupational background is agriculture (42.3%). It is statistically significant and lower than the mother who is involved in daily wages (46.3%), working abroad and others (51.1%), house works (55.6%), and business, service and teaching (58.8%). The results are presented in figure 7.20 as classified by DTA.

Figure 7.20 DTA of mother's occupation and students' achievement in English



Noticeable higher achievement is seen when the mother is having government service or teaching jobs (58.8%). However, result indicates that the mothers who are employed in any kind of job or business are seen beneficial for the achievement of students. The difference is found statistically significant (at $p < 0.001$) for each group.

Figure 7.21 DTA of father's occupation and students' achievement in English



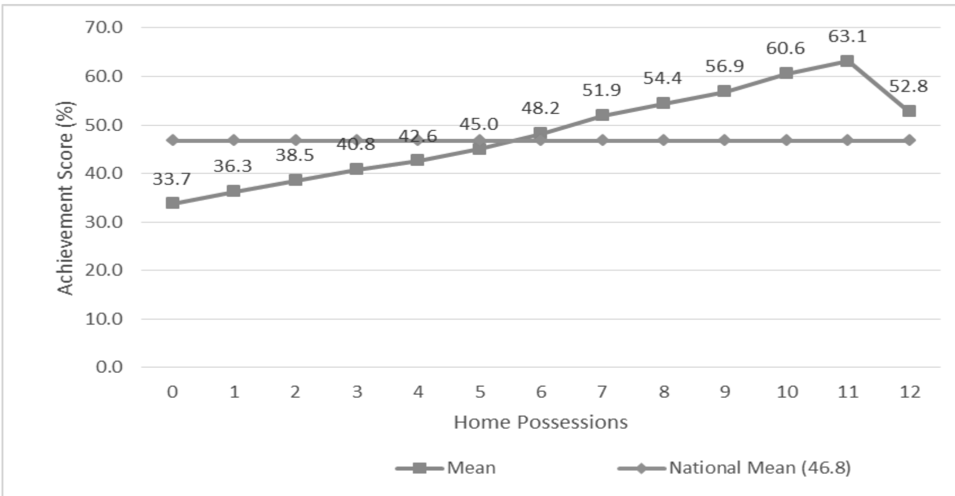
Similarly, in case of father's occupation, the main division is based on whether the father works in agriculture or not. If the father is in a government or other service, the students have performed better (60.1%) than the children of fathers who are working in agriculture (39.3%). Similarly, if the fathers are involved in household works or work on daily wages, students have performed lower (43.3%) than other five groups. These seven groups are created by DTA as the difference between each of the groups is found statistically significant at $p < 0.001$.

Home Possessions and Accessories

Home possessions and accessories are taken into account, considering whether the facilities and resources available at students' home have some effects on the achievement. There were two kinds of home possessions defined in the background information questionnaire for the students. One is related to the facilities that are supportive for the students' study at home; for example, a table for study, a separate room for them, a peaceful place for study, a computer for homework and additional study, software for the computer assisted learning, internet facilities, literary magazines, access to classical literature, poetry books, or artistic things like pictures, dictionary and other books. Other types of home possession include different types of normal home accessories such as the number of mobile phones, televisions, computers and vehicles (cars, motorbikes).

Related to home possessions, there were 12 questions in the background questionnaire for students. Each question was scored 1 if the student had an access to the possession. Adding these items up, the maximum score was 12 indicating that the student had access to all of the possessions and hence lower score means fewer possessions at home. Figure 7.22 shows positive correlation between the home possessions and achievement level, except in the case of the highest category (i.e., all the home possessions). The variation is found about 20% in the achievement level and the factor (r = 0.199) is statistically significant (p < 0.001).

Figure 7.22 Relation between home possessions and achievement in English



The above figure shows positive correlation between the availability of home accessories and achievement score of student. The cut-off for the factors was set at four possessions for calculating the value of home possession. If the students possessed 5 items or more as mentioned in background questionnaire, the student was given 1, otherwise 0. The same pattern – the more accessories, the better results – is found with home accessories (figure 7.22).

The question asked to the students in the background questionnaire about the availability of accessories was: *"How many of the following accessories do you have in your family?"* The availability of home accessories is dichotomized in three groups. After dichotomizing the items individually by using meaningful cut-offs found with ANOVA and DTA (and maximizing the differences in achievement level, see table 7.11), all the three indicators were added together. The maximum score was three indicating that the students possessed a set of all the accessories.

Table 7.11 Dichotomizing the indicators for home accessories

Accessory	Cut-off for 1	Cut-off for 0
Mobile phone	2, 3	0 and 1
Television	1–3	0
Computer	1–3	0
Car and Motorcycle	1–3	0

Figure 7.23 Availability of home accessories and achievement in English

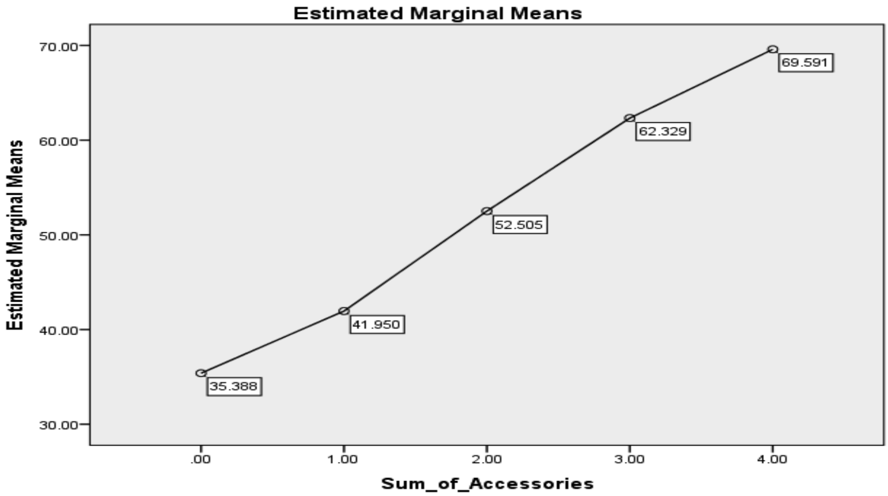


Figure 7.23 demonstrates that when the number of home accessories increases, students' achievement also increases ranging from 35.4% (if none of them are available) to 69.6% (if all of them are available). The results show that there is a positive correlation between the availability of home accessories and student achievement except in the case of availability of vehicles as there is no significant correlation between the availability of car or motorcycle and student achievement. Availability of the stated facilities indicates a high SES of the family. The dataset indicates that either economic or educational ability or both at home support children to raise their educational achievement. If the students come from high SES family (having a large number of home possessions and accessories), their performance is found remarkably higher than the performance of other students.

Socio-economic Status (SES) and Achievement

The value of socio-economic status was calculated based on seven indicators, which were dichotomized first. The variables mother's education, father's education, mother's occupation, father's occupation, home possessions, home accessories and the type of school attended were summed as SES and changed into the percentage of the maximum score (P-SES). Deeper description of the transformations for SES has been given in the methodology chapter. The P-SES represents the percentage of SES of the student's family, in which 100 means that the student has the highest possible SES where all the seven indicators of SES are positive, and 0 refers to the lowest possible SES where all the seven indicators of SES are negative. The analysis of P-SES by using Univariate GLM (the Regression modelling) shows strong relationship between SES and student achievement. Figure 7.24 presents the relationship between SES of the students and their achievement.

Figure 7.24 Relation between the SES and achievement in English

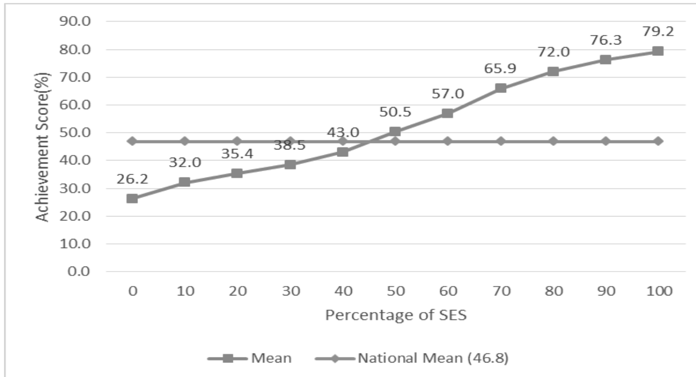
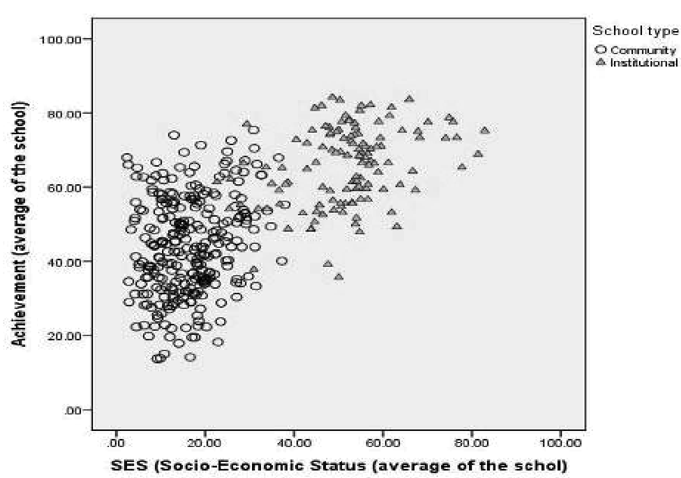


Figure 7.24 shows a positive relationship between SES and achievement, which is statistically significant (p < 0.001) and the effect size is high (f = 0.625). SES explains 28% of the variation in achievement (η^2 = 0.281). The dataset indicates that socio-economic status plays a vital role in student achievement. The difference between the lowest and highest SES groups is remarkable (53 percent). The achievement is even low when the father or both parents are illiterate or both of them work in agriculture.

While plotting the school in terms of mean score and SES in a scatter diagram as in figure 5.25, community and institutional schools (community school in circle and institutional schools in triangle) fall into two distinct groups. Most of the institutional schools, which have high SES value, are among the relatively high performers, whereas even the community schools having low SES value belong to the ranges from the low performing to high performing.

Figure 7.25 Achievement of schools with various levels of SES in English



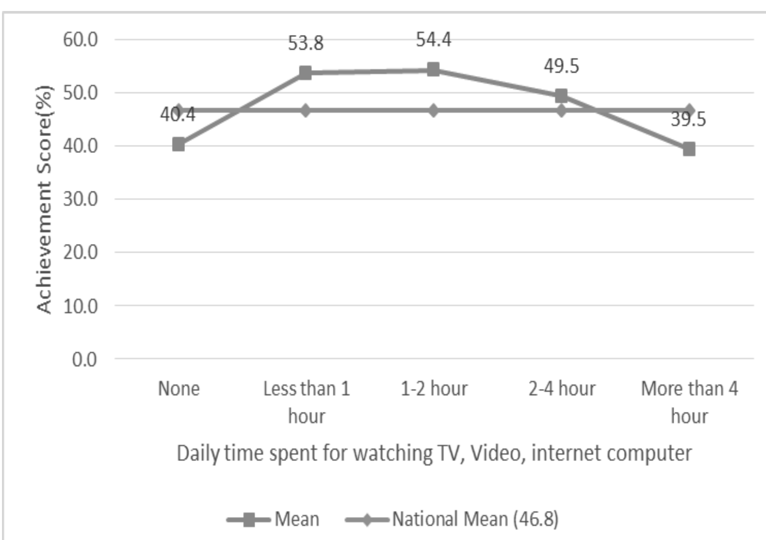
The dataset indicates that most of the students in the institutional schools have high SES, and they have achieved relatively higher score in comparison to the students from community schools. Further, community school students form two distinct groups: high-performing and low-performing schools. Although the variations in achievement among institutional as well as community schools are wide, community schools are concentrated more on the range of 20% to 60% score. Institutional schools are concentrated more on the range of 40 to 80% of score.

Activities Beyond School Hour and Achievement

Seven questions were incorporated in the background questionnaire about the students' activities beyond the school time. The activities included in the analysis were watching TV, playing games, involvement in household chores, studying/doing homework and working in a paid job. The values of the variables are divided into five categories based on the time spent in particular activity: 0 (no at all), 1 (less than 1 hour per day), 2 (1 to 2 hours per day), 3 (2 to 4 hours per day), and 4 (more than 4 hours per day).

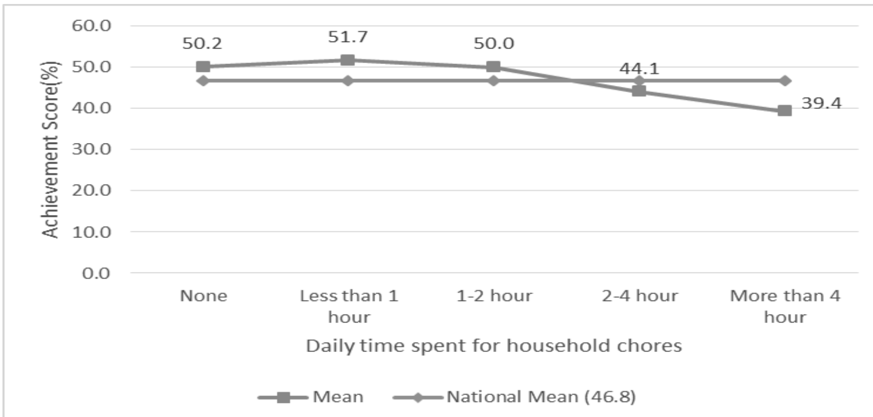
The dataset informs that watching TV 1 to 2 hours a day has positive effect on students' achievement, whereas spending on watching TV more than two hours or not is found to be negatively affecting the students' achievement (see figure 7.26). There is a significant difference in the achievement based on the time spent in watching TV at $p < 0.001$ where the effect size is medium ($f = 0.264$) and explains about 7% variation in achievement ($\eta^2 = 0.065$).

Figure: 7.26 Relationship between watching TV and achievement in English



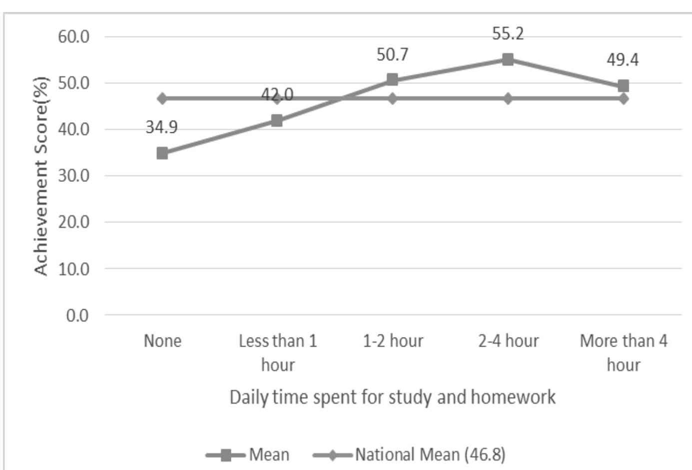
Similarly, less than one hour's involvement in household chores shows higher performance level whereas more than two hours spending on household chores negatively affected the students' achievement (see figure 7.27). There is a significant difference in the time spent on household chores ($p < 0.001$) and the effect size is small ($f = 0.146$) explaining about 2% variations in achievement ($\eta^2 = 0.021$).

Figure 7.27 Relationship between household chores and achievement in English



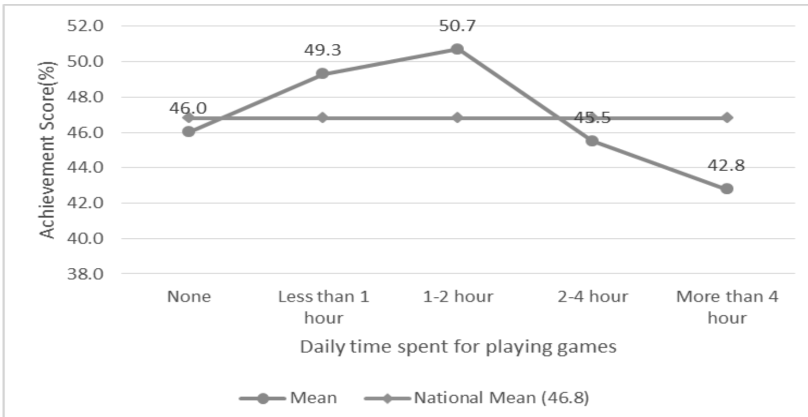
Two to four hours' involvement in study and homework at home shows higher performance level, while spending less than one hour on study and homework negatively affected the students' achievement (see figure 7.28). Similarly, time spent for more than four hours per day at home in studying and doing homework reduces the students' achievement. There is a significant difference in the time spent to study and homework at $p < 0.001$, the effect size is medium ($f = 0.241$) explaining about 5% variation in achievement ($\eta^2 = 0.055$).

Figure 7.28 Relationship between study/doing homework at home and achievement in English



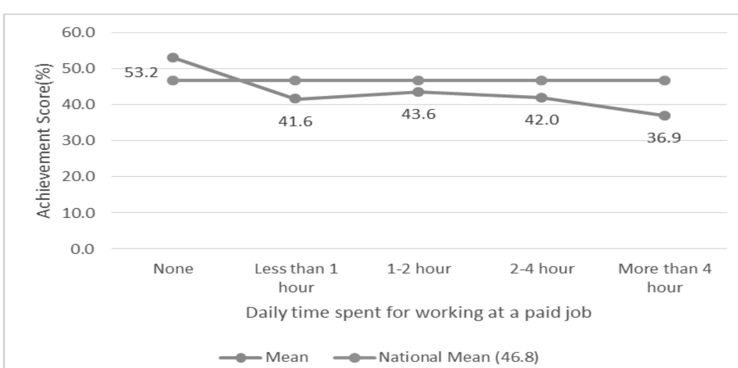
Regarding the time spent for playing game, one to two hours' involvement in playing game shows higher performance level (50.7%) followed by less than one hour (49.3%). But students' playing for more than two hours is seen to be negatively affecting their achievement (see figure 7.29). There is a significant difference in the achievement based on the time spent in playing games at p < 0.001, however, the effect size is very small (f = 0.071) explaining about 0.5% variation in achievement (η^2 = 0.005).

Figure 7.29 Relationship between playing games and achievement in English



The involvement of students in paid jobs is found to be affecting negatively in their achievement scores, whether their involvement is less than one hour or more than four hours a day. The students who are not involvolved in paid job achieved 53.2% score, whereas those working for paid job achieved lower score than the national mean (see figure 7.30). There is a significant difference in achievement based on the working for paid job at $p < 0.001$ where the effect size is medium (f = 0.232) explaining about 5% variations in achievement (η^2 = 0.051).

Figure 7.30 Relationship between paid job and achievement in English

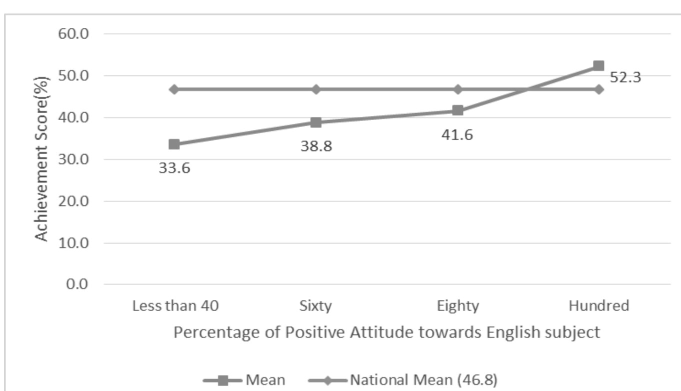


The above figure shows that there is a positive relationship between the activities beyond the school time and student achievement. There is a positive correlation between the activities like playing games, watching TV, doing homework/studying and involving in household chores in a limited time whereas working in any paid job for any duration has negative effects on student achievement.

Students' Attitude towards the Subject of Study and Achievement

The attitude of students towards English language has been explained in terms of what the students think about English subject and its usefulness in their daily life and future. There is some association between the attitude of the students towards the subject and their achievement. The correlation between achievement and attitude towards the subject is widely studied, though the connection is not always clear, (see, for example Metsämuuronen 2012a; 2012b; House & Telese, 2008; Shen & Tam, 2008; Kadijevich, 2006; 2008). In NASA 2013, the same shortened version of Fennema–Sherman Attitude Scales (FSAS, Fennema & Sherman, 1976), as used in several international comparisons like TIMSS and PISA studies are used. The original scale contains nine dimensions but in these international comparisons, only three were used with four items in each dimension and two negative items in each of the first two dimensions (see the detail in the methodology chapter). The names of the factors can be "Liking English", "Self-Efficacy in English", and "Experiencing Utility in English" (compare naming in, e.g., Kadijevich, 2006; 2008). Factor analysis was used to identify the factors of the responses in FSAS and the negative items were reversed to make the whole test unidirectional. As in several countries of Asia, the expected factor structure cannot be found in Nepal (see Metsämuuronen, 2012a; 2012b). Hence, only the total score is used to show the connection between attitude and achievement. The relation between attitudes, as divided into ten groups with somehow an equal number of students, that is, deciles, and the achievement scores are shown in figure 7.31.

Figure 7.31 Relation between students' attitude and achievement in English



There is a positive correlation between the students' attitude and their achievement (r = 0.275). The connection is moderately high (f = 0.295); the division of attitude into ten groups explains about 8% ($\eta^2 = 0.08$) variation in achievement. The difference between the lowest and highest attitude groups is about 19 percent. The connection is clear though it is not known whether the positive attitude is a consequence of high achievement or the other way round. Data indicates that positive attitude towards English subject is positively correlated with higher achievement.

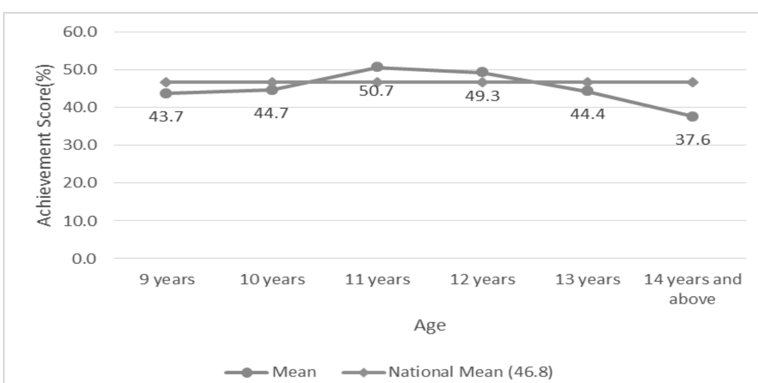
Age of the Students and their Achievement

The age of the students attending Grade 5 varies widely, although the official age for Grade five students is 10 years; and in this assessment period the students' right age was 11 years as the assessment was conducted when the students just completed the 5th Grade. Some of the students have mentioned their age below 10 years and some above 14, but the aged below 9 were encoded as '9 years', and all those having 14 and above 14 were encoded as '14 years and above'. The descriptive statistics of the mean by each age year are presented in table 7.12 and depicted in figure 7.32.

Table 7.12 Students' achievement in different age groups in English

Student's Age	N	Mean	Std. Deviation
9 years	171	43.7	24.48494
10 years	1257	44.7	23.93188
11 years	3451	50.7	25.18619
12 years	4398	49.3	24.32532
13 years	2029	44.4	23.11476
14 years and above	1098	37.6	20.55804
Total	**12404**	**47.3**	**24.33800**

Figure 7.32 Relation between the age of the students and achievement in English



Those students who are at the proper aged for Grade 5 (i.e., 11 years old) have scored 50.7% and the ones who are one year older than the proper age scored 49.3%. The students studying at higher age means, either they have started their study much later, or they have repeated the class.

The achievement level is remarkably lower than the average when the students are 13 years or higher in age. Correlation between the variables is negative (i.e., - 0.089 at $p < 0.001$) indicating small effect size ($f = 0.16$). The ANOVA hints that the age explains 3% ($\eta^2 = 0.026$) of variation in results. The difference in achievement between proper age and over age students is up to 13.6 percent. Dataset indicates that the students studying at proper age have obtained the highest achievement score, and the achievement goes gradually down as the age increases.

Support to the Students for Study and their Achievement

The relation between the supports received by the students for their study and their achievement was analysed based on the information gathered through the question: *"Who supports you when you do not understand what you have read?".* In this question, there were six options and students were instructed to choose only one option. About four hundred students chose more than one option. Thus, none, father, mother, siblings, tuition and more than one supporter are taken as a separate category. The descriptive statistics of the supporters are given in table 7.13.

Table 7.13 Support to the students' study and achievement level in English

Support Providers	Mean	N	Std. Deviation
Father	47.9	1763	25.48
Mother	53.2	1012	24.80
Siblings	49.7	6143	23.52
Teacher	43.0	828	25.07
Tuition	41.6	1749	21.51
More than one	31.5	392	25.65
None	49.5	650	24.50
Total	**47.6**	**12537**	**24.27**

The above table shows that the support from mothers (53.2%) in children's study is most important than supports from others. The support received from more than one person is found weakest (31.5%). The supports from their siblings (49.7%) and father (47.9%) is somehow better but support from teacher (43.0%) and tuition (41.6%) is not found to be contributory to raise the achievement. The dataset indicates the importance of the support provided by the mother. There is a significant difference in support received at $p < 0.001$, the effect size is small ($f = 0.185$) and it explains about 3% variation in achievement ($\eta^2 = 0.033$).

Availability of Textbook and Student Achievement

There were students who did not have the English textbook even up to the end of the academic session. Table 7.14 shows the descriptive statistics of availability of the English textbook and achievement.

Table 7.14 Availability of English textbook and the achievement in English

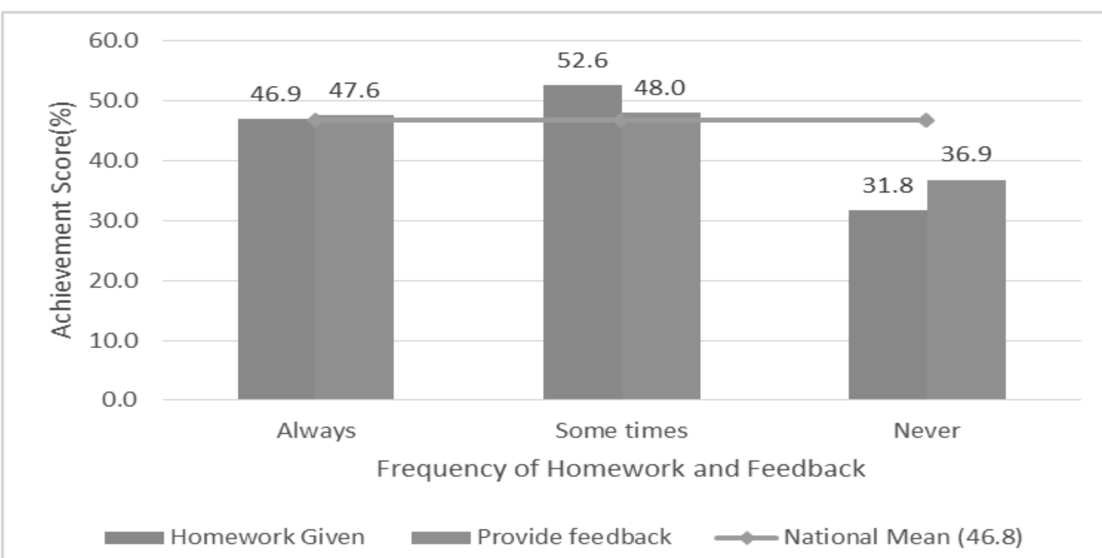
Availability of Textbook	N	Mean	Std. Deviation
Yes	11630	48.3	24.12
No	443	44.3	24.39
Total	**12073**	**48.1**	**24.14**

Out of 12073 students who responded to the question, 443 students (3.7%) did not receive textbook even up to the end of the session. There is a significant difference in achievement based on the availability textbook at $p < 0.001$, but the effect size is small ($f = 0.032$) explaining about 3% variation in achievement ($\eta^2 = 0.033$). The achievement level of students without textbooks is significantly lower than those who have access to the textbooks.

Homework and Achievement

Homework is considered as one of the ways to enhance learning which can be used as drill, exercise and an evaluation tool as well. When homework is regularly given and feedback is provided to the students, it is likely to boost their achievement level. Statistics related to assigning of homework and providing feedback is presented in figure 7.33.

Figure 7.33 Relation between homework and feedback and achievement in English



The above figure shows that if the teacher assigns homework and provides feedback to the students, achievement level becomes higher than in the case of those students who are not assigned homework. When the teacher assigns homework and provides feedback regularly, the differences are statistically significant ($p<0.001$).The homework given sometimes contributed more than the regular type of homework, but higher than in the case of none homework for students. The number of students who are receiving homework and feedback regularly is very big. The dataset explains that if the teacher assigns homework and provides feedback, the achievement is higher than those without assigning home work.

Positive and Negative Activities in School and Student Achievement

The activities of the students and teachers determine the learning environment of the school. Bullying, for example, is one of the hindering incidents for students in the school that affects learning. In the background questionnaire for students, several schools related activities were asked – some positive and some negative. Here, bullying is handled as one of the negative indicators and students' positive impression of school and teacher activities are taken as the example of positive indicators.

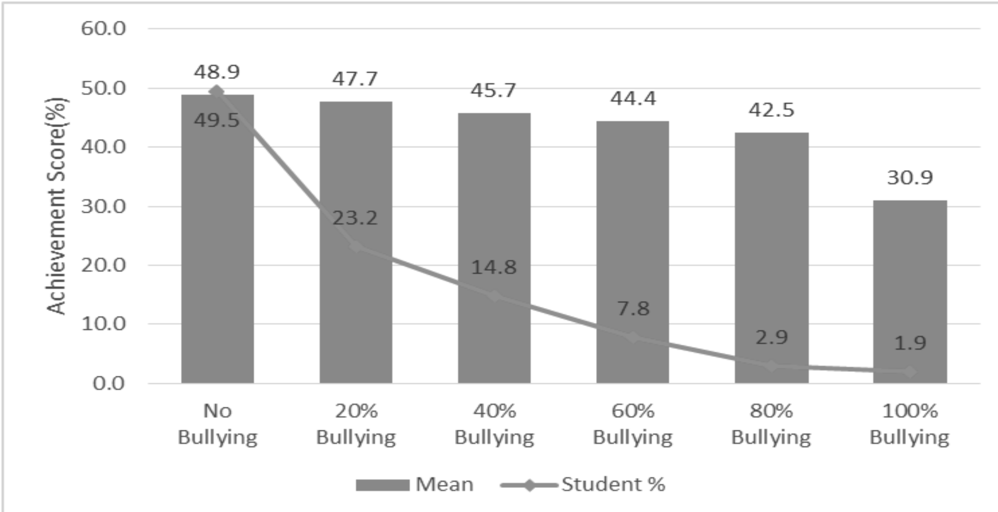
Bullying at School

Bullying is one of the problems in the school that worsens the learning environment. International studies like TIMSS and PISA also emphasize to identify such indicators. In NASA 2015 student background questionnaire, five questions indicate the varieties of bullying that are likely to happen in schools. All the incidences were stemmed by the phrase "*Which of the following activities happened in your school within the last month?*" Students' responses are presented in table 7.15 and depicted in figure 7.34. 'No (%)' indicates the percentage of the students' responses indicating no such activity happened in the school, and 'Yes (%)' indicates the percentage of the students who experienced the particular type of bullying within the last month.

Table 7.15 Bullying and achievement in English

Types of Bullying	No (%)	Yes (%)
Something of mine was stolen.	71.2	28.8
I was hit or hurt by other student(s).	77.1	22.9
I was made to do things I didn't want to do by other students.	85.2	14.8
I was made fun of or called names.	77.2	22.8
Fellow students kept outside without involving me in activities.	86.6	13.4

Figure 7.34 Effect of bullying in student achievement in English



The sum of all items (five) is taken as an indicator of 100% bullying. Figure 7.34 shows the extent of bullying with the percentage of the students and achievement of the students in each category of bullying. If only one activity of bullying is reported, it is categorised as 20% bullying. If all (five) activities are reported, it is categorised as 100% bullying. Knowing that 51.1% of the students did not encounter any bullying during the last month, one can infer that the remaining 48.9% did encounter at least one type of bullying, which is a remarkable number of students. About 1.9% of the students had experienced severe bullying (the sum of 80% and 100% bullying). It is found that achievements are remarkably lower for the students who have encountered more than two types of bullying. Students who did not experience bullying and those who encountered extreme bullying of four or five kinds have 18 percent achievement gap; though there are a few number of the students who reported this kind of bullying. The difference is statistically significant ($p < 0.001$), but the effect size is small ($f = 0.119$). Though extreme cases of severe bullying is small in number (n = 242), incidents of bullying seem to be quite common in schools. The dataset indicates that a large number of students (49.5%) have encountered bullying in schools within a month. The phenomenon tends to affect the learning outcomes in almost all groups of the students who experienced bullying.

Positive Activities in School

The activities that can boost the learning achievement of students are termed as positive activities. Students were asked about such positive activities in the school in two sets of questions listed in table 7.16. The table shows the responses of the students in all four categories, which are in the 4-point rating scale, anchored to "fully agree" and "fully disagree".

Table 7.16 Students' responses towards teacher and school related activities at school

Teachers' and Students' Activities	Respondents(%) in percentage			
	Strongly Disagree	Disagree	Agree	Strongly Agree
BQ_28a School: Students get along well with most teachers.	1.6	1.5	7.9	88.9
BQ_28b School: Most teachers are interested in student's well-being.	1.9	1.4	5.8	90.9
BQ_28c School: Most of the teachers really listen to what I have to say.	2.2	2.9	17.2	77.7
BQ_28d School: If I need extra help, I will receive it from my teacher.	2.0	1.5	8.0	88.5
BQ_28e School: Most of my teachers treat me fairly.	2.3	2.0	8.6	87.1
BQ_29a School: I enjoy to be in the school.	1.5	.6	3.0	94.9
BQ_29b School: Students in my school like me.	1.7	1.4	11.1	85.8
BQ_29c School: Friends of my school want to do their best in the study.	1.9	1.7	6.3	90.1
BQ_29_d School: Teachers expect good results of their students.	2.3	1.2	3.7	92.8

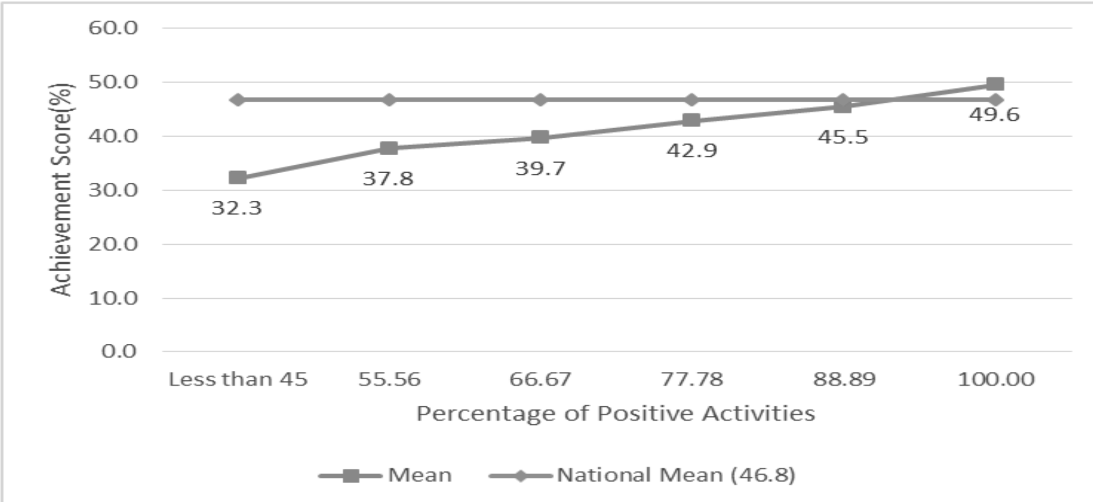
Further analysis was carried out by recoding the variables by grouping them into two categories: 1 for agree, and 0 for disagree. Furthermore, the sum of nine indicators was converted into the percentage of maximum score to analyse the level of positive activities and its relation to the achievement.

DTA finds four attitude groups in the indicator. These boundaries and descriptive statistics are presented in table 7.17 and depicted in figure 7.35. The overall result is that feeling the positive actions in the school associates positively with the student achievement. The correlation between the sum of nine positive activities and achievement is found positive (r = 0.196, p < 0.001).

Table 7.17 Students' response towards teacher and school related activities in schools

Percentage of Positive Actions	Mean	N	Std. Deviation
< 45.00	32.3	760	20.93
55.56	37.8	362	22.69
66.67	39.7	436	23.19
77.78	42.9	715	22.65
88.89	45.5	1371	23.87
100.00	49.6	9258	24.28

Figure 7.35 Relation between positive actions in school and achievement in English



The data shows that there is a positive relation between the positive activities of the school and the achievement. The increase in achievement is directly proportional to increase in the intensity of such activities. After dividing the indicators into four groups based on DTA, the differences between the groups are statistically significant ($p < 0.001$); however, the effect size is medium ($f = 0.20$). The learning achievement is much higher than the average (49.6%) when the students are extremely positive towards school and teachers. The difference between the most positive group and the most negative group is notable (16.3%). The dataset indicates that when students think that the actions of the teachers and schools are ultimately good, the results are higher than the average. At the other extreme of feeling, when such actions are ultimately negative the results are far below the average.

7.5 Summary of Findings

The main findings of student achievement for Grade 5 in English subject are as follows:

Basic results

- The Grade 5 population in English subject is almost normally distributed. A remarkable difference is found in achievement scores among the different performing levels (high, medium and low) of schools.
- Students are performing low in *Reading* and *Writing* compared to other content areas (V*ocabulary* and G*rammar*).
- The students' ability to comprehend and perform with higher ability tasks is quite low in comparison to K*nowledge* and A*pplication* level tasks. Students are much better in the recall

type of questions comparatively than in analytical questions.

- The students are performing well in recognizing the correct answer and in recalling simple facts from the texts, and applying the knowledge in grammatical forms. On the other hand, they are found weaker in comprehending the basic information, writing the texts and logical reasoning.

Diversity factors

- There is a wide difference in achievement among the Districts in English. The difference is very wide (about 44%) between low performing (Bhojpur) to high performing (Kathmandu) Districts in English. Furthermore, the number of Districts performing lower than the national average is large.
- The students from the Kathmandu Valley (68.2%) outperformed those from the other Ecological Zones whereas the students from the Mountain area performed the lowest (36.7%) followed by Hill area (40.3%).
- While comparing the achievements of Development Regions, the highest achievement is found in the Kathmandu Valley (68.2%), which is 20.6 percent higher than the Region having nearest achievement level, that is, the Western Development Region (47.6%). The achievement is the lowest in the Mid-western Development Region (37.5%) followed by Far-western Development Region (38.4%) and Central Development Region (40.8%). The mean achievement of Western and Eastern Development Regions are close to the national mean.
- The achievement gap between the community and institutional schools is remarkable. The average achievement score of institutional schools is 73.1% whereas it is only 38.0% for community schools with a wide difference at 35.1 percent points.
- The achievement level of the students from urban schools (66.0%) is remarkably higher than that from rural schools (39.9%). It should be noted that achievement level of urban schools is raised due to the performance level of institutional schools and the schools from the Kathmandu Valley.
- The students from Newari (63.9%) speaking community have performed higher comparing the students from other language communities including Nepali speakers (51.0%). On the other hand, the students from Rai (34.8%), Magar (38.7%), "Others" (39.0%), Gurung (39.1%), Tharu (40.7%) and Tamang (44.4%) speaking background have performed much lower in comparison to the ones from Nepali speaking communities.
- Among various ethnic/caste groups, Dalit students are performing the lowest (39.5%) in English followed by Madhesi (41.4%) and ethnic Minority (42.2%), which are all below the national average (46.8%).
- There is no significant difference between girls' and boys' national average performance in English. However, there are some communities and Regions where disparity is found in

achievement between girls and boys. The widest differences between the boys and girls are noticed in Madhesi community (6.9%), as the lowest performance level of girls is found in Madhesi community (38.3%). Girls' performance is found better in Chhetri, Janajati, Minority and others communities, whereas boys' performance is better in Brahman and Dalit communities. Girls' performance is better than boys' in the Kathmandu Valley and Eastern Development Region, and boys' performance is found better in the remaining Development Regions. Similarly, girls' performance is better than the performance of boys in institutional schools (3.8%), whereas boys have performed slightly better in community schools.

Selected explanatory factors

- The achievement of students raises as the parents' education increases up to Bachelor's degree. It shows that the students of illiterate parents achieved the lowest. Data shows that mothers' education contributes more than fathers' education in raising the learning achievement of their children. The educational level of parents affects the children's achievement level. Achievement of the students is more influencing if either father or mother is at least SLC passed, or if father has the qualification of Bachelor's level or higher. Higher parental education has resulted in better learning achievement of students.
- Significantly, higher achievement is seen when mothers are having government service and teaching jobs (58.8%). Overall, the mothers who are employed in any kind of job or involved in business are beneficial for the learning achievement of students. Similarly, in case of father's occupation, the main factor is related to whether the father works in agriculture or not. If the father is in a government or other service, the students performs better (60.1%) than the children of fathers who are working in agriculture (39.3%). Similarly, if the students' fathers are involved only in the household and mothers in daily wages, their performance is the lowest (43.3%).
- When the number of home accessories increases, students' achievement also raises ranging from 35.4% (if none of the accessories are available) to 69.6% (if all of them are available). Availability of the stated facilities indicates higher SES of the family. The dataset indicates that either economic or educational ability or both at home support children to raise their educational achievement. If the students come from high SES family (having a large number of home possessions and accessories), their performance is found remarkably higher than the performance of other students.
- Socio-economic status plays a vital role in student achievement. The difference between the lowest and highest SES groups is remarkable (53 percent). Especially the situation is worse to the families where the father or both parents are illiterate or work in agriculture or daily wages.
- Either any paid work or unpaid household work for four hours per day beyond the school has remarkably lowered down the students' achievement. However, a decent amount of household work up to two hours per day does not affect learning negatively in English. It

should be noted that students' involvement in paid work for any duration negatively affects the students' achievement.

- Positive attitude towards the subject is positively correlated with the higher achievement in English.
- The students studying at proper age have obtained the highest score and the achievement lowers down as the age increases.
- Support from mothers (53.2%) in study is more important than the support from other. The support received from more than one person gives weakest results (31.5%). Support from siblings (49.7%) and father (47.9%) is somehow better but support from teacher (43.0%) and tuition (41.6%) is not seen contributory to raise the achievement.
- It is a big challenge for students to learn without textbook, as many as 3.7% of the students did not have textbooks of English even by the end of the academic session. The achievement level of those students lacking textbook is significantly lower (44.3%) than those who have access to the textbook (48.3%).
- If the teacher gives homework and provides feedback to the students, the achievement is higher than in the case of those who are not given homework.
- Students in a large number (49.5%) have encountered at least one kind of bullying in schools within a month. The phenomenon has negatively affected the learning outcomes in almost all groups of students who experienced bullying.
- Learning achievement is much higher than the average (49.6%) only when the students are extremely positive towards school and teachers' behaviours. The difference between the most positive group and the most negative group is notable (16.3%). When students think that the teachers and schools' actions are ultimately good, the results are higher than the average. At the other extreme of feeling, when such actions are ultimately negative the results are far below the average.

Chapter 8: Conclusions and Implications

8.1 Introduction

It is well evident that subject specific knowledge fostered among the youngsters through schooling helps them understanding the social, political, economic and cultural world to lead an informed life. Alongside developing the subject specific knowledge on core subjects, schooling has also long been established for fostering in youngsters the cognitive abilities in the form of improved reading, writing, numeracy and communication skills essential for information retrieving and processing, reasoning and analytical thinking, problem solving capability to lead an informed and a productive life in economic world. These knowledge and skills are fundamental for individual, social and economic development of a country forming the basis of technological adoption, diffusion and innovation that have direct impacts on an individual's success in the job market, economic growth of a nation and promotion of equality in the society. Additionally, the knowledge and skills acquired at early schooling leads success in further level of schooling forming a strong foundation for lifelong learning. Moreover, this being the age of knowledge economy, national development largely depends on the availability of highly skilled labour force, its capability for innovation, and the intellectual property it generates.

With this understanding, nations across the globe urge for ensuring universal schooling with good quality as a central part of development strategy in order to enhance skills and employability of youth, raising national productivity, and reducing poverty. So, it is an important concern for all parents, teachers, governments and general public to know how well the school education systems equip youths with knowledge and skills they need to get decent job for making their lives better, to play a role in building more peaceful and equitable societies (Matsuura, 2004), and to be able face future challenges. Besides, measuring these knowledge and skills during the period of schooling is also essential to tracking development and assessing effectiveness of education policy and practices (Boreman, Hews, Overman & Brown,2003; Hanushek & Rivkin, 2010). With these concerns, measuring and monitoring students learning achievement by means of large scale assessment developed world wide during the last decade of 20th century for determining existing level of learning, finding the gaps as well as for providing feedback for improving quality of learning. Nepal is no exception to adopt national assessments for measuring students' learning achievement in order for monitoring it and providing policy feed back to the system which has been conducting national assessment of student achievement since 1995.

The final chapter of this report first summarises the context and objectives of this assessment, methods and process used in this assessment, basic principles considered while conducting assessment. It, then presents a summary of pragmatic standards utilised in this assessment and major results of assessment. Next two sections of this chapter present conclusions and implications of this assessment.

8.2. Summary of Context, Methods and Results

This section summarises the context and objectives of this assessment; methods and principles used in this assessment. Similarly, it presents pragmatic standards for this assessment and the summary of main results.

Context

Various assessments of student achievement in Nepal were also conducted before 2011 for various Grades (in 1995, 1997, 2001 for Grade 3; in 1997 for Grade 4; in 1998, 1999; 2003, 2008 for Grade 5; in 1999 for Grades 6 and 8 and again in 2008 for Grade 8) aiming at assessing learning outcomes of students to determine level of learning at the respective Grades and to provide policy feed back to the system. However, they all were small scale in nature and could be inadequate to nationally represent. Large scale assessment for the first time came into practice since 2011 after the establishment of the ERO under the MOE. During the implementation of SSRP 2009-2015, four assessments, two rounds for Grade 3 and two rounds for Grade 8, were planned to accomplish. Accordingly, two rounds of large-scale assessments for Grade 8 in 2011 and 2013 and one round for Grade 3 and 5 in 2012 have already been accomplished. Next to them, this is another, the second round of large scale assessment for Grade 3 and 5 conducted in 2015.

Objectives

Similar to other large scale assessments whether conducted at international, regional and national level; this assessment also has been designed to generate accurate, objective and comparable data on student learning in order to assess the health of education system, particularly the primary level of Nepal. At the same time it also aims to describe national levels of learning achievement in key subjects Mathematics and Languages and to determine the extent to which primary Graders have developed a reading, writing, numeracy skills and fundamental understanding on the core subjects. More specifically, this assessment is motivated to:

- Determine the current national level of achievement of Grade 3 and 5 students in Mathematics, Nepali and English languages against the goals set in the curricula;
- Analyze variations in student achievement by region, gender, location, and language of instruction;
- Explore factors that influence student achievement in primary education;
- Compare student learning achievement in the current study with that of the previous studies of Nepal and international ones like TIMSS and PISA;
- Generate evidence based data for monitoring the trends in students achievement for these subjects over the period of time;
- Generate recommendations to feed in policy making in order to improve quality and equity in education.

Methods

This test was administered in 1543 randomly selected schools from 23 sample Districts covering all Ecological Zones and Development regions to assess the learning outcomes of 3rd Graders in Mathematics and Nepali and 5th Graders in Mathematics, Nepali and English. Altogether 74078 students (33863 in Grade 3 and 40015 in Grade 5 from randomly stratified 1543 sampled schools) participated in the assessment. In the sample for Grade three, 16927 students were boys and 16936 were girls. Similarly out of the total sample for Grade five, 19926 students were boys and 20089 were girls. Out of the 75 Districts of Nepal, the dataset represents a random selection of 23 Districts covering all five Development regions (Eastern, Central, Western, Mid-Western, Far-Western), and the Kathmandu Valley, as well as all Ecological Zones (Mountain, Hill, Tarai). In addition, both rural and urban schools as well as community and institutional (private) schools are proportionally represented so that the results of the assessment can credibly be extended to the whole student and school populations of Nepal.

Final items were standardized after pretesting six sets of test papers in each subject in 7724 students (ranging from 1514 to 1565 students in each subject) from 153 schools and 11 Districts representing different strata. Only those items having high discriminatory power were selected for final test by analyzing each of their difficulty level. Based on the pretest results of the items, the difficulty levels of the tests were set around 50–60%. Some linking items from international tests like PIRLS and TIMSS also were used to make the test results comparable to the international standard. All the items were analyzed and equated using IRT modelling.

Three versions of the items in each subject were administered and the final scores were equated by utilizing the IRT modeling. Reliability of the tests was found high and the validity was assured by applying specification grids of the national curriculum developed by the Curriculum Development Center (CDC). From a methodological standpoint, the process and practices of the inquiry has successfully followed the procedures as used in some international level test with some contextualization on them to reflect the reality of the Nepali context. Thus, this test is believed to fulfill the national and international ethical principles, criteria and standard to qualify it to be a credible assessment. The results were linked to the set of results from the 2012 assessments well as to the international item banks of TIMSS and PIRLS.

The tests were administered uniformly at a time in one shot in all the sample schools throughout the country in the scheduled day. Each selected school was assigned to conduct test in one of the selected subjects for each Grade. Thus, the students in a Grade were required to participate in one of the selected subjects assigned to the school. The answer sheets were marked and achievement scores were tabulated using Optical Mark Reading (OMR) machine.

The results are reported mainly as percentages of maximum marks where 100 (%) represents all tasks solved and 0 (%) none. As a result of pre-testing of the items, the difficulty levels of the tests were set at 50−60%. For English and Nepali language assessment, the Common European Framework of Reference for Languages (CEFR) was also used to obtain the level of students from a language achievement point of view.

Basic Principles

Learning achievement test, particularly the large-scale one, needs to be grounded on some ethical principles to ensure its validity, fairness, reliability and generalizability. As opined by Race, Brown and Smith (2005), an ethically good assessment should be just and fair, valid and reliable, transparent, motivating, sufficiently demanding high possibility to show excellence. Based on these premises, this test is characterized as:

- In the entire process from test design,item writing to administration, the items are ensured to be equal, just and fair to all the pupils from the sample schools (see Section 2.2.2) as the test settings in all schools makes it possible that no single school or student is neither favored nor biased (Section 2.2.6). The pupils who participate in the tests have studied, in principle, the same national curriculum all over the country. Hence, they all are supposed to learn the same contents. These core contents are measured with a common test.
- An attempt is also made to maintain the validity and reliability of the indicators at a high level as much as possible. While carrying out the item analysis, it is particularly taken care that the individual questions and the whole tests are of high quality as much as possible.
- The measurement process is transparent in the sense that there is not guided with any hidden objective and hence no surprises or traps are set in the test and in its procedures. The procedures of test, its objectives, test administration schedule, involvement of possible personnel and the like have been made public to all teachers and students earlier the test administration. In the test, only the learning outcomes stated on national curriculum-the public document, are assessed. The test items, however, are not public – except the released ones – in order to guarantee the possibility of using linking items in years to come. This secrecy of the test items, which is also an international practice in all assessments, has been maintained for the items used.
- When compiling the tests, an attempt is made to keep the test versions versatile and motivating to all students. The low-stake role of the assessment and the textbook-independent test items facilitates the fact that the pupils do not need to study additionally for the preparation of the test.
- Further, attempt has always been made to design and administer the tests in a way that all the students are able to show their excellence at their best. On the other hand, the fairness implies that it is important to motivate the weakest students. Hence, always some very easy items are selected to the test batteries. Furthermore, the test batteries include such items that suit to all students to ensure average students demonstrate their performance at best. On these grounds, one can conclude that this test fulfils the ethical principles to be followed in the large scale assessment at the national level to the extent possible.

Pragmatic standards of Assessment

Though overlapping partly with the ethical principles suggested by Race, Brown and Smith (2005) above, Ivernizzi and colleagues (2005) maintains that a good assessment takes into account the technical and pragmatic standards for a good assessment as validity and reliability, standardized administration, scoring and reporting, comparability of the scales, norms and scores, fairness in testing, the test-takers' separate language and culture background. On these grounds, this test makes an attempt to the extent possible meeting the technical and pragmatic standard as:

- The trustworthiness of the measurement instruments – reliability and validity – is secured separately to every test version. The level of accuracy has been tried to rise as high as possible by pretesting the items and by selecting only the highly discriminating ones to the final tests. Validity has been secured by wide content coverage and following the content weights allotted in the curriculum.
- Entire management of test including items writing, finalizing, test administration, scoring and reporting have been centralized in an equal manner for all schools to ensure fairness or no biasness and avoid local influences of any kinds in each step. The centralized marking of the papers by experienced teachers after a thorough orientation on marking procedures makes the scoring objective and fair for all students and schools.
- IRT modeling is used to calibrate the different test versions with each other and, finally, to equate the test score over the versions. In this manner, it is possible to reduce the bias coming from the slightly different difficulty levels of the test versions. Hence, serious attempts are made to keep the scales, norms and scores comparable over the different versions as well as over the different years within the same subject.
- The fairness of the testing is assured in four ways: First, the versions test items were made comparable with each other and with the tests of the past. Second, only the core contents in each sujects are measured in the tests. Third, the tests are based on learning outcomes stated in the public document of national curriculum. Fourth, the test results are openly reported and interpreted by using widely accepted statistical methods. The interpretations concerning the school or the pupil are never based on only one indicator.
- All the test papers administered for Mathematics have been prepared in two language versions in both Nepali and English to avoid confusion, misunderstanding and language bias and to ensure the same level of understanding items among all test takers in the intents of test items. For language testing in Nepali and English, the test papers have been prepared in the repective language. It may be possible that in the future some local languages are also used – especially in the early Grades because officially the government is has been promoting the use of indigenous languages at the early Grades. However, the language issue is challenging in Nepal as more than 120 indigenous languages exist.

Having considered the standards followed in the test, one can conclude that this assessment fulfils also the practical principles to the extent possible in this national level assessment. Though one issue regarding the language of the tests still persists which needs to be thought of critically and resolved reasonably in the assessments to come.

Main Results

The data set shows that learning achievements for both Grades has remained at 52 percent in Nepali and 45 percent in Mathematics for Grade 3 which is 46 and 48 percent respectively for Grade 5 in those subjects. In Grade 5 English, it is 47 percent. At the outset, 3rd Graders' average achievement score is higher than that of 5th Graders in Nepali as well as Mathematics though comparison of two grades based on their average assessment scores may not be preferable, it indicates some trends about the achievement of students. The language proficiency level of 5th Graders is more or less at the same level in both English (47) and Nepali (46). To look at the mean achievement from equality point of view, gender parity has almost maintained in both the Grades in all the subjects. It is notable that girls are ahead of boys by 2 to 3 percent point in Nepali in both Grades and by 2 percent points in Grade 5 Mathmatics. In English too, they are seen equal. However, when seen from other perspectives going deep down into the results on different variables, low levels of learning with high discrepancies are noticed. Having analyzed the database for each subject as described in earlier sections, the following conclusions are drawn as the main results.

I. Clearly divided student population into three distinct groups

The population of schools participating in the test is seen to have divided into different distinct groups from the lowest, low, average to the high and highest achieving. Out of the total, 5 to 5.6 percent schools fall into the lowest achieving even less than 10 percent and 1.4 to 3.2 percent of them fall into the highest achieving more than 90 or above of the average score. From minimum 17.3 to maximum 24 percent schools fall into the low achieving group with average score from 10 to 30 percent. Maximum percentage of schools ranging from 37.3 to 49.1 belongs to the average performing achieving 31 to 59 percent of the score. Among the high achieving, are 26.3 to 38.1 percent schools with average score from 60 to 89 percent.

Not only the schools are divided into various groups, the population of student also is seen to have been divided into two to three distinct groups in terms of the achievement score with remarkable variances: both low (15 percent or below) and high (63 percent) performing students in community schools and mainly high (above than 60 percent) and highest (above than 80 percent) performing students in the institutional schools. In comparison to Mathematics and Nepali for both Grades and English for Grade 5, the 5th Graders population for Mathematics is more normally distributed. In community schools, 50% students belong to the low performing groups below the average forming a large plot whereas in institutional schools more than 50% students lie at more than average achieving population forming its bigger plot with a small plot of low achieving group. It is notable that community schools also are found to have achieved as high as institutional schools but their number is too small which is not enough to raise average achievement. Similarly, institutional schools are achieving as low as community schools whose number is small. Because of the low achieving large population and high achieving small population in community schools and mainly high to higher achieving population in institutional school, the entire system is seen to have shifted towards low performing one indicating that the system is not able to give sufficient support for those students who are lagging behind in the early Grades.

II. An unbalance learning across all curricula contents

Against the expectation of curricula for similar level of learning in all content areas, the dataset is evident that certain contents of the curricula are learnt less effectively than others. For instances, in Grade 3 Mathematics, the achievement level in *Arithmetic* (45) and *Numeracy* (49) is remarkably lower than *Algebra* (57) and *Geometry* (57) whereas in Grade 5 it is lower in *Algebra* (46) and in *Numeracy* (47) than in *Arithmetic* (51) and in *Geometry* (49). Similarly, in Nepali and English, the *Writing* (42 in Grade 3,33 in Grade 5 for Nepali and 40 in Grade 5 English) followed by *Reading* skills (51 in Grade 3, 39 for Grade 5 Nepali and 46 in English) are notably poorer in comparison to the achievement in *Grammar* (58 for Grade 3 Nepali and 53 for Grade 5 English) and *Vocabulary* (76 for Grade 3,52 for Grade 5 in Nepali and 62 in English). From the data set, it is also evident that instead of improving, achievement level in these skills lowers down at the later Grade. Circumscribed with such unbalance learning on some domains, the entire system is shifted towards a low performing making it less effective to yield better results. Such glaring evidence raises an issue that why schooling at upper Grade fails to foster writing and reading skills.

III. Lower level of cognitive ability for the tasks requiring higher ability

Due to poor writing and reading, students are found comparatively poorer in solving the *Subjective* types of items in which average score ranges from 33-48 in Nepali, remains 35 in Mathethematics and 40 English in comparison to 55-63 in Nepali, 54-57 in Mathematics and 54 in English in *Objective* items. From this fact, it is evident that students are found poor in describing in detail the information, elaborating the ideas, explaining the concepts in structured form or to create short paragraphs in the relevant themes in an organized way. Similarly, they are also weaker in producing fluent texts, preparing synthesis and abstracts from the given text.

The dataset is evident that students in all subjects are seen performing comparatively better in *Knowledge* level tasks (56 to 65) whereas they are found very poor in the tasks requiring *Higher* ability which ranges from just 31 at the lowest to 39 at the highest. For instances, as many as 22.8 and 15 percent students from Grade 3 could not solve any of the tasks requiring higher ability in Math and Nepali respectively, whereas such figures of students not solving higher ability tasks is seen to be rather big in Grade 5 which is 25 percent in Mathematics, 9 percent in Nepali and 29.7 percent in English. Likewise, having analysed achievement scores in all levels, average achievement goes lowering down as order of cognitive level are higher up. From this fact it is evident that students are somehow good in the tasks requiring recalling simple facts, recognizing correct answers that are explicit in the texts, comprehending basic information, fundamental thinking, basic description or interpretation of given paragraphs, table, charts in few steps. However, they are seen very poor in demonstrating the ability to solve complex problems, to analyze, deduce logic, generalize, justify an argument or viewpoint, and in the ability to transfer learning from one context to another. Similarly, they are found weaker in solving problems requiring multistep functions or procedures. By types of school, the students in the institutional schools are seen more able to solve problems requiring higher ability in comparison to their peers in community school with a wide variations (for example 27.9-45.7 Nepali in Grade 5). More discouraging is the fact that higher the Grade, the more is the population with lesser higher ability tasks.

In Mathematics, students are able to do basic calculations, but are weak in reasoning, problem solving, proving theory or formula, and in constructing shapes and figures. In many cases, the students did not even attempt to complete the open-ended questions of higher cognitive level. Similarly, in Nepali and English subject, students performed well in the tasks requiring recognizing the correct answer that is explicit in the text, recalling simple facts from the texts, locating the basic information and its interpretation of text read. However, they are much weaker in producing fluent texts or essays, and in preparing synthesis and abstracts from a text. The students tended to attempt open-ended tasks but the skills were not high enough for obtaining the highest marks. In case of applying the gained knowledge in new situation, they are found very poor than the higher level thinking ability.

IV. Wider inequality in learning achievements among students in terms of location and types of school.

Variance in achievement also is seen while analyzing the results in terms of location and types of schools. Students from urban schools are seen performing better than the rural ones. For instances, the achievement score for urban students has reached to 53 to 59.8 in Nepali Grade 3 and 5, 50.9 to 54.8 in Grade 3 and 5 Mathematics and 66 in English; whereas it is 44.4 to 48.9.4 in Nepali Grade 5 and 3,43.5 to 47.4 in Math Grade 3 and 5 and 39.9 in English for rural. Such level of remarkable difference is not justifiable in any grounds from equality point of view.

Besides the variance in terms of location of schools, achievement gap is also seen wider between the types of schools students attend as students' average achievement in institutional school is found notably higher than the others. The average achievement score in institutional schools reaches at 54.1 to 60.8 in Grade 3 and 5 Math, 61.4 to 67.4 in Grade 5 and 3 Nepali and 73 in English, which is 42.8 to 45.9 for Nepali Grade 5 and 3, 41.0 to 44.5 for Math in Grade 3 and 5 and 38 in English in Community schools. One of the reasons behind such variances may be the higher socio-economic status of the family sending children to institutional schools as higher socio-economic status and educational achievement are positively correlated.

It is noteworthy that there are also small number of community schools where the average results are at the same level as in the private schools even though the SES is remarkably lower. In these schools, either the processes are more effective than in the private schools or the students are of the same ability as those in the private schools and are not adversely affected by the processes within the school or their socio economic status. It could be concluded that higher achievement of community schools is not due to the system but because of the individual effort of school, teacher or students themselves.

V. Wider disparity in achievement between Districts, Development regions and Ecological Zones

A comparative analysis of the datasets shows that there is an inequality in learning opportunity for students among Districts, Development Regions and Ecological belts leading to wider variances in the achievement level. In Nepali, the results for Rautahat (33), Bhojpur (38), Kalikot (41), Doti and Banke (42), Bajhang (43), Pyuthan (44) in Grade 3 and for Jajarkot and Bajhang (34), Bhojpur, Morang and Bara (36), Rautahat (37), Doti and Kanchanpu (39), Kalikot and Banke in Grade 5 are relatively poor compared to other high achieving Districts such as Mustang (80), Kathmandu (66), Bhaktapur (65), Jhapa (64) Ilam (62) in Grade 3 and Mustang (59), Jhapa and Bhaktapur (58) Kathmandu (56), Dhading and Nawalparasi (55) in Grade 5. Similarly, the mean score among the highest in Grade 3 Math reaches to 60 in Lalitpur, 59 in Mustang, 58 in Jhapa 56 in Kathmandu, 55 in Dhading and Nawalparasi whereas it lowers down to 27 in Pyuthan, 31 in Bhojpur, 35 in Sankhuwasabha. In Grade 5 too, there is a wider difference among the high performing Districts such as Kathmandu and Mustang (60) and the lowest achieving Bhojpur (36), Pyuthan and Sankhuwasabha (38) with a difference of more than 20 percent points. In English, the difference widens further with more than 34 percent point between Kathmandu (70) and Bhojpur (26). The other high performing Districts in this subject are Bhaktapur and Lalitpur, all from the Kathmandu Valley and low performing are Pyuthan and Jajarkot (33), Kalikot (35) and Rautahat (36).

While looking at the results by Development Regions, the same level of disparity in achievement is evident. After the Valley, Western Development Region demonstrates better level of performance in all subjects (50 and 57 in Nepali and 48 in English) and Grades except in Grade 3 Mathematics (45 and 49) in which Eastern and Central Regions are seen ahead (46 and 48 for Eastern and 46 and 49 for Central). The Valley outperforms the rest in all subjects and Grades (51 at the minimum to 68 at the highest) with a great differences whereas Mid Western is seen the lowest achieving in both the subjects in Grade 3 (37 and 43) including Grade 5 Mathematics (43). Likewise, the achievement in Nepali (37) and in English (38) for Far Western region is found the lowest of all.

To look at it by Ecological belts, the Mountain Region achieves the lowest (37 at the lowest to the 45 at the highest) whereas the Valley again outperforms all the regions (51 at the lowest to the 68 at the highest). In Nepali, Hill (51 in Grade 3 and 46 in Grade 5) and Tarai (49 in Grade 3 and 43 in Grade 5) come after Valley (57 to 64). The same is the situation of Hill (40) and Tarai (47) in English, which is higher than Mountain (37) and lower than Valley (68). In Nepali, Hill is ahead of the Tarai by 2-3 percent points whereas in Math Hill lags by 3-4 percent points behind Tarai.

VI. Influence of caste/ethnicity and home language backgrounds in achievement

From the data sets, it is evident that language spoken at home and ethnicity/caste–the cultural background of the students, among others, has been one of the influential factors in determining student achievement. While lots of efforts have been put to reduce disparity in achievement, differences are still noticed in their achievement by ethnic/caste groups. For instances, Brahamans always are seen highest performing groups (61 and 54 in Nepali, 50 56 in Maths; and 59 in English) whereas Dalits and Madhesis are found performing the lowest which is just 47 and 44 in Nepali, 41 and 43 in Maths and 40 in English for Dalits, and 45 and 41 in Nepali and 41 in English for Madhesis. In Maths Madhesis are seen comparatively better than the Minorities but lower than the group "Others". The Chhetris and Janajatis are above the Dalits and Madhesis but more or less at the similar level to the groups "Others".

In Nepali, the overall difference between the groups is statistically significant ($p < 0.001$) and the effect size is small ($f = 0.17$). Caste/ethnic background explains 2.9% variation in achievement (($\eta^2 = 0.029$). Apart from this, the mean of the various caste/ethnic groups differ from each (except Chhettri and other, Janjati and other) other statistically significant at 95% confidence level. A positive sign from equity point of view is that the Dalit students have performed remarkably better than the national mean (48%) in the Western Mountain (62%) and Eastern (51%) and Western Tarai (49) as well. However, the results are much lower than the average in other regions and very poor in Far-Western and Mid-Western Tarai area (30%). The number of students in certain strata is small, and hence it may be wise not to make the too strong implications of the results. However, in the Kathmandu Valley Dalits performed high (50%).

In Math, the difference between Brahmans/Chhetris (40%) and Dalit (24%) is 16 percent. The overall difference among the groups is statistically significant ($p < 0.001$), but the effect size is moderate ($f = 0.2$). Categorization of students according to their ethnic/caste background explains 4% of the variation in data ($\eta^2 = 0.04$).

In English, Dalit students are performing the lowest (39.5%). The overall difference between the groups is statistically significant ($p < 0.001$) and the effect size is medium ($f = 0.232$). Caste/ethnic background explains 5% variation in achievement (η^2=0.051). The results shows that the lowest achievement status is found of Dalit (39.5%) and it followed by Madhesi (41.4%) and Ethnic Minority (42.2) which is below than the national average (46.8%). The status of Dalit is found better in Kathmandu Valley (57.1%) and lower in Eastern Mountain (27.9%). Except in Madhesi group, in all caste and ethnic group girls' language ability is better than boys.

While comparing the achievements of different language groups, a notable difference is found between the Nepali and other language groups mentioned as "Non-Nepali". In English, it is found that the Newari speaking have performing higher (63.9%) than the students from Nepali speaking community (51.0%). On the other hand, the students from Rai (34.8%), Magar (38.7%), "Others" (39.0%), Gurung (39.1%), Tharu (40.7%) and Tamang (44.4%) speaking performed much lower in comparison to the students from Nepali speaking communities. The difference between the language groups is statistically significant ($p < 0.001$), but the effect size is medium ($f = 0.25$) because the minority groups are small and the division into smaller language groups explains about 6% of the variation in the data ($\eta^2 = 0.059$). However, the results are different in other language groups as their performance is lower than that of students having Nepali as home language. The disparity is found widest (6.9) in Madhesi community, the lowest performance level also is found in the girls of same community (38.3). The girls' performance is found better in Chhetri, Janajati, Minority and others communities, whereas the boys' performance is found better in Brahman and Dalit communities. The girls' performance is found better than the boys in Valley and Eastern region and the boys' performance is seen better in other remaining regions. The girls' performance also is found better than the boys in institutional schools (3.8%) whereas boys' performance is found slightly better in community schools**.**

In Mathematics too, Nepali speakers outperform non-Nepali speakers by 3.4–3.9 percent points. Among all national language speakers, Newari and Gurung speakers achieved the highest score (55.3 and 53.9 respectively) and Rai speakers achieved the lowest score which is 37.9 in Grade 3 and 42.1% in Grade 5. Variation between the highest and lowest speaking student group is medium low.

Among the language groups in Grade 3, Newari (59), Tamang (58) and Urdu (58.9) speakers have performed higher than those Nepali speakers (56.%) in Nepali. On the other, Maithili (39.9),"Other" (43.2), Limbu (48.4) and Tharu (48.0) speaking students have achieved much lower than national average. Newari speakers achieving low in earlier test (54 in Community School and 68 in Institutional School) have raised their position at the highest (59.0) among non-Nepali speaking groups. The difference is statistically significant ($p < 0.001$) but the effect size is medium ($f = 0.18$) because the minority groups are small and the division into smaller language groups explains about 3 percent of the variation in the results ($\eta^2 = 0.031$). In Grade 5 too, the students from Newari (61%) and Sherpa (53%) speaking have performed higher in comparison to the students from Nepali speaking community (52%). On the other hand, the students from Limbu (34%), Tharu (35%), "Others" (36%), and Urdu (39%) speaking performed much lower in comparison to the students from Nepali speaking communities. The difference between the language groups is statistically significant ($p < 0.001$), but the effect size is medium ($f = 0.34$) because the minority groups are small and the division into smaller language groups explains about 10% of the variation in the data ($\eta^2 = 0.105$).When analysing only the minority languages by excluding the Nepali language groups and the group "Others", the effect size is high ($f = 0.40$) indicating a remarkable difference between the highest performing group (Newari, 61%) and the lowest performing group (Limbu, 34%). In this context, it is difficult to predict the result in the Nepali subject based on the different language speaking groups. However, when combined all groups into "Nepali" and "non Nepali", the Nepali speaking students always perform better than the non-Nepali. Despite the given the context, the mother tongue reflects, to a large extent, the ethnic background and hence level of difference need to be taken as a possible source of inequality in learning.

VII. Association of low socio-economic status with remarkably lower learning achievement

Students learning level in Nepal is found dependent largely on the various components of socio-economic status of parents such as parent's education, occupation, home accessories and possessions, available for children's study at home. The difference in achievement due to the high and low SES is seen wide ranging from 27 percent (Grade 5) to 32 percent (Grade 3) in Nepali, 27 percent in Grade 5 Maths and 53 percent in English– the widest of all, for instances.

From the datasets, it is evident that parent illiteracy is still persistent in significant percentage. As reported as many as 12 and 26.5 percent children studying in Grade 3 have illiterate father and mother respectively. Similarly, 14.2 and 32 percent children studying Grade 5 have illiterate father and mother respectively. Datasets show that low achievement is persistent particularly among those children whose parents are illiterate. For instances, the students whose parents are illiterate achieve very low 42-43 in Grade 3 Nepali and 46-47 in Grade 5 Maths in comparison to above than 52 for those whose parents are literate. In English, the children of illiterate parents achieve 40 which go up to 62 for those whose parents are literate. From the data it is also evident that students' achievement increases along with rise in parents' educational level. It is also seen that mothers literacy and education level have contributed more than the father's to raise achievement level.

Similar to the parents' education, availability of *home possession* and *accessories* are also correlated with achievement level. In Maths, the children with 0 to 3 home possessions out of the 8, the achievement level is lower than the national average and those with more than 3 home possessions the achievement is higher than the national average. The average achievement score of students from the families with 7 possessions is the highest (54 – 55). The same pattern shows with home accessories too. Similarly, the results are very poor (36.5-39) when there is none or less than 3, whereas it is remarkably high (51.9-56.8) when there are all of the accessories. In Nepali, the achievement is < 37- < 42 when none or two of the accessories are available whereas it is > 55 - > 51 when more than two are available and it rises further up to > 51–60 if all of them available. The same trend is noticed in English too. For instance, the achievement is 35.4 for the children with no or less than two accessories whereas it 69.6 if all of them are available.

Parents occupation is seen another detrimental factor for student achievement in Nepal. The achievement of the children is low (43-49) whose parents either mother or father works abroad, it is further low (41-47) for the children whose parents works for agriculture, daily wages or unspecified jobs in comparison to the children whose parents are either teacher or business persons or hold government job (52-60) 5 to 7 points advance than the agriculture related occupation or daily wages.

VIII. Effect of homework and support for study in higher achievement

Home work is considered an integral part of teaching and learning for boosting up students achievement if assigned purposefully and feedback provided, which can be used as a means of assessment tools, exercise or drill. Based on the dataset, it is seen that the students receiving homework regularly achieve 47-54 and receiving sometimes which is higher (31-38) than those not receiving it. The difference is statistically significant ($p < 0.001$). Though the number of students not receiving home work is seen very small, and hence the effect size is also small ($\eta^2 = 0.02$ and 0.02) which explains only 0.9 percent variance.

In order to raise the achievement, support for children for study in any form also is seen essential as the students receiving support from parents, their older siblings or teacher are found achieving higher score. In Nepali, the support from mother or from older siblings is seen beneficial (56) whereas teachers support is seen effective in Maths to achieve high score (49-53). In English, those achieve high (50-53) who receive support from their mother or older siblings. However, the children studying independently also are seen achieving as high as 49-54 above than the children receiving support from others. These children achieving better than others without support may be talented. As the ability to afford for support at home reflects better socio economic condition of the family, the children from such family always achieve high.

IX. Effects of availability of textbook in students achievement

In the context of Nepal, textbook has remained only the sole means for students to study and learning due to the limited availability of other reference materials for study. Besides, it has also been a reliable means for teachers as well to conduct teaching at classroom since entire teaching is dependent in textbooks. So, having or not having it in the hands of students and teachers is a matter of great importance to achieve better in schools. Being cognizant of its importance in teaching and learning, government has also been putting its concerted efforts to ensure timely availability of textbooks for all students. However, the dataset reveals the fact that still textbooks are not available for 3.7 to 4.4 students even by the end of academic session which has resulted in varied level of achievement between having it or not. For instance, the achievement in Nepali rises up 48 for the students having the textbooks and which lowers down to 40 for those not having textbook with a difference by 8 percent points. In English, the difference is found at 4 percent points (48 for having textbooks and 44 for not having textbok). The great variance by 10 percent points is found in Maths in which students having textbooks achieve as much as 50 against 40 of the students lacking textbook.

X. Effects of attending schooling at proper ages

Although proper ages of children to study in Grade 3 and 5 are 8 and 10 respectively, the ages of students attending these Grades vary widely from below 7 years to above 13 and even 14 years above. As the ages vary, the achievement also vary from 40.7 (Grade 3 in Maths) for the 7 years old children and to 42.9 for those who are 13 and above years-the lowest, whereas it is 44.7 to 46.8 for those who belong to the age group from 8 to 11. The students who are at the age 10 achieve the highest (46.8) followed by 9 (46.2) and 11 years (46.1). Similarly in Grade 5 too, among the lowest achiever are 9 years (42.9) and above 14 years old whose achievement is 41.9 whereas higher achiever are the students belonging to the age groups from 10 to 13 whose achievements range from 44.5 (for 10 years old) to 49 .2 (for 11 years old) followed by 48.6 (for 12 years old) and 46 (for 13 years old) in Nepali. The same is the trend in other subjects and Grade as well that is lower than average achievement for under and higher aged children and better achievement for proper aged children. The students studying at their overage means that either they started their schooling at later age or repeated the earlier or same Grades many times. However, it is good that these over aged children (35.18 percent in Grade 3 and 10.05 percent in Grade 5) are retained at school and hence need additional support in their study. Given context implies that attending school at proper age leads children towards a higher possibility to achieve better.

XI. Bullying as an impeding factor for learning achievement

Though the phenomena of bullying and its consequences has remained unknown to many of the parents and even to the school teacher, it is found to be rampant in school impeding the learning potential of many children. The obtained datasets show that as many as 46-50 percent (in Nepali) and 51.4-51.9 (in Math) children have experienced at least one types of bullying within a month which is a big number. Out of the total, 8-9 percent (in Grade 3) and 4.6 - 5.3 percent (in Grade 5) children have experienced severe kind of bullying gaining remarkably lower achievement (35.3 - 39.9 in Grade 3 and 32.1-39.5 in Grade 5) against 48.5-56.5 (in Grade 3) and 49.0-50.4 (in Grade 5) for those who never experience bullying. This level of achievement is even lower than 54.0-51.9, 45.7-43.7 (for Grade 3) and 47.2-45.0, 46.2-49.4(for Grade 5) for those who encounter some kind of bullying of more than two. As the phenomenon is seen to be affecting the learning outcomes for many of the students encountering bullying, it is felt essential to root out it from our schooling to promote learning friendly atmosphere in school.

XII. Effect of utilizing beyond school hour time in achievement

Achievement level of students varies widely depending upon the amount of time they spent beyond school hours for household chores, paid work and other activities like watching TV, playing games chatting with friends. Dataset shows that spending times only in playing games, playing and chatting with friends, working for more than 4 hours a day and passing times in entertainment has negative effects on achievement whereas watching TV and engaging in households chores for 1-2 hours a day have a positive effects.

The achievement score in Nepali rises to 57.8 and 57.4 (Grade 3), 51.0 and 51.8 (Grade 5), 53.8 and 54.4 (English) for those spending less than an hour or 1-2 hours a day in watching TV which lowers down to 53 and 42 (Grade 3), from 45.9 and 40.1 (Grade 5) and 49.5 and 39.5 (in English) for those who spend much time more than four hours a day in it. It is interesting that the achievement is remarkably lower (48.8 for Grade 3 Nepali, 44.6 for Grade 5 Nepali and 40.4 for English) for those who spend no time in watching TV. Similar trend is noticed in Maths too in both Grades.

Similarly, the achievement is lower (49 in Grade 3, 46 in Grade 5 Nepali and 40.4 in English) than the national average for the students who do not work at all at home in comparison to (57-53, Grade 3; 48- 50.6 Grade 5 in Nepali; 49- 54.4 in English) those who works for 1 to 2 or more hours a day for household chores. It is notable that as the amount of time for household chores increases to more than four hours, the achievement is seen to be falling down to 47.1 (Grade 3) and to 44.2 (Grade 5) in Nepali 39 .4 in English. From the given facts, one can draw that involvement of children in decent level of household chores for a minimum amount of time that is, 1-2 hours a day, serves a part of socialization process that does not affect negatively in students learning rather gives some positive effects.

Though child labor is prohibited by laws in Nepal, the children studying at primary Grades are reported to have been working for paid job of any form. It is reported that 11 to 13.1 percent students studying in Grade 3 and 12.7 to 16.5 students studying in Grade 5 are involved in paid job of any kinds working for more than an hours a day. Data shows that the students not involved in paid job achieve (49.3-53.1 in Maths, 51.0-57.3 in Nepali) above the national average which is higher than (41.5 in Maths, 48-48 in Nepali) for the students being involved in paid job of any kind even less than an hour. Those working for more than four hours a day achieve the lowest (23-40.1) of all. Working for paid job or involvement of them for more than four hours a day indicates poverty of the family requiring the children to work for earning for pocket money or for livelihood at a minimum subsistence level. So apart from banning the child labor by laws, something more is also required to prevent children from working paid job to address this structural problem persisting within Nepalese socio-economy.

XIII. Stagnant learning achievement over the years

To compare the average achievement in each subject for 2015 with that of previous assessment of 2012, it is seen to be stagnant over the years in all subjects. For instances, the average achievement score in Nepali was 63 (Grade 3) and 60 (Grade 5) in 2012, whereas it has fallen down to 52 (Grade 3) and 46 (Grade 5) by more than 11 to 14 in this assessment. In Maths it has fallen down by 5 to 15 percent points remaining at 45 (Grade 3) and 48 (Grade 5) in 2015 against 60 (Grade 3) and 53 (Grade 5) in comparison with 2012. The same is the trend in English too which has fallen down by 7 percent point from 54 percent in the earlier to 47 percent now.

To look at the content wise results of each subject, the achievement level in *Reading* has fallen down by 13(Grade 3), 17 (Grade 5) in Nepali and 4 percent points in English coming down to 51, 39 and 46 respectively in this assessment from 64, 56 and 50 in the earlier. Similarly in *Writing*, it has fallen down to 42 (Grade 3), 33 (Grade 5) in Nepali and 40 in English from the earlier 54, 58 and 49 with difference by 12, 25 and 9 respectively. Likewise in *Grammar* too, the same level of decrease is evident in both Grades and subjects which has come down to 58, 38 in Nepali and to 53 in English from 65, 64 in Nepali and 56.5 in English by 7,27 percent points in Nepali and 3.5 percent points in English. Exceptions are the 3rd Graders in *Vocabulary* (Nepali) who have improved their achievement by 16 percent points achieving 76 in this assessment against 60 in the earlier which has lowered among 5th Graders from 70 (Nepali) and 62 in 2012 below to the 52 and 57.8 in this assessment. With regards to Mathematics, the achievement level does not rise up to the level of earlier assessment, for example the earlier average score reaching to 61 (Grade 3) and 54 (Grade 5) in *Arithmetic* has come down to 45 and 51 respectively in this assessment. Likewise, it has fallen down to 51 and 49 in this assessment from 60 in the earlier. In *Algebra* the 3rd Graders have raised their average score (by 17 percent points) to 57 in this assessment from 40 in the earlier one whereas it has fallen down by 3 percent points from 49 below to 46 for 5th Graders in this assessment.

Achievement level is seen to be stagnant also in each level of cognitive domains. In Nepali, it has lowered down by 4 to 9 percent points for 3rd Graders and by 6 to 14 percent points for the 5th Graders Compared to the earlier assessment. In comparison to the previous assessment, the highest decline is noticed in *Application* and *Higher* ability by 14 and 11 percent points lower for 5th Graders followed by in *Knowledge* and *Comprehension* by 7 and 9 percent points for 3rd Graders. In Maths, it has lowered down by 19 percent points in *Comprehension* and by 17 percent points in *Application* and *Higher* ability for Grade 3 whereas followed by 15 percent points in *Application* for Grade 5. In the rest, the decline is not so big, which has lowered down by 3 percent points in *Higher* ability and by 8 and 7 percent points respectively in *Knowledge* and *Comprehension*.

While comparing the results in terms of latent ability (θ), the 3rd Graders have improved their ability to $\theta = 0.236$ in 2015 in comparison to the earlier $\theta = 0.11$ in Maths. In case of 5th Graders, it has lowered slightly down to $\theta = 0.003$ in this test against the earlier $\theta = 0.008$. From this fact, it is evident that the latent ability has also remained stagnant (See, ERO, 2015b for further details).

XIV. Nepalese students lag behind the international standards in achievement

When compared the achievement level of Nepalese Grade 5 students with that the achievement level of Grade 4 students participating in international tests like TIMSS and PIRLS, Nepalese 5th Graders are found either at the level of international Grade 4 or even below that. In Math, Nepalese 5th Graders have been able to achieve $\theta = 0.391$ in total (ranging from the lowest $\theta = 0.388$ in *Geometry*, $\theta = 0.390$ in *Arithmetic*, $\theta = 0.391$ in *Algebra* and highest $\theta = 0.396$ in *Numeracy*) above than the average achieved by international students of Grade 4 in TIMSS, indicating that they are just above their international Grade 4 colleagues. However, this level of achievement is not considered enough to be equal with the international 5th Graders.

On the contrary in Nepali, the achievement level of 5th Graders are seen to be lagging far behind the international 4th Graders participating in PIRLS test by $\theta = -1.08$ points below. In comparison to the PIRLS for Grade 4, they are below by $\theta = -1.08$ in *Reading*, by $\theta = -1.00$ in *Writing* and by $\theta = -1.00$ in *Grammar* and *Vocabulary*. This level of achievement of 5th Graders hints that they lag behind even the international Grade 4 students falling off two years or more behind in learning to reach the level of their international peers. The Same level of underachievement was noticed in the previous assessment too. This level of achievement is not justifiable in any grounds so requires strong emphasis on support for these students.

XV. Persisting inequality in learning level among the regions

Disproportionately distributed achievement level between and among geographical and Development Regions persisted in the previous test has continued in this test too perpetuating a wider inequality among them. In earlier test, the varied level of achievement gap, remaining at 53-57 in Nepali, 49-55 in Math and 46 in English for the Tarai–the lowest of all the Ecological belts and reaching to 78-80 in Nepali, 68-77 in Maths and 79 in English– the highest of all for the Valley, has also continued in this test too from 43-49,45-49 and 47– the lowest for Tarai and 57-

64, 51-58 and 68 the highest of all for Valley, showing a wider inequality in learning opportunity for children between them. Similarly, the earlier achievement gap with 21-24 percent points in Nepali, 5-7 percent points in Maths and 2 percent points in English between Mountain and Hill with the Valley by 20-21 percent points in Nepali, 17-20 percent points in Maths and 30 percent points in English have remained more or less at the same level in this test too. In comparison to the earlier in when the Mountain was performing better than Hill with difference by 1-7 percent points higher, in this test the Hill is seen to be outperforming the Mountain with differences by 2-5 percent points higher.

While compared the Development Regions, the Western Region followed by the Far Western outperformed the rest with significant differences in the earlier. In this test too, the Western region has been able to maintain its previous position- the best performing of all while Far Western has fallen off down to among the lower achieving Regions lagging behind the Central. In this test, the Eastern Region has recorded the second highest achieving Region improving its position from the lowest achieving one in the earlier. In comparison to the earlier, Mid Western has consistently remained among the low achieving after the Central Region.

Though the achievement gap between rural and urban schools is found to have reduced to some extent, still rural schools lag behind. In the earlier, the gap between the two ranged from minimum 15 percent points in Mathematics, 16-18 percent point in Nepali to maximum 25 percent points in English, it has reduced to 8 percent points in Mathathematics for Grade 5 and 9-17 in Nepali against the earlier. Rather, the gap has widened in this test further to 17 percent points in Grade 3 Mathematics from the previous 15 percent points and to 26 percent points in English against the earlier 25 percent points.

Given the context a couple of points are worth noting. First, there is wider disparity in learning level among the children because of the geographical, regional and locational diversity. Second, inequality in learning opportunity is continuously persisting over the years among the different geographical and Development Regions including locations across the country that limits the children's opportunity to learn for reaching the preset goals of curriculum. Third, these low performing children concentrate in certain pockets particularly in rural settlements of remote and outreach areas of the country always remaining in a disadvantage circumstances requiring concerted interventions for improvement. Fourth, the wide differences in achievement not only reflect inefficiency of the education system but also wider social inequality which has been perpetuating in the nation.

8.3 Implications

The low level of learning perpetuating over the years which is compounded with a wide gap in achievement among and between rural-urban, community-institutional schools, caste/ethnicity, social groups across all Developmental Regions, Ecological belts and Districts are neither a good sign for equality and nor the positive indication for the system. Similarly, achievement variations of students among various subjects, content areas within the subjects, and poor achievement in the domain of application and higher ability urge for the systemic improvement in educational delivery process and practices. Given the context, the following would be the main implications for the system to improve the achievement level of Nepali students.

I. Reducing inequality in achievement

As the result confirms a wide inequality in achievement level, which continues to persisting over the years, between students from rural and urban locations, among various language speaking and ethnic/castes groups, and across Districts, Regions and Ecological belts; it has been an imperative for the policy makers, curriculum planers, teacher educators as well as education managers to look for the ways for enhancing the capacity of current delivery system to ensure an equal level of learning opportunities for all children irrespective of caste/ethnicity, social and language groups, family in which one is grown up and the school types on attends. Although a lot of efforts have been put into the system to reduce the inequality, the persisting gap is still demanding further measures that actually reduce the inequality in practice. In addition, reason behind the low performance of students from community schools and the students of particular communities should be explored with micro level studies going deep down into the root causes. Simultaneously, discussions need to initiate with the teachers of low performing schools, parents of the low achieving students on reducing the inequality and improving low level of achievement that is persisting.

The causes of persisting inequality in achievement between the students from community and institutional schools would be another area to further explore whether it is due to the teacher effect or more rigorous teaching learning practice at institutional school, or socio-economic background of the family that provides students attending institutional schools additional supports. If the rigorous teaching learning practice and caring environment available at private school is found only the contributing factor to raise achievement, then educational managers and policy makers need to find out the ways to encourage community school teachers for providing caring environment and rigorous teaching learning practices. One of the measures to raise and reducing achievement gap would be introducing the performance based incentives to the schools and teachers for raising achievement and reducing achievement gap of the low performing students through benchmarking and setting the target to achieve in the given timeframe against the target given.

II. Ensuring balance learning in all content areas across the subjects

Besides problems of low achievement level, Nepalese students are not found to have developed similar levels of knowledge, skills competencies equally over all content areas within and across subjects. Little learning in one area of subject impedes to acquire expected level of proficiencies in other areas leading ultimately towards underachievement in all content areas within and across all subjects. As shown by the results, *Reading* and *Writing* in Nepali and English, *Arithmetic, Algebra* and *Numeracy* in Mathematics are poorly learned areas in comparison to others.

Given the context of such unequal level of learning, curriculum planners of the respective subjects first need to give serious consideration over the little learned contents in order to find the answer to the questions as to why students are not able to learn them and how curriculum planning and designing in given contents would be restructured in order to enable learning. Similarly, root sources of low level of learning in the given areas and innovative ways to facilitate students learning in those contents also need to be explored with further researches and studies in the days to come. Existing teacher training courses and packages also require revisiting in order to sensitize and prepare teachers for further facilitating students learning in the areas identified.

III. Improving reading ability

As confirmed by the dataset reading proficiency of Nepali students is seen poor, which has not improved over the years. The low level of reading proficiency among the students has kept them in a weak state of comprehending implied meaning, solving complex problem, abstracting of deeper ideas, producing open-ended text not only in Nepali subjects but also in other subjects like Mathematics and Science leaving them less able to perform the tasks demanding higher cognitive ability. Such state of reading proficiency necessitates developing new instructional design for classroom practice with more reading activities, exercises on comprehension, introducing varieties of text with different genres by spelling out specifically the standards on fluency, accuracy, performance and problems solving in curriculum and textbooks across the subjects. Another measure for fostering reading ability would be determining certain amount of texts to be read by students for each Grade in addition to the text from the textbooks. The curriculum planners, textbook writers, teacher educators as well as the classroom teachers need to sensitized, oriented and trained in such instructional design. Allocating additional time for reading activity in school hours through curricula provision would also be one option for which the CDC needs to initiate discussions and dialogues with curriculum planners, textbook writers as well as with teachers. Similarly, ongoing teacher education courses and packages also need to be redesigned incorporating required skills and competencies for teacher to organize reading skill promoting activities in classroom.

In addition, existing student assessment and examination practice needs to be revisited incorporating the reading tasks in assessment activities across all subjects. Generally, existing student assessment and examination practice tend to either ignore or give less emphasis to assess skills on listening, speaking, oral presentation, reading fluency and accuracy. So, teachers are also likely to be reluctant on such skills. Student assessment frameworks and guidelines also need to be revised incorporating tasks and activities to assess the skills.

IV. Fostering higher cognitive ability

Although cognitive skills are seldom taught explicitly in schools, various researches indicate that schooling through teaching knowledge and skills in language, Mathematics and science need to promote cognitive ability. In this regard, fundamental goal of education is to equip students to think critically, solve complex problems and succeed in the society and economy of 21st century (Fin et al, 2014). The subject wise datasets reveals that Nepalese students are good at lower level of cognitive skills such as knowledge and recognizing, but found poor across all subjects in higher level ability such as application, analysis, synthesis and evaluation required for solving the novel problems, information processing and applying the knowledge and skill learned one context into the new one. Provided situation leads to imply that either our teaching learning activity deviates from organizing activity to engage in students in the tasks requiring higher ability or reading materials including textbooks lack adequate exercises or practice that foster given ability.

Our education system needs to be well aware of the reasons behind why schooling is not promoting desired higher cognitive skills among eighth Graders even after completing eight years school education. Curriculum planners and textbook writers need to pay enough attention to address the issue of low performance level in higher ability tasks and initiate discussion on how to design new curricula and textbooks so as to keep engaged students in tasks demanding higher mental ability. Similarly, teacher educators and training modules designers also need to look for the possible measures to incorporate skills and competencies required for teachers in organizing classroom activity to foster higher ability among students. Teacher and test item writers too are to be reoriented on developing assessment tools to assess these skills. Future researches and studies need to concentrate their focus on this issue to find out the lacking part whether it is because of curricula or classroom practices or teacher preparation.

V. Improving student assessment system

Modern learning theory conceives student assessment for learning rather than assessment of learning, which implies that assessment should focus on both assessing the processes of learning i.e. inquiring, independent learning, use of generic skills, reflections and the products of learning e.g. knowledge/concepts, problem solving capabilities adopting different varieties of methods such as oral test for oral communication, discussion for collaboration, presentation/performance for creativity, tests and examination for knowledge etc. However, the datasets show that a notable number of students are found to have been not used to answering open ended, problem solving and application types of subjective items. Given circumstances lead one to conclude that our assessment practice and test items have not engaged students in solving novel problems, producing creative works, dealing with open ended tasks which demands an immediate reform to make it able to assess both the process and products. For this to happen, whole assessment mechanism and practice from classrooms at the lower unit to public level examinations at the national level need to overhaul reform.

To bring reform in the classroom assessment practice, teachers require desired capacity in designing and using various assessment tools including standardized test items against learning objectives set in the curricula. Schools should be reoriented on the tools and process of assessing students learning process and products. Furthermore, existing ceiling of 32 percent marks to pass the Grade and level wise examination should also be raised at least to 50 percent to raise the expectation and study habits among the students. At the same time, public examination at District and national level need to ensure that students are assessed by means of standardized test items, which corresponds the curriculum objectives set for to go through.

VI. Rooting out the incidences of school bullying

The phenomenon school bullying, as recognized in other parts of the world, is not so much familiar to parents and teacher so has gone unnoticed at schools; but it is found rampant in Nepalese schools in some forms negatively affecting their learning potentials for those children who experience. As results of this study confirm that it is found to have been associated with low level of learning for a notable percentage of school children (from minimum 4.5% experiencing severe forms to maximum 31% experiencing with 20% of bullying). Given situation alarms concerned teachers, education managers and even to the parents to be aware of the phenomenon and to expedite the possible preventive measures that help schools at least minimizing such incidences.

One of the measures for this would be making teachers aware of the issue and its possible consequences to the school children. Students themselves are to be sensitized with the phenomenon and its effects on their colleagues through mobilizing child clubs that are functioning at schools. Child right activists also would be supportive to lead the process. Similarly, child friendly school framework being implemented under the initiative of the DOE would also need to incorporate possible indicators regarding the measures that discourage school bullying. Curricula and packages on teacher training also need to consider these issues seriously in their future revision or repackaging.

VI. Raising parents' educational level

Low level of parents education is found to have been directly associated to the under achievement of their children at school which further has been the main sources of perpetuating social inequality and disparities. Despite the lots of effort put and investments made for raising literacy level of parents, youth and adult literacy rate of nation (84.72 in total with only 80.16 for women among youth population and 59.6 in total with only 48.8 for women among adult population respectively (CBS, 2013)) has not been improved to the desirable extent which also has resulted in the low level of learning of those children whose parents are illiterate.

The results of the study confirm that parents' education level in general and the mother in particular is found major determinant for their children achievement level at the school. So, raising parents literary level especially targeting to mothers is seen an urgent need not only for achieving higher level of students achievement at school but also for achieving social equality and a just society. One of the probable measure for raising parents literacy level would be devising a policy and program that encourage schools to make those illiterate parents literate residing within the catchment area and whose children are attending the particular school. Another measure would be devising and introducing, as implemented in some parts of the world (e.g. Mauritania, in some Latin American countries), the family literacy program through schools. Such provision would help raise literacy level of the parents and improve achievement level of their children on the one and develop strong school community relations that ultimately increases parental involvement in school development endeavors on the other.

VIII. Ensuring timely availability of textbooks to all students

Despite the government efforts for the delivery of textbook on time, the data reveals that a notable number of students are compelled to complete the Grade without textbooks. In this regard, educational managers from District to central level need to look for the further ways to strengthen the existing delivery mechanism and make school accountable for ensuring on time availability of textbooks to all students. One of the measures towards this would be enforcing the provision to earmark certain amount fund by each school every year or devising a mandatory provision to establish a book corner with certain set of textbooks in each. Similarly, designing multi-year usable textbooks could also be one of the measures for textbooks distribution. In this context, the existing mechanism of textbooks development, printing and distribution should be reviewed based on earlier studies or with some additional studies to ensure timely availability of textbooks with good quality. One of the options for ensures timely availability of textbooks with good quality is to introduce multiple textbooks system. One of the crucial factors regarding the use of textbooks is enhancing teachers' capacity to use multiple materials including printed, electronic and online material in order to deliver effectively the curricular competencies.

IX. Catching up with international standards

Having compared the datasets of English, Mathematics and Nepali with the TIMSS and PIRLS standards, Nepalese students are seen 1 to 1.5 years behind their international counterparts of the same Grade level which will likely handicap them to be competitive in global world of job market. Taking the issue into consideration, further researches and studies need to concentrate in exploring the reasons behind it as whether it is due to curricula or classroom practice. Similarly, the MOE also needs to initiate dialogues and discussions with academia, curriculum planners and policy makers on how to raise the performance of Nepalese students to the international level in order to make them competitive in the global world.

Summing up

This study has not only determined the achievement level of primary Grades in each of three subjects, but it has also analysed variations in student achievement in terms of Development regions and Ecologicalbelts, gender, location and type of school, language of instruction, caste ethnicity of students, socio economic, backgrounds and activity at schools. While doing so, it has also compared the achievement of Nepalese students with international level tests like TIMSS and PIRLS.

Having analysed the results, the dataset shows some gaps within the system. First, there is a wide inequality in learning opportunity for children between regions, types of schools and their location, regions and Ecologicalbelts, language groups, caste and ethnicity. Second, inequality is also seen in learning of contents and domains across the curricula. Third, there is low level of proficiency in higher order skills in all the subjects. Moreover, it has also confirmed that the achievement of Nepalese students is lower than international standard lagging 1 to 1.5 years behind their international colleagues while compared to the TIMSS and PIRLS results. Given the context, further reform strategies and programs require to concentrate towards narrowing inequality gaps between regions, Districts, locations, social and language groups; enhancing higher level cognitive skill; and raising the standard of Nepalese education to the international level.

The study has generated rich data on different variables which do not only serve the purpose of bench marking for future but also provides numerous insights for policy planning, program designing and identifying the areas for reform. As it has generated micro level data of each individual school, specific support and intervention can easily be identified and targeted to the poor performing schools in order to support them raise the standard. More research is needed not only on the processes and practice of schools but also on how the achievement levels have changed over time. Moreover, the dataset can also be used as a basis for generating knowledge on educative process of best and poor performing schools for comparison.

References

Adams, R. & Wu, M. (2002). *PISA 2000 Technical Report.* Paris: OECD Publishing.

American Psychological Association (APA). (2007). *Report of the APA Task Force on Socioeconomic Status.* Washington, DC: Author.

Béguin, A. (2000). *Robustness of Equating High-Stake Tests*. Enschede: Febodruk B.V.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956).*Taxonomy of Educational Objectives: The Classification of Educational Goals,Handbook I: Cognitive Domain.* New York: Longmans, Green.

Boreman, G. D., Hewes, G.M., Overman, L.T. & Brown, S. (2003). Comprehensive school reform and achievement : a meta-analysis. *Review of Educational Research, 73(2):125-230*.

BPEP. (1995). *Basic and Primary Education Project Achievements Study*. Kathmandu: Author.

----- (1997). *The Effect of New Curriculum on the Achievement of Grade 4 Student*. Kathmandu: Author.

----- (1998). *The Effect of New Curriculum on the Achievement of Grade 5 Student*. Kathmandu: Author.

Bradley, R. H. & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53*: 371–399.

Brooks-Gunn, J. & Duncan, G. J. (1997). The effects of poverty on children. *Future Child, 7(2)*: 55–71.

Bryk, A. S. & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin, 101(1)*: 147–158.

Central Bureau of Statistics (CBS). (2012). *National Population and Housing Census 2011.* Kathmandu: Author.

----- (2013).*National Population and Housing Census 2011. (National Report)*. Kathmandu: Author.

CERID. (1993). *Performance Levels of Public and Private Primary Schools. A Comparative Study.* Kathmandu: Author.

----- (1998). *Performance Levels of Grade V Students.* Kathmandu:Author.

----- (1999). *Assessment of Learning Achievement of Lower Secondary Children: Grades 6 and 8.* Kathmandu: CERID and MOES.

CERSOD. (2001). *District Level Assessment, 2001.* Sanothimi:DOE.

Cochran, W. G. (1977). *Sampling Techniques* (3rd ed.). New York: Wiley & Sons.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. (2nd Ed.).Hillsdale, NJ: Erlbaum.

Curriculum Development Center (CDC). (2063 B.S.). *Primary Education Curriculum Grade 1-3.* Sanothimi: Author.

Department of Education (DOE). (2014). *Flash I Report 2071(2014/2015)*. Sanothimi: Author.

DeVellis, R. F. (1991). *Scale Development: Theory and Applications (Applied Social Research Methods Series, Vol. 26)*. Newbury Park: Sage.

EDSC. (1997). *National Assessment of Grade 3 Students.* Kathmandu *:* BPEP.

----- (1999). *National Assessment of Grade 5 Students.* Kathmandu: BPEP.

----- (2001). *National Assessment of Grade 3 Students.* Sanothimi: DOE.

----- (2003). *National Assessment of Grade 5 Students.*Sanothimi: DOE.

----- (2008). *National Assessment of Grade 8 Students.* Sanothimi*:DOE.*

Education Review Office(ERO). (2013). *Where are We Now? Student Achievement in Mathematics, Nepali and Social Studies in 2011*. Kathmandu: Author.

-----(2015a). *National Assessment of Student Achievement (NASA) 2013 (Grade 8: Mathematics, Nepali and Science*. Sanothimi: Author.

-----(2015b). *National Assessment of Student Achievement(NASA)2012 (Grade 3 and 5)*. Sanothimi: Author.

Fennema, E. & Sherman, J. A. (1976). Fennema-Sherman mathematics attitude scales: instruments designed to measure attitudes toward the learning of mathematics by females and males. *Journal for Research in Mathematics Education, 7(5)*: 324- 326.

Fink, A. (1995). *The Survey Handbook*. Thousand Oaks, CA: Sage Publications.

Finn, A.S., Kraft, M.A., West, M.R., Leonard, J.A., Bish, C.E., Martin, R.E., Sheridan, M.A., Gabrieli, C.F.O & Gabrieli, J.D.E. (2014). Cognitive Skills, student achievement tests and schools. *Psychological Science, 25(3):*736-744.

Fullbright Consultancy. (2008). *A Final Report on National Assessment of Grade V Students.* Sanothimi*:* DOE.

Hanushek, E.A. & Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review, 100(2)*: 267-271.

Hair, J. F. Jr., Anderson, R. E., Tatham, R. L., & Black. W. C. (1998). *Multivariate Data Analysis,* (5th Ed.). Upper Saddle River, NJ: Prentice Hall.

House, J. D. & Telese, J. A. (2008). Relationships between student and instructional factors and algebra achievement of the students in the United States and Japan: an analysis of TIMSS 2003 data. *Educational Research and Evaluation*,*14(1)*:101- 112.

Invernizzi, M. A., Landrum, T. J.,Howell, J. L. & Warley, H. P. (2005). Toward the peaceful coexistence of test developers, policymakers, and teachers in an era of accountability. *The Reading Teacher, 58(7)*: 610–618.

Kadijevich, Dj. (2006). Developing trustworthy TIMSS background measures: a case study on Mathematics attitude.*The Teaching of Mathematics*, *9(2)*: 41-51. Available at: http://elib.mi.sanu.ac.yu/journals/tm/17/tm924.pdf.

----- (2008). *TIMSS 2003: Relating Dimensions of Mathematics Attitude to Mathematics Achievement.* Available at:

http://www.ipisr.org.rs/Dokumenta/zbornici%20PDF/Zbornik%202008%2040- 2/Zbornik-08-2-006.pdf.

Krieger, N., Williams, D. R & Moss, H. W. (1997). Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annual Review of Public Health, 18*: 341–78.

Liberatos, P., Link, B. G., Kelsey, J. I. (1988). The measurement of social class in epidemiology. *Epidemiology Review,10*: 87–121.

Lord, F.M. (1980). The relationship of the reliability of multiple-choice test to the distribution of item difficulties. *Psychometrika,18*: 181–194.

Martin, M. O., Mullis, I. V. S, & Kennedy, A. M. (2006). *PIRLS 2006 Technical Report.* Chestnut Hill,MA: TIMSS& PIRLS International Study Center: Lynch School of Education, Boston College.

Matsuura, K. (2004). Foreword. In *UNESCO: EFA Global Monitoring Report 2005. Education for All. The Quality Imperative.* Paris:UNESCO.

McLoyd, V. C. (1997). The impact of poverty and low socioeconomic status on the socio emotional functioning of African-American children and adolescents: mediating effects. In R. D.

Taylor& M. Wang (Eds.), *Social and Emotional Adjustment and Family Relations in Ethnic Minority Families, 7-34*. Mahwah, NJ: Erlbaum.

Metfesser, N., Michael, W. B., & Kirsner, D. A. (1969). Instrumentation of Bloom's and Krathwohl's taxonomies for the writing of behavioral Objectives. *Psychology in the Schools, 6*: 227–231.

Metsämuuronen, J. (2009). *Methods Assisting the Assessment*. Oppimistulosten arviointi 1/2009. The Finnish National Board of Education. Helsinki: Yliopistopaino. [In Finnish].

----- (2012a). Comparison of mental structures of eighth-graders in different countries on the basis of Fennema-Sherman test. *International Journal of Psychological Studies,4(4)*.

----- (2012b). Challenges of the Fennema-Sherman test in the international comparisons. *International Journal of Psychological Studies, 4(3).*

Metsämuuronen, J., Kuosa, T. & Laukkanen, R. (2013). Sustainable leadership and future oriented decision making in the educational governance–a Finnish case. *International Journal of Educational Management, 27(4)*. http://www.emeraldinsight.com/journals.htm? Articleid =17083274.

Mullis, I. V. S., Martin, M. O., & Foy, P. (with Olson, J. F., Preuschoff, C., Erberber, E., Arora, A. & Galia, J.) (2008). *TIMSS 2007 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at Fourth and Eighth Grade.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Available at: http://pirls.bc.edu/timss2007/mathreport.html.

Nunnally, J. C. (1978). *Psychometric Theory (2nd ed.)*. New York: McGraw-Hill.

Oaks M. (2011). Measuring Socio-economic status. In, e-Source. *Behavioral & Social Science Research*. Avaialble at:

http://www.esourceresearch.org/Portals/0/Uploads/Documents/Public/Oakes_FullChapter.pdf.

PEDP. (1998). *Primary Education Development Project: An Assessment.* Kathmandu. Author.

Phyak. P. (2012). *Language Issues in Educational Policies and Practices in Nepal: A Critical Review.* Kathmandu: Asian Development Bank.

----- (2006). *PISA Released Items-Reading* . OECD Publishing.

----- (2009). *Take the Test. Sample Questions from OECD's PISA Assessments.* OECD Publishing.

Race, P., Brown, S. & Smith, B. (2005). *500 Tips on Assessment*, 2nd *Edition*. London: Routledge

Falmer.

Shen, C. & Tam, H. P. (2008). The paradoxical relationship between student achievement and self-perception: a cross-national analysis based on three waves of TIMSS data. *Educational Research and Evaluation, 14(1)*: 87—100.

TIMSS. (2007). *TIMSS 2003 Science Items. Released Set, Eight Grade.* Chestnut Hill,MA: IEA, TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

----(2009a). *TIMSS 2007 User Guide for the International Database. Released Items, Mathematics – Eight Grade.* Chestnut Hill, MA: IEA, TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

----(2009b). *TIMSS 2007 User Guide for the International Database. Released Items,Science – Eight Grade.*Chestnut Hill, MA: IEA, TIMSS & Pirls International Study Center, Lynch School of Education, Boston College.

Verhelst, N. D., Glas, C.A.W., Verstralen, H. H. F. M. (1995). *One-Parameter Logistic Model (OPLM)*. Arnhem: Cito.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*: 461–81.

Wright, B. D. (1968). Sample-free test calibration and person measurement. *Proceedings of the 1967 Invitational Conference on Testing Problems.* Princeton NJ: Educational Testing Service.