National Assessment of Student Achievement 2018

MAIN REPORT

(Report on the National Assessment of Student Achievement in Mathematics and Nepali for Grade 5)

Government of Nepal
Ministry of Education, Science and Technology
Education Review Office (ERO)
Sanothimi, Bhaktapur

2019 (2076 BS)

Main Report

(Report on the National Assessment of Student Achievement in Mathematics and Nepali for Grade 5)

Written by

Prof. Dr. Basudev Kafle with Shyam Prasad Acharya and Deviram Acharya

Data analysed by

Dr. Bishnu Khanal with Shyam Prasad Acharya, Deviram Acharya

Language Edited by

Prof. Dr. Basudev Kafle and Ms Kunti Adhikari

Government of Nepal
Ministry of Education, Science and Technology
Education Review Office (ERO)
Sanothimi, Bhaktapur

2019 (2076 BS)

National Assessment of Student Achievement 2018: Main Report

(Report on the National Assessment of Student Achievement in Mathematics and Nepali for Grade 5)

This main report of NASA 2018 is written by Prof. Dr. Basudev Kafle with Shyam Prasad Acharya and Deviram Acharya
Data analysed Dr. Bishnu Khanal with Shyam Prasad Acharya, Deviram Acharya

Year of Publication: 2019 (2076BS)

Contact:

Education Review Office
Sanothimi, Bhaktapur, Nepal
Phone: +977-1-6636518, 6634362
Email: eronasa@gmail.com
Website: www.ero.gov.np

Printed at:

FOREWORD

National Assessment of Student Achievement (NASA) communicates the status of student achievements and suggests the measures for improving the students' learning outcome. The assessment provides evidence to the policy makers to formulate practical and implementable educational policies in national and sub-national level.

In Nepal, the Education Review Office (ERO) started NASA since 2011. The first national assessment was carried out in Grade 8 for Nepali, Mathematics and Social Studies. In later years, Grade 3, 5 and 8 assessments have been conducted in a periodic basis. During the School Sector Reform Plan (2009-2015), two rounds of assessment (Grade 3, 5, and 8) were administered. During the School Sector Development Plan (2016-2022/23) two rounds of assessment (Grade 5, 8 and 10) will be administered. NASA 2018 is the first and baseline assessment administered during the SSDP period for grade 5.

NASA is a curriculum-based systematic evaluation of student learning outcomes. An analysis of the curriculum to develop an assessment framework is the first step while carrying out this assessment. This framework works as an assessed curriculum that helps to operationalize the abstract learning objectives into practical sense.

This is the main report of NASA 2018 for Grade 5 in Mathematics and Nepali subjects. The assessment was conducted in the national representative sample of 28381 students from 1400 schools of Nepal with an almost equal number of schools and students in each of the three subjects, considering seven provinces as explicit strata. Three versions of standardized tests together with the background information questionnaire to the sample students, teacher questionnaire to subject teachers and school survey questionnaire to the head teachers were administered in each school. Data were analysed in such a way that along with presenting overall mean score and proficiency levels, the relation between the achievement scores and various influencing factors was demonstrated with the use of the background information questionnaire. Analysis and comparison of the results were done using the Item Response Theory (IRT) and the parameters of linking items. Results are presented in a transformed scale of student latent ability (θ) with 500 mean and 50 standard deviation. The results presented in this report are the generalized results over the defined population and they provide the evidence of the level of learning. Generalization of the results was possible with the use of the multi-stage probability sampling, that is Probability Proportional

to Size (PPS) sampling method, which is widely appreciated method in educational research.

Throughout the process of tools development, test administration, data analysis and report writing, several stakeholders, teachers, experts and researches have contributed in different many ways. I acknowledge the collaboration and participation of the subject teachers, experts, subject committee members, and assessment committee members. My sincere thanks go to the previous ERO head Dr. Lekhnath Paudel, undersecretary Mr Hemraj Pokhrel and Prahlad Aryal, Mathematics subject committee chaired by Professor Dr Hari Upadhyaya, Nepali subject committee chaired by Professor Dr. Parasmani Bhandari. This report appeared in this form after the rigorous efforts of the officers at ERO Mr Shyam Prasad Acharya and Mr Deviram Acharya who have been heavily involved in the research tasks from tools development and working for report writing. I would like to thank the consulting firm, Centre for Educational Research and Social Development as well as Council and Council for their contribution. I am grateful to all the ERO staffs including undersecretary Mr Hari Aryal, Mr Uttar Kumar Parajuli and the officers Mr Prakash Kumar Kharel and Mr Lavdev Bhatta who were directly and indirectly involved in various phases of this assessment.

I highly appreciate the contribution of Professor Dr Basudev Kafle and Technical Officer Ms Kunti Adhikari for the language edit. I would like extend my thanks to the personnel from various Central Level Agencies and the Ministry of Education, Science and Technology for their contribution during the program and budgeting phase, monitoring of test administration and tools development. I acknowledge the role of the World Bank in supporting the capacity development of the human resource involved in this large-scale assessment.

I commend this report to education policy makers, programme designers, teachers, educators and community members and researchers for their reference in improving students' learning. I hope this report will be a milestone for improving quality of education at school level in Nepal.

Mr. Tek Narayan Pandey
Director General
Education Review Office

EXECUTIVE SUMMARY

Context

In the beginning of 2018, the Education Review Office assessed the learning outcomes of grade 5 in Mathematics and Nepali. The prime objective of this assessment were to prepare the baseline data for School Sector Development Plan (SSDP) and compare the learning achievement of 2018 with the previous cycle of NASA (2015) to ensure quality education in school system. Altogether 28381 students, 1400 teachers, 1400 head teachers from 1400 schools participated in this assessment. National assessment has been well accepted as a means of measuring quality of education (TIMSS & PIRLS, 2008) that provides both quantitative and descriptive form of information on student achievement which after is considered as an output of the teaching learning process and its quality (World Bank, 1996). It provides basic information for policy makers, politicians, and the broader educational community and informs policy makers about the key aspects of the system" (Greaney & Kellaghan, 2008b, p. 7, ERO, 2013). In this context, ERO has its roadmap to conduct two round NASA of grade 5, 8 and 10 to assess and ensure the quality of education and trends of learning achievement within SSDP period. This NASA 2018 is the first cycle assessment of grade 5 of Mathematics and Nepali subjects in SSDP where as it is a third cycle after the establishment of ERO in 2010.

Objectives of NASA 2018

The main aim of NASA is to provide policy feedback through the assessment of learning and identify the trends of learning over the time. NASA 2018 has the following objectives:

a. To identify the current level of Grade 5 students' achievement in Mathematics and Nepali,
b. To identify variations in student achievement by gender, province, identity with geography, types of school, ethnicity, home language, socio-economic status,
c. To explore the factors that influence student achievement,
d. To identify trend in student learning and produce the baseline data for comparison in the future,
e. To strengthen the capacity of the education system in conducting national assessment,

f. To provide the Ministry of Education, Science and Technology with recommendations for policy formulation to improve quality and ensure equity, particularly in school education.

Methodology

Three set of questions with background information questionnaire were asked in each subject. All sets were linked with anchor items. The ERO has used Item Response Theory to assess the latent ability of students using various contextual variables to explain those latent traits of the students. NASA 2018 has used advanced procedure to bring rigor to data analysis by generalizing the results in national level and province levels through 7 explicit strata and various other implicit strata. Use of Replicate Module for estimating the population parameters and Weighted Likelihood Estimation (WLE) for analysis of individual student level and reporting are the instances of the advancement. Furthermore, the advancement of procedures has also been noticed in the sampling methods. A Probability Proportional to Size (PPS) sampling procedure has been used for selecting the schools as Principal Sample Unit (PSU), the school clusters. Student achievement at province level and national level is reported in a transformed scale with mean 500 and standard deviation 50 by using the formula:

*Average scale score = 500 + plausible value * 50*

*or, Average scale score = 500 + logit * 50*

Assessment results have shown that the national average achievement is 500 in both subjects. However, it does not mean that both subjects have been equally learnt. This report presents the results in terms of what the students can and cannot perform, the existing gap between the written curriculum and the achieved curriculum, and the number of students who have developed their ability in a minimum competency level.

Major Findings and Recommendations

1. **Huge mass of students is at the underperforming level:** As the study indicates, 32 out of 100 students fall below *basic level* (Pre-basic) in Mathematics achieving only 5% of the tested curriculum and the *basic level* (level 1, about 40%) students have achieved only 28% of the tested curriculum. More than 70% students have achieved only below 28% of the tested curriculum in Mathematics indicating a huge mass of students underperforming in this subject. The *proficient level* (level 2, 24%) students achieved 62% of the tested curriculum and *advance level* (level

3, 4%) students achieved 96% of the tested curriculum. It seems that only 28% of the students have adequate knowledge and skills in Mathematics curriculum. The gap in the achievement of curriculum between below basic level (5%) and advance level students (96%) is 91% indicating remarkably high inequality in the classroom. Similarly, in Nepali, 20% students who are at below basic level (Pre-basic) achieved only 18% of the tested curriculum and 35% students achieved only 38% of the tested curriculum in Nepali, who are at basic level (level 1). As 30% students fall in proficient level and 15% in advance level, altogether 45 of 100 students have adequate knowledge and skills of the tested curriculum in Nepali. This shows that 55% of the students, a big mass of the students in Nepali, represent the underperforming group. And 45 of 100 students have adequate knowledge and skills of the tested curriculum as 30% students fall in proficient level and 15% in advance level. Altogether, 55% of the students, a big mass of the students in Nepali, represent the underperforming group. The proficient level students achieved 60% of the tested curriculum and the advance level students achieved 88% of the tested curriculum. The gap in the achievement of the curriculum between below basic level (18%) and advance level students (88%) is 70% indicating high inequality in the classroom.

Recommendation: As basic level is assumed to be a minimum competency level, a campaign of "no child should be left behind basic level" should be initiated effectively to develop minimum competency level in the students. Teachers should provide the students with many opportunities to learn in many ways and through various means. Existing mis-match between the written curriculum and achieved curriculum urges the need to review the national curriculum, teaching methods, teacher motivation system, learning environment and the evaluation system.

2. **Huge gap between the provinces and districts has been noticed. Such gap increases disparity in learning achievement among the groups of students.**

Recommendation: All the community schools should provide equal opportunity to the students for learning. A minimum standard for physical infrastructure, learning opportunities, resources, incentives and retention of good teachers under teacher management should be set to bring uniformity in the achievement level of the students. Retention of good teachers has relation with increased learning achievement of students. Learning difficulties of students in all schools should be identified and then addressed by remedial teaching. Regular follow up support and monitoring mechanism should be strengthened to enhance learning.

3. **Students can perform better irrespective of their SES and home language:** The Socio-Economic Status of student's family has low effect in Mathematics and medium effect in Nepali Language. Many students have performed better despite their unfavourable low socio-economic status; this indicates that the socio-economic background of the students does not exclusively determine the learning achievement. Province 2 has majority of students with low SES, but they have achieved the higher position in Mathematics compared to other provinces. And similar is the situation with Nepali language as well.

   Recommendation: As students can perform better irrespective of their SES and home language, focus on the study and practice on the part of student is a principal measure to boost their learning achievement. In all community schools, minimum learning materials, library facility and students' clubs, numeracy promotion program and reading programs should be made available to promote students' performance.

4. **There is a minimum gap in learning achievement between boys and girls.**

   Recommendation: Reduced and minimum gap between boys and girls should be maintained with more focus on providing girls with equal opportunity. The existing gap can best be addressed through affirmative action such as scholarship, girls friendly environment, and receptive teacher behaviour.

5. **Wide gap between the type of schools has been noticed:** There is a wide gap in learning achievement between community and institutional schools. The students from institutional schools have out-performed the community schools with a gap of 30 scale score in Mathematics and 34 in Nepali.

   Recommendation: Upgrading of community schools to increase their academic performance should be initiated as a regular targeted intervention by the government. Co-curricular activities should include Mathematics or number games, reading and writing, literature, contemporary issues and challenges as part of the curriculum for all community schools. Remedial teaching and individualized instruction should also be implemented for the targeted students.

6. **Bullying is affecting learning in the schools:** A high number (53%) of the students were bullied in the school by their peers and others. The performance of the bullied students was found lower than that of those who were not bullied. The gap is 31 scale score in Nepali and 22 in Mathematics.

   Recommendation: Local governments together with the schools should

regularly monitor the school to maintain peace, discipline and regulations as well as cohesive environment among the students. "No bullying" movement should be initiated by the school in consultation with and involvement of the local community. Active focal person in the school, psycho-social counselling service and child club activity against bullying are the other activities that the school can provide to minimize bullying in the schools.

7. **Feedback to the students in homework has positive relationship with learning achievement:** The students who were provided with regular homework and feedback from the teacher performed better than those who were not provided homework and feedback in both subjects. About 7% community school teachers never provided homework and feedback to the students in Nepali subject.

   Recommendation: Teacher's performance evaluation should be strongly linked with students' performance. Providing homework and feedback in certain subjects (perceived to be difficult) in the form of scaffolding in the schools, continuous assessment service, criterion-based assessment for teaching-learning, formative evaluation system and regular communication with parents about their children's performance, some instructional activities directly influence students' learning achievement.

8. **Students at right age performed better:** Students studying in grade 5 at appropriate age (11 years) performed better than the under or over age students. The gap was found 9 scale score in Mathematics and 11 in Nepali. Similar trend was noticed in the previous study as well. This further means that age wise grade or grade wise age or age appropriate enrollment ensures higher learning achievement.

   Recommendation: Net enrollment practice at the basic level education should be maintained and increased by admitting appropriate age students in the appropriate grades in schools. Age appropriate level of students can learn smoothly with their peers of equal age.

9. **After-school activities determine student learning achievement:** Students' involvement in study during their after-school time has a relation with their performance. The students' who spent about one to two hours' after-school time in the activities like household chores, watching TV, and playing performed better. On the other hand, the students who spent more than two hours of their after-school time in those activities other than study performed relatively lower. The gap was found 17 scale score in Nepali and 16 in Mathematics.

Recommendation: All schools should make their parents aware of the influence of children's involvement in study with their performance. Besides being involved in entertainment and household chores, the students should be encouraged to study not only the textbooks based on the curriculum but also the literature books, reference books and other extra-curricular materials as well.

10. **Many students do not have a sense of Mathematics:** In Mathematics, students below minimum learning level have quite limited knowledge and skills in Mathematics. Around 50% of them do not have any sense of reading and writing numbers and number operation. With this limited knowledge and skills, they cannot calculate and solve Mathematical problems. Some of them are able to choose correct answers when options are given in the MCQ items. Mostly, they are unable to perform any Mathematical subjective calculations independently.

    Recommendation: Students experiencing difficulty with learning the numbers and Mathematical calculation should be provided with learning opportunity to developing minimum Mathematical skills and competencies. The teachers should be made accountable to their students' performance. An action needs to be taken if the teachers fail to develop their students' minimum level of competencies. Not only the teachers' performance but also the performance of the Principals of such schools with below minimum learning level of students should be linked up with their career path. The schools with such lower achievement level should be made responsible to develop short term and long-term learning improvement strategies and actions. To monitor the progress of the implementation of the planned strategies, regular standardized tests should be administered and local level monitoring and supervision mechanism should be ensured. Besides, the primary level teachers should be equipped with national level short-term practical, school based professional training to advance their content knowledge and pedagogical skills.

11. **More than half of the students are struggling with the simple Mathematical calculations:** Basic level (level 1) students have superficial knowledge and skills in most of the Mathematical contents. However, they are struggling with calculations. They are able to identify the ordered pair of a point, square pattern in dots, the sum of decimal numbers, the place of a digit in numbers, the relation of kilogram and gram and the numbers, mixed fraction and type of angles. But they also have a limited knowledge of formula for volume and area, they cannot estimate the angle shown in figure and cannot identify the relation between decimal and fraction. They can read the table and bar graphs to take simple information

but cannot draw conclusion by comparing the data. They can solve very simple problems of unitary methods, subtract small same denominator fractions, round numbers in the nearest tenth of a decimal number. They can also recognize limited square numbers and express Mathematical sentence in Mathematical language to calculate. They can also subtract a univariate one-degree Algebraic term from another, find the value of x in one variable equation and subjective, can simplify Algebraic expression (univariate) in Algebra. However, they cannot perform grade level Mathematical calculations independently.

Recommendation: The knowledge of basic level contents is essential to be able to grasp the Mathematical content. Therefore, the teachers should focus on the development of basic level contents among all students. In addition to these skills, the students should be prepared to solve simple grade appropriate Mathematical calculations independently. To improve learning achievement, an emphasis should be laid on the underperforming students instantly through the application of problem-solving method.

12. **In Nepali language, the students below level 1 can read only a few words or sentences but they cannot write the sentences independently.**

    Recommendation: Such underperforming (below level 1) students should be involved more on the activities such as reading and writing words and sentences and describing familiar events independently. Language teaching should focus on meaningful reading and comprehension exercises rather than on reciting the paragraphs in the textbook and rote learning the answers.

13. **Decreasing trend of students' performance in Mathematics:**

    Recommendation: A diagnostic study about the challenges in teaching and learning culture should be carried out. The factors responsible for reduced learning achievement should be identified and disseminated. The involvement of parents and community members should be ensured in making the schools accountable for their student's low performance level.

14. **The achievement and gap related results of NASA 2018 are quite similar to the results of NASA 2012 and 2015 :** The consistent recurring results not only proves the reliability of the NASA study, but also indicates that interventions were not sufficient in improving the quality of learning in the school level.

    Recommendation: MOEST should review the existing plan and policies from the quality concerns.

15. **Overall Recommendation:** As a final step of National Assessment, MOEST should initiate Post-NASA policy review and intervention plan at the national level, sub-national level, and implementing agency level of the education system.

Table of Contents

List of Table

List of Figure

ACRONYMS/ABBREVIATIONS

CDC	:	Curriculum Development Centre
CEHRD	:	Centre for Education and Human Resource Development
CR	:	Constructed Response
CRT	:	Criteria Referenced Test
CI	:	Confidence Interval
CTT	:	Classical Test Theory
DEO	:	District Education Office
DOE	:	Department of Education
EDCU	:	Education Development Coordination Unit
EMIS	:	Education Management Information System
ERO	:	Education Review Office
ICC	:	Item Characteristic Curve
ID	:	Identification
IEA	:	International Association for the Evaluation of Education
IRT	:	Item Response Theory
MC	:	Multiple Choices
MLE	:	Maximum Likelihood Estimation
MOS	:	Measure of Size
N cases	:	Number of cases/students in the population,
NASA	:	National Assessment of Student Achievement
SE	:	Standard Error,
NRT	:	Norm Referenced Test
NU cases	:	Number cases/students in the sample,
NU psu	:	Number of Primary Sample Units (schools)
OECD	:	Organisation for Economic Cooperation and Development
OMR	:	Optical Mark Recognition
One PM	:	One Parametric logistic Model
PCAP	:	Pan-Canadian Assessment Program
PCM	:	Partial Credit Model

PISA	:	Programme for International Student Assessment
PPS	:	Probability Proportionate to Size
PRC	:	Printing Ready Copy
PSU	:	Primary Sample Unit
PV	:	Plausible Value
RWGT	:	Replicable Weight
SE	:	Standard Error
SPSS	:	Statistical Package for Social Science
SR	:	Selected Response
SRS	:	Simple Random Sampling
SSDP	:	School Sector Development Programme
TIF	:	Test Information Function
TIMSS	:	Trends in International Mathematics and Science Study
WLE	:	Weighted Likelihood Estimation

CHAPTER I

AN OVERVIEW OF THE NATIONAL ASSESSMENT OF STUDENT ACHIEVEMENT, 2018

1. Introduction

In this report, chapter 1 presents an overall introduction of the National Assessment of Student Achievement (NASA), its historical overview and objectives. In chapter 2, methodological chapter, the report has attempted to explore the contextual variables, tools and technologies used during the overall study and also explains the contextual variables used like geography, ethnicity, gender, language and economic status etc. Chapter 3 presents the basic result by contextual variables, chapter 4 compares the NASA 2018 cycle with the previous 2015 cycle and trends of learning achievements. Finally in chapter 5, a summary of findings, conclusion and recommendations are presented in the form of an executive summary.

This is a report on the national assessment of Grade five students in Mathematics and Nepali subjects conducted by the Education Review Office (ERO) in 2018. The report of the assessment is based on the curriculum-based standardized test. A comparative presentation is made in all the sub-chapters focussing on province wise results as explicit strata and other variable specific results as implicit strata like the type of schools, gender, ethnicity, language, in a disaggregated form.

The assessment was conducted in 24 sample districts, 1400 schools and 32262 students. The major aim of NASA is to provide valid and reliable information on student learning achievement at grade five of basic education level with policy feedback to the Ministry of Education, Science and Technology. Specifically, NASA provides feedback to the teachers, schools, curriculum developers, program and policy executing agencies for the needed reform. A repeated cycle of NASA provides information on the trend of student learning and other contextual variables that provide pathways for the review and design for policy and program.

More specifically, the assessment answers the questions like: How well are the students learning? Is there an evidence of particular strengths and weaknesses in students' leaning? Do certain sub-groups of students perform poorly? What factors are associated with student achievement? Do the achievements of students change over the time? (Grenaney & Kellaghan, 2007). This report has highlighted the related

issues and problems, and made some recommendations to the policy makers and other stakeholders.

1.1 National Assessment of Student Achievement

Globally, it has been well accepted that the means of measuring the quality of education is the students' achievement (TIMSS & PIRLS, 2008). The national assessment provides both quantitative and descriptive form of information on student achievement, which after is considered as an output of the teaching learning process and its quality (World Bank, 1996). National assessment thus provides basic information for policy makers, politicians, and the broader educational community (ERO, 2013). Further, "it provides data for a type of national education audit carried out to inform policy makers about the key aspects of the system" (Greaney & Kellaghan, 2008b, p. 7, ERO, 2013). It is argued that the achievement of the students in a curriculum area be aggregated to provide an estimate of the achievement level in the education system as a whole at a particular age or grade level (Greaney & Kellaghan, 2008b; NASA, 2013). NASA is also a popular means of determining the achievement of curriculum and finding the gaps between the written curriculum and the taught curriculum. So, it is useful for making policy decisions especially when decisions are to be made in relation to the optimum utilisation of resources (EDSC, 2008). It provides evidence for policy makers on availability of textbooks, class size, and number of years of teacher training. Therefore, every country has accepted that it is "systematic, regular measure of learning achievement in a country that is designed to assist policy making" (Lockheed et al. cited in EDSC, 2008, pp. 19, NASA 2013).

1.2 Evolution of NASA in Nepal

Assessment practice is found to have started from the last years of the decade of 1980s in Nepal. However, the Ministry of Education has formally started the National Assessment since 1995 and continued it up to 2010 in a small scale. Large scale NASA was administered under the Ministry of Education since 2011 AD. Four NASA cycles were completed during the School Sector Reform Plan (SSRP) and two including NASA 2018 were completed during the School Sector Development Plan (SSDP). In both the plans, NASA is considered as a tool to measure the quality of education for making the educational institutions accountable to achieving the educational goals.

NASA studies are conducted for both backward and forward-looking purposes. The backward-looking purpose is concerned mainly with building a database to

analyse both the strengths and weaknesses of educational policies and practices that affect students' learning achievement (ERO, 2018).

The assessments completed so far and the upcoming assessments as per the designed NASA roadmap have been presented in table 1.

Table 1.1 NASA Cycles Completed and Planned

SSRP					SSDP					
2011	2012	2013	2015	2017	2018	2019	2020	2020	2021	2022
Grade 8	Grade 3 and 5	Grade 8	Grade 3 and 5	Grade 8	Grade 3 and 5	Grade 10	Grade 8	Grade 5	Grade 10	…
✓	✓	✓	✓	✓	✓	Progressing..				

A complete NASA cycle goes over a period of 3 years. In the first year, all items development, pre-testing of the items and item analysis are completed. In the second year, final test administration is conducted and finally, in the third year, activities like report writing, dissemination of the report and policy informing are done at the end of the NASA cycle.

The ERO follows globally accepted practices of conducting national assessments. Although the context of each country is different, there are some common practices to national assessments in most of the countries (ERO, 2018). Building on the comprehensive review of national assessments from various countries, ERO has adopted the following procedures:

- The Ministry of Education, Science and Technology (MOEST) selects an implementing agency either from within the MOEST system or an independent external consulting organization. In case of Nepal, Education Review Office (ERO) within the MOEST system is solely responsible for the national assessment.
- The MOEST or implementing agency develops policies and frameworks for assessment in consultation with (and with participation of) key stakeholders such as subject experts, teachers and policy makers.
- The MOEST identifies the Grade level and determines the area (e.g., literacy or numeracy) to be assessed.
- The implementing agency (ERO in Nepal) defines and describes the areas of achievement testing in terms of both content and cognitive skills and develops test items along with supporting questionnaires and manuals for test administration.

ERO

- Pilots the test items with the support of external experts and reviews their validity, appropriateness and sensitivity in terms of gender, ethnicity and culture.
- Ensures that the assessment instruments are reliable and valid.
- Selects the samples schools, arranges for printing the test papers and other relevant materials; and communicates with the schools and teachers for test administration.
- Orients the test administrators (focal persons, head teachers and teachers), and then administers the test and survey questionnaires in the selected schools.
    - Collects test scores and other necessary information, cleans the data as needed and analyses them.
    - Prepares draft report/s which is/are reviewed by relevant subject committees and external experts.
    - Prepares and disseminates final report/s through various means such as publication and the mass media.
- Finally, the MOEST, implementing agency and relevant stakeholders study the report/s of national assessment and identify major areas for policy reforms (ERO, 2017, 2018).

1.3 NASA Cycle

ERO has adopted the following cycle to conduct the national assessment of Grade 5 students in Mathematics, Nepali and Science.

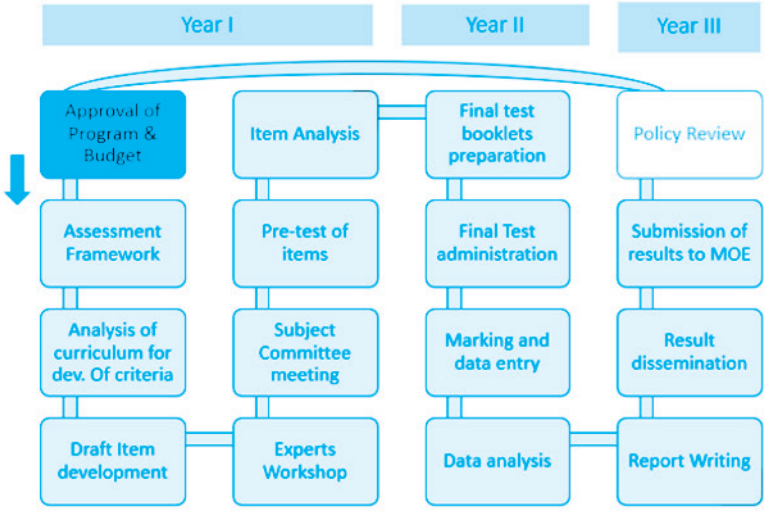


Figure 1.1 NASA process cycle

The above figure presents the major steps taken in planning, designing, administering and reporting of the assessment. NASA process cycle begins with an approval of the required budget and programme and goes through the series of assessment procedures: development of the assessment framework, criteria and standards, items and questionnaires; piloting, analysing and selecting the items; designing the test booklets; administrating the test; scoring and preparing data; calibrating items and equating the tests; analysing and setting proficiency levels; and reporting and disseminating the results.

1.4 Objectives of NASA 2018

The purpose of this assessment is to provide feedback to the Ministry of Education, Science and Technology to improve the quality of school education. This assessment does not report individual students' performance, nor does it compare the proficiencies of each individual student and school. Rather, it provides the national and provincial level results as well as the differences in the achievement scores in relation to various influencing factors such as socioeconomic status, home language, and identity with geographical region. More specifically, NASA 2018 has the following objectives:

a. To identify the current level of Grade 5 students achievement in Mathematics and Nepali.
b. To identify variations in student achievement by gender, province, identity with geography, types of school, ethnicity, home language, socio-economic status.
c. To explore factors that influence student achievement.
d. To identify the trend in student learning and produce the baseline data for the future for comparing.
e. To strengthen the capacity of the education system in conducting national assessment.
f. To provide the Ministry of Education, Science and Technology with recommendations for policy making to improve quality and equity, particularly in school education.

1.5 Distinct Features of NASA 2018

The ERO has used Item Response Theory to assess the latent ability of students using various contextual variables to explain those latent traits of the students. This assessment has used advanced procedure to bring rigor to data analysis by generalizing the results in national level and province levels through 7 explicit strata and various

other implicit strata. Use of Replicate Module for estimating the population parameters and Weighted Likelihood Estimation (WLE) for analysis of individual student level and reporting are the examples of its advancement. Furthermore, the advancement of procedures has also been noticed in sampling methods. A Probability Proportional to Size (PPS) sampling procedure has been used in selecting the schools as Principal Sample Unit (PSU), the school clusters. Reporting of student achievement at province level and national level is done in a transformed scale with mean 500 and standard deviation 50 by using the formula:

*Average scale score = 500 + plausible value * 50*

*or, Average scale score = 500 + logit * 50*

If readers want to extract the WLE of latent ability, they can use:

$$\textit{average latent ability (logit) of any group} = \frac{\text{average score - 500}}{50}$$

The distinct features of this report are:

1. A comparative presentation of NASA 2018 and 2015 by using IRT methods and rigorous process
2. Trends of the results over all NASA cycles of grade 5
3. Learning level descriptors prepared through a rigorous analysis
4. A gap in learning between the written curriculum and the taught curriculum in the form of achieved curriculum is presented.

Chapter 2
Methodology

This chapter presents the process adopted for sampling, assessment framework, tools development, setting contextual variables and determining the reliability and validity of the tools. It also presents the statistical tools and techniques used in data analysis and comparing the NASA 2018 data to previous NASA 2015. Moreover, various formula, symbols and techniques used in data analysis and reporting are described in greater details in this chapter.

2.1 Sampling

2.1.1 Target Sampling Frame

Sampling is a process of selecting a set of data from the population by using a defined procedure. In this assessment, the multi-stages sampling process was adopted. In the first step, a list of all 27464 schools to be included in the assessment, with their unique ID (school EMIS code) provided by Department of Education-DOE (now Centre for Education Human Resource Development - CEHRD) was listed. This list was considered as the target population for developing the sampling frame. In addition to the name, location (provincial, district, geography and municipality) and ID (code) of each school, public and private categories, the total number of students, with gender categories, in each school was taken as the sampling frame. These data are available from the EMIS of CEHRD, which are collected through the national census of schools every year. The target sampling frame for this assessment was thus prepared on the basis of the school data of 2018 with 710499 students as the target population.

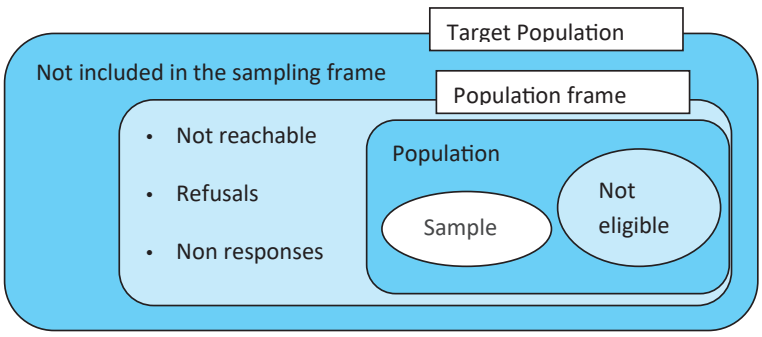


Figure 2.1 Conceptual diagram of population for sampling frame

2.1.2 Population

The population of the study is the schools running the classes up to Grade 5. However, some of the schools did not report the number of students (zero students) and such schools were excluded from the population frame. After the exclusion of non-student school, the schools with less than 10 students were also excluded as the non-eligible schools. Then the population of this assessment reached 21093 schools at national level. Sample cluster schools were selected from those schools, by using Probability Proportional to Size (PPS) sampling method. Thus the population for this assessment covered all students enrolled at Grade five taken randomly from primary sampling units (PSUs). The exclusion of the schools was defined by following criteria:

- Schools having less than 10 students
- Students who did not respond the test items (during data cleaning)
- Schools at very remote distance or unreachable at the time of assessment
- Schools which do not have students in Grade 5

2.1.3 Sample Size

The educational survey research studies suggest that the sampling precision requirements should be satisfied by a simple random sample (SRS) of 384 students for the main criterion variable. This size of simple random sample of students yields 95% of confidence interval for the student-level estimate with 3% of confidence interval (Margin error). However, a perfect random sampling is not an easy task in such a large-scale national assessment. The sampling design includes the combination of different sampling techniques in different stages, including stratification, clustering and random selection of students. For this, the design effect due to the multi-stage sampling has to be calculated and adjusted while selecting the sample size.

In this assessment, actual sample size of this assessment was calculated in multi-stage sampling methods. Intra-class correlation was taken from the recently administered survey of grade 3. Taking intra-class correlation *r = 0.25 (NASA, 2015) and school cluster size (C) equal to 25, the design effect (deff) was calculated by using the formula given:*

Where: Deff = Design effect
C = the size of the cluster (number of students within the school who will be assessed in a subject)
r = Intra-class correlation

Now, to calculate the Clustered Sample Size (CSS), using the formula

Where Effective Sample Size (ESS; that is equivalent to a Simple Random Sample size). Taking 384 as sufficient sample, CSS is equal to 2688. Now by adjusting the student non-response by 3% and school non-participation by 3%, total number of students has become 2856. This number of students is equal to 114 school clusters and 25 PSU size.

To produce sufficient number of samples, at least 114 schools per province and even more than this number can be included in some strata (provinces) as sample. Hence altogether, 7*114 = 799 schools are determined to be valid number for the assessment. However, intra-class correlation is more than 0.25 in grade 5 or above grades. So, taking design effect = 10, sample cluster number becomes 163 by the same formula. By 7*163 = 1142 number of schools are defined as an adequate sample of schools. For assuring sufficient school size and addressing the diversity of community and institutional schools, 1400 schools were taken as the sample PSU.

There are other methods as well for calculating the sample size. One of the other formula for taking the required sample size is

$nc = n^* \times deff$,

where nc is the required sample size, n* is the effective sample size for simple random sampling (srs), and Deff is the design effect. Deff = 1 + (C —1), where C is the population size, (Rho) = Intra-class correlation.

Putting the value of design effect, Deff = 10, as a multi-stage cluster sampling at the national level might have a design effect of 10 or higher (Murphy & Schulz, 2006), the minimum sample size of students is estimated as:

$nc = 400 \times 10 = 4000$, implies, 4000/25 = 160 schools. For seven provinces, 7 * 160 = 1120 schools.

From the above both methods, a minimum number of PSU does not exceed 1120 schools. This number is also very close to the above calculation. Hence 1400 PSU is a sufficient number also for this National Assessment.

2.1.4 Sample Design and Stratification

The sample design for NASA 2018 Grade 5 assessment was a multi-stage by the selection of schools from each explicit stratum (province). In Nepal, seven provinces are politically divided entities of the country, which govern educational administration within their region in their own. A sufficient number of samples taken from the provinces will ensure the generalizability or the results. The selection of districts from each geographical location was done randomly to incorporate Himal, Hill and Terai areas as far as possible. The Primary Sampling Unit (PSU) schools (clusters), were selected within the district by using PPS method. The selected 24 districts from all 7 provinces are presented in the following figure 2.2:

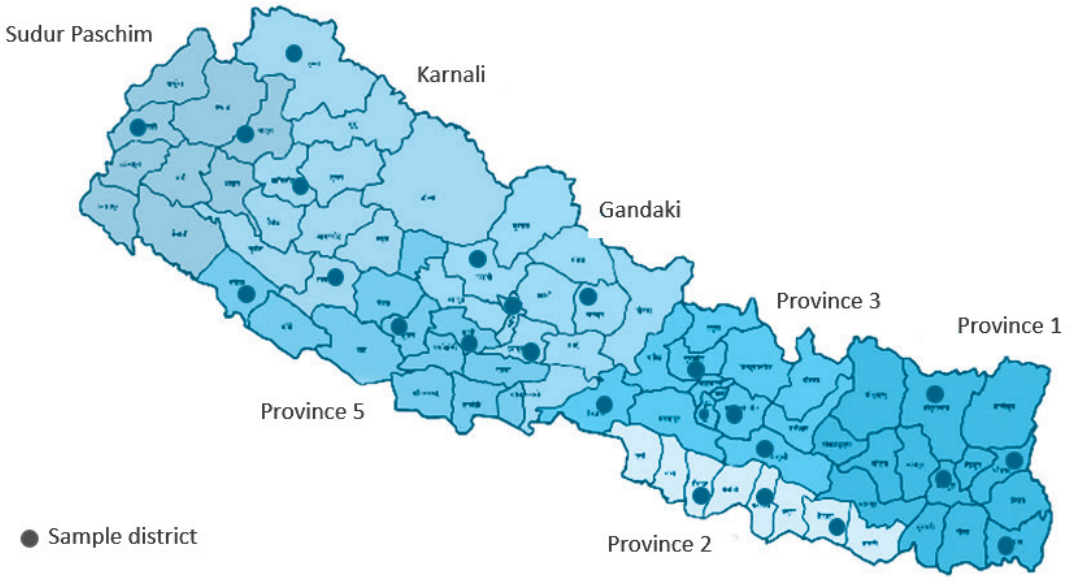


Figure 2.2 Sampling districts

Among the 77 districts, five districts from Province 1: Bhojpur, Dhankuta, Okhaldhunga, Jhapa and Panchthar; three districts from Province 2: Mahottari, Rautahat, Siraha; five districts from Province 3: Sindhuli, Kavre, Chitwan, Nuwakot, and Lalitput; three districts from Gandaki province: Syangja, Lamjung, Parvat; three districts from Province 5: Bardiya, Pyuthan, Gulmi; three districts from Karnali province: Kalikot, Salyan, Humla and two districts from Sudur Paschim province: Bajura and baitadi were selected randomly.

2.1.5 Selection of the Schools and the Students

From the population, a total 32262 students were taken as the sample. However, 28,381 students participated in the assessment. The minimum sample for the province having the smallest number of student population was fixed to be 1060 per subject. Viewing the different sizes of schools, the maximum sample size was fixed to be 25 per school, which is called Measure of Size (MOS).

In the case of a sample school having more than 25 students, the students were selected by using a random sampling method otherwise all the students were taken as the samples with defined number of students. More specifically, the number of students sampled from each of the selected schools was of two different ways: (i) If the size of the students was less than or equal to the expected sample size (MOS), all the students were sampled. (ii) When the size of the students was greater than the expected size, the required number of the students was selected randomly. The probability of selection of a particular student from schools was always the same.

Because of school replacement and student non-response adjustment, calculation sample weight by PPS sampling methods was completed. In the raw database, some records were background information only and some were subjective test item response only with unidentified unique ID or school were deleted from the database. So, finalized and cleaned data by removing duplicate cases, outliers and invalid entries was as per given in following table 2.1.

Table 2.1 Sample students from community and institutional schools

Province	Math			Nepali			Grand Total
	Community	Institutional	Total	Community	Institutional	Total	
Prov. 1	1847	984	2831	2040	755	2795	5626
Prov. 2	1470	269	1739	1602	211	1813	3552
Prov. 3	2335	1609	3944	2224	1672	3896	7840
Gandaki	829	770	1599	799	818	1617	3216
Prov. 5	1608	335	1943	1617	297	1914	3857
Karnali	1045	26	1071	995	117	1112	2183
Far-western	992	55	1047	994	66	1060	2107
Total	10126	4048	14174	10271	3936	14207	28381

In the above table 2.1, it can be seen that the number of schools selected from every province was not the same. While selecting the schools, they were selected as the primary sampling unit (PSU) from each province with a minimum of 54 schools from

each province for one subject. However, provinces with higher number of schools and students have proportionally greater number of sample schools and students. School selection within a district was done by using a PPS sampling method by sorting out the list of schools from the selected districts. Sorting of institutional and community schools within the district, was an implicit stratum. In addition to preparing the list of sample schools, a list of replacement schools was also prepared. During the orientation programme about NASA administration to the district focal persons, the final list of the sample schools was prepared, choosing the schools from the list of replacement schools upon requirement. The number of school clusters sampled from seven provinces are listed in table 2.2.

Table 2.2 Province wise number of sample schools

Provinces	Math		Nepali		Total
	Community schools	Institutional schools	Community schools	Institutional schools	
Prov. 1	96	45	111	32	284
Prov. 2	68	13	72	9	162
Prov. 3	125	68	123	72	388
Gandaki	52	34	49	36	171
Prov. 5	74	15	72	14	175
Karnali	54	2	50	5	111
Far-western	51	3	52	3	109
Total	520	180	529	171	1400

2.1.6 School Weight

School level base weights were calculated using the formula:

$$BW^{i}_{sc} = \frac{N_{pop}}{n_{sc} \times N^{i}_{mos}}$$

where Npop was the population size (students), nsc was the total number of schools sampled within each explicit stratum; and Nimos was the measure of size (MOS) assigned to the school (i). School level base weights were calculated for all sampled schools that satisfied the condition for eligible students actually participating in the study. For example, in mathematics, altogether 650 schools were sampled, out of which 1 school did not participate in testing due to some unavoidable circumstance. For this, a school-level non-response adjustment was calculated separately for each explicit stratum, using the formula:

$$Sc_{adj} = \frac{n_{sc}}{n_{psc}},$$

where nsc is the total number of originally sampled schools; and npsc was the number of schools that actually participated.

The final school weight was then calculated with non-participation adjustment to the base school weight. The final school weight was then equal to the product of the school base weight and non-participation adjustment,

$$W_{sc} = BW^{i}_{sc} \times Sc_{adj}$$

2.1.7 Student Weight

For schools with 25 grade five students, student base weight was 1; and for schools with more than 28 students and fewer, the base weight was calculated using the formula:

$$BW_{st} = \frac{N_{st}}{n_{st}},$$

where Nst was the total number of students at Grade 5 in the sampled school and nst as the number of sample students from the class.

A student non-participation adjustment was calculated for any school that had at least one student who was sampled and was eligible to do the test but did not participate for some reason. This was calculated with the formula:

$$St_{adj} = \frac{n_{st}}{n_{pst}},$$

where nst was the number of sample students and npst was the number of students who participated in the particular school.

The final student weight of a particular school (say, ith school) is then equal to the product of the student base weight and non-participation adjustment: $W^{i}_{st} = BW^{i}_{st} \times St_{adj}$

The final weight is thus the adjustment between the product of the school and student final weights: $W_i = W^{i}_{sc} \times W^{i}_{st}$. For example, see the sample weight calculated in figure 2.3

Sch_code		Province	Sample_size	Sch_pi	Sch_wt	Within_sch_prob	Within_sch_wt	Final_st_prob	Final_st_wt	Sum_of_final_wt	School_base_wt_adj	Sch_non-response_adj	Student_base_wt_adj	Stu_non-response_adj	Final_st_WT_adj	Final_Sch_WT_adj	Final_Sampling_Wt_adj
160840007	.....	2	25	0.04	25.92	0.66	1.52	0.03	39.39	1103.04	25.92	1.00	39.39	1.00	39.39	25.92	1021.00
240210004		3	25	0.03	36.48	0.93	1.08	0.03	39.39	1103.04	36.48	1.00	39.39	1.00	39.39	36.48	1436.97
580140001		5	25	0.05	18.58	0.47	2.12	0.03	39.39	1103.04	18.58	1.00	39.39	1.00	39.39	18.58	732.04
460500011		5	25	0.03	36.48	0.93	1.08	0.03	39.39	1103.04	36.48	1.00	39.39	1.00	39.39	36.48	1436.97
550420005		6	25	0.04	23.45	0.60	1.68	0.03	39.39	1103.04	23.45	1.00	39.39	1.00	39.39	23.45	923.76
370360001		4	25	0.05	18.58	0.47	2.12	0.03	39.39	1103.04	18.58	1.00	39.39	1.00	39.39	18.58	732.04
20290028		1	25	0.10	9.85	0.25	4.00	0.03	39.39	1103.04	9.85	1.00	39.39	1.00	39.39	9.85	387.98
404100013		5	25	0.03	39.39	1.00	1.00	0.03	39.39	1103.04	39.39	1.00	39.39	1.00	39.39	39.39	1551.92
460320007		5	25	0.03	39.39	1.00	1.00	0.03	39.39	1103.04	39.39	1.00	39.39	1.00	39.39	39.39	1551.92
520240008		5	25	0.03	39.39	1.00	1.00	0.03	39.39	1103.04	39.39	1.00	39.39	1.00	39.39	39.39	1551.92
670170010		7	25	0.03	37.88	0.96	1.04	0.03	39.39	1103.04	37.88	1.00	39.39	1.00	39.39	37.88	1492.23
70310011		1	25	0.03	36.48	0.93	1.08	0.03	39.39	1103.04	36.48	1.00	39.39	1.00	39.39	36.48	1436.97
250270133		3	25	0.03	36.48	0.93	1.08	0.03	39.39	1103.04	36.48	1.00	39.39	1.00	39.39	36.48	1436.97

Figure 2.3 Example of Sample weight calculation

2.2 Test Administration and Supervision

Test administrators for the NASA 2018 were appointed from Resource persons, School Supervisors, and Headteachers. The appointed administrators were trained to administer standardized National Assessment as per the NASA test administration guidelines. For the support and inspection of the test administration, a teacher from the schools who were not teaching the assessed subject in the particular school were also appointed. For other support, two other support staffs were assigned in the test administration in a school.

For monitoring and supervision of the NASA test administration, three types of monitors were mobilized. Some from civil servants at central level agencies of the Ministry who were appointed by ERO and some were appointed by EDCU. A team for supervisors was mobilized for immediate support and monitoring of the process in every sample district. In bullet points, the test administration process followed has been summarized below:

- One school participated in only one subject.
- Subject teachers were not allowed in the test administration hall, rather they were assigned to provide response on Teacher's Background Information Questionnaire.
- Test administration centre head oriented their students, support staffs and invigilator to ensure smooth test administration.
- Clear instruction to the students was provided to write with their full efforts in a low-stake environment.
- After the test administration the head teachers also responded on the background

information questionnaire provided to them.

- To maintain the confidentiality of the test items, no one was allowed to copy the papers, take the pictures of the paper, or keep the test papers in the school.
- After the test administration was over, booklets were collected in the EDCU by consulting firm. Each school submitted their monitoring report, test administrator's report and list of participated and non-participated student/school list as well.

2.3 Analysis Methods

The data analysis methodology consists of two parts. The first part is item analysis and the second part is data analysis and interpretation. In the first part, SPSS 23 was used to code, recode and clean the database. During the data cleaning, duplicate cases, outliers, and unidentified cases were cleaned. All the background variables were recoded to make them readable for ACER ConQuest 4 software. Also, dummy variables were prepared for conditioning the run in ConQuest.

ACER ConQuest 4 software was used to analyse the items to generate item parameters in set by set manner. Later, joint file was prepared by combining all three sets of a subject and the joint run of all three sets was useful to generate item level parameters viz. difficulty parameter, discrimination parameters, item fit parameters, distractor analysis, ICC plots, TIF plots. From the joint run, item parameters in the form of logits were generated and those parameters were fixed for case analysis. After the case analysis, ".wle" file was generated with case estimation that was used for conditioning the run. The overall data analysis process is presented in the figure 2 below

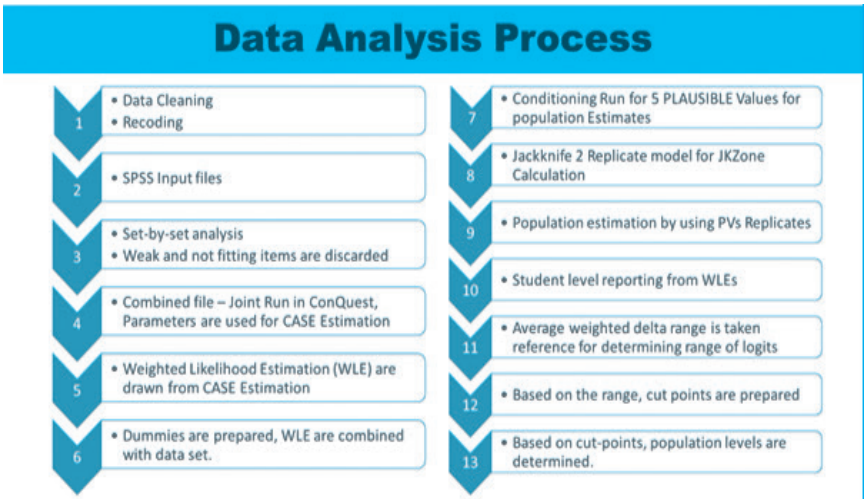


Figure 2.4 Data analysis process of NASA 2018

After the estimation of WLE and Plausible values (5 PVs), a process of replicate weights was used to estimate standard errors of population estimates. The figure below shows an example of Replicate module used in NASA 2018. *[See ERO (2017) for detail process and formula used]*

A sample of front end of estimating population parameters from replicate module is presented in figure below:

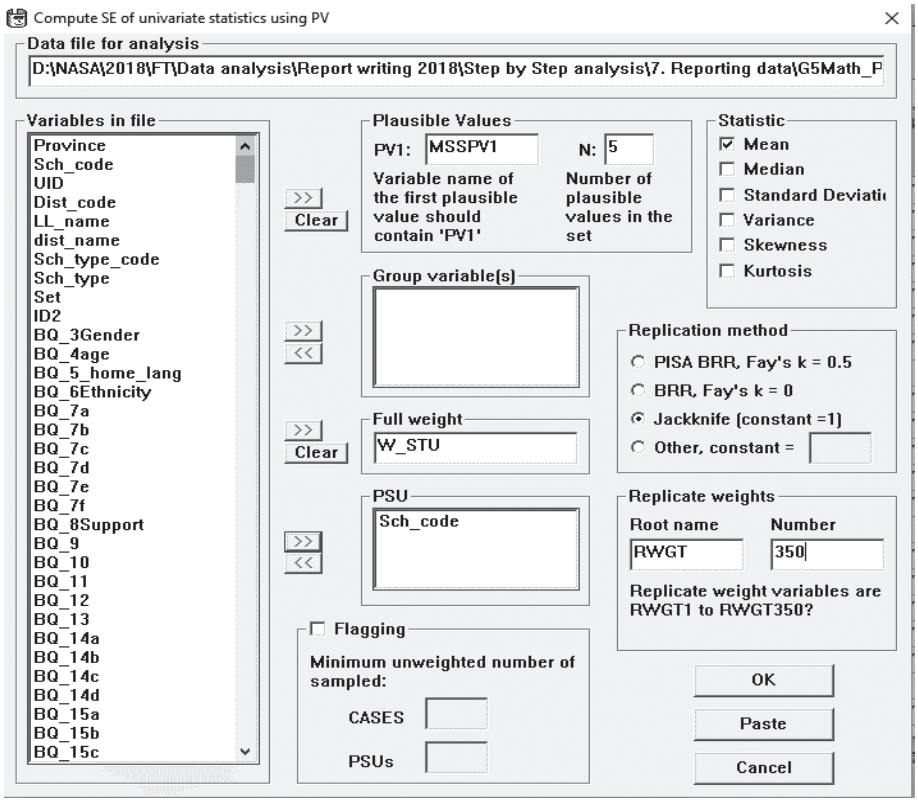


Figure 2.5 Replicate module used to calculate the Standard Estimate of Univariate Statistics using PVs in Mathematics

2.4 Tools Development, their Reliability and Validity

2.4.1 Assessment Framework

Curriculum based test items were developed based on the Assessment Framework. The assessment framework is a plan of content, item type, content domain and proportion of test items to be included. It is a blueprint of whole standardized assessment of NASA.

The assessment framework was developed before designing the test and developing the test items. The assessment framework was developed to:

> provide a clear guideline for a sound assessment approach to inform policy makers and the other concerned stakeholders on quality of education. It includes domains to be assessed, the statement of criteria together with standards, specification of items, framework for contextual variables to be considered while conducting an assessment and brief guidelines for assessment design. (ERO, 2017)

The assessment framework has identified and described the domains and constructs to be assessed in Mathematics and Nepali subjects. It has also proposed a framework for designing background questionnaires for students, teachers and head teachers. In addition, it has presented a brief guideline on overall methodological approach to be adopted for the assessment.

2.4.2 Framework for Mathematics

The objectives, general in nature, are elaborated in each grade for making them more specific. The following are the detail learning objectives specified in the national curriculum for grade 5 (CDC, 2008):

1. Measure the angles from 00 to 1800 in the difference of 15 degree.
2. Measure the angles and sides of a given triangle and quadrilateral.
3. Classify triangles according to the measures of sides and angles.
4. Identify cub, cuboidal and cylindrical solids.
5. Count and read the numbers greater than a crore.
6. Identify/recognize and write the place value of digits used in the number greater than a crore in Nepali.
7. Read and write the number up to million in English.
8. Identify prime and composite numbers from 1 to 100.
9. Rounding of numbers up to the nearest thousands.
10. (a) Calculate the square of 1 to 10 and the cube of 1 to 5;
    (b) Calculate the square root of the square numbers from 1 to 100 and the cube root of the cubic numbers from 1 to 125.
11. Factorize 3-digit numbers by using a prime factor method and construct prime factors tree.
12. Solve the numerical and daily life problems by using any of two operations among -, +, × and ÷ together with two brackets {( )}.
13. Multiply and divide the units of time and solve related verbal problems.

14. Solve the verbal problems related to rupees and paisa.
15. Multiply and divide distance and solve the related verbal problems.
16. Estimate the length, breadth, height of the given object, and estimate the distance of places from the surroundings of house or school.
17. Calculate the perimeter of rectangles (by using formula).
18. Calculate the area of rectangular objects and solve the simple verbal problems.
19. Multiply and divide litre and millilitre and solve the related verbal problems.
20. Calculate the volume of cuboid by using formula.
21. Multiply and divide gram and kilogram and solve the weight related simple problems.
22. Estimate the weight of any object and identify the relation between kilogram and quintal.
23. Convert mixed numeral and improper fractions into each other.
24. Add and subtract mixed numerals (up to two terms) and find the product of simple fractions.
25. Solve the verbal problems related to addition and subtraction of fraction.
26. Convert fraction and decimal from each other (up to three digit from decimal)
27. Add and subtract decimal numbers (three numbers after decimal) and solve the simple problems related to daily life.
28. Round off decimal numbers to the nearest place.
29. Convert fractions and percentage into each other and solve the simple daily life related problems on percentage.
30. Solve the problems by using a unitary method.
31. Calculate the simple interest by using a unitary method.
32. Read and prepare simple bills.
33. Identify the information from a family budget description.
34. Derive information and conclusions from tabulated information.
35. Present information in simple bar graphs.
36. Plot ordered pairs in a graph (first quadrant only).
37. Present a given set by using set notation { }.

The content domains and their weightage drawn from the curriculum are presented in Table below. To reduce the number of content domains, some of the content areas included in the curriculum are reorganised and regrouped in the table. The numbers of content domains are regrouped and reorganised from 9 to 6. For example, perimeter, area and volume from the areas of measurement are all included within the domain of geometry.

Table 2.3 Content domains for Mathematics in Grade 5

Content Domain	Elaborated Content Domain	Teaching hrs		Weightage (%)
Geometry	Geometry	17	27	14
	Measurement-perimeter, area, volume	10		
Numeracy	Numeracy	35	35	18
Arithmetic	Basic operation	27	59	31
	Fraction and decimal and percentage	24		
	Unitary method and simple interest	8		
Time, money and measurement	Time,	6	26	14
	Money	6		
	Measurement- Distance, capacity and weight	14		
Bill, budget and statistics	Bill and budget	9	18	9
	Statistics	9		
Sets and Algebra	Sets	9	27	14
	Algebra	18		
	Total	192	192	100

The major content areas mentioned in the above table are further expanded into the following content domains:

Content domain #1: Geometry

1. Angles: Construction of angles 0-1800 in the difference of 15 degrees.
2. Triangle and quadrilateral:
   - Measurement of the angles and sides of triangle and quadrilateral.
   - Classification and identification of triangles based on the measurement of angles and sides.
   - Identify cub, cuboidal and cylindrical solids.

3. Measurement:
   - Calculation of perimeter of rectangular shapes by using formula and solution of related simple problems.
   - Calculation of area of rectangular shapes by using formula and solution of related simple problems.
   - Volume of cuboid by using formula.

Content domain #2: Numeracy

1. Counting and writing the place value of the numbers greater than crores.
2. Counting and writing the numbers in English up to a million.
3. Identifying prime and composite numbers from 1 to 100.
4. Rounding off a number in required places.
5. Squaring numbers (1 to 10) and cubing numbers (1 to 5) and their roots.
6. Factorization of numbers (the number up to three digits) into their prime factors.

Content domain #3: Arithmetic

1. Four simple problems: Numerical and verbal problems involving any two operations among +, -, × and ÷; and two brackets {( )}.
2. Fraction and decimal:
    - Conversion of mixed numbers and improper fractions into each other.
    - Addition and subtraction of mixed numbers (two fractions only).
    - Multiplication of fraction (only two fractions).
    - Solution of simple problems related to addition and subtraction of fractions (involving only two fractions).
    - Conversion of decimal and fraction into one another (up to three places of decimal).
    - Addition and subtraction of decimal numbers (up to three places of decimal).
    - Solution of simple problems involving addition and subtraction of fractions.
    - Rounding off a decimal number to the required places.

3. Percentage:
    - Understanding the meaning of percentage and conversion of fraction and percentage to each other.
    - Solution of simple problems related to percentage.

4. Unitary method and simple interest:
    - Calculation of simple household transactions by using a unitary method.
    - Solution of simple problems on simple interest by using a unitary method.

Content domain #4: Time, money and measurement

1. Multiplication and division of the units of time and solving related simple problems.
2. Addition, subtraction, multiplication and division of money and the solution of related problems.
3. Multiplication and division of litre and millilitre and solution of the related simple problems.

4. Multiplication and division of gram and kilogram and solution of the related simple problems.
5. Multiplication and division of distance and solution of the related simple problems.
6. Estimation of length, breadth and height of objects.
7. Conversion of kilogram and quintal to each other.
8. Estimation of weight of various goods and objects.

Content domain #5: Bill, budget and statistics

1. Reading simple bills.
2. Drawing conclusions from the tabulated information.
3. Presenting simple information in a bar graph.
4. Identifying a point using ordered pairs (only in the first quadrant).

Content domain #6: Sets and Algebra

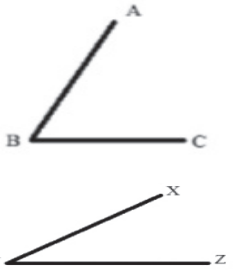
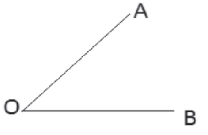
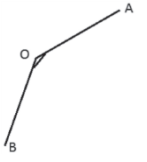
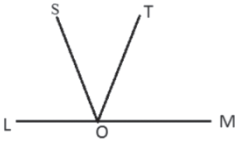
1. Writing sets by using set notations and words.
2. Conversing simple verbal problems involving addition and subtraction into algebraic expressions (involving only two terms).
3. Solving simple linear equations of one variable by using axioms of equality.

Criteria and Standards for Grade 5 Mathematics

As mentioned earlier, four standards pre-basic, basic, proficient and advanced have been identified in a hierarchical order of the level of competency (from lower to higher). Accordingly, levels 1, 2, 3 and 4 are defined in a hierarchy of depth or complexity of knowledge, skills and application in each of 36 criteria for Grade 5 in Mathematics. Table 2.2.3 shows the content areas, criteria and standards. The test items for NASA 2018 were developed using these criteria and standards.

[See: Assessment framework 2017 for grade 5 for detail]

Table 2.4 Content areas, criteria and standards for Mathematics, Grade 5

Content Area	Criteria	Standards			
		Below basic (Level 1)	**Basic (Level 2)**	**Proficient (Level 3)**	**Advanced (Level 4)**
	1. Measurement of the angles between 00 to 1800 in the difference of 15 degree	Compare acute angles by estimating. Example: Which of the following figures show bigger angle?  .	Measure the angles (less than 900) in the difference of 15 degree. Example: Using a protractor, measure ∠AOB and write the measure in degree. 	Measure the angles of measurement which is greater than 900 and less than 1800 in the difference of 15 degree. Example: Measure the following angle in the given figure. 	Measure the angles given in different positions. Example: (a) In the adjoining figure, measure ∠ TOM and ∠SOM.  (b) Which angle is greater, ∠TOM or ∠ SOM?
……	……	……	……	……	……

(see: "NASA 2018: Assessment Framework for Grade Five in Mathematics and Nepali" forl detail)

Table 2.5 Representation of various cognitive domains in the test for Grade 5 in Mathematics

Cognitive Domain	Weightage
Remembering	15%
Understanding	40%
Applying	30%
Reasoning	15%
	100%

2.4.3 Specification of Items

The following specification table suggests a tentative weightage (percentage), number and types of items, allocation of marks, and distribution of test-items in each content domain and standards.

Table 2.6 Table of specification for item selection

Content domain	Criteria No.	Weightage	Marks allocation	Weightage for items of various standards
Geometry	1-3, 14-15, 17	14%	7	The weightage of items in each set should be around as follows: Pre-basic: 15%, Basic 2: 35% Proficient: 35%, and Advance: 15%.
Numeracy	4-8	18%	9	
Arithmetic	9-, 20-28	31%	15	
Time, money and measurement	10-13, 16, 18-19	14%	7	
Bill, budget and statistics	29-33	9%	5	
Sets and Algebra	34 -36	14%	7	
Total		100%	50	

Note:

1. Type of items will be SR (selected response items - MCQ), CR (constructed response – fill in the blanks and very short answer) items carrying 1 mark and CR items each carrying 2 marks.

2. While selecting the items for each content domain it is necessary to select both SR and CR items with a reasonable ratio.

2.4.4 Framework for Nepali Subject

विषयक्षेत्र पहिचान *(Defining the content domain)*

कक्षा ५ का विद्यार्थीको सक्षमताले आधारभूत तहको कक्षा पाँचसम्मको सक्षमता जनाउने भएकाले विषयवस्तुको क्षेत्र पहिचान गर्दा तल्ला कक्षादेखिको उर्ध्व (upward) सक्षमताको समेत विचार गर्नुपर्ने हुन्छ । यही कुरालाई विचार गरेर विषय क्षेत्रको पहिचान गर्नका लागि प्राथमिक तहको पाठ्यक्रमले निर्धारण गरेको नेपाली विषयको तहगत साधारण उद्देश्य उल्लेख गरिएको छ । पाठ्यक्रममा प्राथमिक तहको अध्ययन पूरा गरेपछि नेपाली भाषामा विद्यार्थीहरू निम्नलिखित कार्य गर्न सक्नेछन् भनी तहगत साधारण उद्देश्यहरू उल्लेख गरिएको छ :

(क) सुनाइ सीप

- विभिन्न प्रकारका निर्देशनहरू सुनेर सोहीअनुसार गर्न ।
- कुनै विषयवस्तुको वर्णन, छलफल आदि ध्यानपूर्वक सुनेर तिनको अर्थ र आशय बुझी प्रतिक्रिया जनाउन ।
- बालकथा, बालगीत, कविता आदि आनन्द लिने गरी सुन्न ।

(ख) बोलाइ सीप

- शब्दहरू शुद्ध र स्पष्टसँग उच्चारण गरी स्वाभाविक गतिका साथ बोल्न ।
- देखेसुनेका र अनुभव गरेका कुरा सिलसिला मिलाई बताउन ।
- कुराकानी, छलफल आदिमा आफ्ना कुरा धक नमानी शिष्ट ढङ्गले राख्न ।

(ग) पढाइ सीप

- नेपाली भाषामा लेखिएका सरल पाठ्यसामग्रीहरू गति र यति मिलाई स्पष्ट रूपमा पढ्न ।
- स-साना सूचना विवरण पढेर आशय बुझ्न ।
- सरल बालसाहित्यका रचनाहरू रुचि लिई पढ्न ।

(घ) लेखाइ सीप

- ठीक दूरीमा बान्की मिलाएर अक्षर लेख्न ।
- देखेसुनेका र अनुभव गरेका कुराहरू सिलसिला मिलाएर शुद्धसँग लेख्न ।
- आफ्नो अनुभवका सेरोफेरोका विषयमा छोटा लिखित रचना तयार गर्न ।

पाठ्यक्रमअनुसार कक्षा ५ को नेपाली विषयमा समावेश भएका विषयवस्तुहरू निम्नअनुसार छन् :

Table 2.7 कक्षा ५ को नेपाली विषयको क्षेत्र र विषयवस्तु

क्षेत्र (Domain)	विषयवस्तु (Contents)
सुनाइ	■ निर्देशन, अनुरोध ■ प्रश्नोत्तर, कुराकानी, छलफल, वर्णन ■ बालगीत, कथा ■ विद्युतीय सञ्चार साधनका कार्यक्रम आदि
बोलाइ	■ स्वाभाविक बोलाइ ■ शिष्टाचारका शब्द ■ कुराकानी, सोधपुछ, प्रश्नोत्तर, छलफल, ■ वर्णन, संवाद, अभिनय, उद्घोषण ■ बालगीत, बालकथा, गाउ खाने कथा,चुट्किला आदि
पढाइ	■ साधारण सूचना विवरण, चिठीपत्र, ■ निवेदन, समाचार ■ पोस्टर, भित्तेपात्रो ■ सरल बालसाहित्य आदि
लेखाइ	■ चित्र र वस्तु तथा अनुभवको वर्णन ■ अनुलेखन ■ हिज्जे ■ श्रुतिलेखन ■ निर्देशित रचना ■ स्वतन्त्र रचना ■ चिह्नको प्रयोग
कार्यमूलक व्याकरण	■ पद सङ्गति ■ क्रियाका काल

पाठ्यक्रममा सुनाइ, बोलाइ, पढाइ र लेखाइ गरी चारओटा भाषिक सीपबाहेक कार्यमूलक व्याकरणसमेत गरी जम्मा पाँचओटा विषय क्षेत्रहरू समावेश गरिएको छ । पाठ्यक्रमका यी पाँच क्षेत्रमध्ये यस परीक्षणमा पढाइ र लेखाइ मात्र समावेश गरिएको छ भने कार्यमूलक व्याकरणलाई अलग्गै क्षेत्रका रूपमा नराखी लेखाइकै अङ्गका रूपमा समावेश गरिएको छ । यसप्रकार यस परीक्षणले कक्षा ५ का विद्यार्थीको पढाइ र लेखाइ सीपको परीक्षण गर्ने छ । कक्षा ५ को नेपाली विषयको पाठ्यक्रमका आधारमा पढाइ र लेखाइ सीपअन्तर्गत निम्नलिखित सिकाइ उपलब्धि हासिल हुनुपर्ने देखिन्छ ।

पढाइ

१. सामान्य गद्य र पद्य पढ्न र बोध गर्न ।
२. पोस्टर, भित्तेपात्रो, हस्तलिखित र विज्ञापनका सामग्री आदि पढ्न र बोध गर्न ।
३. चिठीपत्र, समाचार र अन्य विवरणहरू पढ्न र बोध गर्न ।
४. बाल साहित्यहरू मौन रूपमा पढ्न र बोध गर्न ।

लेखाइ

१. चित्र र अन्य उपयुक्त विषयवस्तुको लिखित वर्णन गर्न ।
२. हिज्जे मिलाएर शब्द लेख्न ।
   - छ्य, क्ष, व-ब, श-ष-स को सही प्रयोग गर्न ।
   - संयुक्त वर्णको सही प्रयोग गर्न ।
   - चन्द्रबिन्दु र शिरबिन्दुको प्रयोग गर्न ।
   - परिचित शब्दहरू ह्रस्वदीर्घ मिलाएर लेख्न ।

३. शुद्ध र सफासँग अनुलेखन गर्न ।
४. सरल अनुच्छेदको श्रुतिलेखन गर्न ।
५. निर्देशनका आधारमा र स्वतन्त्र रूपमा रचना तयार गर्न ।
   - निर्देशनका आधारमा शुभकामना पत्र, चिठी विद्यालयको निवेदन दैनिकी, वर्णनात्मक प्रबन्ध र कथा लेख्न ।
   - स्वतन्त्र रूपमा शुभकामना पत्र, चिठी, विद्यालयको निवेदन दैनिकी, वर्णनात्मक प्रबन्ध र कथा लेख्न ।
   - समूहमा भित्तेपात्रो लेख्न र सम्पादन गर्न ।

६. पूर्णविराम, अल्पविराम प्रश्नवाचक, उद्धरण र विस्मयादिबोधक चिह्नको प्रयोग गर्न ।

कार्यमूलक व्याकरणअन्तर्गत वचन, लिङ्ग, पुरुष र आदरसँग क्रियापदका मेल गराएर लेख्ने सामान्य वर्तमानकाल, सामान्य भूतकाल र सामान्य भविष्यत्काल प्रयोग गरी लेख्ने, तथा शब्दभण्डार वृद्धि गर्ने सिकाइ उपलब्धिलाई अलग्गै क्षेत्रका रूपमा नराखी लेखाइ सीप मै समवेश गरी प्रश्नहरू बनाउँदा भाषिक सीप विकासको परीक्षण अर्थपूर्ण हुन्छ ।

कक्षा ५ को नेपालीका लागि स्तर र तिनको व्याख्या (Standards and their Descriptors for Grade 5 in Nepali)

पढाइ (Reading)

Table 2.8 नेपाली विषयको आधार, स्तर र तह

क्र स	आधार (Criteria)	स्तर र तह (Standard and levels)			
		न्यून आधारभूत सक्षमता (Pre-basic) - तह १	आधारभूत सक्षमता (Basic) -तह २	प्रविणता (Proficiency) -तह ३	विशिष्ट (Advance) -तह ४
१	पोस्टर, भित्तेपात्रो, विज्ञापनका सामग्री पढी बोध गर्न	दिइएका सामग्रीबाट सामान्य खालको सूचना बोध गर्न	दिइएका सामग्रीबाट घटना, विवरण, सूचनाको आंशिक बोध गर्न	दिइएको सामग्रीबाट घटना, विवरण सूचनाको आशय पूर्ण बोध गर्न	दिइएको सामग्रीबाट घटना, विवरण सूचनाको मुख्य सन्देश बोध गरी सोको व्याख्या गर्न
...	...	...	...	...	...

(see: "NASA 2018: Assessment Framework for Grade Five in Mathematics and Nepali" for detail. You can download from www.ero.gov.np)

संज्ञानात्मक क्षेत्र (Cognitive domain)

शिक्षण सिकाइको प्रक्रियाद्वारा विद्यार्थीहरूमा संज्ञानात्मक क्षेत्रका विभिन्न तहका क्षमताहरू विकास भए नभएको मापन गर्न परीक्षण साधन तथा प्रश्नहरूले संज्ञानात्मक क्षेत्रका सबै तहलाई समेटेको हुनुपर्दछ। ब्लुम (Bloom) ले संज्ञानात्मक क्षेत्रका सामर्थ्यहरूलाई ज्ञान, बोध, प्रयोग, विश्लेषण, संश्लेषण र मूल्याङ्कन गरी ६ तहमा वर्गीकरण गरेका थिए। शैक्षिक गुणस्तर परीक्षण केन्द्रले सञ्चालन गरेका विगतका विद्यार्थी सिकाइ उपलब्धि राष्ट्रिय परीक्षण (२०११, २०१२, २०१३ र २०१५) मा उल्लिखित ६ वर्गीकरणलाई आधार मानी ज्ञान, बोध र प्रयोग, विश्लेषण, संश्लेषण र मूल्याङ्कनलाई उच्च दक्षतामा राखी चार तहका प्रश्न निर्माण गरी नतिजा पनि तदनुरूप विश्लेषण गरिँदै आएको छ। यसमा ब्लुमको परिमार्जित वर्गीकरणका ६ तह (सम्झाइ (Remembering), बुझाइ (Understanding), प्रयोग (Applying), विश्लेषण (Analysis), मूल्याङ्कन (Evaluating) र सृजना (Creating) तह (हे. Aderson & Karthwohl, 2001)) मध्ये सम्झना, बोध (बुझाइ), प्रयोग गरी तीन तहलाई यथावत समावेश गरी बाँकी तीन तहलाई तार्किक क्षमताका रूपमा वर्गीकरण गर्ने र सोही ४ तहका परीक्षण साधन तथा प्रश्न निर्माण गर्नेगरी प्रस्ताव गरिएको छ। यिनै वर्गीकरणलाई आधार मानी सिकाइ उपलब्धिको राष्ट्रिय परीक्षणमा निम्नानुसार ४ तहका निम्नलिखित भारअनुसारका प्रश्नहरू प्रयोग गर्नु उपयुक्त हुनेछ।

कक्षा ५ को नेपालीमा संज्ञानात्मक क्षेत्रका अङ्कभार (Cognitive domain and weightage of Nepali subject in Grade 5)

Table 2.9 कक्षा ५ नेपालीमा संज्ञानात्मक क्षेत्र र अङ्कभार

संज्ञानको तह	अङ्कभार
सम्झना/प्राप्ति *(Remembering/ Retrieving()*	१५ %
बोध/एकीकरण *(Understanding/ Integrating)*	४० %
प्रयोग/व्याख्या *(Applying/Interpreting)*	३० %
तार्किक क्षमता/प्रत्यावर्तन *(Reasoning/ Reflecting)*	१५%
जम्मा	**१०० %**

प्रश्नहरूको विशिष्टीकरण (Specification of Items)

तल दिइएको विशिष्टीकरण तालिकामा विषयवस्तुको क्षेत्र, मापदण्ड, भार प्रतिशत, प्रश्नका सङ्ख्या र प्रकार, अङ्कको विभाजन र विभिन्न ४ स्तरमा प्रश्नको विभाजन प्रस्तुत गरिएको छ ।

कक्षा ५ को उपलब्धि परीक्षणका लागि प्रश्नहरूको विशिष्टीकरण

Table 2.10 कक्षा ५ को उपलब्धि परीक्षणका लागि प्रश्नहरूको विशिष्टीकरण

विषयवस्तुको क्षेत्र (Content domain)	मापदण्ड सङ्ख्या (Criteria No.)	भार (Weightage)	जम्मा पूर्णाङ्क (Total Marks)	विभिन्नस्तरमा अङ्क विभाजन (Weightage for items of various standards)
पढाइ	1-6	60%	36	प्रत्येक स्तरको भार देहायको प्रतिशतको नजिक हुनेछ . Level 1: 15%, Levels 2: 40% Level 3: 30% Level 4: 15%
लेखाइ	1-8	40%	24	
जम्मा		100%	60	

द्रष्टव्य :

प्रति प्रश्न १ अङ्क आउने उत्तर छनोट गरिने (SR) बहुवैकल्पिक प्रश्नहरू २० देखि २४ ओटा र प्रति प्रश्न १ अङ्क आउने उत्तर अति छोटो उत्तर आउने रचना गर्नुपर्ने प्रश्न (CR items) ६ देखि १० ओटा हुनेछन् भने २ वा २ भन्दा बढी अङ्क आउने उत्तर आउने रचना गर्नुपर्ने प्रश्न (CR items) १० देखि १५ ओटा हुनेछन् ।

यहाँ विचार पुऱ्याउनु पर्ने कुरा के छ भने तालिकामा दिइएको विभिन्न स्तरको भार प्रारम्भिक मात्र हो । वास्तविक भारको गणना विद्यार्थीको उत्तरसमेतलाई आधार मानी स्तर निर्धारणसम्बन्धी विधिहरूमध्ये

कुनै एक विधि प्रयोग गरी प्रत्येक स्तरको आधार अङ्क (Cut score) बाट निर्धारण गर्नुपर्नेछ । माथि सुझाव गरिएको वर्गीकरण र भारले एकातिर प्रत्येक स्तरका लागि प्रश्नहरू छनोट गर्न सहयोग गर्दछ । प्रश्नपत्र विकास र छनौट गर्दा ४ ओटा स्तरका साथै संज्ञानात्मक क्षेत्रको समेत प्रतिनिधित्व हुनु आवश्यक छ ।

2.5 Item Development and Selection

2.5.1 Item development workshop

Item development process began with an one-day orientation to the well trained item writers on test items development followed by workshop to write draft items by school and university teachers. After computer setting of those developed items, expert workshop was organized. Experts in the workshop reviewed the items to ensure their alignment with curriculum framework and also checked the level and appropriateness of the items.

After the experts' workshop edited the items, the subject committee workshop finalized the test item booklets. After the subject committee workshop, final language editing and layout design were done before printing them in a secured press.

2.5.2 Pre-test of Test Items

To generate item parameters for every items, they were pre-tested on 300 students. Altogether six versions were pretested in the pre-test sample districts and schools. The pre-test was done in the following number of schools and students:

Table 2.11 Number of schools and students participating in pre-test

S. No.	Subject	No. of sets piloted	No. schools piloted	No. of students participated
1	Mathematics	6	80	1800
2	Nepali	6	80	1800
Total		12	160	3600

After the pre-test, items were analysed to produce item parameters. Those parameters were

- difficulty,
- item-rest correlation,
- internal consistency and
- distractors analysis

Table 2.12 Example of item analysis and decision in the pre-test of Mathematics items

Item_code	Set	criteria	Level	Area/ Domain	Item description/ content	Cognitive level	Item type	Maximum	Mean	P value	Item rest correlation	Full marks	Key
1.1.1.4	D	1	1	Geometry	Measurement of angle	Understanding	SR	1	0.50	0.50	0.28	1	B
3.3.1.3	E	3	1	Geometry	Classification of triangle	Understanding	SR	1	0.52	0.52	0.31	1	C
3.4.2.2	F	4	2	Numeracy	National place value system	Remembering	SR	1	0.59	0.59	0.40	1	B
6.7.3.5	F	7	3	Numeracy	Square and cube number	Applying	SR	1	0.35	0.35	0.30	1	B
7.6.1.2	B	6	1	Numeracy	Rounding off	Understanding	SR	1	0.52	0.52	0.35	1	D
…	…	…	…	…	…	…	…	…	…	…	…	…	…

2.5.3 Item Booklet, Scoring Key and OMR Design

Selected items, in each subject, were arranged into three booklets with some linking items between the booklets. Scoring keys for SR items and scoring schemes for CR items were prepared for each booklet. Based on the booklets and scoring schemes, OMR sheets were designed to use for data generation and entry process.

2.5.4 Preparation of the scoring scheme and guidelines

A group of teachers and experts of the respective subjects worked for compilation, review and finalization of the scoring schemes for each subject. For multiple choice and other selected response (SR) types of items, answer keys were reviewed

and reconfirmed. For created response types of items, the possible answers as well as marks to be provided in each step were reviewed and confirmed. For dichotomised items, conditions for 0 and 1 credit were clearly specified. For CR items with partial credit conditions, each of the credits 0, 1, 2 and so on were clearly mentioned. Along with the preparation of scoring scheme for each subject, some guidelines for scoring were also prepared. Rubrics were developed including score distribution, in various skills of writing and levels of proficiency.

2.5.5 Review of test booklets and scoring schemes

At the final stage of item selection and item booklet preparation, subject committee of each subjects reviewed the items and item booklets by editing the items, confirming the data, and formatting the items. The committees prepared the final test booklets which were then sent for preparing Printing Ready Copy (PRC). The subject committees also reviewed the scoring schemes.

While selecting the items and preparing the test booklets for the final test, the following criteria were considered:

- Curriculum based
- Coverage of all contents areas
- Proper representation of various cognitive domains
- Assessing the various levels of proficiencies
- Items having a range of difficulties from p-value 0.15 to 0.90
- Proper discrimination power of the items, item rest correlation r>0.02
- Comparability with previous NASA and TIMSS

2.5.6 Preparation of item Register

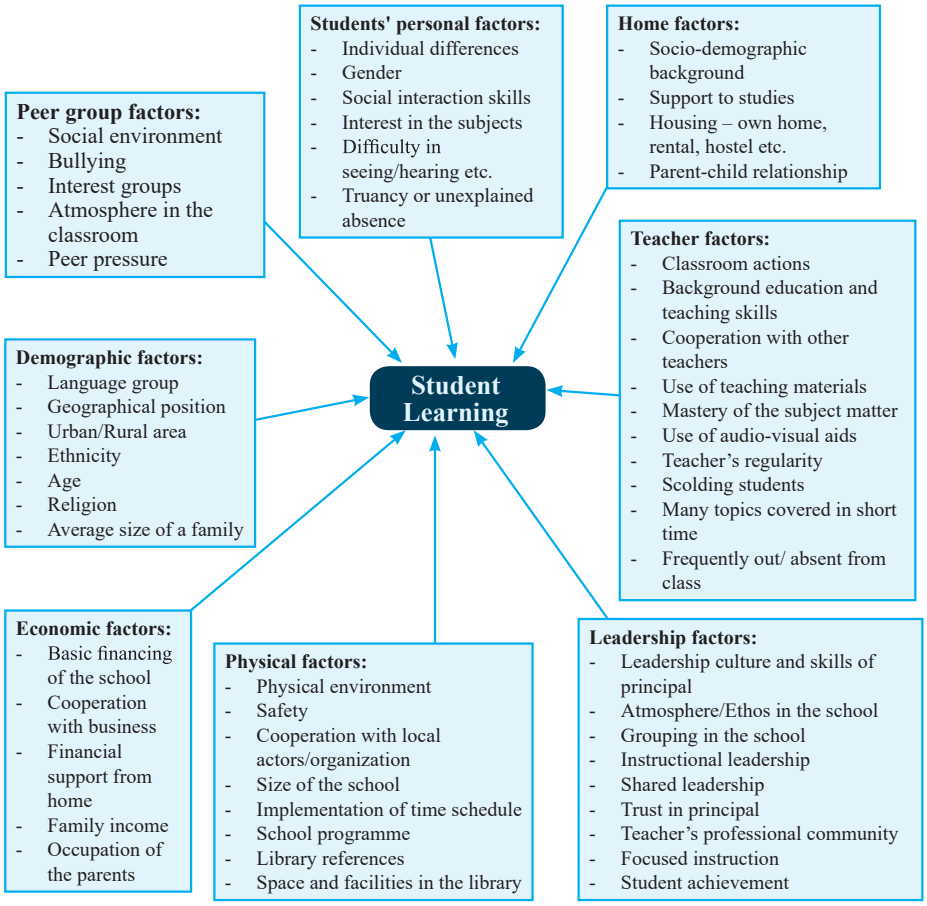
Working with subject experts, ERO prepared an item register in each subject in an excel sheet. Item ID (unique), item descriptor for each item and scoring keys for MC items and various credits as well as description of each credit of CR items were included in the item register. The following is the example of an item register:

Table 2.13 Example of Item register of Mathematics

Item_ code	Set	criteria	Level	Area/ Domain	Item description/ content	cognitive	Item type	Maximum	P- value	Item rest correlation	Full_marks	Set A	Set B	Set C	Key	Label	Label	…
1.1.1.4	D	1	1	Geometry	Measurement of angle	Understanding	SR	1	0.50	0.28	1	1		1	B	M1SRQ1	M1SRQ1	…
3.3.1.3	E	3	1	Geometry	Classification of triangle	Understanding	SR	1	0.52	0.31	1	2			C	M1SRQ2	M1SRQ2	…
3.4.2.2	F	4	2	Numeracy	national place value system	Remembering	SR	1	0.59	0.40	1	3			B	M1SRQ3	M1SRQ3	…
6.7.3.5	F	7	3	Numeracy	square and cube number	Applying	SR	1	0.35	0.30	1	4		7	B	M1SRQ4	M1SRQ4	…
7.6.1.2	B	6	1	Numeracy	rounding off	Understanding	SR	1	0.52	0.35	1	5			D	M1SRQ5	M1SRQ5	…
22.7.2.2	A	7	2	Numeracy	square and cube number	Understanding	SR	1	0.32	0.45	1	6			C	M1SRQ6	M1SRQ6	…
9.10.1.1	A	10	1	Time, money and measurement	Time	Remembering	SR	1	0.88	0.23	1	7			C	M1SRQ7	M1SRQ7	…
11.18.2.5	A	18	2	Time, money and measurement	weight	Applying	SR	1	0.28	0.42	1	8			C	M1SRQ8	M1SRQ8	…
10.20.1.5	C	20	1	Arithmetic	Fraction	remembering	SR	1	0.41	0.42	1	9		9	A	M1SRQ9	M1SRQ9	…
12.17.3.2	F	17	3	Geometry	Volume	Applying	SR	1	0.32	0.35	1	10			D	M1SRQ10	M1SRQ10	…
13.22.3.2	A	22	3	Arithmetic	Fraction	Higher ability	SR	1	0.29	0.21	1	11			C	M1SRQ11	M1SRQ11	…
13.23.2.3	C	23	2	Arithmetic	Decimal	Remembering	SR	1	0.54	0.39	1	12			D	M1SRQ12	M1SRQ12	…
14.24.2.2	C	24	2	Arithmetic	Decimal	Understanding	SR	1	0.32	0.37	1	13			A	M1SRQ13	M1SRQ13	…
14.20.3.6	F	20	3	Arithmetic	Fraction	Understanding	SR	1	0.47	0.52	1	14			B	M1SRQ14	M1SRQ14	…
16.27.2.2	E	27	2	Arithmetic	Unitary method	Understanding	SR	1	0.63	0.49	1	15			C	M1SRQ15	M1SRQ15	…
15.25.2.2	C	25	2	Arithmetic	Decimal	Remembering	SR	1	0.61	0.34	1	16	22	17	A	M1SRQ16	M1SRQ16	…
9.19.1.1	C	19	1	Time, money and measurement	weight	Remembering	SR	1	0.51	0.15	1	17			B	M1SRQ17	M1SRQ17	…
Old 9.4		9	4	Arithmetic	Basic operation	Higher ability	SR	1			1	18	21		B	M1SRQ18	M1SRQ18	…
	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…

2.6 Background Variables

The ERO has developed a framework for collecting background information through the questionnaires after studying students, teachers, and school survey instruments used in various international assessments such as Programme for International Student Assessment (PISA), Trends in International Mathematics and Science Study (TIMSS), and Pan-Canadian Assessment Program (PCAP) together with the tools used in previous NASA conducted by ERO in timeline with some discussions with academicians, practitioners, parents, teachers and the students. Besides, student attitude scale used in previous NASA was revised and used. The following figure shows the overall framework adopted for background information questionnaires used in the study.



Source: ERO, 2018, p. 84

Figure 2.6 Framework for the background information for NASA 2017

Following student background variables were included in the assessment

Table 2.14 Student background variables/variable blocks

School Id	Attitude of student towards subject
Location of School	Student's subject related activities in classroom
Student's gender	Mother's education
Student's age	Mother's occupation
Language spoken at home	Father's education
Caste/ethnicity	Father's occupation
Identity with geography	Number of family members
Time spent beyond school time	Home possession and accessories
Support for study at home	Activities in leisure time at school
Availability of textbooks	Frequency of extra activities at school
Time to reach school	Frequency of participation in extra activities
School opening and attendance days in last month	Attitude towards teacher
Homework and feedback	Attitude towards school
Student's future aim	Bullying at school

Teacher Questionnaire

Teacher questionnaire was used to collect the following information:

- Gender, age, first language
- Teaching conditions including class size, access to resources, percentage of students having textbooks, access to substitute teachers in case of absence
- Educational experience, teacher qualifications and teaching experience
- Teaching-learning practice and conditions at school
- Professional engagement with learning, such as access to and interest in professional development, interest in teaching, and time spent on preparation for classes
- Availability of instructional support such as classroom visits and feedback by head teacher, school supervisor
- Teaching methodology, such as medium of instruction, use of assessment, and style of teaching
- Satisfaction with working conditions, such as tenure, pay rate, and level of supervision
- Relationship between the school and community, such as interactions with parents, involvement in school committees
- Attitude of cooperation from students

Headteacher Questionnaire

Questionnaire for headteachers was used to collect the following information:

- Gender and age
- Educational and management experience and qualifications
- School environment, including the quality of buildings and facilities, as well as availability of resources
- School records, such as fluctuations in student number, student and teacher absenteeism
- Professional engagement of school leadership, such as access to and interest in professional development and interest in education
- Leadership style and use of time
- Assessment of teachers' work
- Satisfaction with working conditions
- Relationship with the community

Students' Attitude Survey

In order to find the relation between attitude of students towards the subject and achievement, the attitude survey questionnaire was administered. The questionnaire was adapted from shortened version of FSMAS, Fennema Sherman Mathematics Attitude Scales (Fennama & Sherman, 1976). The attitude survey questionnaire was included in the students' background information questionnaire. The following are the statements used to identify the attitude of students towards the subject:

Self-confidence

1. Studying Mathematics makes me feel nervous.
2. I am always under a terrible strain in a math class.
3. I am able to solve Mathematical problems without much difficulty.

Value

4. Mathematics is important in everyday life.
5. Mathematics is one of the most important subjects for people to study.
6. High school math courses will be very helpful to me no matter what I decide to study.

Enjoyment

7. I have usually enjoyed studying Mathematics in school.
8. Mathematics is dull and boring.

9. I am happier in a Mathematics class than in any other classes.

Motivation

10. I would like to avoid using Mathematics in college.
11. I am willing to take more than the required classes of Mathematics.
12. I plan to take as much courses of Mathematics as I can during my education.

Socio-economic Status (SES) Survey

The questionnaire to assess the socio-economic status of the family was included in the students' background questionnaire. The aggregate of the students' responses to the questions on the following seven factors indicates the SES of the student's family.

- Two variables related to parental education, including mother's and father's education;
- Two variables related to parental occupation, including mother's and father's occupation;
- Availability of various home accessories;
- Availability of home possessions; and
- Type of school (public or private) attended by student.

2.7 Test Administration

Preparation for test administration begins with printing, packing and delivery of test items and background questionnaires. ERO conducted a one-day orientation on test administration and test booklet collection process to the head teachers of each sample school in 26 districts. With the help of two teachers, the head teacher of each sample school administered the test. Subject teacher and head teacher of the sample school (in which test was administered) filled teachers' and head teachers' questionnaires respectively. Then students' answer sheets as well as teacher's and head teacher's responses were collected in the scoring centre in Kathmandu. The process followed for the purpose of test administration is described in this section.

For completion of the works, some of the tasks of test administration were outsourced to a consulting firm, while others were carried out by the DEOs and the schools. The sub-headings that follow deal with the tasks and process adopted to accomplish the work of test administration of NASA 2018.

Following activities were completed before administering the test to maintain peace and security in a low-stake environment.

- Delivery of head teacher guidelines for test administration. These guidelines mentioned every steps of test administration on what to do and what not to do .
- Delivery of test booklets in the sample schools.
- Orientation to the district level head teachers and monitors.
- Test administration arrangement by allocating monitoring team to the centres, scheduling test administration.
- Random selection of the test taker students where the number of sample students was less than the number of students required in the sample class.

Test administration had two parts: in the first part, background questionnaire of students, head teacher and corresponding subject teacher was administered. After completion of the background information questionnaire response by the students, a ten-minute break was scheduled. After the 10 minutes break, two hour test administration was completed. To ensure proper administration of the test, monitoring and sample school visits were conducted by different agencies during the test administration. Educational Development Coordination Unit - EDCU (the then District Education Office) not only managed the whole process of test administration, but also monitored the administration process at school level. The ERO also sent at least one person to each district to facilitate and monitor the administration of the test. Besides, the consulting firm also monitored the process of test administration. After the test administration, the consulting company collected the booklets and delivered them to the Kathmandu centre. After collecting the answer sheets in the Kathmandu centre, data preparation was completed by adopting the following steps:

- Optical Mark Reader (OMR) sheet development and printing
- Answer sheet coding and marking and scrutiny
- OMR input of the scores and cleaning the data
- Submission of clean data and marked answer sheets to the ERO

2.8 Item Parameter Estimation, Item Review and Calibration

Item parameter estimation of each item was carried out and the items were reviewed accordingly. During the analysis, decisions were made on whether or not to use any particular item in the analysis. Classical as well as IRT parameters were estimated to review the items. Item parameters in IRT were used not only for item selection but also to estimate students' latent ability. Based on the item parameters of linking items, three versions of tests were calibrated and these three sets were integrated into single set for analysis. Item parameter estimation, item review and calibration of

the test were some of the key processes of IRT analysis from which students' ability was estimated and data were further analysed.

2.9 Reliability and Validity

The validity of the test items was assured by using the assessment framework. The item level parameters and set level reliability of Mathematics and Nepali subjects are given below:

2.9.1 Reliability

Table 2.15: Reliability of item booklets in Mathematics

S. No.	Booklet	Reliability
1	Math set 1	0.86
2	Math set 2	0.88
3	Math set 3	0.89
4	Math Combined set (Joint)	0.856

All the sets were highly reliable (reliability>0.82).

Table 2.16: Reliability of item booklets in Nepali

S.No.	Booklet	Reliability
1	Nepali set 1	0.901
2	Nepali set 2	0.886
3	Nepali set 3	0.877
4	Nepali Combined set (Joint)	0.888

All the sets were highly reliable (reliability > 0.82).

2.9.2 Validity

Since all the test item were developed and standardized according to National curriculum and the assessment framework, NASA 2018 test sets were considered to be valid. In the test, there were a total of 110 items from all three sets.

Table 2.17: Number of items asked in various content area based on the weightage given by Mathematics curriculum

S. No.	Content area in Mathematics	Number of items
1	Arithmetic (basic operation, fraction, decimal, percentage, unitary methods, simple interest)	32

S. No.	Content area in Mathematics	Number of items
2	Data (Bill, budget, statistics)	10
3	Geometry	18
4	Numeracy (Concept of numbers and naming)	21
5	Sets and algebra	14
6	Measurement (time, money, distance, area, volume)	15
	Total number of items	110

Altogether, 110 items were asked and full marks of joint sets was calculated to 128 maximum.

Table 2.18: Number of items asked in various content area based on the weightage given by Nepali curriculum

S. No.	Content area in Nepali	Number of items
1	Writing with Grammar	22
2	Reading with Vocabulary	82
	Total	104

In Nepali, 63 items were selected response and 41 items were constructive response items. Among them items carrying 51 marks were asked from writing with grammar and items carrying 83 marks were asked from reading with vocabulary.

2.10 Item Parameters

In the table, the "Avg Delta" represents the IRT parameter of difficulty. The Remaining ones are classical parameters.

Table 2.19: Example of item parameters (Mathematics)

Item	N	Facility	Item-Rest Cor	Item-Total Cor	Wghtd MNSQ	Avg Delta
item:1 (M1SRQ1)	4484	68.69	0.31	0.36	1.03	-1.16
item:2 (M1SRQ2)	3990	58.55	0.32	0.37	1.05	-0.56
item:3 (M1SRQ3)	4335	68.97	0.37	0.42	0.98	-1.16
item:4 (M1SRQ4)	8892	37.58	0.29	0.35	1.07	0.34
item:5 (M1SRQ5)	4502	48.67	0.34	0.39	1.04	-0.16
item:6 (M1SRQ6)	4502	45.36	0.36	0.41	1.03	0.00
item:7 (M1SRQ7)	4516	80.14	0.33	0.37	0.97	-1.87
…	…	…	…	…	…	…

(For more detail, see appendix.)

Item analysis was carried out by using ConQuest software that generated various item level statistics and curves as well. Some examples of item analysis output are given below:

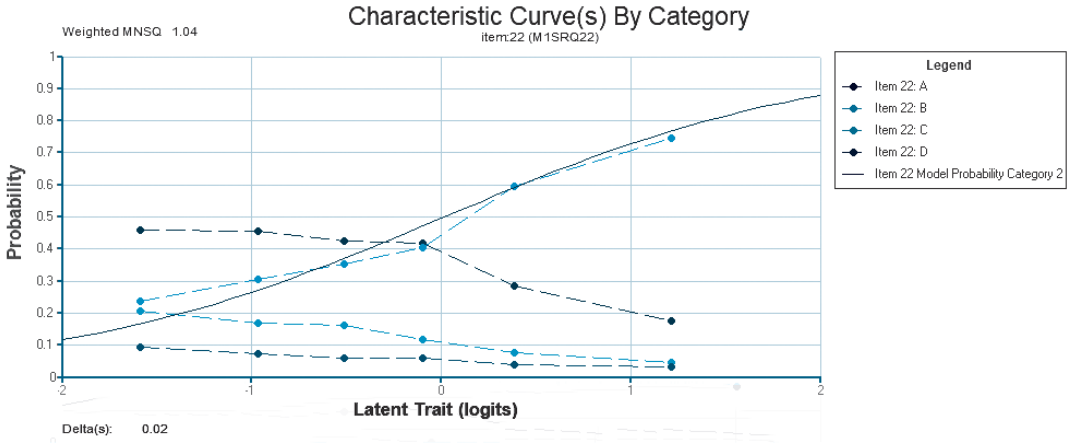


Figure 2.7 Example of item characteristics curve in Mathematics

In the figure, B is the correct option (answer), other options A, C and D are the distractors.

Similarly, an example of ICC in Nepali subject is presented in figure 2.8.

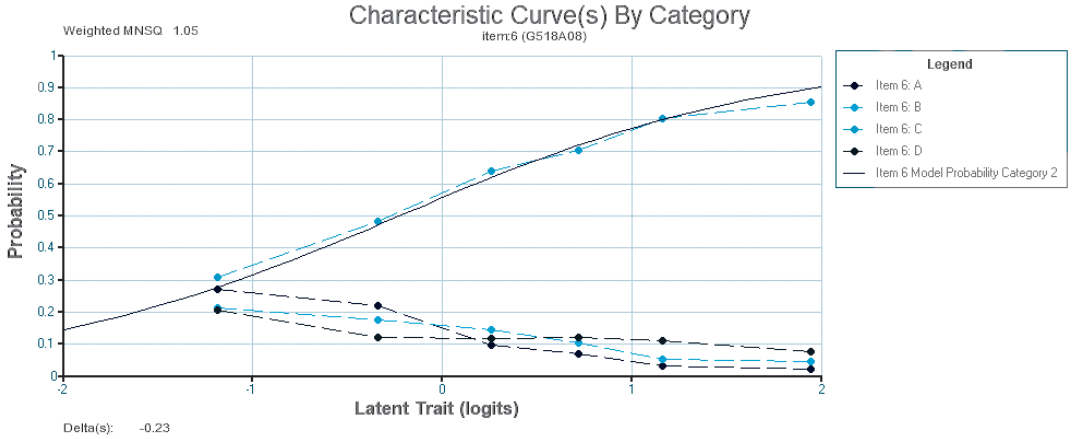


Figure 2.8 Example of item characteristics curve in Nepali

In the item analysis, the acceptable range of Item-rest correlation was taken r > = 0.2 and Weighted MNSQ was considered from 0.8 – 1.20 acceptable. Facility index was used as it was because all the items are already standardized. Two items in

Mathematics from set-2 (item number 3 and 4) were discarded because of out of range MNSQ and negative item-rest correlation. The produced ICCs were used to analyse the item's appropriateness in the model.

2.11 Plausible Values (PVs)

Plausible values (PVs) improve precision of prediction ability for the population estimates. The PVs are calculated with conditioning background variables and some school related index. Conditioning provides unbiased estimates for modelled parameters. In this assessment, five plausible values (PV1 – PV5) were used to estimate population ability.

In this context, Yamamoto & Kulick (2000) mention that the PVs approach "uses students' responses to the items together with all background data in order to estimate directly the characteristics of student populations and sub-populations" (cited in Laukaitytė, 2016, p. 9). But, PVs are not individual test scores; they are the measures of the performance of population.

It produces unbiased estimate of population parameters if assumption of scaling is reasonable, but it is not fair to use it for level of student ability.

The following inputs were prepared to generate the PVs:

- Case estimation using weighted likelihood estimation (WLE)
- Provinces
- School type
- Group with highest frequency is set to zero before using conditioning.
- School mean index of WLE, etc.

The following table is an example of plausible values in Nepali subject drawn by conditioning run:

Table 2.20 PVs and RWGT (example from Grade 5 Nepali)

School ID	UID	SCH_M Index	MPV1	MPV2	MPV3	MPV4	MPV5	RWGT1	RWGT2	RWGT3	RWGT4	…	W_STU	MSSPV1	MSSPV2	…
70300006	23506	-0.89	-0.32	-0.59	-0.96	-0.93	-0.73	39.39	39.39	39.39	39.39		39.39	483.84	470.74	
440120007	20796	-1.10	-1.38	-0.05	-1.37	0.47	-1.05	54.71	54.71	54.71	54.71		54.71	430.98	497.27	
640170004	22900	-0.74	-0.70	0.54	-1.79	-1.66	-0.96	39.39	39.39	39.39	39.39		39.39	464.86	527.11	
740500001	20017	-1.42	-0.43	0.90	-0.97	0.01	0.12	39.39	39.39	39.39	39.39		39.39	478.32	545.16	
550130004	26140	-1.93	-2.84	-3.75	-3.75	-3.22	-3.01	39.39	39.39	39.39	39.39		39.39	357.89	312.72	
580270006	26907	-2.09	-0.72	1.27	-0.08	-0.96	-0.59	39.39	39.39	39.39	39.39		70.35	463.96	563.32	
200310002	30394	-1.26	0.36	-0.10	-1.34	0.51	0.61	70.35	70.35	70.35	70.35		39.39	517.79	495.17	
440120007	28616	-1.10	-0.82	-1.40	1.00	-0.17	-0.72	54.71	54.71	54.71	54.71		54.71	458.87	429.78	
20050001	29496	-0.93	-1.65	-0.56	-0.43	-1.07	-0.49	39.39	39.39	39.39	39.39		39.39	417.64	472.14	
460140005	27785	-1.43	-3.08	-4.04	-4.45	-3.32	-3.32	44.77	44.77	44.77	44.77		44.77	345.93	297.92	
40100014	23965	1.44	-0.86	0.13	-0.12	0.32	1.00	39.39	39.39	39.39	39.39		39.39	456.99	506.25	
40170008	24035	1.82	2.94	1.91	1.66	2.33	3.10	39.39	39.39	39.39	39.39		39.39	646.88	595.30	
40170008	24037	1.45	2.78	3.31	3.21	3.16	2.22	39.39	39.39	39.39	39.39		39.39	639.09	665.61	
240660014	24692	0.69	2.61	3.58	2.14	3.41	2.26	39.39	39.39	39.39	39.39		39.39	630.36	679.22	
370540009	21538	0.63	2.69	4.00	2.15	3.17	2.87	39.39	39.39	39.39	39.39		39.39	634.34	699.76	
740410005	14159	1.81	2.44	3.07	3.13	2.23	2.92	54.71	54.71	54.71	54.71		39.39	622.09	653.52	
240660014	18604	1.10	2.93	3.56	2.31	2.96	2.74	39.39	39.39	39.39	39.39		54.71	646.48	677.92	
460740008	23619	0.96	3.21	2.49	2.55	3.33	2.82	39.39	39.39	39.39	39.39		39.39	660.74	624.45	
250270035	16729	0.60	3.75	2.84	4.74	3.46	3.54	46.90	46.90	46.90	46.90		51.83	687.60	642.05	
160950007	15967	1.4	3.44	4.22	3.83	4.22	4.22	51.83	51.83	51.83	51.83		46.898	672.15	710.97	

2.12 Comparing the Overall Results of NASA 2018 with the Results of 2018

To find the trend of learning achievement, linking items' percentage of correct answer is reported in the first step. Then, for the IRT based reporting, linking items of NASA 2015 and 2018 are separately calibrated to generate the item parameters. Separate item parameters are generated by using one parameter Rash Model. From those parameters latent ability of both years' students was calculated. Then comparison by using latent ability (WLE) and item parameter difficulty of those items were separately used to find the achievement gap. By using WLE and item parameters, achievement gap is found to determine consistency in the result.

Trends of NASA cycle results over the years are presented based on the following methods:

1. A comparative item wise classical parameter (percentage of correct answer) is reported between both years.
2. Average score of WLE 2015 Vs WLE 2018 is reported.
3. Shift in mean by using mean shift calculation by difficulty parameters is used.

2.13 Provincial Results

Provincial results are prepared separately in each subject. The provincial results provide the opportunity of comparing the results in major variables. In each subject, provincial report begins with comparing overall mean scores of provinces followed by the mean scores in relation to various influencing variables on the achievement of students.

RESULTS AT NATIONAL LEVEL

CHAPTER 3
RESULTS IN MATHEMATICS AT NATIONAL LEVEL

3.1 Introduction

In this chapter, the basic results of population estimates drawn from the responses of 28381 students from 24 districts and 1400 schools are presented. Population estimates presented in this chapter are based on the five plausible values drawn from WLE and conditioning variables like school mean index, student background variables, student weights, provinces, gender. The population mean/achievement score is presented in all basic results with either standard error or in confidence interval (CI). In most of the bar-charts, the confidence interval of the population mean is represented by a line with cap in both ends. Such population estimates do not represent the individual level results. Thus, all the achievement scores reported are the weighted mean scores weighted by adjusted student weights, and the difference is reported at confidence level of 95%. The standard errors and confidence intervals were estimated to identify whether the difference in mean was statistically significant.

The students' ability scores were transformed into mean 500 and standard deviation of 50. This reporting has always national mean score fixed at 500 points to compare any two or more groups. The formula for transforming the student ability (logits or θ) was:

Average score = 500 + logits * 50

Variation of average score comes from the variation in the logits (latent ability of students/WLE). The five PVs are also generated based on the logits.

3.2 Wright-map of student ability and item difficulty in Mathematics

The Wright-map is organized as two vertical histograms. The left side shows candidates and the right side shows the items. The left side of the map shows the distribution of the measured ability of the candidates from most able at the top to least able at the bottom. The items on the right side of the map are distributed from the most difficult at the top to the least difficult at the bottom. In the following figure, student ability (θ) in the left and NASA 2018 items to the right are plotted in the same scale. When a person and an item lies at the same level, probability of responding that item by the particular person is 50%. Figure 3.1 presents the NASA 2018 Mathematics Wright-map.

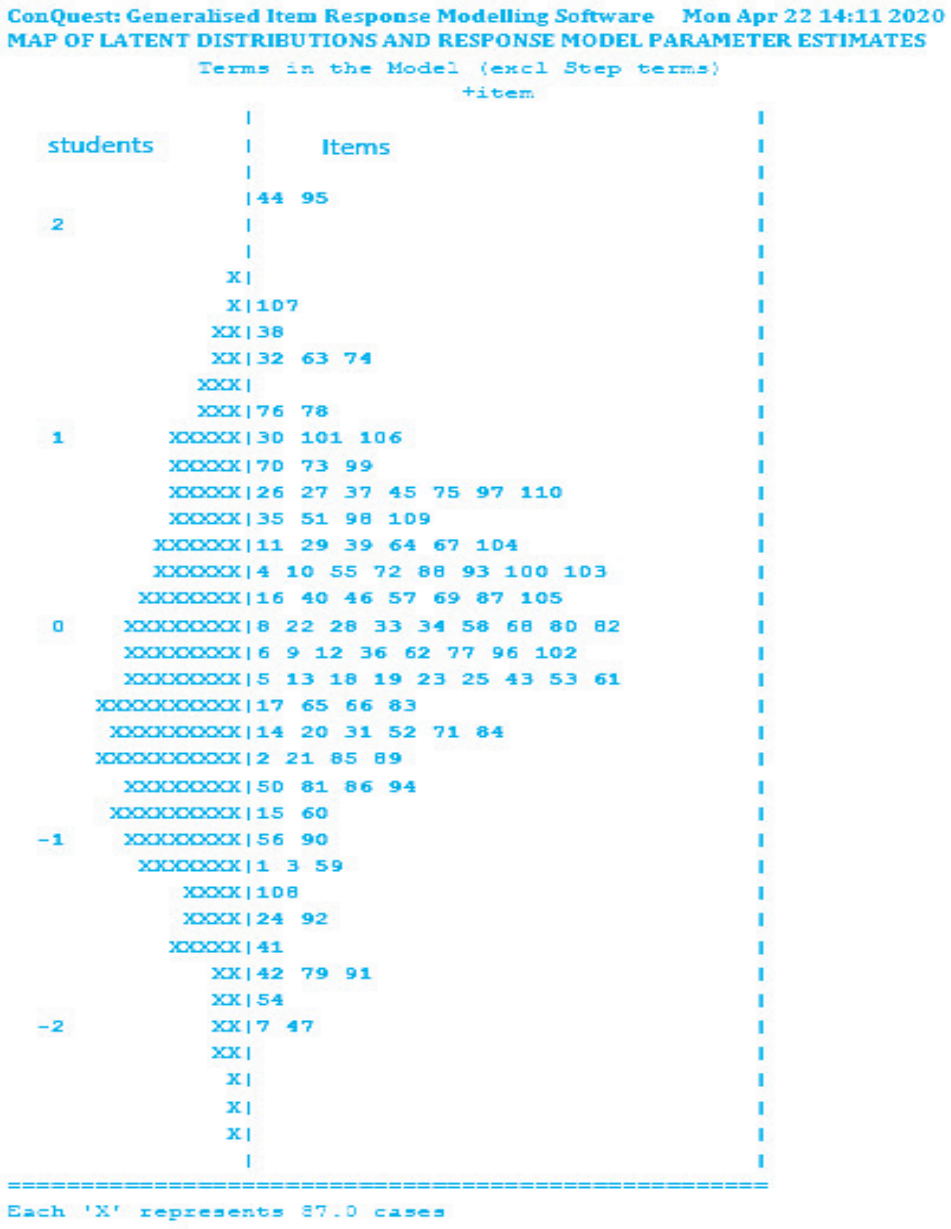


Figure 3.1 Wright map showing person and item in the same scale

To the left side, an 'X' represents 87 students, their latent ability is given in the logit scale ranging from -2 or less to +2 or more. The distribution of students against the items asked (item numbers are shown to the right side) reveals that most of the items were difficult for the students. Although items were pre-tested and based on the grade 5 curriculum, most of the students are lagging behind below the average latent

ability '0'. This indicates that items were difficult for the participant students. This further indicates that performance level of the students was achieved not as expected by the curriculum.

3.3 Plausible Values, their Mean and Standard Error

After estimating the student ability (θ) in the form of WLE, five plausible values (PV1 to PV5) were generated by conditioning the data with student background variables and school mean index. Those plausible values are transformed in to a scale of mean 500 and standard deviation 50. Those values were weighted by student full weight and using 350 replicates (just half number of number of schools taken in the sample for Mathematics). After all, MSSPV1 to MSS PV5 were calculated to report the population estimates. The mean and standard error of five plausible values are presented in table 3.1.

Table 3.1 Standard Error of five plausible variables in Mathematics

Plausible values	N (Sample)	Mean	SE
MSSPV1	14174	500.606	1.296319
MSSPV2	14174	500.682	1.282064
MSSPV3	14174	500.803	1.281243
MSSPV4	14174	500.719	1.270628
MSSPV5	14174	500.384	1.273071

3.4 Defining Proficiency Levels in Mathematics

Assessment framework for NASA 2018 recommends to set performance level into four levels. For this, three cut-points for proficiency levels were decided by dividing the range of 204 (maximum 612 – minimum 406) by the interval of 52. Thus, four proficiency levels cut-points were 446, 497 and 549 decided. Table 3.2.1 shows how proficiency levels are determined.

Table 3.2 Proficiency levels and the score range in Mathematics

Proficiency Level	Score
Level 3 (Advanced)	561 above
Level 2 (Proficient)	509 - 561
Level 1 (Basic)	458-509
below level 1 (Pre-basic)	below 458

Based on the descriptions of items that correspond to each of the above proficiency levels in item-person map in Mathematics together with subject experts' judgment, the descriptions of students' four level proficiency have been defined. These descriptions of four proficiency levels in Mathematics for Grade 5 indicate what a student at particular competency level can do in Mathematics.

Internationally, students who cross 67% of their achievement are considered as *Minimally Accepted Candidate*. Replicating the same concept in determining the minimum acceptance level of learning in those four proficiency levels is possible. However, in this assessment, around 50% items were objective and almost equal weightage was given to subjective items. So, in this analysis 50% correct answer was supposed to be threshold of minimum accepted proficiency for any of the four levels. From this point of view, student response on every item was analysed to find the response rate of those four level students. For this, at the first step, below level 1 (pre-basic) items were detected then, level 1, level 2 and level 3 respectively. In such a rigor, all the items were assigned to different four levels to draw proficiency descriptors. Table 3.3 specifies the minimum proficiency level of all four level students in a descriptive form.

Table 3.3 A summary of minimum proficiency level in all four levels.

Level	Score	What students can typically do
Below level 1: pre-basic	below 458	Below level 1 students have quite limited knowledge and skills in Mathematics. Around 50% of them do not have any sense of reading and writing number and number operation. Around half of them could have superficial knowledge of number and number operation. However, they cannot calculate and solve Mathematical problems. Only half of them have some knowledge of time like hour and minutes, day and hour, month and years. Some of them are able to choose correct answer when options are given in MCQ items. Almost all of them can not perform any Mathematical subjective calculations independently.
Level 1: Basic	509 - 458	Basic level students have superficial knowledge and skills in most of the Mmathematical contents, however, they are struggling with calculations. They can identify the ordered pair of a point, square pattern in dots, sum of decimal numbers, place of a digit in numbers, relation of kg and gram and recognize

Level	Score	What students can typically do
		numbers, mixed fraction and type of angle. They also have limited knowledge of formula of volume and area, can estimate angle shown in figure and relation between decimal and fraction. They can read the table and bar graphs to take simple information but cannot draw conclusion by comparing the data. They can solve very simple problems of unitary methods, subtract small same denominator fractions, round numbers in the nearest tenth of a decimal number. They can also recognize limited square numbers, express Mmathematical sentence in Mmathematical language to calculate. They can also subtract a univariate one degree algebraic term from another, find the value of x in one variable equation and subjective; can simplify algebraic expression (univariate) in algebra. However, they cannot perform Mmathematical calculations of their grade level independently.
Level 2: Proficient	561 - 509	Proficient level students have wide range of knowledge and skills expected by the curriculum of grade 5. Most of them have adequate knowledge in the following areas: • They can round numbers into nearest tens and hundreds, compare numbers, find cube number up to 5 and vice versa, solve simple word problems involving basic operations (+,- and x) and time related addition, identify the factors of numbers below 20, know the relation of quintal and kg. • They can identify right angled triangle, measure side of geometric figures, find perimeter, area and volume in geometry and measurement. • They can identify proper fraction, select simple interest when principal and rate is given, solve very simple unitary methods problems, compare length in decimal numbers, find fraction of a number (eg. 1/4th of 12), have the concept of converting tenth fraction into decimal and place of digit in decimal number; can change percentage into fraction; can solve word problem of addition and subtraction of fraction with same denominator; can change kilometer in decimal into grams and convert fraction in percentage (with 100 in denominator).

Level	Score	What students can typically do
		• In Algebra, they can represent a set in set notation {}, select the sum of a number and a variable, can solve one-variable linear equation and simplify, add Algebraic expressions having one degree terms. However, some (less than 50%) of them have limited ability in reasoning, problem solving and finding the relationship between two variables.
Level 3: Advance	561 above	Advance level students have almost all of the abilities as expected by the curriculum. They can independently calculate and solve Mathematical problems of their grade level. They have abilities of thinking critically, reasoning and finding the relationships among variables.

Note: although some students of lower level (for example: below level 1) have also answered few items of upper level (for example: level 1) correctly, those items were located in upper level (level 1) because rate of correct answer of those items was less than 50% in lower level (below level 1).

3.5 Distribution of Students by Proficiency Levels

The student achievement scores based on 5 plausible values (PV1 to PV5) were analysed in terms of four proficiency levels of students' achievement. Level wise descriptors are presented in section which also presents the number of students falling in those four levels from population estimate. The standard error of the percentage of students is also presented in table 3.1.3a.

level 3. Table 3.4 shows the weighted percentage of students distributed over four proficiency levels.

Table 3.4 Distribution of the students in various proficiency levels and their Standard Error

Proficiency level	% of students	SE	N_cases	NU_cases	NU_psu
Below level 1 (pre-basic)	32.166	1.037	220496.8	4594	567
Level 1: Basic	39.577	0.833	272024.3	5609	677
Level 2:Proficient	24.006	0.837	163676.9	3320	595
Level 3: Advance	4.252	0.473	28300.73	570	187

NB: SE = Standard Error, N cases = Number of cases/students in the population, NU cases = Number cases/students in the sample, NU psu = Number of Primary Sample Unit (schools).

The Below level 1 (Pre-basic) indicates the lowest ability of students who are struggling in the classroom where as level 3 shows the highest level of proficiency that even crosses the grade level. Figure 3.2 shows how students are distributed over those levels visually.

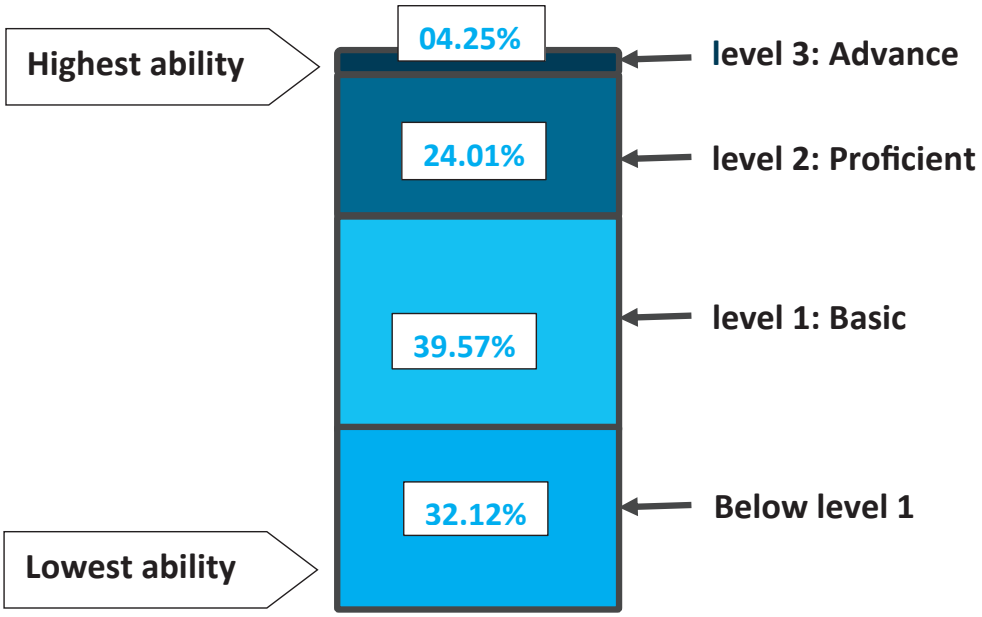


Figure 3.2 Distribution of number of students (%) in different levels

Figure 3.1.4 shows that 32% of the students fall in *below level 1*, 39.57% in level 1-basic, 24% in proficient level and 4.25% in advance level. Below level 1 (32%) students can not even write numbers and do Mathematical basic operation. They are struggling in grade 5 with no grade level minimum ability as expected by curriculum. Basic level (level 1) students can not perform Mathematical calculations independently. However, they have limited basic knowledge in various concepts of Mathematics. Proficient level (level 2) students have adequate knowledge and skills in their grade level where as advance level (level 3) students have advance level of Mathematical ability with ability to logically solve mathematical problems. They can find relationship between two quantities and think critically.

From total number of students, 28.26% of them have adequate mathematical knowledge and skills who lie in proficient level (24.01%) and advance level (4.25%). Students in basic level (39.56%) have limited basic mathematical concept however they can not perform Mathematical calculations of their grade independently. 32 out of 100 students do not get sense of what is taught in the classroom of the total as they lie in below basic level (32.12%). They are struggling with Mathematical concepts and are left behind. Hence, more than 70% of the students are below the level of competencies expected by the curriculum.

3.6 Minimum Level of Achieved Curriculum

The assessed curriculum is that which is reflected by the assessment or evaluation. It can be either formative or summative evaluation of the students. Assessed curriculum is a tested curriculum by school, national or international organization based on the written curriculum/intended curriculum. It is valuable because it enables the educational organizations and stakeholders to evaluate the impact of written and taught curriculum upon students. It determines the level of the learned curriculum. Research (e.g. Berliner, 1984; Turner, 2003) indicates that the mismatch between assessed and taught curricula has serious consequences (cited in MeshGuide). This section presents the level of learning in the form of achieved curriculum in terms of percentage. In this analysis, it is assumed that every test item is equivalent in the sense that each of them represents a learning objective mentioned in the written curriculum.

As mentioned in sub-section *3.4*, 67% correct response can be considered as cut-score for being minimally proficient at any level. As in this assessment, around half number of items were objective type (MCQ) and half of them were subjective. 50% correct responses are considered as the threshold of minimum level of accepted proficiency at any of the four levels. Hence, test items were organized in terms of at least 50% correctly answered items or more at each level of students. Based on this criterion, all the items were re-allocated into four levels. From this rigorous analysis, performance descriptors were developed.

In every level of the proficiency, there are ranges of students from being very weak performers to the highest performers. Considering 50% as the threshold of minimum proficiency of any four levels, percentage of learning was Mathematically calculated based on the number of items answered correctly. Mathematical value of achieved curriculum is thus given in table 3.5.

Table 3.5 Mathematical presentation of the achieved curriculum

Performance level	Achieved curriculum
Below Basic level (32% students)	5% of the curriculum
Basic level (39.57% students)	28% of the curriculum
Proficient level (24% students)	62% of the curriculum
Advance level (4% students)	96% of the curriculum

Table 3.5 reveals that 32 students out of 100 fall below basic level who have achieved only 5% of the tested curriculum (% of items). Similarly, only 28% curriculum is achieved by basic level students. Altogether, 71% students are under the 28% achievement of the tested curriculum in Mathematics. This huge mass of the students represents underperforming group. The proficient level students achieved 62% of the tested curriculum and advance level students achieved 96% of the tested curriculum. It shows that only 29% of the students have achieved adequate knowledge and skills in Mathematics. The gap in achievement against curriculum objectives between below basic level (5%) and advance level students (96%) is 91%. This reveals that inequality in learning in the classroom is remarkably high.

3.7 Overall Mean Score by Province

In the Federal context, Nepal is divided into seven provinces and 753 local government units. While selecting the schools as Principal Sample Units (PSU), provinces were treated as strata. The average scores reported in this section are the transformed/scale score at 500 national average. National mean is taken as a reference to compare the provincial mean. Those provinces whose average score is above the mean score are recognized as better performing provinces whereas below 500 are assumed to be low performing.

The mean score of achievement reported here is based on the plausible values as mentioned in introduction chapter. In the figure 3.1.4, a vertical dotted line represents the national mean score of achievement and horizontal bars represent the achievement scores by province.

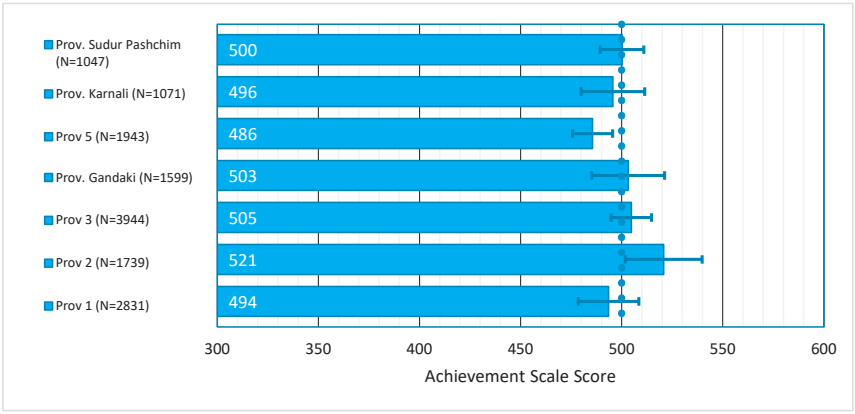


Figure 3.3 Provincial level mean achievement scores in Mathematics

The data reveals that student achievement in Mathematics in province 2 was found highest (521) among the seven provinces, while student achievement in province 5 was found to be the lowest (486). The difference between the provinces ranges by 35 scale score. Learning achievement of *province 1, 5 and 6 is lower than the national average.*

To identify the reasons how province 2 achieved the highest score in Mathematics, a district wise mean score was calculated. Figure 3.1.4b presents the mean score in green bars (dotted) when it is above the national mean (500) and the red bars when it is below national mean.

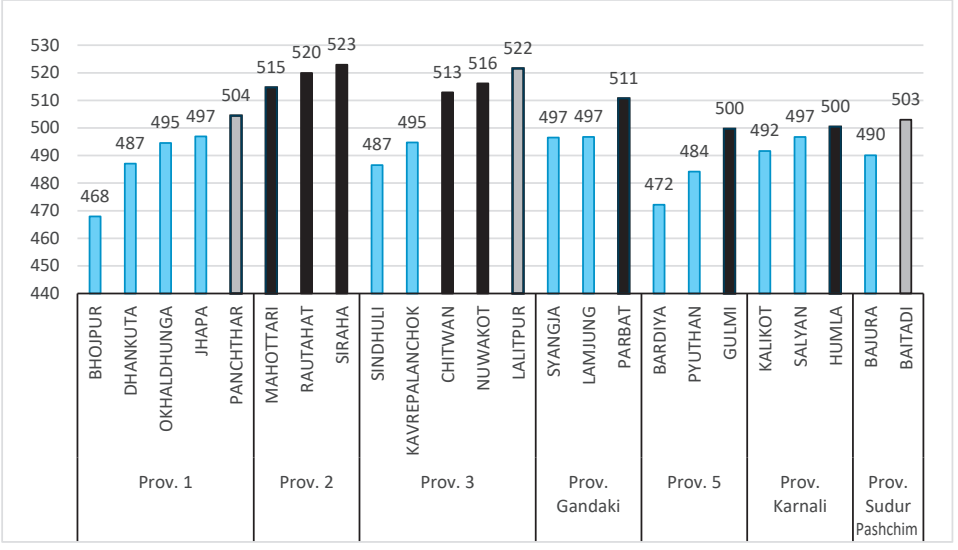


Figure 3.4 District wise mean achievement score

Figure 3.4 shows that out of 24 districts, only 10 districts' mean score was above the national mean (500). In province 2, all three districts had higher mean score than the national mean, Mahottari (512), Rautahat (520) and Siraha (523). The lowest achieving province was province 5 in which mean score of Bardiya (472), Salyan (484) were below national mean and only Gulmi (500) was just equal to national mean. Moreover, the lowest achiever district comes from province 1 that is Bhojpur (468) which was nearly equal to the *below level 1*. This indicates that Bhojpur district students were struggling with Mathematics learning. Similar was the situation in Bardia from Province 5 (474 score). Sindhuli from Province 3 had the lowest performance (487 score). Variation in achievement between the districts was high (55 scale score).

3.8 Results by Gender, Ethnicity and Home Language

From equity concern of learning, gender, ethnicity and home language are considered to be the important variables in learning. In this section, achievement scores of various groups have been compared.

a. Achievement by Gender

For the equal level of learning to take place, girls and boys should have equal opportunity and support in their study. In the background questionnaire, students had reported their gender. In this data, there were 6855 (48.4%) boys, 6978 (48.6%) girl and 2.4% students did not reveal their gender. Based on this sample, the weighted mean score is presented in figure 3.5.

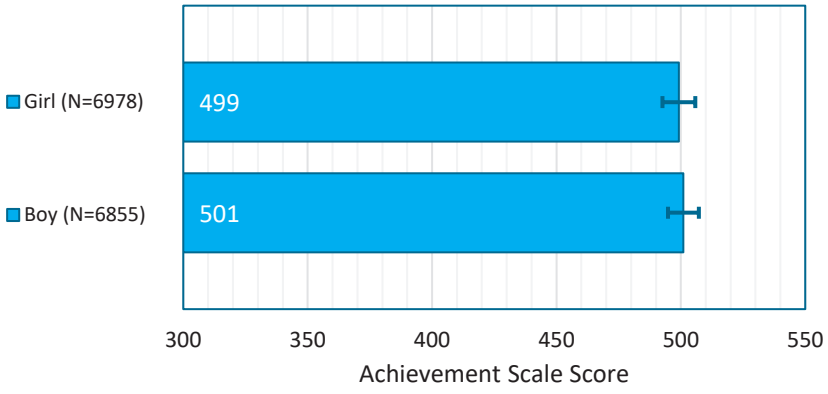


Figure 3.5 Mean score in Mathematics by gender at national level

The analysis of data from a gender lens, demonstrates that boy students' achievement score (501) was higher than the mean score of girl students (499). Thus,

boys have outperformed girls by 2 scale score. Moreover, the achievement of girl students was 1 score lower than the national mean (500) and boys have 1 score above the national mean. Although the difference of scores between boys and girls was found statistically significant at $p<0.05$, the effect size Cohen's $f = 0.003$) confirmed that the difference was very narrow. From the equity perspective, *such a narrow effect size in difference in learning performance of boys and girls was very close to gender parity level.*

When the data was disaggregated at the provincial level and replicate module was run, the difference in the performance was found. The performance of the students by gender at province level is presented in table 3.6.

Table 3.6 Performance of the students in Mathematics at province level by gender.

Province	Gender	Mean	N	Std. Error of Mean	Sig.	Eta squared	Effect size
Prov. 1	boy	496	66325	0.188	0.000	0.002	0.045
	girl	491	69910	0.184			
Prov. 2	boy	525	35498	0.288	0.000	0.007	0.084
	girl	516	40051	0.247			
Prov. 3	boy	504	95307	0.159	0.000	0	0.000
	girl	505	92005	0.163			
Prov. Gandaki	boy	503	41889	0.258	0.000	0.001	0.032
	girl	500	40784	0.252			
Prov. 5	boy	484	43828	0.221	0.000	0	0.000
	girl	485	41761	0.235			
Prov. Karnali	boy	497	27369	0.300	0.000	0	0.000
	girl	495	26258	0.335			
Prov. Far-western	boy	503	25190	0.298	0.000	0.003	0.055
	girl	497	26244	0.274			

The above table shows that there was difference in achievement between boys and girls in all the provinces. However, the difference was very narrow in all the provinces, the effect size shows that province 2 had comparatively wider difference in learning achievement between the boys and the girls.

b. Achievement by Caste/Ethnicity

From equity concern, ethnicity is an important variable. Various cultural qualities are embedded in ethnicity as it influences diversity, social integration, time spent in the study. All these factors have their effect on learning (Virginia, 2017). The caste/ethnicity wise comparative result is presented in figure 3.6.

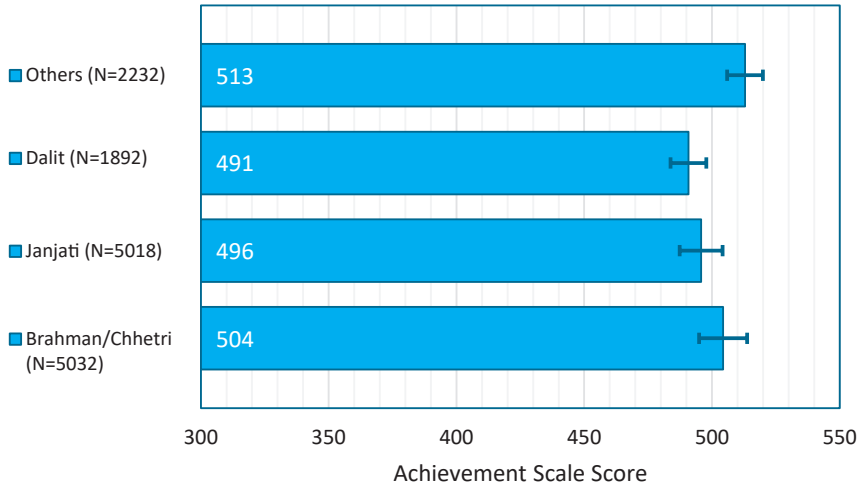


Figure 3.6 Mean score in Mathematics by ethnicity

The analysis of data from a caste/ethnicity lens indicates that the students from Brahman/Chhetri caste have higher (504 score) mean achievement score than the national mean (500 score). However, in the student's inter-ethnicity category "others" achieved the highest (513 score) in all the caste/ethnicity categories. Students from Dalit and Janjati achieved lower than the national mean. Although the difference of scores between the categories was statistically significant at $p<0.05$, the difference was very narrow (effect size : Cohen's $f = 0.03$) indicating that there was *only 3% variance explained by caste/ethnicity when it is considered as the only variable. Since, 'others' category was a mix of different castse/ethnicities, it was difficult to recognize in which caste/ethnicity did those students belong.*

c. Achievement by Home Language

Students were asked "Which language do you speak most in your home?". Their response revealed that of the total, 63% of the students spoke Nepali language in their home whereas 30.3% students reported "other" languages. Few (4%) students did not report anything. Later the missing values were recorded within "other" categories. Based on the student response, achievement score is presented in the figure 3.7.

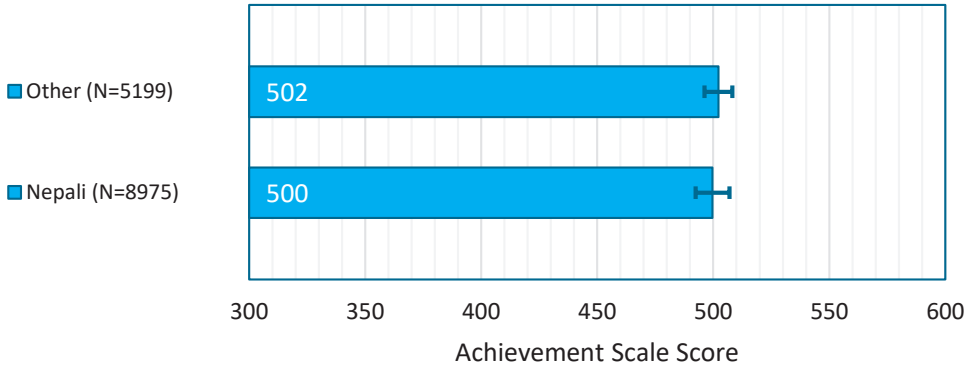


Figure 3.7 Mean score in Mathematics by home language

The dataset reveals that students speaking *Nepali* at their home had performed lower (500 score) than those who spoke *other* than Nepali language (502). The difference in mean between those two groups' score was significant at $p<0.05$. However, the difference was very narrow (effect size, Cohen's $f = 0.003$), indicating that variation in achievement based on home language was very low. *This result revealed that home language is not the determinant factor in Mathematics learning.*

3.9 Results by Types of Schools

As there were randomly selected 520 community schools and 180 institutional schools in the sample, most of the institutional schools are concentrated in the urban areas where as community schools are distributed all over the geographical locations. Comparative analysis of the community and institutional schools is presented in figure 3.8.

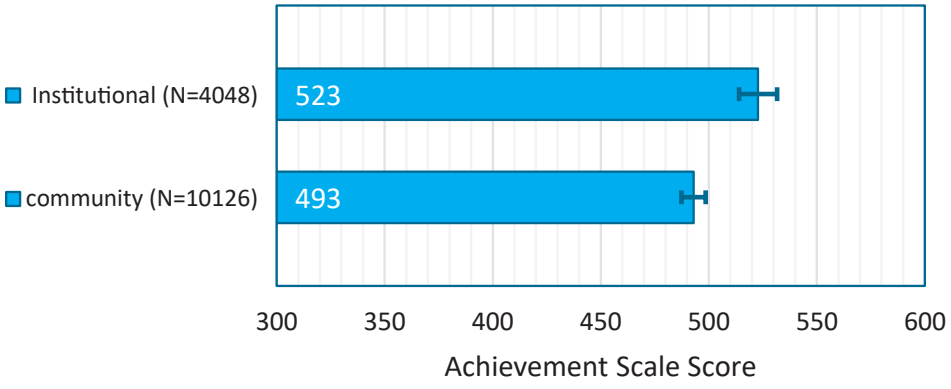


Figure 3.8 Mean score in Mathematics by types of schools

The mean score was found 493 in the case of community schools and 523 in institutional schools. The score of community schools had remained significantly lower than both the national mean and institutional schools whose achievement was distinctly above the national mean. With the difference of 30 scale score, the gap between the two types of schools was significantly different at $p<0.05$. The difference was alarming as the gap between community and institutional schools was very high (effect size, Cohen's $f = 0.26$).

3.10 Results by Various Influencing Factors

Educational researches inform that various contextual variables influence learning of students. In this research also many contextual variables were incorporated. The variables like student's personal variables, home related variables, socio-economic variables, attitude scales, school and teacher related variables were included and administered. The comparative analysis results are discussed below under respective headings.

a. Parents' Education and Students' Learning Achievement

Students were asked about their parents' education level, specifically whether they are *illiterate, just literate, Grade 10, Grade 12, Bachelor's, and Master's or above degree holders*. In this analysis, mother's educational level was related with achievement level as presented in figure 2.9.

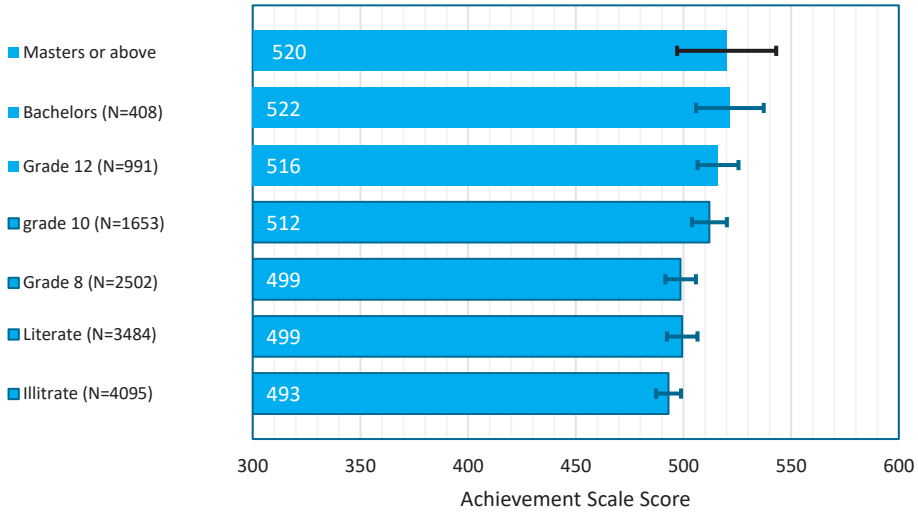


Figure 3.9 Mean score in Mathematics by level of mother's education

The dataset reveals that mother's education has positive effect on student learning achievement. Illitrate mother's children has the lowest achievement (493) whereas Bachelors degree mother's children have the highest achievement score (522). The difference in achievement associated with mother's education was significant at $p<0.05$. The difference was medium (effect size, Cohen's $f = 0.13$).

Similarly, father's education was also positively related with learning achievement. The educational level of father corresponding with their children's learning achievement is presented in figure 2.10.

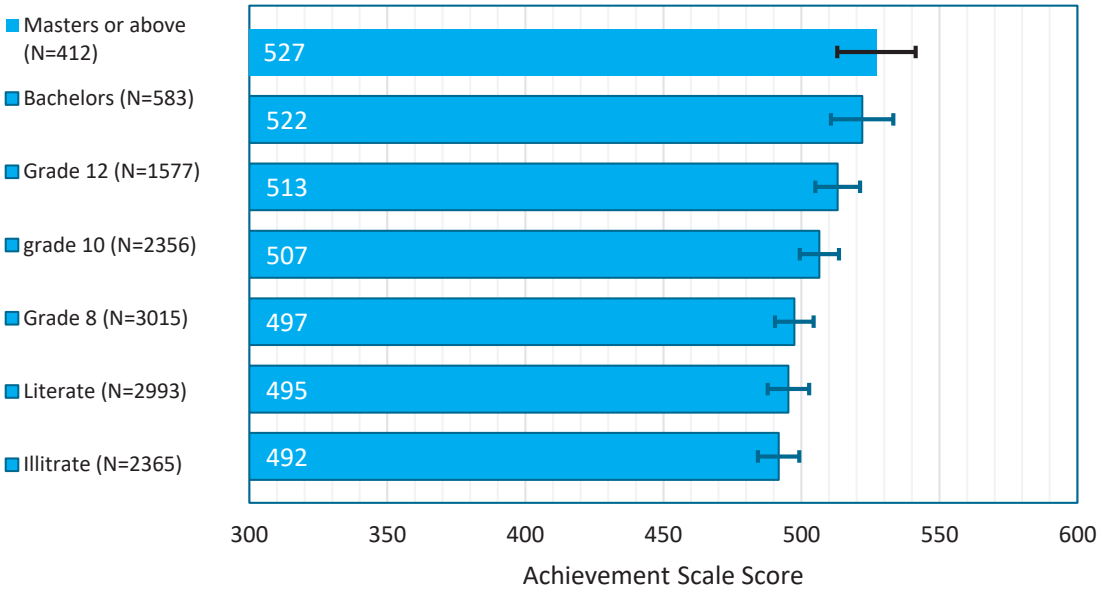


Figure 3.10 Mean score in Mathematics by level of father's education

The above figure reveals that father's education had positive effect on student's learning achievement. Illiterate father's children had lowest achievement (492) whereas Master's degree holder parents' children had the highest score (527). The difference in achievement based on different levels of father's education was significant at $p<0.05$ and the difference was moderate (effect size, Cohen's $f = 0.19$).

b. Students' Learning Achievement by Age Group

As a background variable, students were asked to report their age. The reported ages were grouped into six age groups— 9 or below, 10, 11, 12, 13 and 14 or above. The majority of students (4545) were of 12 years old. The smallest group of students was of the age of 10 and 9 or below. The influence of age on learning achievement can

be seen in the data presented in Figure 3.11.

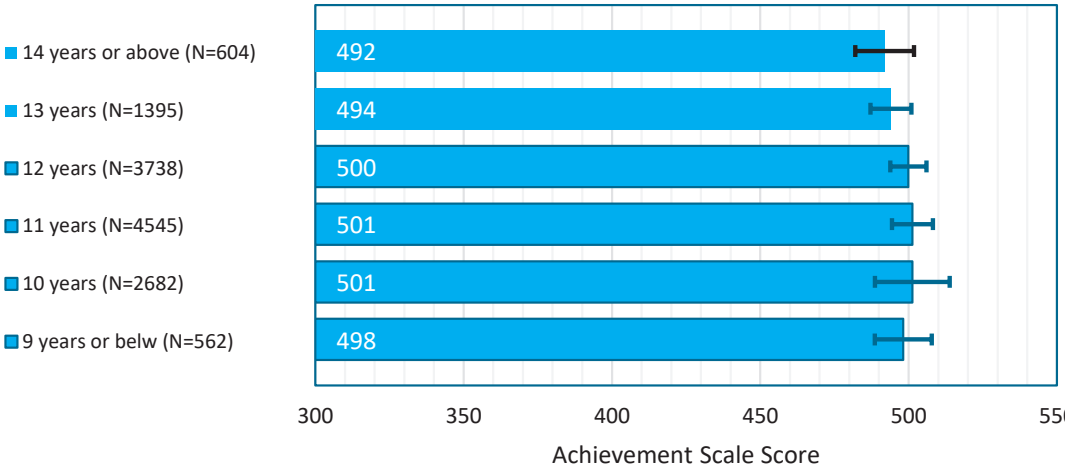


Figure 3.11 Mean score in Mathematics by age group

The dataset informs that students with appropriate age group performed better than the under or above age students. The highest level of achievement was found at the age of 10 and 11 years (501 score) which is higher than the national average (500 score). The lowest achievement was found at 14 years or above age (492 score). The difference in achievement between the highest and lowest groups was significant at 95% confidence level. *Such result was repeatedly found in other NASA results (NASA 2013, NASA 2015, NASA 2017) as well. This result reveals that providing opportunity to study in right class at right age is an appropriate strategy to maximize the learning.*

c. Results by Parents' Occupation

Socio economic status of a family is connected with parents' occupation. To analyse the impact of parents' occupation on students' learning achievement, the students were asked to report their parents' occupation on multiple options (agriculture and household work, household work only, work in others' houses, labour, foreign country employment, teaching, business, government job, and other jobs). The achievement scores are analysed by considering the students' parents' (mother's and father's) employment separately.

The impact of mother's occupation was noticed on students' learning achievement as presented in Figure 3.12.

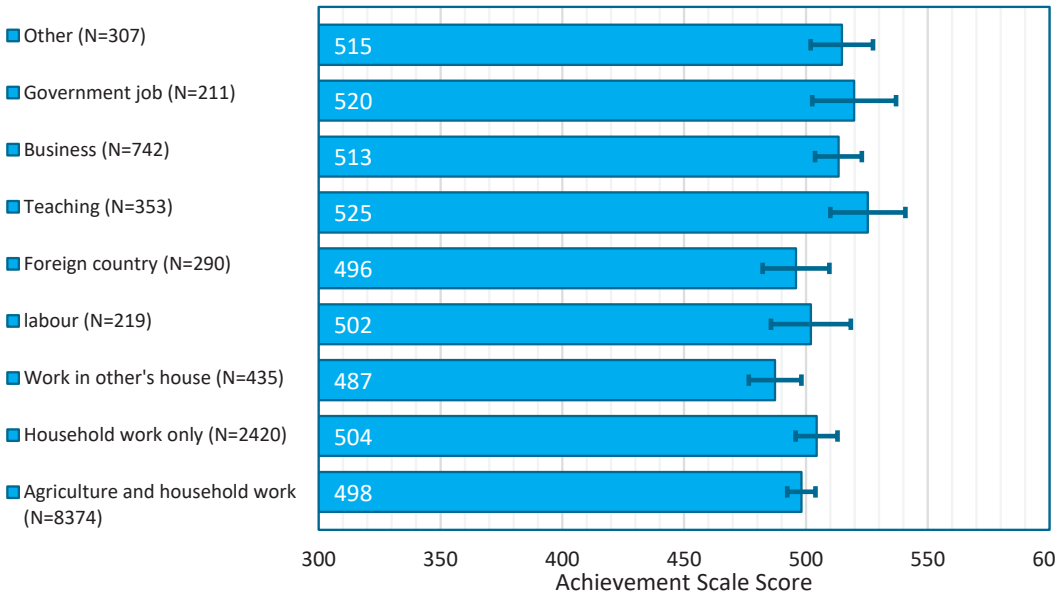


Figure 3.12 Relationship of learning achievement with mother's occupation

The student's mother with occupation of teaching, government job and business have achievement score 525, 520 and 513 respectively. Conversely, students whose mothers work in other's home have achieved 487 and whose mother work in foreign country have achieved 496. There was a relationship of mother's occupation in student's learning achievement.

Since mother's occupation also indicates the level of income of the family, this comparison reveals that regular income yielding professions of parents have better effect on student learning. Overall, a remarkable difference between the lowest and highest scoring variables (with 38 point, which is of considerable performance) is noticed and this difference is statistically significant at 95% confidence level.

d. Father's Occupation

Like mother's occupation, the impact of father's occupation on learning achievement was compared. The figure 3.13 presents the comparative results of parents' occupation student's achievement.

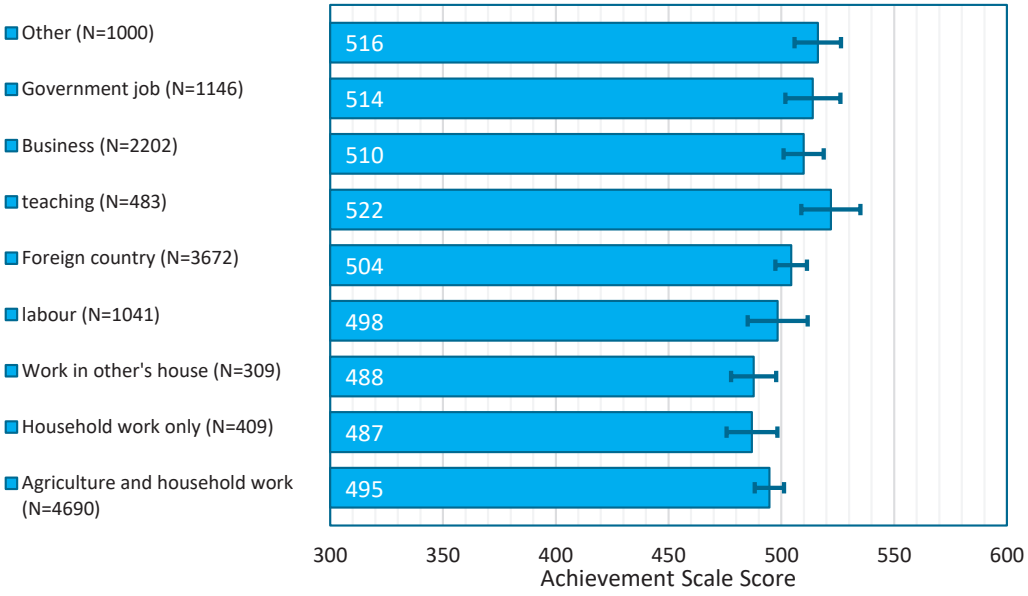


Figure 3.13 Mean score in Mathematics by father's occupation

The students' learning achievement in Mathematics was found lower than the national mean (500) in the case of their fathers' occupation being agriculture plus household work (495), only the household work (487), work in others' houses (488), and labour (498). The students of fathers of the remaining categories of occupation have scored higher than the national mean. The highest mean score went to the students with their fathers in teaching occupation (522). Considering the mean scores from the highest point towards the bottom, the data depicted that the second highest score was 516 (in the case of fathers having the profession of "other" than the listed) followed by 510 (in the case of government job), and 504 (in the case of foreign employment).

A remarkable difference is noticed between the lowest and highest mean scores (44 points) which is very high. Similarly, the influence of father's occupation was found strong in grade 5 students' learning achievement in Mathematics. The mean score of the students with father's occupation other than agriculture, household work and working in other's house was statistically significant at 95% confidence level than the other occupations.

e. Relationship of After- school Activities with Achievement

Students were asked how they spent time at home after the school activities. Various five activities were included in the questionnaire, namely, involvement in TV, internet and computer; play and talk with friends; home chores; homework and study;

work for wages; read and study other books. The time intervals were given as: *a. I don't give time, b. less than one hour, c. upto 3 hours d. more than 3 hours*. Based on the students' response categories in x-axis and the group achievement in y-axis, their engagement is plotted.

Figure 3.14 shows the relationship of time spent in after school activities with their relationship with the learning achievement.

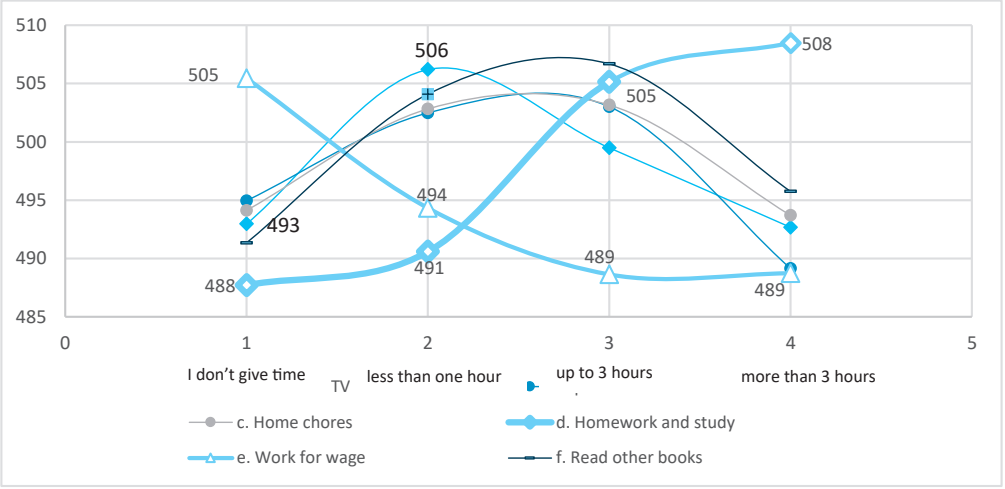


Figure 3.14 Achievement by after school activities

The above figure presents the fact that students who spent up to 3 or less hours in any activities do not have much variation in learning achievement. The students who spent more time in wage making work have a decreasing achievement with the time. For example, those who did not spend time in wage making scored 505 while those who spent 3 or more hours in wage making scored 489.

To improve the learning, the data suggest that spending up to 2 hours in the activities such as playing, watching TV, using computer, internet, chatting is beneficial. But the students should focus on study and homework without being involved in other activities after two hours..

f. Results by the After-school Support in Study

Students require after-school support for increasing their learning. Based on this assumption, the students were asked about the person who supports them the most in the after-school activities. The percentage of students' response on those who provide them support is presented in figure 3.7.

Table 3.7 Percentage of support received from various sources

Who supports in study at home?	% of students		
	Community	Institutional	
1 father	25.4	19.6	*
2 mother	11.4	17.1	
3 brother/sister	50.6	42.7	
4 tuition	6.4	15.4	*
5 friend	2.9	1.9	
6 any other	1.2	1.6	
7 none	2.0	1.8	

Table 3.7 indicates that the students from both types of schools: community and institutional get some kind of support in their the study after school. Major difference seems to be in the extra tuition received as 15% of the students from institutional schools get extra tuition whereas only 6% students from the community schools receive such extra tuition as a support to their study. Extra tuition is one of the influencing factors in increasing learning achievement. The influence of after-school support in the learning achievement is presented in figure 3.15.

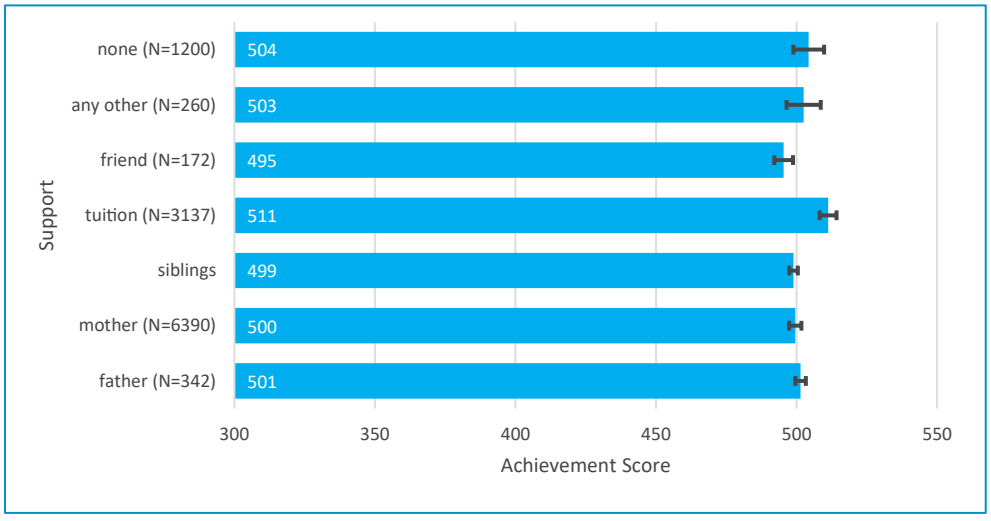


Figure 3.15 Achievement by after school support

Students who received extra tuition support achieved the highest score (511) and the achievement score of the students who received support from mother, father, any other person was 511, 501, and 503 respectively. Students who did not take any support also achieved above the national average score. However, the lower achievement was

related with support taken from the friends (495). The effect of support received from siblings was also quite close to national mean (499).

Most students who took extra tutorial support achieved the highest, and mostly such students were from institutional schools.

g. Students' Experience with Bullying at School and Learning Achievement

Through the background questionnaire, the students in this assessment were asked to respond whether they had bullying experience during one month's period (in the previous month from the time of assessment). Seven types of bullying experiences were recorded from students' response. The percentage of students who experienced bullying in seven categories is presented in figure 3.16 below:

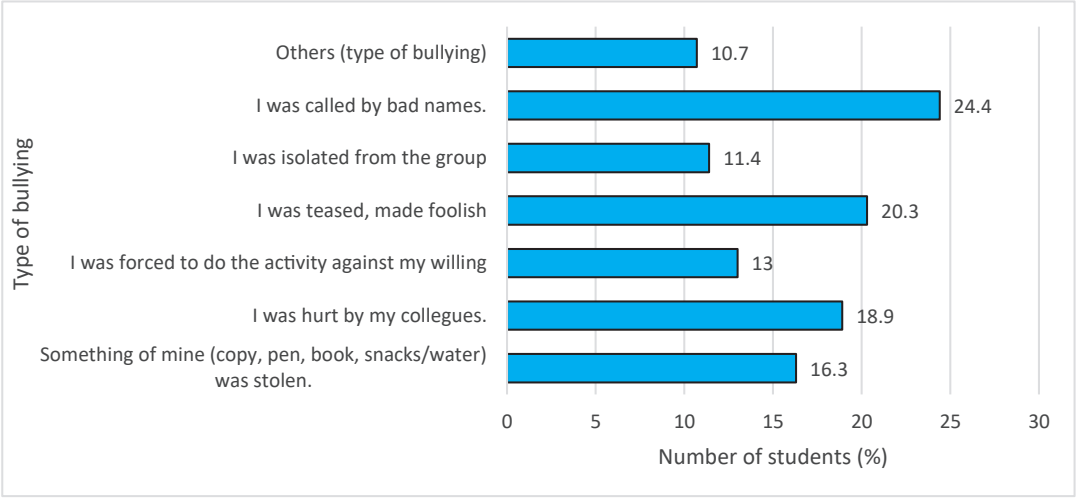


Figure 3.16 The percentage of students who experienced different types of bullying

From the above figure, it is clear that most of the common types of bullying experienced by the students in their school are: being called with bad names, forced to do unwanted activities, teased, hurt, and finding their belongings stolen.

Although 47% students did not experience any type of bullying; almost half of the students experienced some type of bullying and a few students experienced all of the above listed bullying activities. This total percentage of experience with bullying with corresponding achievement is plotted in the line graph given in figure 3.17 below:

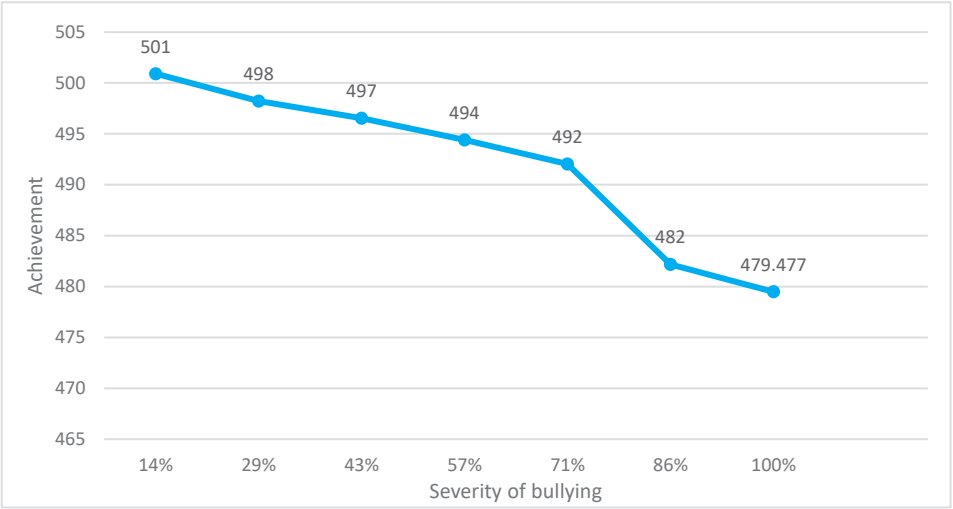


Figure 3.17 Influence of the experience with bullying on learning achievement

The above figure 3.1.7g presents that as bullying increases, achievement decreases. The frequency of experience with bullying and the learning achievement has negative relationship. *There should be a good strategy to minimize bullying in schools to increase the learning achievement of the students.*

h. Results by the Availability of Textbook

In this study, students were asked whether they had Mathematics textbook. Their response showed that the majority of students in this study had textbooks while some of them (652) reported that they did not have Mathematics textbook. The response of students and their corresponding achievement score is presented in figure 3.18 below:

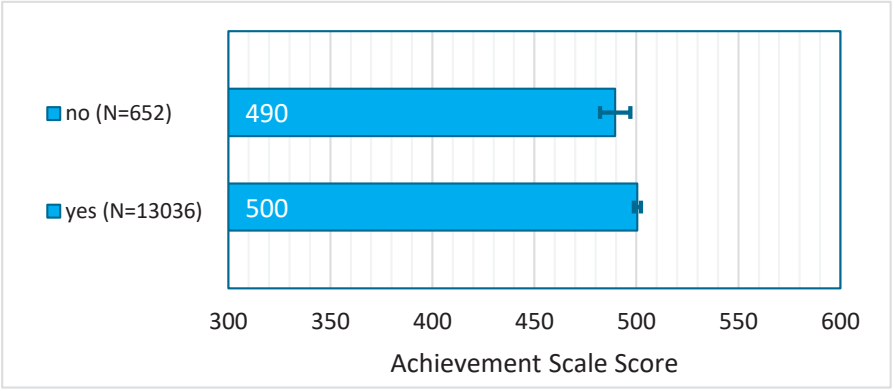


Figure 3.18 Mean score in Mathematics by the availability of textbook

The results, as presented in Figure 3.1.7h, show that the students who did not have textbooks achieved 490 mean score which looks remarkably lower than the national mean score. But the students having access to the textbook were able to achieve 500, which is equal to the national mean score. The difference between the mean scores of these two groups is found statistically significant. Out of 652 students who did not have text book, 441 students were from community schools. Although, only few students did not have textbook, it should be ensured that all students have a textbook with them.

i. Results by Feedback Provided on Students' Homework

The students were asked to mention how often their teachers provide them homework and how often the teachers provide them feedback on their homework. 51.2% students reported that their teachers provide homework occasionally and 42.7% reported that teacher provide homework regularly. Some students (6%) did not respond on this question. The students were asked how often their teachers provide them feedback on their homework. In the response, 66% reported "occasionally" and 34% reported "regularly". Figure 3.18 presents the comparison of achievement score of the students based on the feedback provided by their teachers on their homework.

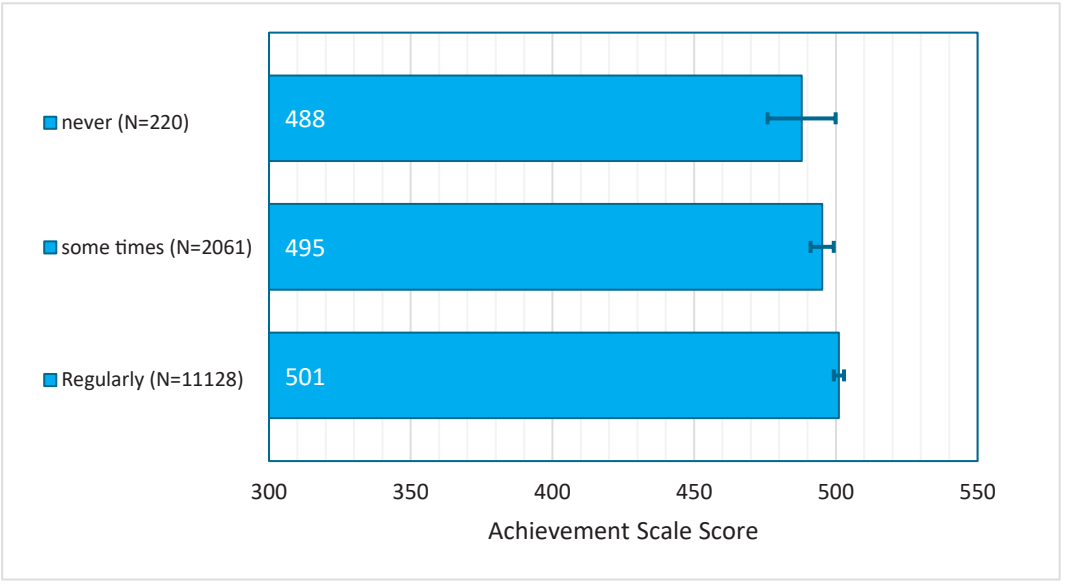


Figure 3.19 Teacher's feedback on students' homework and achievement

The data presented in Figure 3.18 shows that the mean score of the students who received feedback regularly on their homework was higher (501) than of those who

did not receive regular feedback (488), with the difference of mean score of 13 points. The difference is significant ($p<0.05$).

3.11 Relation Between Socio Economic Status and Schools' Performance

In this analysis, school's mean score was calculated by aggregating the scores that the students scored. In the same way, social economic status (SES) of students was also estimated by aggregating seven variables: mother's education, father's education, mother's occupation, father's occupation, home possessions (8 items: study table, separate room for study, peace study place, computer for study, story books, picture books, internet facility), home accessories (TV, computer, motorcycle, car, permanent house) and students attending private schools as given below:

Table 3.8 Dummy variables prepared from summed background variables

Variables	Effective variable
Mother's education dummy:	Grade 10 or above = 1, else 0
Father's education dummy:	Grade 10 or above = 1, else 0
Mother's profession dummy:	Teaching, business, government job, other = 1, else 0
Father's profession dummy:	Teaching, business, government job, other = 1, else 0
Home facilities for study:	Home possession 5 or higher = 1, else 0
Home accessories:	Home accessories 4-9 = 1, else 0
Institutional schools:	Institutional schools = 1, else 0

A regression model was developed in order to identify relationship between school mean of SES and mean achievement score in Mathematics based on the first plausible value (PV1) which is presented in figure 3.20 below:

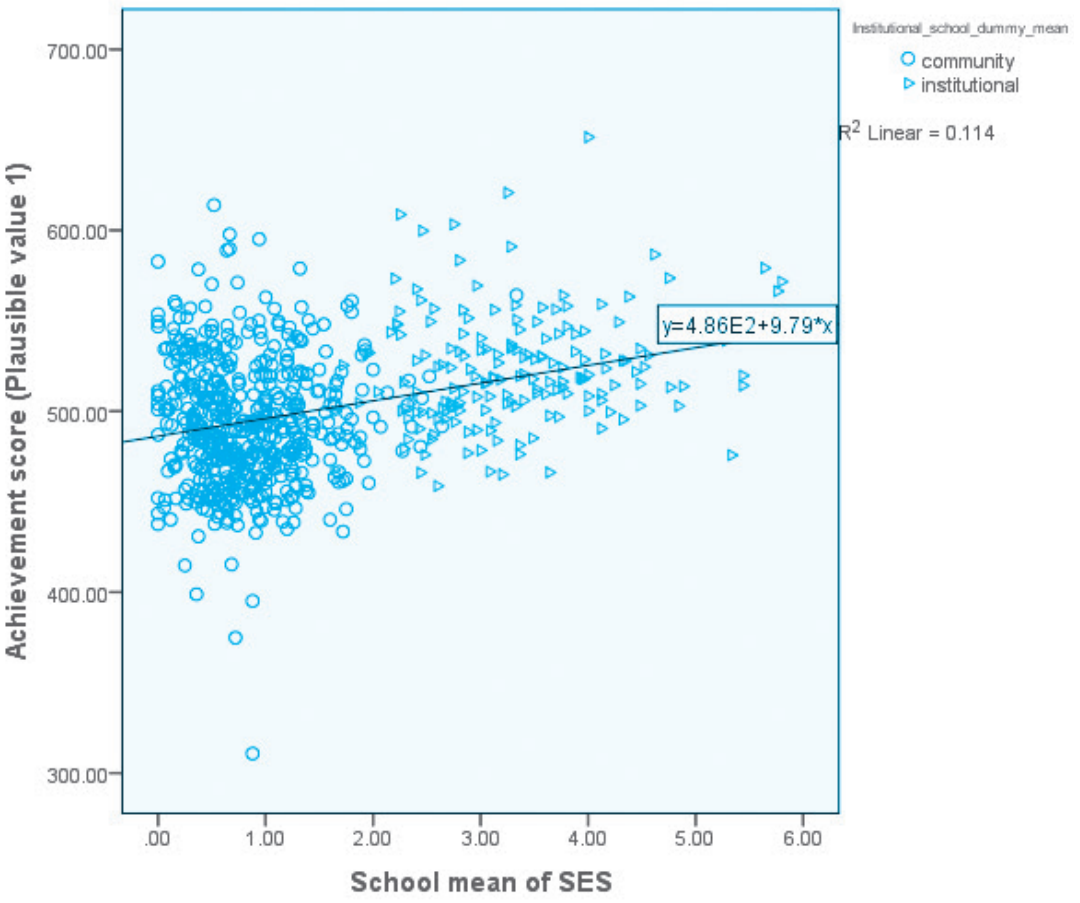


Figure 3.20 Relation between SES and Schools' Mean Score in Mathematics

The scattered plot in the above Figure 3.1.8 shows that the schools with high SES (average SES value of students) concentrated more on relatively high mean score and the schools with low SES concentrated more on relatively low mean score. However, there are some cases of high SES schools having relatively low mean scores. Similarly, there are some cases of low SES schools with relatively a high mean score.

This plot further indicates that most of the institutional schools have high SES students compared to community schools. It is however to be noted that SES explains only 11% of the variation in student achievement in the schools.

Chapter 4.

RESULTS IN NEPALI AT NATIONAL LEVEL

4.1 Introduction

The actual number of students who participated in NASA 2018 in Nepali was 14207. The collected data was analyzed by ACER ConQuest 4 software for IRT and other regular analysis was done by using SPSS and Excel tools. Results are presented in the form of proficiency levels, their descriptions and comparative form. Comparisons are made on the basis of groups formed from background information such as provinces, socioeconomic status, ethnicity, gender, home language, school type, identity with geography, school and classroom practices, and home environment.

The students' ability scores were estimated by IRT methods and these ability scores were transformed into a scale score with the national mean score of 500 and 50 standard deviation. Hence, the national mean score was fixed at 500 point for the purpose of analysing the overall test scores.

The test scores were first drawn from the sample students and analysed considering the sample weight. Population parameters were then estimated by using replicate module to generalize in whole population of grade 5 students in the country at 95% confidence level. The standard errors and confidence intervals of mean scores were estimated during the analysis and the confidence intervals were estimated to identify the statistical significant of comparable means.

4.2 Wright-map of Student Ability and Item Difficulty in Nepali

The Wright-map is organized into two vertical histograms: the left side shows candidates and the right side shows the items. The left side of the map shows the distribution of the measured ability of the candidates from most able at the top to least able at the bottom. The items on the right side of the map are distributed from the most difficult at the top to the least difficult at the bottom. In the following figure, student ability (θ) in the left and NASA 2018 items to the right are plotted in the same scale. When a person and an item lie at the same level, probability of making that item by particular person is 50%. Figure 4.1 presents the NASA 2018 Nepali Wright-map.

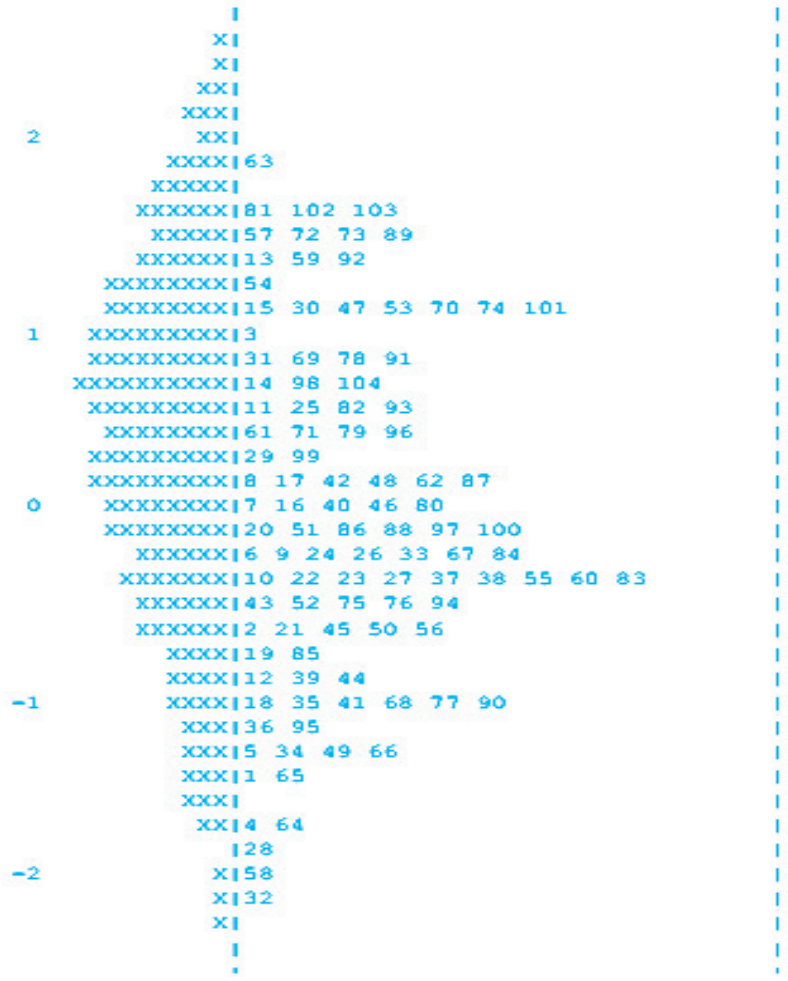


Figure 4.1 Wright map showing person and item in the same scale

To the left side, an 'X' represents 76.9 students and their latent ability is given in the logit scale ranging from -2 or less to +2 or more. The distribution of students against the items asked to students (item numbers are shown on the right side) reveals that most of the items were difficult for the students. Although the items were based on grade 5 curriculum, many students are lagging behind below the average latent ability '0'.

4.3 Plausible Values, their Mean and Standard Error

After estimating the student ability (θ) in the form of WLE, five plausible values (PV1 to PV5) were generated by conditioning the data with student background variables and school mean index. Those plausible values were transformed in to a

scale of mean 500 and standard deviation 50. Those values were weighted by student full weight using 350 replicates (just half number of numbers of schools taken in the sample for Mathematics). After this, all MSSPV1 to MSS PV5 were calculated to report the population estimates. The mean and standard error of five plausible values are presented in table 4.1.

Table 4.1 Standard Error of five plausible variables in Nepali

Plausible values	N (sample)	Mean	SE
MSSPV1	14207	499.9463	.42193
MSSPV2	14207	500.1478	.41977
MSSPV3	14207	499.7415	.41806
MSSPV4	14207	500.3016	.41872
MSSPV5	14207	499.8628	.41895

4.4 Defining Proficiency Levels for Nepali

The assessment framework for NASA 2018 recommends to set performance level into four levels. For this, three cut-points for proficiency levels were decided by dividing the range of 204 (maximum 599 – minimum 395) to divide student performance into 4 levels. Thus, four proficiency levels cut-points were 446, 497 and 548. The table 4.2 shows the range of proficiency levels scores.

Table 4.2 Proficiency levels and the score range in Nepali

Proficiency Level	Score
Level 3 (Advance)	548 and above
Level 2 (Proficient)	497 — 548
Level 1 (Basic Level)	446 — 497
below level 1 (Pre-basic)	Below 446

Based on the description of items that correspond to each of the above proficiency level in item person map in Nepali together with subject experts' judgment, the description of students' four levels of proficiency has been defined. These descriptions of four proficiency levels in Nepali for Grade 5 indicate what a student at particular competency level can do in Nepali. Table 3.2.1 specifies the competencies of students at various levels of proficiency in Nepali.

Basically, students who cross 67% of the achievement in a level are considered

as *Minimally Accepted Candidate*. Replicating the same concept in determining the minimum acceptance level of learning in those four-proficiency levels is possible. However, in this assessment, around 50% items were objective and almost equal weightage was given for the subjective items as well. So, in this analysis, 50% correct answer in the particular item was supposed as threshold of minimum accepted proficiency in particular item in particular proficiency levels. From this point of view, student response to every item was analyzed to find the response rate of those four levels of students. For this, at the first step, below level 1 (pre-basic) items were detected then, level 1, level 2 and level 3 items were detected respectively. In such a rigor, all the items were assigned to different four levels to draw proficiency descriptors. Table 4.3 specifies the minimum proficiency level of all four level students in a descriptive form.

Table 4.3 Description of six proficiency levels in Nepali for Grade 8

Level	Score range	Description
Below level 1: pre-basic	<446	छोटा अनुच्छेदमा भएको साधारण सूचना तथा तथ्य पहिचान गर्न, नक्सामा स्थान र दिशा पहिचान गर्न तथा परिचित विषयमा साधारण दुई तीन वाक्य लेख्न सक्छन्। जस्तैः बगैँचामा के को फूल फुलेको थियो। सरस्वती पूजामा के गरिन्छ ?
Level 1	446 - 497	छोटा अनुच्छेदको साधारण आसय पहिचान गर्न, अनुच्छेदमा प्रयुक्त परिचित शब्दको अर्थ पहिचान, आफूले गरेका काम एक वाक्यमा लेख्न, परिचित वाक्यमा उपयुक्त क्रियापद छनौट गर्न सक्छन्। जस्तैः फूलले किन घमण्ड गर्न छोड्यो ? हिजो गरेको एउटा काम लेख।
Level 2	497 - 548	अनुच्छेदको मुख्य आशय वा सन्देश पहिचान गर्न, अनुच्छेदबाट साधारण कारण पहिचान गर्न, नक्साबाट जानकारी लिन, साधारण निवेदन लेख्न र अनुच्छेदमा प्रयुक्त शब्दको अर्थ पहिचान गर्न सक्छन्। जस्तैः सूचना के विषयसँग सम्बन्धित छ ? कक्षा शिक्षकलाई बिदाको निवेदन, निःशुल्क शब्दको अर्थ ?
Level 3	548 >	अनुच्छेदको सूचनाका आधारमा तर्क गर्न, अनुमान गर्न, केही लुप्त सूचना बोध गर्न, नाम, सर्वनाम पहिचान गर्न र अनुलेख गर्न सक्छन्। जस्तैः यस्तो घटना घट्न नदिन के गर्नुपर्छ ? चिठी किन लेखिएको हो ? मुख्य विषय के हो ? तलको अनुच्छेद जस्ताको तस्तै सार्नुहोस्। तर, यस तहका विद्यार्थीले पनि अनुच्छेदको उच्चतम बोध र कारण पहिचान, स्वतन्त्र अनुच्छेद लेखन, चिह्नको उपयुक्त प्रयोग तथा साधारण वाक्यमा शुद्धाशुद्धि स्पष्टसँग मिलाउन भने सक्दैनन्।

4.5 Distribution of students by Proficiency Levels

The student achievement scores based on 5 plausible values (PV1 to PV5) were analysed in terms of four proficiency levels of students' achievement. Level wise descriptors are presented in the above section. This sub-section presents the number of students falling in those four levels from population estimate. The standard error of the percentage of students is also presented in the table 4.4

Table 4.4 Distribution of the students in various proficiency level and its Standard Error

Proficiency level	% of students	SE	N_cases	NU_cases	NU_psu
Below level 1 (pre-basic)	20	0.7043	132684	2710	625
Level 1: Basic	35	1.0098	236162	4808	686
Level 2: Proficient	30	1.1544	201517	4169	687
Level 3: Advance	15	0.6076	103989	2198	525

NB: SE = Standard Error, N cases = Number of cases/students in the population, NU cases = Number of cases/students in the sample, NU psu = Number of Primary Sample Units (schools).

The below level 1 indicates the lowest ability of students who are struggling hard in the classroom where as level 3 shows the highest level of proficiency even crossing the grade level. Figure 4.2 shows how students are distributed over those levels visually.

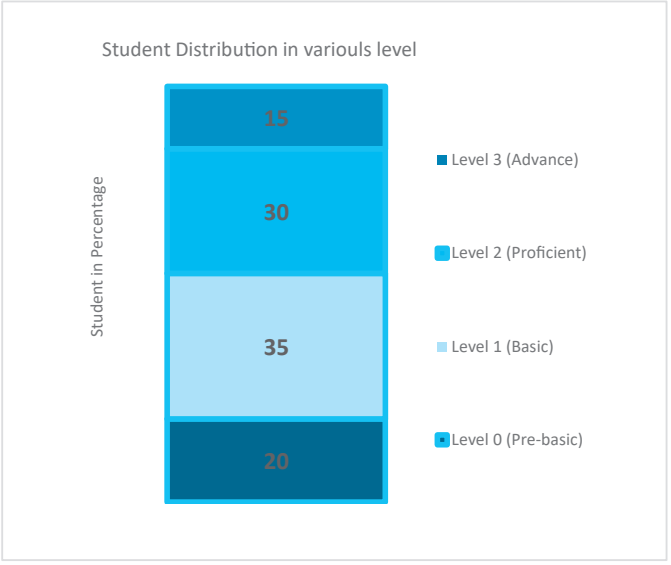


Figure 4.2 Student percentage by proficiency levels in Nepali

Figure 4.2 shows that 20% of the students are in *below level 1*, 35% in level 1-basic, 30% in proficiency level and 15% in advance level. Below level 1 (20%) students can pick some direct information from very short text but can not understand the meaning of the text well. There are 35%, 30% and 15% students at basic level, proficient level and advance level respectively. Basic level students have minimum level (below average) of understanding required to grasp the knowledge and skills taught in the classroom.

4.6 Minimum Level of Achieved Curriculum

The assessed curriculum is reflected by the assessment or evaluation. It can be formative or summative evaluation of the students. Assessed curriculum is a tested curriculum by school, national or international organization based on the written curriculum/intended curriculum. It is valuable because it enables the educational organizations and stakeholders to evaluate the impact of written and taught curriculum on students. It determines the level of the learned curriculum. Research (e.g. Berliner, 1984; Turner, 2003) indicates that the mismatch between Assessed and Taught Curricula has serious consequences (cited in MeshGuide). This section therefore presents the level of learning in the form of achieved curriculum in terms of percentage. Every test item is equivalent in the sense that each of them represents a learning objective mentioned in the written curriculum.

Practically, 67% correct responses can be considered as cut-score for being minimally proficient at any level. As in this assessment, around half number of items were objective type (MCQ) and half of them were subjective, 50% correct response are considered as the threshold of minimum level of accepted proficiency at any of the four levels. Hence, test items were organized in terms of at least 50% or more correctly answered in each level students. Based on this norm, all the items were re-allocated into four level. From this rigorous analysis, performance descriptors were developed. In every level of the proficiency, there are ranges of students from being very weak performers to highest performers. Thus considering 50% correct answer as a threshold of minimum proficiency of any four levels, percentage of learning was Mathematically calculated based on the number of items answered correctly. Mathematical value of achieved curriculum is thus given in the table 4.5.

Table 4.5 Mathematical presentation of the achieved curriculum in Nepali.

Performance level	Achieved curriculum (%)
Below Basic level (20% students)	18 % of the curriculum
Basic level (35% students)	38 % of the curriculum
Proficient level (30% students)	60 % of the curriculum
Advance level (15% students)	88 % of the curriculum

In Nepali, assessment framework covers reading and writing which can be measured from the paper-pencil tests. This test does not cover listening and speaking skills at all. Grammar and vocabulary are embodied in reading and writing test itself. Based on the items asked, in Nepali, 20 out of 100 students achieved only 18% of the tested curriculum and 38 out of 100 students achieved only 38% of the tested curriculum. 45 out of 100 students have adequate knowledge and skills of tested curriculum as 30% students fall in proficient level and 15% students fall in advance level. Altogether 55%, a big mass of the students represent underperforming group. The proficient level students achieved 60% of the tested curriculum and advance level students achieved 88% of the tested curriculum. *The gap in achieved curriculum between below basic level (18%) and advance level students (88%) is 70%. This indicates that inequality in the classroom is very high. Overall, 61% students are below 40% of the learning achievement as per the tested curriculum. Hence, about 55% of the students are below the expected level of the competencies defined by the curriculum in reading and writing in Nepali.*

4.7 Overall Mean Score by Province

In the Federal context, Nepal is divided into seven provinces. The PPS sampling method was adopted to select the PSU and provinces were treated as the strata. The average scores reported in this section are the transformed/scale score at 500 national average. National mean is taken as a reference to compare the provincial mean. Those provinces whose average score is above the measure are recognized as better performing provinces whereas those below 500 are assumed to be low performing. In the figure 3.1.4, a vertical dotted line represents the national mean score of achievement and horizontal bars represent the provincial level achievement scores. Figure 4.3 presents the comparison of student achievement in mean scale score.

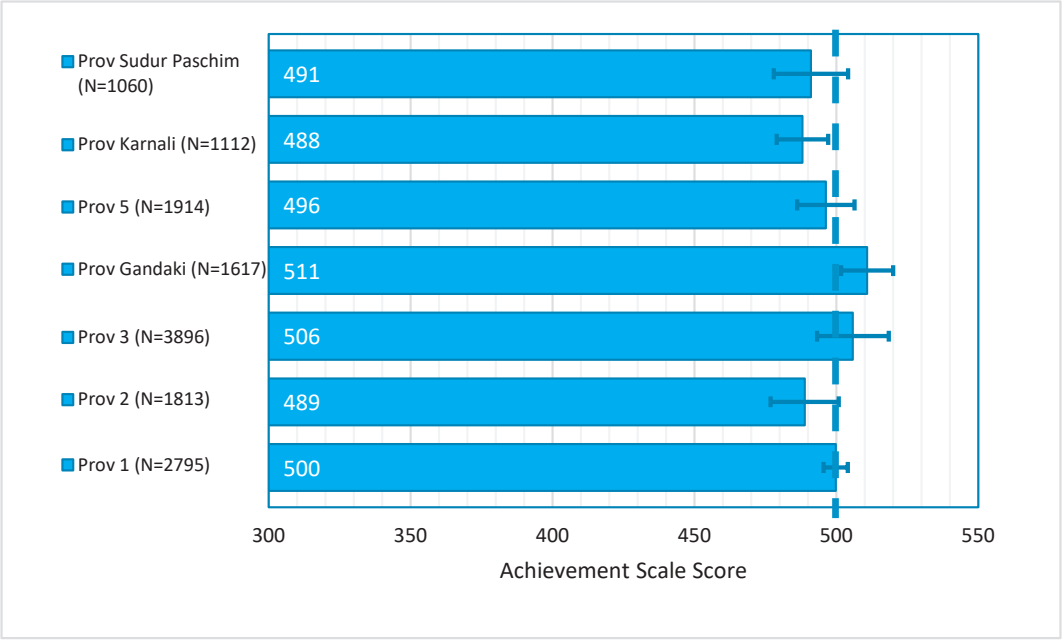


Figure 4.3 Provincial level mean achievement scores in Nepali

The data show that students' achievement in Nepali subject in province 4 is the highest among the seven provinces. The mean score of this province is 511, which is higher than the national mean score by 11 points. Students' achievement in Nepali in province 6 is the lowest with the mean score of 488 followed by province 2 with the score of 489.

Figure 4.3 also shows that the mean score of province 4 (511) is higher than the national mean score. Likewise, the mean score of province 3 (506) is also slightly higher than the national mean. Students' achievement in province 2, 5, 6 and 7 is below the national mean score (489, 496, 488 and 491 respectively). Similarly, the mean score of province 1 was equal to national level. However, out of the seven provinces the performance of four provinces (2, 5, 6 and 7) is below the national mean.

Overall, the scores show that only three provinces – 3, 4 and 5 – had the mean scores above the national mean. The mean score of other four provinces is lower than the national mean score. This indicates that much is to be done to raise the performance of the students in Nepali.

Further district-wise analysis is presented in the bar figure 4.4.

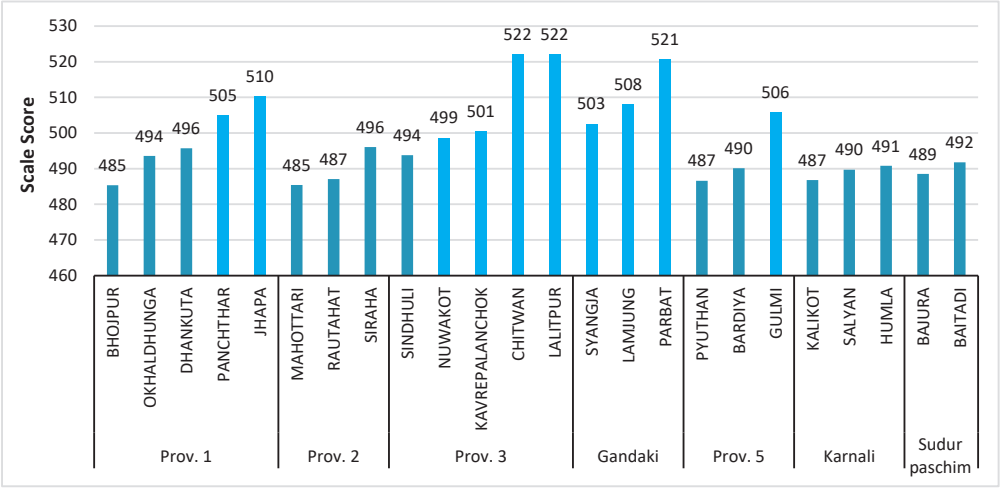


Figure 4.4 Mean achievement score by districts within the provinces

Figure 4.4 presents that all the districts of Gandaki province: Syangja (503), Lamjung (508) and Parbat (521) had the score above the national mean (500); hence this province stood in the highest position. However, all the districts of Karnali Province and Sudur-paschim province had lower score than the national mean score.

The above result shows the wide differences among the districts and provinces in learning achievement. Moreover, differences by gender was also seen in the provinces. Table 4.6 presents the comparation of average achievement score (mean) by gender within the provinces.

Table 4.6 Comparison of achievement within the provinces

Provinces	Gender	N	Mean	SE	Sig	Eta sq	effect size
Prov. 1	Boys	1308	499	1.327	0.052*	0.001	0.0316
	Girls	1451	503	1.321			
Prov. 2	Boys	806	490	1.777	0.696	0.000	0.0000
	Girls	953	489	1.544			
Prov. 3	Boys	1798	506	1.195	0.009*	0.002	0.0448
	Girls	2061	510	1.153			
Gandaki	Boys	801	506	1.801	0.03*	0.003	0.0549
	Girls	781	512	1.859			
Prov. 5	Boys	915	493	1.627	0.126	0.001	0.0316
	Girls	971	497	1.610			

Provinces	Gender	N	Mean	SE	Sig	Eta sq	effect size
Karnali	Boys	500	490	2.072	0.384	0.001	0.0316
	Girls	600	488	1.884			
Sudur Paschim	Boys	507	492	2.045	0.235	0.001	0.0316
	Girls	546	489	1.950			

* *indicates difference is significant.*

From the figure 4.6 reveals that provinces 1, 3 and Gandaki have significant difference in mean ($p<0.05$) between the achievement of boys and girls. In those provinces, girls outperformed boys in achievement. However, difference in mean is very narrow as shown by the effect size (effect size varies 0.03 to 0.05). Such narrow difference shows that boys and girls had almost equal performance.

4.8 Results by Gender, Ethnicity and Home Language

From the equity concern of learning; gender, ethnicity and home languages are considered to be the important variables. In this section, achievement scores of various groups have been compared.

a. Achievement by Gender

For the equal level of learning to take place, girls and boys should have equal opportunity and support in their study. In the background questionnaire, students had reported their gender. In this data, there were 6635 (46.7%) boys, 7363 (51.8%) girls and 209 (1.5%) students did not report their gender. Based on this sample, the weighted mean score is presented in figure 4.5.

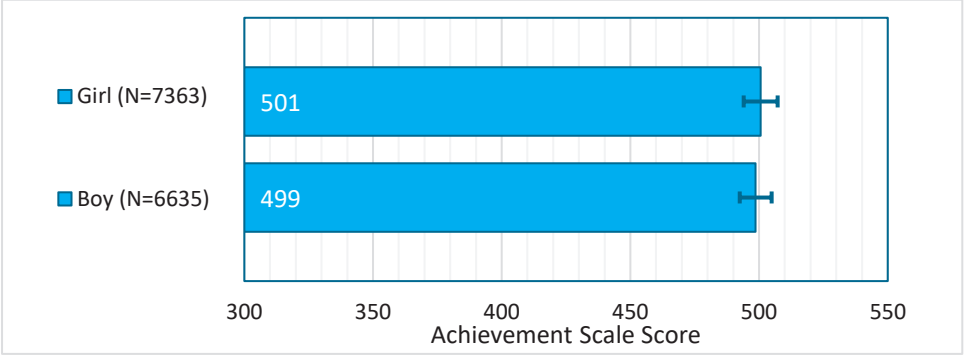


Figure 4.5 Mean score in Nepali by gender

The analysis of data in terms of gender shows that girl students' mean achievement

score was higher than the mean score of boy students in Nepali. The girl students have achieved (501), as presented in Figure 4.5, which was slightly higher than the national mean score, but there was a significant difference in the mean scores between boys and girls. The achievement of boy students was 499, which was also slightly less than the national mean.

b. Achievement by Caste/Ethnicity

From equity concern, ethnicity is also an important variable. Various cultural qualities are embedded in ethnicity as it influences diversity, social integration, time spent in the study. All these factors have their effect on learning (Virginia, 2017). The caste/ethnicity wise comparative result is presented in figure 4.6.

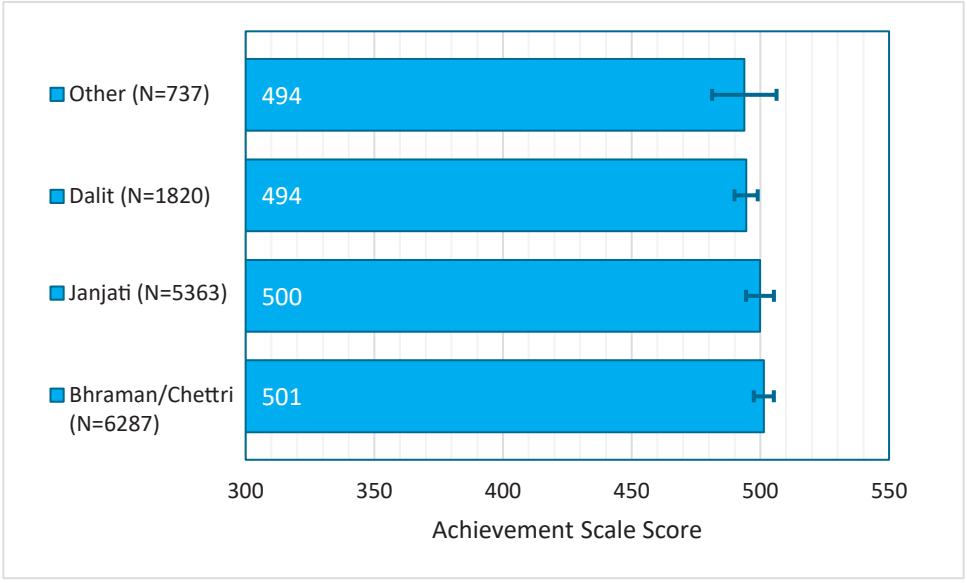


Figure 4.6 Mean score in Nepali by ethnicity

When ethnicity was considered for comparison among the students, the mean score achieved by Brahmin/Chhetri students was found to be slightly higher than the score achieved by the students belonging to other ethnic groups. As presented in Figure 4.6, there was a slight difference in mean score of Brahmin/Chhetri students (501) and Janajati students (500). The mean achievement score of Dalit students was 494 only. Interestingly, the students from Janjati performed as equal to national mean and it was just 1 score below the Brahman/Chettri students. There is no significant difference between the achievement of Bhraman/Chettri and Janjati students. However, difference is found significant at p value 0.05 for other groups.

c. Achievement by Home Language

Students were asked, "Which language do you speak most in your home?" Their response reveals that of the total, 64.3% of the students spoke Nepali language in their home whereas 29.4% students reported "other" languages. Few (6.3%) students did not report anything. Based on the student response, achievement score is presented in the figure 4.7.

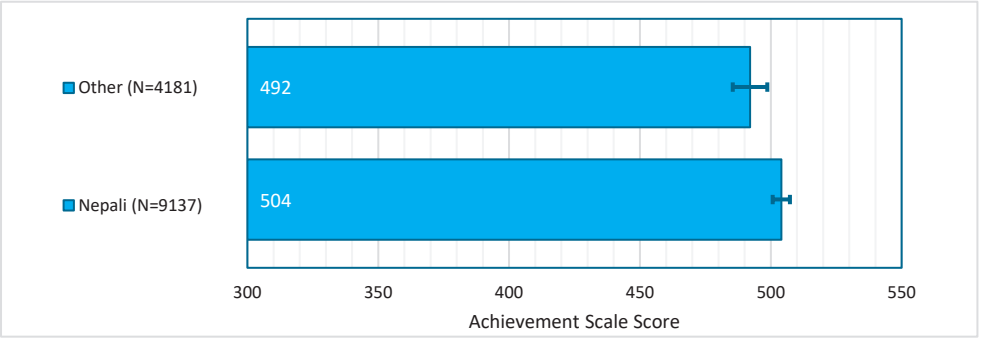


Figure 4.7 Mean score in Nepali by home language

The data show that there is a variation in the mean score in terms of students' linguistic background as well. The students with Nepali as their home language achieved higher than the students whose home language was other than Nepali. As seen in Figure 4.7, the mean score of the students who spoke Nepali as their home language is 504, which was higher than the national mean score. But the mean score of the students speaking other languages at home was 492 only, which is low compared to the national mean score. Significance test for difference of mean shows that the difference in the achievement of the Nepali speaking students was remarkably higher than that of other students.

4.9 Results by Types of Schools

There were 10271 sample students who participated in the testing program from community schools and 3936 students from institutional schools. The achievement score of types by schools is presented in figure 4.8.

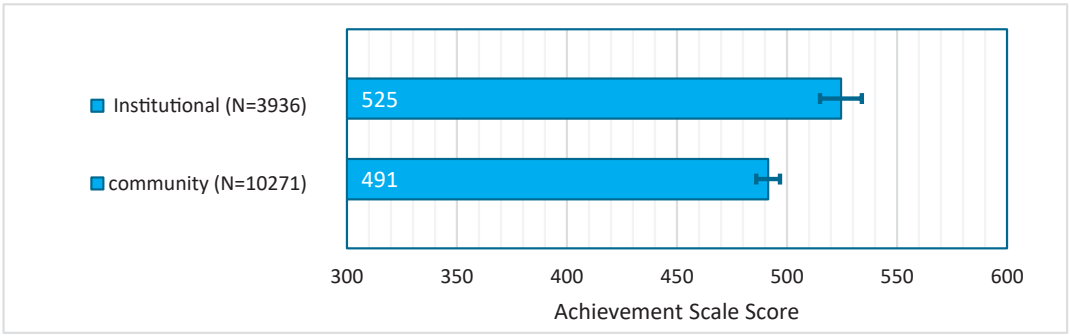


Figure 4.8 Mean score in Nepali by type of schools

The mean scores of the students from these two types of schools were found to be 491 and 525, respectively. The score of community schools remained lower than the national mean, while that of institutional schools was distinctly above the national mean. As the difference between the two means was 34 score, it was statistically significant at $p<0.05$.

4.10 Results by Various Influencing Factors

In order to understand the influence of various factors on students' achievement, the test scores were analysed in terms of parents' education and occupation, students' age, time spent in household chores, support they received in studies, effect of bullying, and the use of free time.

a. Parents' Education and Students' Learning Achievement

In order to identify the impact of parents' education on students' learning achievement, they were asked (in the background questionnaire) to report their parents' education by choosing a response from multiple options (illiterate, just literate, Grade 8, Grade 10, Grade 12, Bachelor's, and Master's degree holders or above).

Out of 14207 students, 3278 reported that their mothers were 'just literate' while 4097 of them reported their mothers to be 'illiterate'. Similarly, 1601 of them mentioned that their mothers had a qualification of Grade 10 while 1100 of them reported their mother's qualification being Grade 12. Likewise, the mothers of 437 students had Bachelor's and the mothers of 180 of them had Master's level educational qualifications.

As presented in Figure 3.8, mother's education had an impact on students'

learning achievement. The mean score of the students whose mothers were 'illiterate' was 491, which was lower than the national mean score. The students' achievements whose mothers were 'just literate' and Grade 10 qualification, had achieved 494 and 514 score respectively. Likewise, the mean scores of the students whose mothers have Grade 12 and Bachelor's qualifications were 516 and 525 respectively. The students whose mothers have the educational qualification of Master's degree and above have achieved 527, and interestingly this was not any better than the students of mothers with Bachelors degree. This implies that mothers having Bachelor's degree qualification have remarkably contributed to students' learning achievement in the case of Nepali subject in Grade 5.

The difference in mean achievement was statistically significant among the student groups based on the level of mother's education at 95% confidence level.

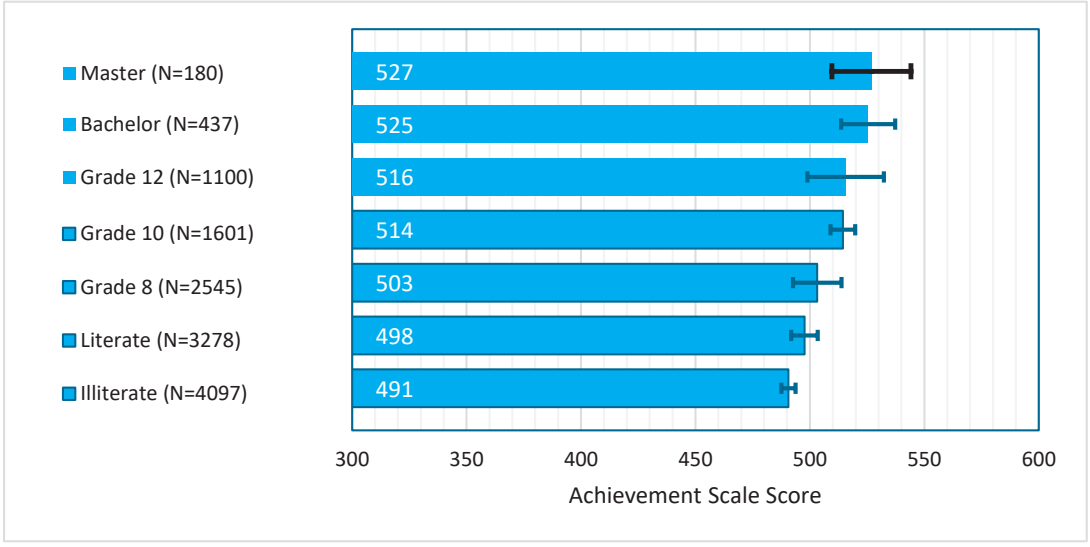


Figure 4.9 Mean score in Nepali by mother's education

The data further show that 2275 students' fathers were "illiterate", and 2942 students' fathers were "just literate". Likewise, 3081, 2248, and 1632 of the students' fathers had the qualification of Grade 8, Grade 10, and Grade 12, respectively. Altogether 612 and 428 students had fathers with Bachelor's and Master's degree qualifications, respectively.

As presented in Figure 4.10, fathers' education influenced the mean score of the students in some notable ways. As the data show, the students whose fathers were illiterate had achieved 487 mean score, which is lower than the national mean score.

The mean scores of the students whose fathers were 'just literate' and the students whose fathers have the qualification of Grade 8 scored below the national mean. The highest score was achieved by the students whose fathers had a Master's degree (527) which was higher than the mean score of the students whose fathers had Grade 12 (511) and Bachelor's degree qualifications (519).

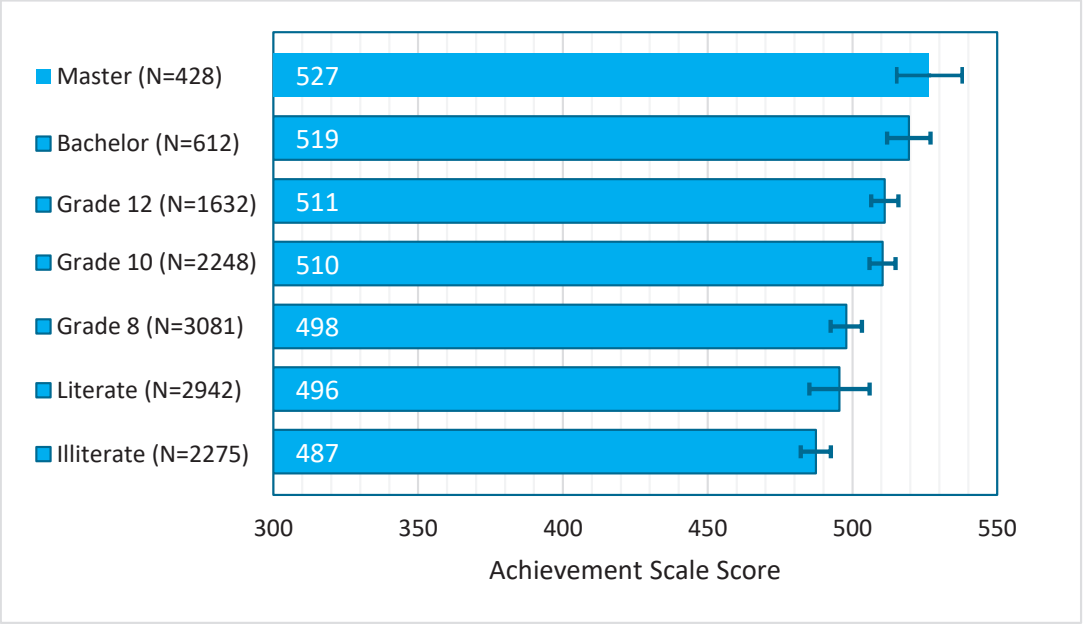


Figure 4.10 Mean score in Nepali by father's education

There are statistically significant differences between the mean scores of students from all of the groups at 95% confidence level. The comparison of the influence of parental education went more in favour of father's education than mother's education.

b. Students' Learning Achievement by Age Groups

The students were also asked to report their age for the analysis of the relationship between their age and their learning achievement. As the data show, the students were grouped into six age groups: 9 and below, 10, 11, 12, 13 and 14 and above. Majority of students (4656) were 11 years old while 4075 were 12 years old. Similarly, 2427 and 548 students were of 10 and 9 years and below respectively. Students' learning achievement by age groups is presented in figure 4.11 below:

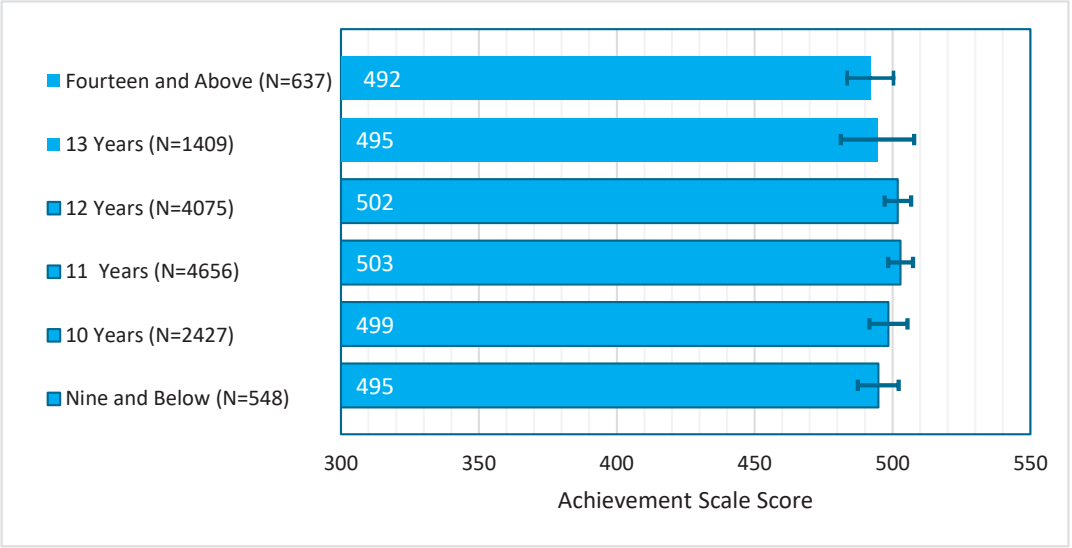


Figure 4.11 Mean score in Nepali by age groups

Regarding the influence of age in learning achievement, the data indicated that the students who are younger and older than the Grade level age (11-12 years) achieved lower mean score compared to the scores of students with a grade appropriate age. The students of fourteen and above age group achieved lower compared to all other age groups.

As Figure 4.11 indicates, the students who are 13 and 14 years and above achieved 495 and 492 mean scores respectively. Likewise, the students who are 9 years old and below achieved 495 mean score, which is lower than the national mean score. The data further show that the students with grade appropriate age achieved higher mean score. Students who are 11 and 12 years old achieved 503 and 502, respectively. Both the scores are higher than the national mean score. The statistical difference in mean score among the groups was significant at 95% confidence interval.

c. Results by Parents' Occupation

In order to assess students' learning achievement in terms of their parents' occupation, they were asked, in the background questionnaire, to report the occupations of their parents. They were given nine categories of occupations to choose from the list, which included: agriculture; household chore; work at others' home; wage labourer; business; foreign employment; government job; teaching; and others. The data show that majority of students' parents had agriculture (mothers = 8568 and fathers = 4767) as their occupation, 3218 students' fathers were in foreign employment; and 1299

students reported that their fathers owned a business. Likewise, 2353 and 699 students mentioned that their mothers had household responsibility and business as their occupation respectively.

As presented in Figure 4.12, the test scores of the students whose mothers were farmers and worked at others' homes were lower than the national mean score. The mean scores of the students in these categories were 497 and 485 respectively. The students whose mothers were teachers and government job holders had achieved higher mean score than the students of mothers in other categories. The scores for the students of these categories are 528 and 518 respectively.

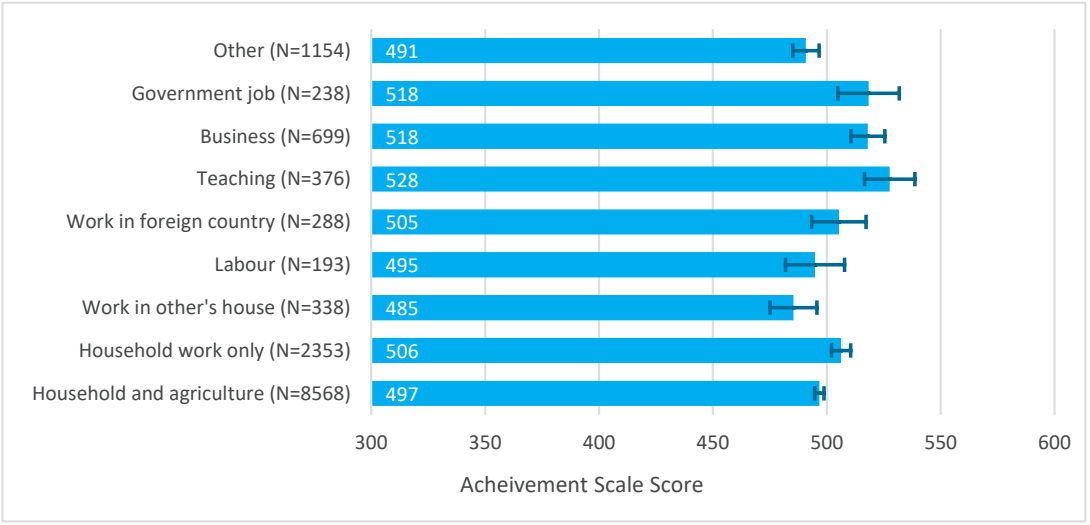


Figure 4.12 Mean score in Nepali by mother's occupation

Figure 4.12 shows that the students with their mothers as wage labourers achieved 495. Likewise, the students whose mothers were either in business or in government job achieved equal score of 518.

d. Father's Occupation

Like mother's occupation, the relation between the student's father's occupation and learning achievement was compared. Figure 4.13 presents the comparative results of the impact of father's occupation on student's achievement.

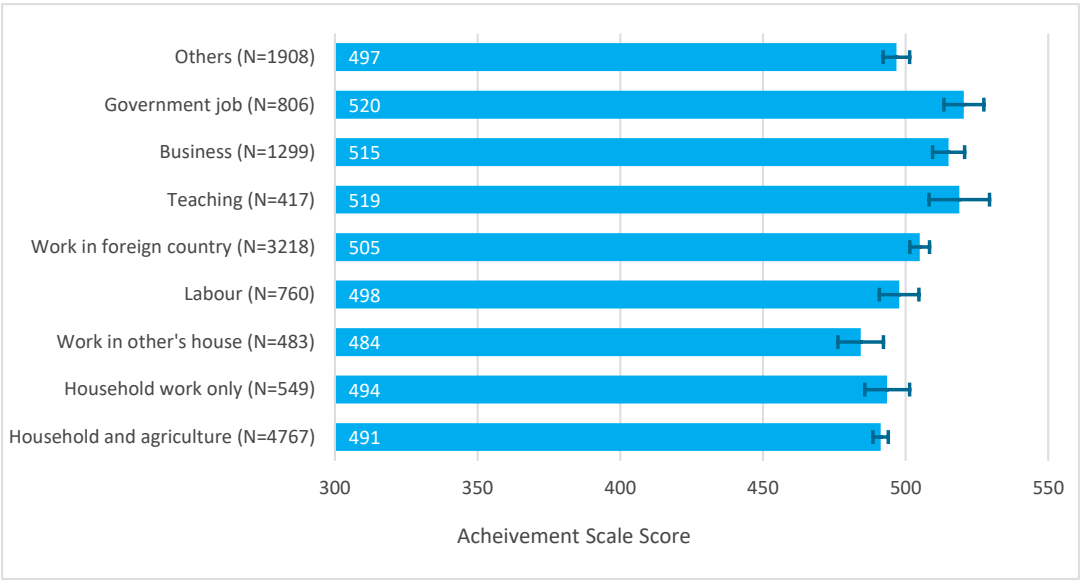


Figure 4.13 Mean score in Nepali by father's occupation

Figure 4.13 shows that the test scores of students whose fathers are farmers; work at others' home; and have household responsibilities, are lower than the national mean score. The mean scores of the students in these categories were found to be 491, 494, and 484 respectively. The students whose fathers are teachers and government job holders, achieved higher mean score (519 and 520) than the students whose fathers belong to other occupational categories (497).

Figure 4.13 further shows that the students whose fathers are in foreign employment achieved the mean score 505. Likewise, the students whose fathers are business persons achieved the mean score 515.

e. Relationship of After-school Activities with Achievement

The students were asked how they spent time at home. Various four activities were included in the questionnaire, namely, involvement in TV, internet and computer; play and talking with friends; home chores; homework and study. The time intervals were given as: a. *I don't give time*, *b. less than one hour, c. up to 3 hours. more than 3 hours.* Based on the students' response categories in x-axis and the group achievement in y-axis, the result is plotted in figure 4.14.

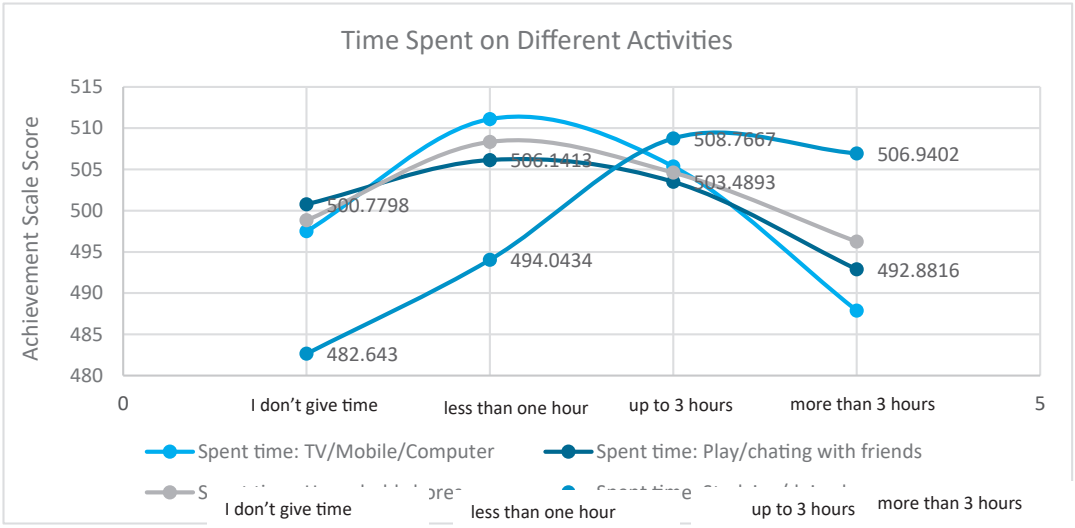


Figure 4.14 Relationship of time spent in various activities and their relationship with learning achievement.

The above figure presents that the students who spent up to 3 or less hours in any of the after-school activities did not have much variation in learning achievement. However, the students who spent more time (more than 3 hours) in the activities like playing and chatting with friends have lower achievement. Those who spent less after- school time in homework and study had the lowest achievement score of 483 whereas those who spent more after- school time in homework and study (more than 3 hours) had their achievement score of 507. It shows that the after-school time (more than 3 hours) spent in any activity other than study has a negative impact on learning achievement of the student.

f. Results by After- school Support to the Students

Students require after-school support for increasing their learning achievement. Based on this assumption, the students were asked about the person who supports them most in the after-school activities. The number of students (6529) receiving support from their siblings for their study was higher than the students who received support from other persons. A small number of students received support from extra tuition classes (985) and friends (497). Likewise, some students received support from their fathers (2954) and mothers (1877) as well (see Figure 4.15). The data presented in disaggregated form below shows the difference in the proportion of support received from different sources among the students.

Table 4.7 Percentage of support received from various sources

Support Received	Community Schools	Institutional Schools
Father	74%	26%
Mother	61.5%	38.5%
Brother/Sister	76%	24%
Tuition	52%	48%
Friend	84%	16%
Any other	65%	35%
None	75%	25%

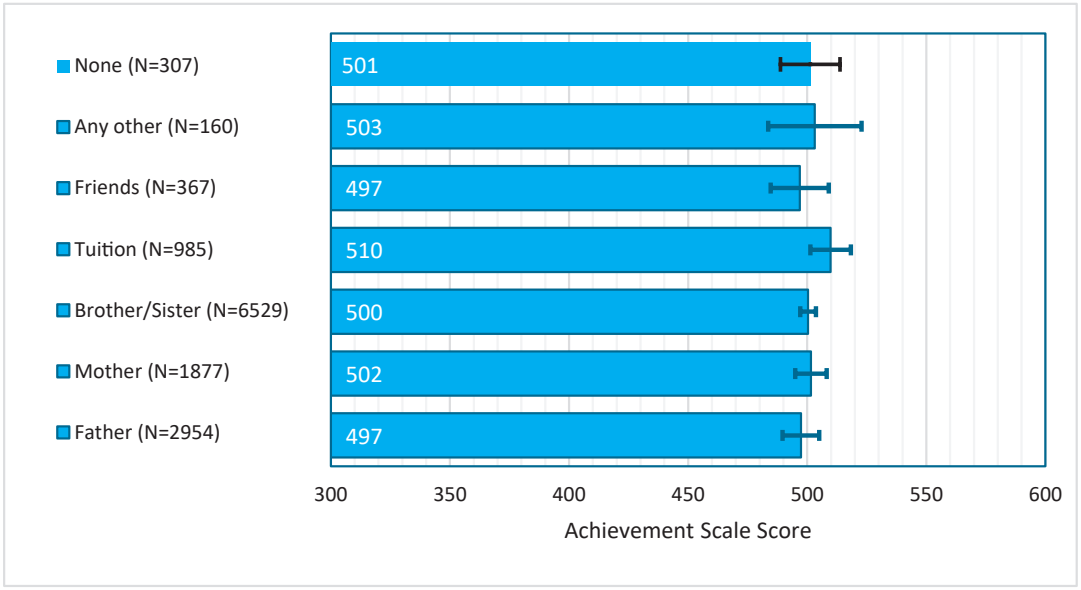


Figure 4.15 Mean score in Nepali by support received by students for their studies

Regarding learning achievement, the students who received support from extra tuition classes achieved high mean score (510) though the support received from mother and any other person was not less important; both types of support were positive showing achievement higher than the national mean score. The mean scores of the students who received support from their fathers and friends was equal (497) which was below the national average.

g. Students' Experience with Bullying at School and Learning Achievement

Through the background questionnaire, the students in this assessment were asked to respond whether they had bullying experience during one month's period (in the previous month from the time of assessment). Bullying experience has influenced student's achievement score. Seven types of bullying experiences were recorded from students' response. Figure 4.16 presents the percentage of students experiencing different types of bullying and those without such experience. The figure depicts that 52% of students experienced at least one type of bullying in their school. When the mean scores of the students experiencing different types of bullying was compared, it was found that bullied students had a poorer learning achievement than those who were not bullied. It indicates that experience with bullying negatively influences students' achievements.

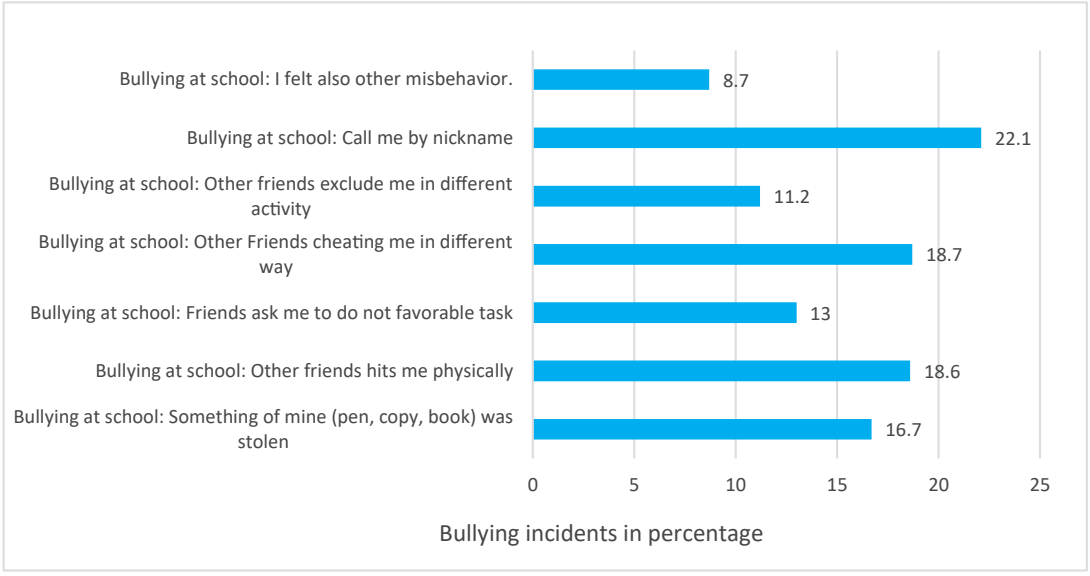


Figure 4.16 The percentage of students who experienced different types of bullying

There exists negative relation between the incidents of bullying experienced by the students and their achievement. The students who did not experience any bullying incidents scored 503 while the students who experienced 5 incidents of bullying scored up to 491. However, the percentage of students experiencing 100% bullying was very low (1.06%).

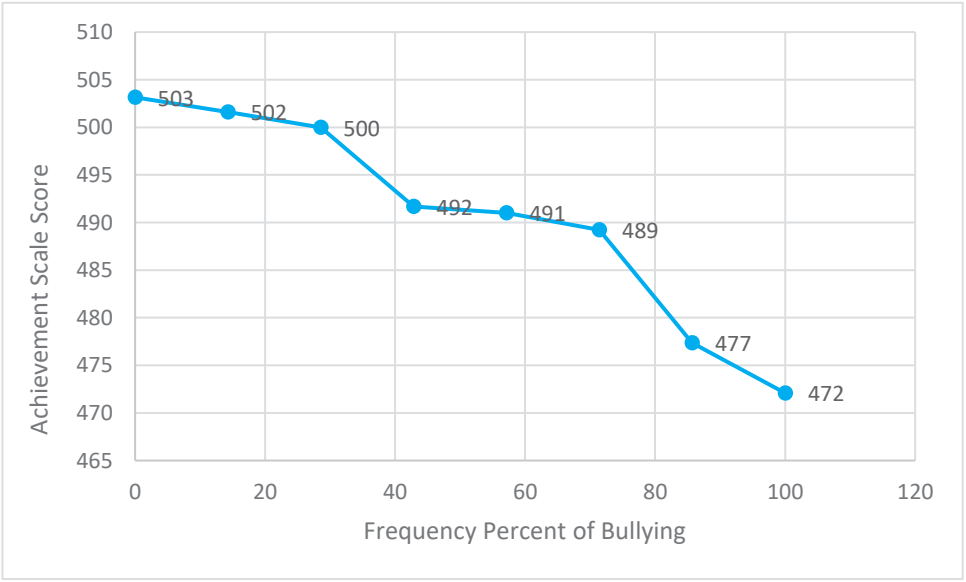


Figure 4.17 Relationship of extent of bullying and achievement

h. Results by the Availability of Textbook

The majority of students in this study had Nepali textbook. Some students (513) reported that they did not have textbook. The result presented in Figure 3.16, show that the students who did not have textbook achieved 485 mean score, which looks remarkably lower than the national mean score. But the students having access to the textbook were able to achieve 501, which is higher than the national mean score. The difference between the mean scores of these two groups is found statistically significant. Out of 513 students who did not have textbook, 378 students were from community schools. It should be ensured that all the students have a textbook with them as shown in figure 4.18.

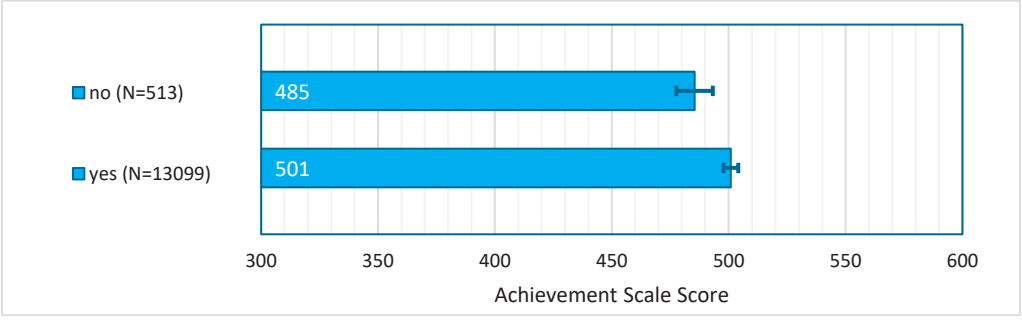


Figure 4.18 Mean score in Nepali by the availability of textbook

i. Results by the Use of Leisure Time at School

The students in this assessment were also asked to report how they used their leisure time at school. As seen in Figure 4.19, the students who are involved in the group work in leisure time achieved the highest score (508). Similarly, students spending time on classwork/homework during their leisure time achieved 502. However, the students who spent their leisure time playing achieved 490, which is lower than the national mean score.

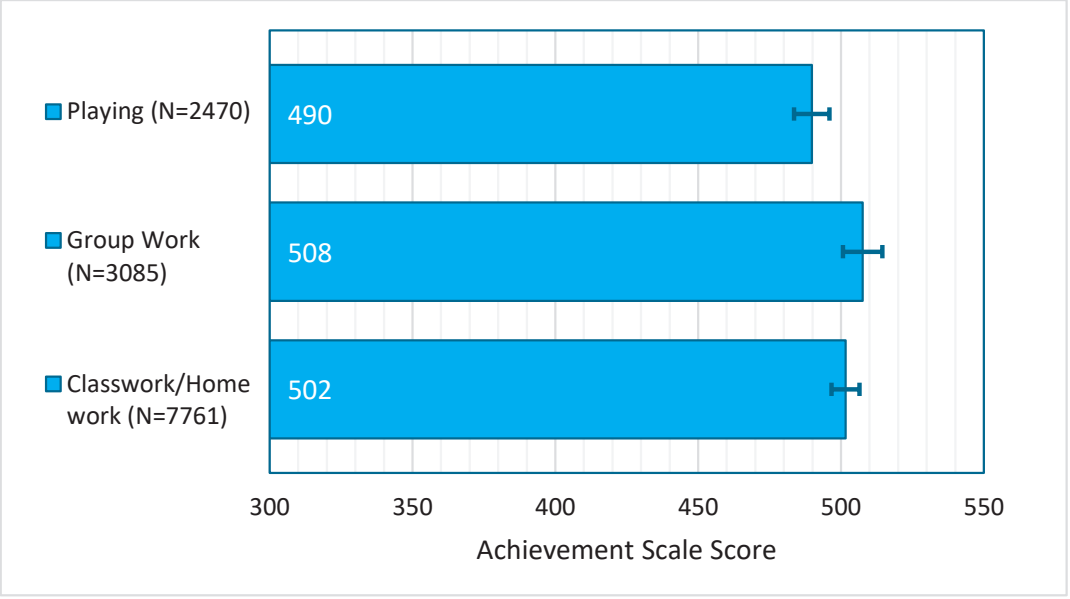


Figure 4.19 Mean score in Nepali by the activities in the leisure time at school

Total 2470 students reported that they played during leisure time. Out of 2470 students 78% students were from community school.

j. Results by Feedback Provided on Students' Homework

The students were also asked to mention how often their teachers provide them homework and how often the teachers provide feedback to them on their homework. The results indicate that there is a remarkable achievement gap between the students who were provided homework and who were not. The achievement scores of the students receiving regular feedback on their homework is higher than the scores of the students who did not receive such feedback from their teachers.

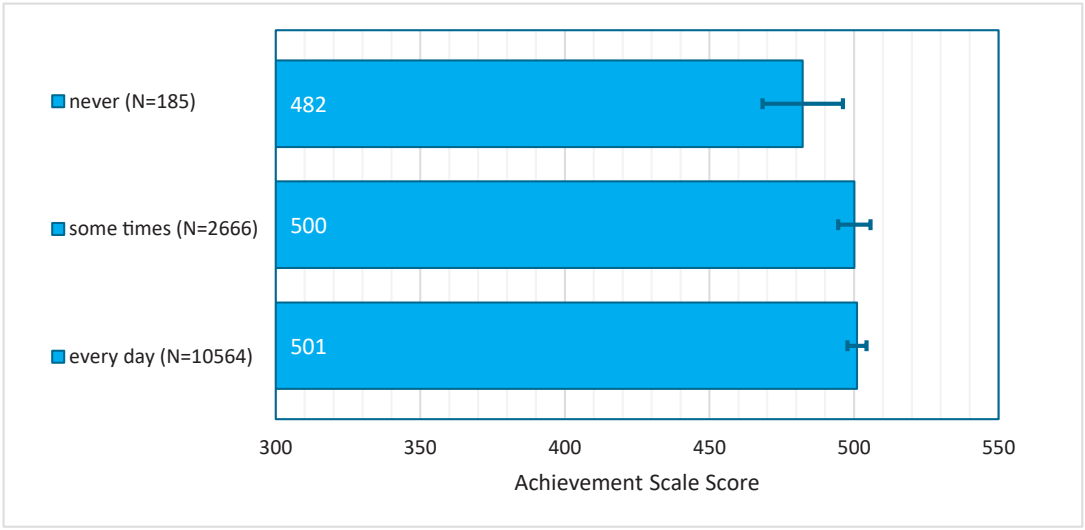


Figure 4.20 Mean score in Nepali by teacher's feedback on students' homework

The data presented in Figure 4.20 show that the mean score of the students who received feedback regularly is higher (501) than who did not receive regular feedback (482), with the difference of mean score of 19 point.

4.11 Relation between SES and Schools' Mean Score

Social economic status of students (SES) was estimated by aggregating the seven variables: mother's education, father's education, mother's occupation, father's occupation, home possessions, home accessories and students attending private schools. A regression model was developed in order to identify relation between mean of each SES variables and the students' mean score in Nepali. The R^2 value was found 0.34, means the school performance explained by their SES at 34 %. The following two-dimensional scatter (Figure 4.21) plot has been generated by plotting the school mean of SES and school mean of achievement.

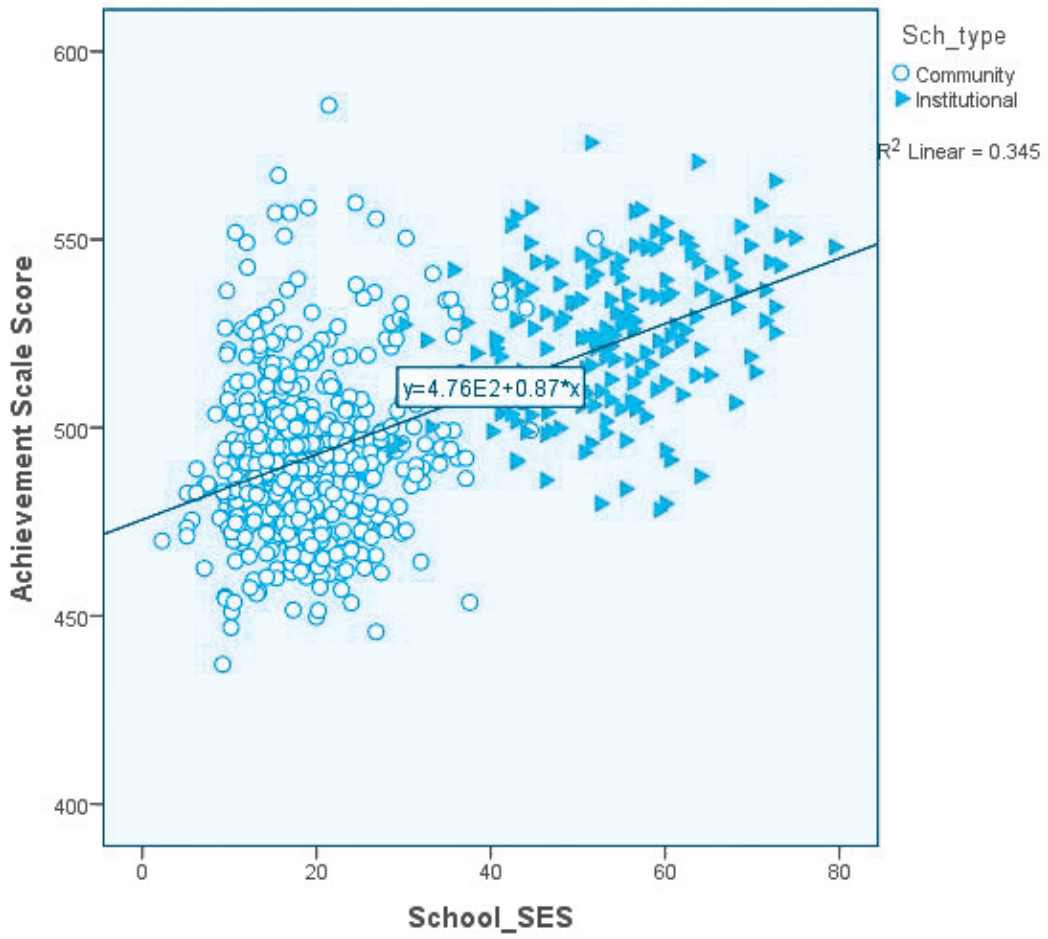


Figure 4.21 Relation between SES and schools' mean score in Nepali

The scatter plot shows that the schools with a high SES (average SES value of students) concentrated more on relatively medium and high mean scores, but the schools with low SES spread from low score to high scores. This is somehow different from the results in Mathematics. However, there are some cases of high SES schools having relatively low mean scores. This plot also indicates that most of the private schools have high SES students compared to the community schools.

Chapter 5
TRENDS OF LEARNING ACHIEVEMENT

5.1 Introduction

This chapter presents the trends of learning achievement in grade 5 over the years of NASA cycles. A short review of trends in learning achievement in grade 3, 5 and 8 in previous cycles is made in this subsection. In the next section, a comparison of learning achievement in NASA 2018 with NASA 2015 is presented.

Educational plans SSRP and SSDP have recognised NASA as a means of determining the progress in learning (SSRP, 2009-2016; SSDP, 2016). Since 2011 ERO has been conducting assessment in Nepali, English, Mathematics, Science and Social Studies (depends on year and grade). NASA is a valid tool to examine progress over the years. In all the NASA cycles from 2011, Item Response Theory was used to analyse the data and compare the results. However, in 2017, comparison was based on the Classical Test Theory because of technical limitations. In NASA 2018, comparison was made by using IRT – two methods: mean shift method and student average achievement score (latent ability). For the general audience, a table of classical parameter is also presented.

5.2 Review of Previous NASA Trends (2011-2017)

The Ministry of Education, Science and Technology (MoEST) implemented various programs during SSRP and SSDP to improve access, equity and quality in education. As a result, Nepal has achieved a remarkable progress in access and equity in education. Net enrolment rate of basic level (Grade 1-5) has reached 97.2. Efforts have been made during SSRP and SSDP to ensure the quality. A comparative presentation of NASA results of grade 3, 5, and 8 of different subjects is shown in table 5.1

Table 5.1 Comparative presentation of all NASA from previous cycles

	SSRP Period achievement in percent (%)						SSDP
Grade/year ➔	Grade 3		Grade 5		Grade 8		Grade 8
Subjects	NASA 2012	NASA 2015	NASA 2012	NASA 2015	NASA 2011	NASA 2013	NASA 2017*
Nepali	63	52	60	46	49	48	500Ÿ
Math	60	45	53	48	43	35	500
English	-	-	45	47			
Science	-	-	-	-	-	41	500
Social	-	-	-	-	49	-	-

Note: The symbol '↓' indicates decreased, '↑' indicate increased, '*' indicates that scores are expressed in transformed scale. The shaded column indicates the base year's achievement.

This table 5.1 summarizes the trends of results from previous cycle of national assessments. Except in Nepali subject in Grade 8 (2017), the achievement in Mathematics, Science and Nepali subject has dropped. Repeated results of NASA indicate that during the SSRP period, learning achievement level has not improved, although the enrolment rate has increased.

5.2.1 Methods of Trend Analysis from NASA 2015 to 2018

To find the trend of learning achievement, linking items' percentage of correct answers are reported in the first step. Then, for the IRT based reporting, linking items of NASA 2015 and 2018 are separately calibrated to generate the item parameters. Separate item parameters are generated by using one parameter Rash Model. From those parameters, latent ability of both years' students are calculated. Afterwards, comparison by using Weighted Latent Estimates (WLE) and item parameter "difficulty" of those linking items are separately used to find the achievement gap. By using WLE and item difficulty, achievement gap was found to be consistent in the result.

More specifically,

1. A comparative item wise classical parameter (percentage of correct answer) is reported between both years.
2. Average score of WLE 2015 vs WLE 2018 is reported.
3. Shift in mean by using mean shift calculation is done by difficulty parameters.

This comparison was made based on only one set of 2015 and 2018 assessments. NASA 2015 was linked with NASA 2018 Mathematics and Nepali datasets. This comparison was made easy by making almost equal number of students in the samples of both tests which made comparison more accurate.

5.3 Trends in Mathematics

In Mathematics, six anchor items of NASA 2015 were placed in NASA 2018. Among those six items, two items were 2 marks items. Hence, the IRT based linking had eight parameters to estimate the student latent ability (θ) from WLE methods and item difficulty (delta). At the beginning part, classical parameters (ie, percentage of correct answer) is presented in the table 4.3.1. The letter "p" represents percentage of correct answer.

Table 5.2 The classical parameter (percentage of correct answer) in Mathematics

Items description	Max	2015		2018	
		N	p	N	p
SetA_18 (2018) Round table multiplication item	1	4426	48%	4820	45%
SetA_19 (2018) sum of x and 4	1	4498	76%	4820	48%
SetA_20 (2018) Solve algebraic equation	1	4426	57%	4820	51%
SetA_23 (2018) Area problem	1	4426	48%	4820	45%
SetA_37 (2018) Volume (filling water)	2	4590	36%	4820	21%
SetA_40 (2018) Verbal problem of two Cloths length	2	4426	30%	4820	29%

The table 5.2 shows that percentage of correct answer in each common items was found different. Unexpected number was found in SetA_19 item where 76% of the students answered correctly in 2015 but only 48% students answered in 2018. Such large gap makes comparison unreal. However, based on the p value of both years, the percentage of correct answer in all items of two years varied and fell down in 2018. This roughly indicates that ability of the students in 2018 was lower than that in 2015.

To get the more accurate results, IRT based comparison is required. In IRT based methods, students ability (θ) was calculated by using the weighted likelihood estimation methods. The result of this analysis is presented in table 5.3.

Table 5.3 Comparison of students' ability (q) between NASA 2015 and 2018 in Mathematics

Test Year	N	Minimum	Maximum	Mean	Std. Error
NASA 2015	4590	-1.38	1.6	0.0045	0.01075
NASA 2018	4820	-1.66	1.52	-0.4492	0.01186
Difference				-0.4537	

The average latent ability of the students based on the analysed items was 0.0045 and -4429 in the year 2015 and 2018 respectively. The average ability of the students in NASA 2018 was declining than in NASA 2015 by 0.4537. In the table, the negative sign in the 'difference' row indicates the declining ability.

The third method of comparing student ability between the years was done by using the difficulty of anchor items. As difficulty declines, it indicates that ability is raised and vice versa. Table 5.4 presents the status of anchor item difficulty to diagnose the mean shift in anchor items.

Table 5.4 Status of anchor item difficulty parameters and shift of the mean in Mathematics

Item description	Average delta (NASA 2015)	Average delta (NASA 2018)
item SetA_19_SetB24	0.19937	-0.56359
item SetA_20	-1.24176	-0.62139
item SetA_23	-0.21803	0.26618
item SetA_37_Set3Q31	0.17756	0.43703
item SetA_40 (1)	0.5338	0.1374
Item SetA_40 (2)	0.49798	0.36539
item SetB_11 (1)	0.83371	0.33952
item SetB_11 (2)	0.57826	1.00755
Mean difficulty	1.36089	1.36809
Mean difficulty	0.30242	0.30402
Mean Shift		- 0.0016

In the table 5.4 shows the difficulty (average delta) of linking items with partial credits of NASA 2015 and 2018 in two columns. Comparing the average difficulty of each item, change is fluctuating by items. However, in average, delta (difficulty) in the

NASA 2018 has increased by 0.0016 than in the difficulty in 2015. This small increase in average difficulty of the test items indicates that ability of the students has slightly decreased in 2018 than in 2015 in Mathematics.

All the three methods have presented a consistent decrease in learning achievement from 2015 to 2018. This difference in student ability (-0.4537) is significant ($p<0.05$), but the difference is moderate (effect size = 0.29). It reveals that there is a notable decrease in student ability from 2015 to 2018 based on the sample data.

5.4 Trends in Nepali Language

In Nepali reading and writing, eight anchor items of NASA 2015 were placed in NASA 2018. Among those eight items, two items were partial credit marks items with maximum score 4 and 3. Hence, the IRT based linking has 14 parameters to estimate the student latent ability (θ) from WLE methods and item difficulty (delta). In the first, classical parameters (ie, percentage of correct answer) are presented in the table 5.5. The letter "p" represents percentage of correct answer.

Table 5.5 The classical parameter (percentage of correct answer) of anchor items in Nepali

Items description	Max	NASA 2015		NASA 2018	
		N	p	N	p
G5LQ2a	1	4526	77%	4847	77%
G5LQ2b	1	4526	76%	4847	73%
G5LQ2c	1	4526	66%	4847	58%
G5LQ2d	1	4526	71%	4585	65%
G5LQ6a	1	4517	38%	4673	69%
G5LQ6b	1	4517	54%	4673	51%
G5LQ12	4	4517	33%	3471	30%
G5LQ14	3	4312	37%	3009	45%

The table 5.5 presents the percentage of correct answer in all items which shows that the percentage of correct answer varies from item to item. So, this does not present the clear picture of trends of rise and fall in the results.

In first IRT based method, student's ability (θ) was estimated by using the weighted likelihood estimation, which is presented in 5.6.

Table 5.6 Comparison of students' ability (q) between NASA 2015 and 2018 in Mathematics

Test Year	N	Minimum	Maximum	Mean	Std. Error
NASA 2015	4526	-0.57	1.32	0.36	0.00575
NASA 2018	4847	-0.80	1.78	0.41	0.00657
Difference				+ 0.04	

The average ability of the students in NASA 2018 was raising from the average ability in NASA 2015 by 0.04 (round figure).

According to the mean shift method, as difficulty declines, it indicates that the ability is raised and vice versa. The table 5.7 presents the status of anchor item difficulty to diagnose the mean shift in anchor items.

Table 5.7 Item wise difficulty of linking items

Item description	Average delta (NASA 2015)	Average delta (NASA 2018)
G5LQ2a	-0.927	-0.936
G5LQ2b	-0.836	-0.677
G5LQ2c	-0.324	0.071
G5LQ2d	-0.598	-0.240
G5LQ6a	0.901	-0.455
G5LQ6b	0.195	0.348
G5LQ12	0.855	1.283
G5LQ12(1)	0.440	-1.657
G5LQ12(2)	-0.533	-0.045
G5LQ12(3)	0.063	0.433
G5LQ14(1)	0.925	-0.767
G5LQ14(2)	-0.685	0.039
Mean difficulty	-0.044	-0.217
Mean Shift		+0.173

Average delta (difficulty) of the NASA 2018 has decreased by 0.173. A very slight decrease in average difficulty of the test items indicates that the ability of the students has increased slightly in 2018 than their ability in 2015 in reading and writing skills in Nepali language.

All the three methods have presented a consistent decrease in learning achievement

from 2015 to 2018. This difference in student ability (+0.173) is significant ($p<0.05$). However, the difference is very small (effect size = 0.0.005). This evidence reveals that there is a very small increase in student ability in 2018 than in 2015 based on the sample data.

5.5 Overall Trends in NASA Results of Mathematics and Nepali

Various factors were found influencing the learning outcomes of the students. The factors include: learning time at home, parent's education level, learning environment provided at home and the school, attitude towards the school, attitude towards the teacher and the subject, socio-economic status of the family, discipline and school environment, bullying, type of school, district and provinces. Among these, some factors have maintained the trend while others have some fluctuations in the results. Table 5.8 presents those trends over the cycles of assessment in grade 5 on the assessed subjects.

Table 5.8 Trends in effective variables among three cycles of NASA in grade 5

Description of consistently recurring results	Trends		
	2012	2015	2018
There is a wide gap in the achievement of community and institutional schools.	✓	✓	✓↓
Huge variation was also found in the achievement of the students from different districts (for example in 2018, Kathmandu 71 and Udayapur 37).	✓	✓	✓
Brahmin/Chhetri students performed better than the Dalit and Janajati students. Dalit 50 and Brahmin 61	✓	✓	✓
Students whose mothers have Bachelor degree and the students whose fathers have Master degree performed better than the students who have illiterate and just literate parents. (50 to 74).	✓	✓	✓
Students whose parents have been involved in the professions like business, government job and teaching performed better than those whose parents have been involved in agriculture and household work.	✓	✓	✓

Description of consistently recurring results	Trends		
	2012	2015	2018
Students with the more possessions and accessories such as computer, television, and dictionary in their home performed better than the others who have less possessions.	✓	✓	Math ✓↓ Nepali✓
A huge gap exists between the achievement of the students who spent more after-school time in wage making activities and those who spent more time in their study.	✓	✓	✓
Students who studied Grade 5 in appropriate age (11 and 12 years) performed better than those who studied in later or earlier age. (46 to 55 gap in 2018)	✓	✓	✓
Students taking support from extra tuition classes performed better than those without such support.	✓	✓	✓
Students (3.8%) responded that they did not have textbook until the end of the academic session. The performance gap between those with and without textbook was 44 to 54.	✓	✓	✓
Overall, SES had a significant impact on learning.	✓	✓	Math✓↓ Nepali✓
The students who speak Nepali language at home performed better in Nepali subject than in Mathematics.	✓	✓	Math ✓↓ Nepali✓
Boys and girls' performance in grade 5 was close to equal.	✓	✓	✓
Gap in the achievement of boys and girls in Terai region is comparatively higher.	✓	✓	Math✓ Nepali✓↓

Note: ✓ represents high gap and same trend, ✓↓ represents reduced gap.

The above consistent recurring results not only proves the reliability of the NASA study, but also indicates that interventions of educational system were not effective in improving the quality of learning in the school level, urges to review the plan and policies from the quality concerns.

Chapter 6

PREDICATOR OF LEARNING ACHIEVEMENT

6.1 Regression analysis

Regression analysis enables to predict 'the specific value of one variable when we know or assume values of the other variable(s). It is a way of modelling the relationship between variables (Cohen and Holliday 1996). To calculate the effects of two or more independent variables on a dependent variable. Multiple regression enables us to predict and weight the relationship between two or more explanatory or independent variables and an explained or dependent variable. The Beta weighting (β) gives us an indication of how many standard deviation units will be changed in the dependent variable for each standard deviation unit of change in each of the independent variables. (Cohen and Holliday 1996).

In this analysis, student latent ability (WLE/θ) in the form of MSS-WLE variable that reflects the student ability was taken as dependent variable and remaining independent variables about 100 variables were taken into account. By using SPSS 23, at the first stage, 'Enter' method was used to detect the removable variables. After this diagnosis, 'Remove' method was used to find the effective predicators. All the variables were changed into positive scale and then converted into z-score so that all the predicators are measured into same scale. After data preparation, regression analysis was run to build a model.

6.2 Regression Analysis in Mathematics

The model summary and regression coefficients for Mathematics are presented in the table 6.1

Table 6.1 Model summary regression in Mathematics

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.492a	0.242	0.242	45.92837
2	.000b	0	0	52.74911
Predicators: Constant and all variables included in the model				
b Predictor: (constant)				

Table 6.2 Regression coefficients – unstandardized and standardized in Mathematics

SN	Predicators (converted into z-score)	Unstan. Coeff.		Stan. Coeff.	t	Sig.
		B	SE	Beta		
1	(Constant)	479	0.142		3378.4	0.000
2	Gender (1 = boys, 2 = girls)	-3.42	0.133	-0.065	-25.79	0.000
3	Student age (9years to 16 or above)	0.59	0.140	0.011	4.23	0.000
4	Home language (1 = Nepali, 2 = other)	1.42	0.148	0.025	9.58	0.000
5	Geographical identity (1 = Madhesi, 2 = hilly, 3 = Himali)	-3.25	0.146	-0.058	-22.27	0.000
6	Time spent in home: watch TV, internet, mobile, computer	1.54	0.148	0.027	10.42	0.000
7	Time spent in home: play and talk with friends	1.66	0.142	0.030	11.72	0.000
8	Time spent in home: Household chores	-0.60	0.145	-0.011	-4.13	0.000
9	Time spent in home: Study and do homework	6.50	0.156	0.110	41.82	0.000
10	Time spent in home: Work for wages	-3.75	0.160	-0.063	-23.49	0.000
11	Time spent in home: Read other books	1.58	0.144	0.028	11.00	0.000
12	After school support for study in home	0.47	0.136	0.009	3.48	0.001
13	Do you have the text book of Mathematics?	2.35	0.149	0.040	15.79	0.000
14	Home to school distance	-2.40	0.134	-0.045	-17.91	0.000
15	Does your teacher give homework?	0.71	0.152	0.013	4.64	0.000
16	Does teacher give feedback?	0.85	0.148	0.015	5.77	0.000
17	Your aim (1 = teacher, 2 = govt service, 3 = private job, 4 business, 5 foreign job, 6 farmer 7 other)	3.26	0.134	0.066	24.26	0.000
18	Perceived utility of Math: learning Math is useful for my household calculations.	0.89	0.156	0.015	5.71	0.000
19	Perceived utility of Math: Learning Math helps in larning other subjects	0.46	0.148	0.008	3.10	0.002
20	Perceived utility of Math: I like to practice Math	2.09	0.161	0.036	12.94	0.000
21	Perceived utility of Math:I have to do better in Math for job and work	0.07	0.155	0.001	0.47	0.641

SN	Predicators (converted into z-score)	Unstan. Coeff.		Stan. Coeff.	t	Sig.
		B	SE	Beta		
22	Attitude towards Math: Mostly, I do good in Math	1.07	0.156	0.019	6.87	0.000
23	Attitude towards Math: I like to learn Math more	0.31	0.174	0.006	1.81	0.070
24	Attitude towards Math: I enjoy learning Math	0.10	0.168	0.002	0.61	0.544
25	Attitude towards Math: I can learn Math quickly	2.39	0.156	0.044	15.35	0.000
26	Attitude towards Math: I feel Math difficult	-1.97	0.134	-0.037	-14.68	0.000
27	Classroom activities: We do exercise given in the text book	6.40	0.159	0.106	40.24	0.000
28	Classroom activities: We practice math mutually with friends	-0.006	0.145	0.000	-0.04	0.968
29	Classroom activities: We try to solve mathematical problem ourselves	1.08	0.140	0.020	7.73	0.000
30	Classroom activities: We use geometric box materials to learn Math	1.03	0.141	-0.019	-7.27	0.000
31	Classroom activities: We start to do our homework in the class	-6.09	0.140	-0.115	-43.56	0.000
32	Classroom activities: We ask questions to the teacher when we don't know	-0.46	0.149	-0.008	-3.11	0.002
33	Mother education (1 = illiterate, .. 7 = Masters or above)	1.29	0.172	0.025	7.50	0.000
34	Mother occupation (1 = agriculture & home, 2 = household, 3 = other's house, 4 = labour, 5 = foreign job, 6 = teacher, 7 = business, 9 = govt job, 9 = other)	0.41	0.135	0.008	3.01	0.003
35	Father education (1 = illiterate, .. 7 = Masters or above)	2.46	0.179	0.047	13.73	0.000
36	Father occupation (1 = agriculture & home, 2 = household, 3 = other's house, 4 = labour, 5 = foreign job, 6 = teacher, 7 = business, 9 = govt job, 9 = other)	2.86	0.150	0.055	19.03	0.000

SN	Predicators (converted into z-score)	Unstan. Coeff.		Stan. Coeff.	t	Sig.
		B	SE	Beta		
37	Family size (1 to 13 or above)	-0.51	0.138	-0.009	-3.72	0.000
38	Home possessions: study table (0 = no, 1 = yes)	0.10	0.139	0.002	0.71	0.478
39	Home possessions: separate room for study (0 = no, 1 = yes)	-1.17	0.130	-0.023	-8.94	0.000
40	Home possessions: peace place for study (0 = no, 1 = yes)	0.95	0.136	0.018	6.96	0.000
41	Home possessions: computer for school work (0 = no, 1 = yes)	-2.63	0.146	-0.054	-18.01	0.000
42	Home possessions: children story, poem, picture books (0 = no, 1 = yes)	1.73	0.154	0.034	11.20	0.000
43	Home possessions: picture books (0 = no, 1 = yes)	-0.22	0.145	-0.004	-1.49	0.135
44	Home possessions: internet facility (0 = no, 1 = yes)	0.81	0.136	0.017	5.96	0.000
45	Home accessories: TV (0 = no, 1 = yes)	3.82	0.156	0.072	24.40	0.000
46	Home accessories: computer (0 = no, 1 = yes)	2.28	0.174	0.044	13.09	0.000
47	Home accessories: motercycle (0 = no, 1 = yes)	1.21	0.155	0.023	7.81	0.000
48	Home accessories: car (0 = no, 1 = yes)	-3.35	0.176	-0.060	-18.98	0.000
49	Home accessories: permanent house (0 = no, 1 = yes)	1.54	0.143	0.029	10.77	0.000
50	What do you do in leisure period? (1 = classwork/homework, 2 = group work, 3 = play)	-0.92	0.133	-0.017	-6.90	0.000
51	How often do extra-curricular activities are happen? (1 = regularly, 2 = occosinally, 3 = never)	-1.14	-0.137	-0.022	8.29	0.000
52	How often do you participate in expra-curricular activities? (1 = regularly, 2 = occosinally, 3 = never)	-1.98	0.142	-0.037	-13.90	0.000
53	teacher: Teachers treat lovely while teaching	1.06	0.166	0.018	6.39	0.000

SN	Predicators (converted into z-score)	Unstan. Coeff.		Stan. Coeff.	t	Sig.
		B	SE	Beta		
54	teacher: Teachers do not scold us	1.59	0.153	0.029	10.43	0.000
55	teacher: Teachers don't use corporal punishment	0.50	0.143	0.009	3.47	0.001
56	Teachers care us equally	1.11	0.180	-0.019	-6.19	0.000
57	Teachers answer while asking confused questions	-1.13	0.178	-0.019	-6.34	0.000
58	Teachers give homework	-1.03	0.183	-0.018	-5.62	0.000
59	Teachers provide feedback by checkign homework	0.34	0.179	0.006	1.88	0.060
60	Teachers teach full time in the class.	0.63	0.160	0.012	3.94	0.000
61	I like to come and stay in school	1.17	0.186	0.018	6.27	0.000
62	Students in my school try to do their best	-0.76	0.160	-0.014	-4.76	0.000
63	There is facility to play, drinking water and toilet	-0.07	0.175	-0.001	-0.42	0.675
64	I participate in child club and children programs	-1.85	0.142	-0.035	-13.03	0.000
65	Bullying: my belongings (copy/pen/book/water pot) was stolen.	0.68	0.161	0.011	4.20	0.000
66	Bullying: I was hit or hurt by other student(s)	-2.63	0.155	-0.046	-16.97	0.000
67	Bullying: fellow students kept outside without involving me in activities	-3.95	0.169	-0.064	-23.32	0.000
68	Bullying: I was made fun of or called names	0.69	0.162	0.012	4.23	0.000
69	Bullying: fellow students kept outside without involving me in activities	0.21	0.181	0.003	1.18	0.240
70	Bullying: I was made to do things I didn't want to do by other students.	-1.97	0.148	-0.037	-13.27	0.000

Note: unstan Coeff. = Unstandardized coefficients, SE = Standard Error, Stan. Coeff = Standardized coefficients

The table 6.2 reveals many predicators (positive and negative) of learning achievement in Mathematics.

6.3 Regression Analysis in Nepali Subject

Like in chapter 6.2 Nepali background variables were converted into z-score before running regression analysis. The model summary of the regression in Nepali subject is given in table 6.3.

Table 6.3 Regression analysis model summary in Nepali subject.

Model Summary				
Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**
1	.409a	.168	.157	46.69798

Predicators: constant and independent variables.

Exactly in similar manner as in Mathematics, the predicators and regression coefficients of Nepali achievement are presented in table 6.4

Table 6.4 Regression coefficients – unstandardized and standardized in Mathematics

SN	Predicators (z-score)	Unstan Coeff.		Beta	Stan. Coeff.	t
		B	SE			
1	(Constant)	500.44	0.72		696.76	0.000
2	bq3: Students' gender	1.79	0.83	0.03	2.16	0.031
3	bq4: Students' age	0.71	0.72	0.01	0.99	0.320
4	bq5: Language spoken at home	-1.76	0.73	-0.03	-2.41	0.016
5	bq6: Ethnicity	0.06	0.65	0.00	0.10	0.924
6	bq7a: Time spent on TV, Internet, Mobile, Computer	-5.69	0.81	-0.10	-7.03	0.000
7	bq7b: Time spent on playing and chatting with friends	-4.75	0.90	-0.08	-5.29	0.000
8	bq7c: Time spent on household chores	-1.48	0.91	-0.03	-1.63	0.104
9	bq7d: Time spent on studying/doing homework	0.44	0.82	0.01	0.53	0.593
10	bq8: Support for study beyond school time	0.52	0.75	0.01	0.70	0.482
11	bq9: Availability of Nepali textbook	1.58	1.00	0.02	1.57	0.116
12	bq10: Time to reach school	-2.29	0.89	-0.04	-2.57	0.010
13	bq11: Homework provided by Nepali teacher	0.42	1.02	0.01	0.41	0.681
14	bq12: Feedback provided in homework	2.01	1.06	0.03	1.90	0.057
15	bq13: Student's' future aim	1.90	0.76	0.03	2.50	0.012

SN	Predicators (z-score)	Unstan Coeff.		Beta	Stan. Coeff.	t
		B	SE			
16	bq14a_Utility: Learn Nepali help me to read & write	2.24	0.98	0.03	2.28	0.023
17	bq14b_Utility: I should do well in Nepali for other subject learning	-2.81	1.19	-0.04	-2.35	0.019
18	bq14c_Utility: I like story and poem in Nepali	3.52	1.27	0.05	2.77	0.006
19	bq14d_Utility: I need to be competent in Nepali to get a good job	1.51	1.21	0.02	1.25	0.212
20	bq15a: Attitude towards Nepali Learning: Often performed good in Nepali	-0.59	1.05	-0.01	-0.57	0.572
21	bq15b: Attitude towards Nepali Learning: Want to learn Nepali more at school	-1.30	1.19	-0.02	-1.09	0.278
22	bq15c: Attitude towards Nepali Learning: I Enjoy learning Nepali	0.62	1.12	0.01	0.56	0.577
23	bq15d: Attitude towards Nepali Learning: I Can learn Nepali quickly	2.22	1.01	0.03	2.21	0.027
24	bq15e: Attitude towards Nepali Learning: Nepali is difficult for me	3.42	0.78	0.06	4.39	0.000
25	bq16a: Classroom activities In Nepali: Do exercise as given lesson	0.64	1.13	0.01	0.56	0.573
26	bq16b: Classroom activities In Nepali: Read/write in small group of friends	-1.86	1.15	-0.03	-1.63	0.104
27	bq16c: Classroom activities In Nepali: Make sentence Our self	0.69	1.15	0.01	0.60	0.547
28	bq16d: Classroom activities In Nepali: Read the lesson with aloud	-1.02	1.07	-0.02	-0.95	0.341
29	bq16e: Classroom activities In Nepali: Start homework at the classroom	-2.00	0.97	-0.03	-2.06	0.039
30	bq16f: Classroom activities In Nepali: Ask the teacher for difficult task	5.55	1.03	0.08	5.38	0.000
31	bq17: Mother's education	2.79	0.86	0.05	3.24	0.001
32	bq18: Mother's occupation	1.18	0.79	0.02	1.48	0.138
33	Zscore: bq19: Father's education	1.46	0.89	0.03	1.64	0.101
34	Zscore: bq20: Father's occupation	4.15	0.78	0.08	5.33	0.000
35	bq21: Number of family member	-0.93	0.64	-0.02	-1.45	0.146
36	bq22a: Home possession: Table for reading	0.76	0.70	0.02	1.09	0.276

SN	Predicators (z-score)	Unstan Coeff.		Beta	Stan. Coeff.	t
		B	SE			
37	bq22b: Home possession: Separate room	-0.01	0.66	0.00	-0.01	0.993
38	bq22c: Home possession: Peace place to read	1.21	0.69	0.02	1.75	0.080
39	bq22d: Home possession: Computer	-1.20	0.72	-0.03	-1.66	0.096
40	bq22e: Home possession: Children story, magazine and poems	0.69	0.76	0.01	0.91	0.362
41	bq22f: Home possession: Pictorial book	0.03	0.76	0.00	0.03	0.974
42	bq22g: Home possession: Other reference book	2.48	0.74	0.05	3.34	0.001
43	bq22h: Home possession: Internet facility	2.55	0.75	0.05	3.41	0.001
44	bq23a: Home accessories: Television	1.87	0.90	0.04	2.07	0.039
45	bq23b: Home accessories: Computer	-1.65	1.04	-0.03	-1.58	0.115
46	bq23c: Home accessories: Motorcycle	-0.10	0.94	0.00	-0.10	0.918
47	bq23d: Home accessories: Car	-3.09	0.97	-0.06	-3.20	0.001
48	bq23e: Home accessories: Permanent house	1.68	0.85	0.03	1.99	0.047
49	bq24: Activities in leisure time at school	-1.38	1.01	-0.02	-1.36	0.174
50	bq25: Frequency of extra activities at school	-0.71	1.07	-0.01	-0.66	0.507
51	bq26: Frequency of participation in extra activities	-1.10	1.06	-0.02	-1.04	0.297
52	bq27a. teacher: Teachers treat lovely while teaching	0.25	0.86	0.00	0.29	0.771
53	bq27d. teacher: Teachers care us equally	2.11	0.86	0.04	2.46	0.014
54	bq27e. teacher: Teachers answer while asking confused questions	2.06	0.92	0.04	2.25	0.025
55	bq27f. teacher: Teachers give homework	1.27	0.89	0.02	1.43	0.153
56	bq27g. teacher: Teachers provide feedback by checking homework	0.84	0.89	0.02	0.95	0.345
57	bq28c. School: There is facility to play, drinking water and toilet	1.97	0.80	0.03	2.48	0.013
58	bq29a. bullying: my belongings (copy, pen, book, water pot. etc) was stolen.	-0.46	0.76	-0.01	-0.60	0.548
59	bq29b.Incidents: I was hit or hurt by other student(s)	-1.22	0.80	-0.02	-1.52	0.128
60	bq29c. Incidents: fellow students kept outside without involving me in activities	-3.01	0.85	-0.05	-3.56	0.000

SN	Predicators (z-score)	Unstan Coeff.		Beta	Stan. Coeff.	t
		B	SE			
61	bq29d. Incidents: I was made fun of or called names	-0.56	0.85	-0.01	-0.65	0.515
62	bq29e. Incidents: fellow students kept outside without involving me in activities	-1.86	0.87	-0.03	-2.14	0.032
63	bq29f. Incidents: I was made to do things I didn't want to do by other students	0.75	0.86	0.01	0.88	0.379
64	bq29g: Other	-0.66	0.77	-0.01	-0.85	0.396
31	bq17: Mother's education	2.79	0.86	0.05	3.24	0.001
32	bq18: Mother's occupation	1.18	0.79	0.02	1.48	0.138
33	bq19: Father's education	1.46	0.89	0.03	1.64	0.101
34	bq20: Father's occupation	4.15	0.78	0.08	5.33	0.000
35	bq21: Number of family member	-0.93	0.64	-0.02	-1.45	0.146
a Dependent Variable: MSSPV1						

Table 6.4 presents both negative and positive predicators of Nepali subject. Negative sign indicate the negative association of the predicator with the achievement and vice versa.

Chapter 7
FINDINGS, CONCLUSION AND RECOMMENDATIONS

7.1 Introduction

Globally, National Assessment of Student Achievement (NASA) has been well accepted as a means of measuring quality of education through students' achievement (TIMSS & PIRLS, 2008). National assessment provides both quantitative and descriptive form of information on student achievement which is considered as an output of the teaching learning process and its quality (World Bank, 1996). National assessment thus provides basic information for policy makers, politicians, and the broader educational community (ERO, 2013). Further, "it provides data for a type of national education audit carried out to inform policy makers about the key aspects of the system" (Greaney & Kellaghan, 2008b, p. 7, ERO, 2013). It is argued that the achievement of the students in a curriculum area be aggregated to provide an estimate of the achievement level in the education system as a whole at a particular age or grade level (Greaney & Kellaghan, 2008b; NASA, 2013). NASA is also a popular means of determining the achievement of curriculum and finding the gaps between written curriculum and taught curriculum. So, it is useful for making policy decisions especially when decisions are to be made in relation to the optimum utilisation of resources (EDSC, 2008). It provides evidence for policy makers on availability of textbooks, class size, and number of years of teacher training. Therefore, every country has accepted that it is "systematic, regular measure of learning achievement in a country that is designed to assist policy making" (Lockheed et al. cited in EDSC, 2008, pp. 19, NASA 2013).

In Nepal, the national assessment practice is found to have started from the last years of the decade of 1980s. However, the Ministry of Education formally started National Assessment since 1995 and continued it up to 2010 in a small scale. Large scale NASA was administrated under the Ministry of Education since 2011 AD. Four NASA cycles were completed during the School Sector Reform Plan and two under the School Sector Development Plan including NASA 2018. In both the plans, NASA is considered as a tool to measure quality of education and making educational institutions accountable to achieving the educational goals.

NASA studies are conducted for both backward and forward-looking purposes. The backward-looking purpose is concerned mainly with building a database to

analyse both the strengths and weaknesses of educational policies and practices that affect students' learning achievement (ERO, 2018).

A complete NASA cycle goes over a period of 3 years. In the first year, all items development, pre-testing of the items and item analysis are completed. In the second year, final test is administered and finally, the activities including report writing, dissemination of the report and policy informing are done during the third year of NASA cycle.

The ERO follows globally accepted practices of conducting national assessments. Although the context of each country is different, there are some common practices to national assessments in most of the countries (ERO, 2018). Building on the comprehensive review of national assessments from various countries, ERO has adopted the following procedures:

- The Ministry of Education, Science and Technology (MOEST) selects an implementing agency either from within the MOEST system or an independent external consulting organization. In case of Nepal, Education Review Office (ERO) within the MOEST system is solely responsible for the national assessment.
- The MOEST or implementing agency develops policies and frameworks for assessment in consultation with (and with participation of) key stakeholders such as subject experts, teachers and policy makers.
- The MOEST identifies the Grade level and determines the area (e.g., literacy or numeracy) to be assessed.
- The implementing agency (ERO in Nepal) defines and describes the areas of achievement testing in terms of both content and cognitive skills and develops test items along with supporting questionnaires and manuals for test administration.

ERO

- Pilots the test items with the support of external experts and reviews their validity, appropriateness and sensitivity in terms of gender, ethnicity and culture.
- Ensures that the assessment instruments are reliable and valid.
- Selects the samples schools, arranges for printing the test papers and other relevant materials; and communicates with the schools and teachers for test administration.
- Orients the test administrators (focal persons, head teachers and teachers), and then administers the test and survey questionnaires in the selected schools.

- Collects test scores and other necessary information, cleans the data as needed and analyses them.
- Prepares draft report/s which is/are reviewed by relevant subject committees and external experts.
- Prepares and disseminates final report/s through various means such as publication and the mass media.

- Finally, the MOEST, implementing agency and relevant stakeholders study the report/s of national assessment and identify major areas for policy reforms (ERO, 2017, 2018).

7.2 Objectives of NASA 2018

The purpose of this assessment is to provide feedback to the Ministry of Education, Science and Technology to improve the quality of school education. This assessment does not report individual students' performance, nor does it compare the proficiencies of each individual student and school. Rather, it provides the national and provincial level results as well as the differences in the achievement scores in relation to various influencing factors such as socioeconomic status, home language, and identity with geographical region. More specifically, NASA 2018 has the following objectives:

a. To identify the current level of Grade 5 students' achievement in Mathematics and Nepali
b. To identify variations in student achievement by gender, province, identity with geography, types of school, ethnicity, home language, socio-economic status
c. To explore the factors that influence student achievement
d. To identify trend in student learning and produce the baseline data for the future for comparison
e. To strengthen the capacity of the education system in conducting national assessment
f. To provide the Ministry of Education, Science and Technology with recommendations for policy making to improve quality and ensure equity, particularly in school education.

7.3 Distinct Features of NASA 2018

The ERO has used Item Response Theory to assess the latent ability of students using various contextual variables to explain those latent traits of the students. This assessment has used advanced procedure to bring rigor to data analysis by generalizing the results in national level and province levels through 7 explicit strata and various

other implicit strata. Use of Replicate Module for estimating the population parameters and Weighted Likelihood Estimation (WLE) for analysis of individual student level and reporting are the examples of its advancement. Furthermore, the advancement of procedures has also been noticed in sampling methods. A Probability Proportional to Size (PPS) sampling procedure has been used in selecting the schools as Principal Sample Unit (PSU), the school clusters. Reporting of student achievement at province level and national level is done in a transformed scale with mean 500 and standard deviation 50 by using the formula:

$$Average\ scale\ score = 500 + plausible\ value * 50$$

$$or, \quad Average\ scale\ score = 500 + logit * 50$$

If readers want to extract the WLE of latent ability, they can use:

$$average\ latent\ ability\ (logit)\ of\ any\ group = \frac{\text{average score - 500}}{50}$$

The distinct features of this report are:

1. A comparative analysis of NASA 2018 and 2015 by using IRT methods and rigorous process is adopted.
2. Trends of results over all NASA cycles of grade 5 are presented.
3. Learning level descriptors are prepared through a rigorous analysis.
4. A gap in learning between written curriculum and taught curriculum in the form of achieved curriculum is presented.

7.4 Findings

The major findings of National Assessment of Student Achievement study 2018 are as follows:

1. In Mathematics, 32 students out of 100 fall below basic level who have achieved only 5% of the tested curriculum and only 28% curriculum is achieved by basic level students. More than 70% students are under the 28% achievement of the tested curriculum in Mathematics. This huge mass of the students represents underperforming group. The proficient level students achieved 62% of the tested curriculum and advance level students achieved 96% of the tested curriculum. The data informs that only 29% of the students have adequate knowledge and skills in Mathematics curriculum. The gap in achieved curriculum between below basic level (5%) and advance level students (96%) is 91%. This indicates that inequality in the classroom is remarkably high.

Similarly, in Nepali, 20 out of 100 students achieved only 18% of the tested curriculum and 40 out of 100 students achieved only 38% of the tested curriculum. 45 out of 100 students have adequate knowledge and skills of tested curriculum as 30% students fall in proficient level and 15% students fall in advance level. Altogether, 55% of the students, a big mass represents underperforming group. The proficient level students achieved 60% of the tested curriculum and advance level students achieved 88% of the tested curriculum. The gap in achieved curriculum between below basic level (18%) and advance level students (88%) is 70%. This indicates that inequality in the classroom is very high.

2. Students' learning is found questionable. Twenty% students in Nepali and 32% in Mathematics have achievements below the basic level. The level below basic means the students are not able to answer even very easy questions satisfactorily.
3. Learning disparity is another critical finding of the study. There is huge gap between the achievements of the students from different provinces. Students from Province 1 and Karnali Province performed lower than the others in Mathematics and Nepali subjects. In case of Province 2, the achievement in Nepali was lower than the achievement in Mathematics.
4. Comparison of the results between boys and girls shows that boys' performance was higher than that of girls in Mathematics, but in Nepali, girls performed slightly better than the boys. The gap is found consistent with the previous NASA study. The gap between boys' and girls' performance is found to be the highest in Province 2 than in other provinces.
5. Comparison of the student's performance by the type of schools shows that the students from institutional schools performed better than the students from community schools. The gap was found to be 30 scale score in Mathematics and 34 in Nepali.
6. Other studies (NASA 2011, 2013) have shown a strong correlation between the SES and student achievement. However, the result of NASA 2018 has shown that majority of Province 2 students in spite of their lower SES achieved the highest in Mathematics among the students from all 7 provinces. This is an encouraging finding that shows the possibility of students to perform better despite their low socio-economic status.
7. About 4 % of the students did not receive the textbooks even towards the end of the academic session. Students without textbooks performed lower than those with the textbooks in the class.
8. A high number of (52%) students were bullied in the school by their peers and

others. The performance of the bullied students was found to be lower than that of those who were not bullied. The gap was 31 scale score in Nepali and 22 in Mathematics.

9. The students receiving support in their study from home were found to have diverse result. The students with extra tutorial support outside the school performed better in Mathematics and Nepali than those who did not receive such support.
10. The students who were provided with regular homework and feedback on regular basis from the teachers performed better than those who were not provided homework and feedback. About 7% community school teachers never provided homework and feedback to the students in Nepali subject.
11. Parental educational qualification is one of the determining factors to boost students' performance. The mothers with Bachelor's degree and fathers with Master's degree qualification had their children perform quite better in both the subjects than those having lower educational qualification. The gap between an illiterate father and a Master's degree holding father was 35 scale score in Mathematics.
12. Parental profession is also another significant influencing factor in students' performance. Parents involved in business, teaching and government job had their children perform better than those engaged in agriculture, household work and any other profession. The gap among the parental groups was 38 scale score in Mathematics and 37 in Nepali.
13. The students who studied grade 5 at appropriate age (11years) performed better than the under or over age students. The gap was found 9 scale score in Mathematics and 11 in Nepali. It indicates that repeating the grade does not help the students perform better.
14. The students' time spent on study has direct relation with the improvement of their performance. The students who spent about 1-2 hours of after-school time in a day in the activities such as household chores, watching TV, and playing performed higher while the students spending more than 2 hours of after-school time in a day in the activities other than study, performed comparatively lower. The gap was found 17 scale score in Nepali and 16 in Mathematics.
15. The achievement and gap related results of NASA 2018 are quite similar to the results of NASA 2012 and 2015. The consistent recurring results not only proves the reliability of the NASA study, but also indicates that interventions of educational system were not effective in improving the quality of learning in the school level, urges to review the plan and policies from the quality concerns.

7.5 Conclusion

The prime objective of the national assessment is to inform the policy makers for setting realistic agenda and ensuring implementation, monitoring and evaluation of educational policies. The policies have focused on curriculum, pedagogy, teaching learning practice and assessment. Using learning assessment result for betterment of policies and their implementation is not a straightforward activity; it demands rigorous discussion at various levels of policy making and implementation with evidence to support it.

National Assessment of Student Achievement results are indicative of several crucial issues that warrant concerted efforts to improve entire teaching learning practices with a focus on the classroom delivery. Efforts need to be made by all stakeholders and the entire government system from their respective areas to ensure the placement of all students in minimal acceptance level (level 1). No one should be left behind the learning level 1 so that further progression from their level can also be ensured.

There exists a wide gap in the learning achievement of the students from among different province, level, district, school type, language (Nepali), and socio-economic status. The students are lagging behind in learning many of basic mathematical concepts and the similar situation is noted in Nepali. Differentiated instruction, individual support and learning ensuring strategies are some of the measures that can be applied in the classroom teaching learning practice.

The teachers are the key actors to improve the classroom teaching learning practice. They are expected to be conversant with different strategies to facilitate learning and teaching that should be meaningfully used to enhance students' learning. Similarly, teacher capacity development program/training (pre-service as well as in-service) should have a balanced focus on the content as well as pedagogy. A total of 32 % of grade 5 students being below the basic level in Mathematics and 20% in Nepali is quite revealing. Easy tasks like recognizing number, performing basic operations in Mathematics and simple text comprehension as well as dictation in Nepali are such simple areas of learning where the students have been lagging behind. This further raises concerns toward the teaching learning being practised in our schools.

There are disparities in learning achievement between different provinces. Province 1 and Karnali province students, for instance, have low performances than their counterparts in other provinces. In the case of province 2, performance in Mathematics was found better than in Nepali. This suggests that provincial government and local

governments have to make contextually strategic policies to reduce learning disparity. Monitoring learning on regular basis and providing adequate support to schools (financial as well as physical) can ensure better performance. Provincial governments, for instance, can develop their own contextual policies and programs for improving learning status of the province.

Students' learning is influenced by various factors and is contingent on a number of elements. The collected information from the students in this study informs that assigning homework and providing feedback, utilization of leisure in school, students' engagement in group work, no bullying in the school, and providing more time for study at home are the favourable factors to improve performance in both the subjects. In particular, home language has a visible influence on Nepali and taking support from extra tuition classes in Mathematics has been found to be meaningful. Parents' occupation and education have influenced student's achievement differently in different subjects. Parents in teaching and business profession have better learning students as they can provide meaningful support and multiple opportunities for their children. In such a case, the schools need to focus on students from low socio-economic status and the disadvantaged groups in order to ensure their equitable access to quality education.

The study findings reveal that there are a number of issues to be addressed to improve classroom learning. The urge is there for all the stakeholders to be aware about required change in the existing classroom practices. Local governments can set learning achievement targets and take appropriate actions to meet them. These strategies are possible through planned policy change that aims at transforming the schools into real learning organizations.

7.6 Recommendations

1. **Huge mass of students is at the underperforming level:** As the study indicates, 32 out of 100 students fall below basic level (Pre-basic) in Mathematics achieving only 5% of the tested curriculum and the basic level (level 1, about 40%) students have achieved only 28% of the tested curriculum. More than 70% students have achieved only below 28% of the tested curriculum in Mathematics indicating a huge mass of students underperforming in this subject. The proficient level (level 2, 24%) students achieved 62% of the tested curriculum and advance level (level 3, 4%) students achieved 96% of the tested curriculum. It seems that only 29% of the students have adequate knowledge and skills in Mathematics curriculum. The gap in the achievement of curriculum between below basic level (5%) and

advance level students (96%) is 91% indicating remarkably high inequality in the classroom. Similarly, in Nepali, 20% students who are at below basic level (Pre-basic) achieved only 18% of the tested curriculum and 35% students achieved only 38% of the tested curriculum in Nepali, who are at basic level (level 1). As 30% students fall in proficient level and 15% in advance level, altogether 45 of 100 students have adequate knowledge and skills of the tested curriculum in Nepali. This shows that 55% of the students, a big mass of the students in Nepali, represent the underperforming group. And 45 of 100 students have adequate knowledge and skills of the tested curriculum as 30% students fall in proficient level and 15% in advance level. Altogether, 55% of the students, a big mass of the students in Nepali, represent the underperforming group. The proficient level students achieved 60% of the tested curriculum and the advance level students achieved 88% of the tested curriculum. The gap in the achievement of the curriculum between below basic level (18%) and advance level students (88%) is 70% indicating high inequality in the classroom.

Recommendation: As basic level is assumed to be a minimum competency level, a campaign of "no child should be left behind basic level" should be initiated effectively to develop minimum competency level in the students. Teachers should provide the students with many opportunities to learn in many ways and through various means. Existing mis-match between the written curriculum and achieved curriculum urges the need to review the national curriculum, teaching methods, teacher motivation system, learning environment and the evaluation system.

2. **Huge gap between the provinces and districts has been noticed. Such gap increases disparity in learning achievement among the groups of students.**

   Recommendation: All the community schools should provide equal opportunity to the students for learning. A minimum standard for physical infrastructure, learning opportunities, resources, incentives and retention of good teachers under teacher management should be set to bring uniformity in the achievement level of the students. Retention of good teachers has relation with increased learning achievement of students. Learning difficulties of students in all schools should be identified and then addressed by remedial teaching. Regular follow up support and monitoring mechanism should be strengthened to enhance learning.

3. **Students can perform better irrespective of their SES and home language:** The Socio-Economic Status of student's family has low effect in Mathematics and medium effect in Nepali Language. Many students have performed better

despite their unfavourable low socio-economic status; this indicates that the socio-economic background of the students does not exclusively determine the learning achievement. Province 2 has majority of students with low SES, but they have achieved the higher position in Mathematics compared to other provinces. And similar is the situation with Nepali language as well.

Recommendation: As students can perform better irrespective of their SES and home language, focus on the study and practice on the part of student is a principal measure to boost their learning achievement. In all community schools, minimum learning materials, library facility and students' clubs, numeracy promotion program and reading programs should be made available to promote students' performance.

4. **There is a minimum gap in learning achievement between boys and girls.**

   Recommendation: Reduced and minimum gap between boys and girls should be maintained with more focus on providing girls with equal opportunity. The existing gap can best be addressed through affirmative action such as scholarship, girls friendly environment, and receptive teacher behaviour.

5. **Wide gap between the type of schools has been noticed:** There is a wide gap in learning achievement between community and institutional schools. The students from institutional schools have out-performed the community schools with a gap of 30 scale score in Mathematics and 34 in Nepali.

   Recommendation: Upgrading of community schools to increase their academic performance should be initiated as a regular targeted intervention by the government. Co-curricular activities should include Mathematics or number games, reading and writing, literature, contemporary issues and challenges as part of the curriculum for all community schools. Remedial teaching and individualized instruction should also be implemented for the targeted students.

6. **Bullying is affecting learning in the schools:** A high number (53%) of the students were bullied in the school by their peers and others. The performance of the bullied students was found lower than that of those who were not bullied. The gap is 31 scale score in Nepali and 22 in Mathematics.

   Recommendation: Local governments together with the schools should regularly monitor the school to maintain peace, discipline and regulations as well as cohesive environment among the students. "No bullying" movement should be initiated by the school in consultation with and involvement of the local community. Active

focal person in the school, psycho-social counselling service and child club activity against bullying are the other activities that the school can provide to minimize bullying in the schools.

7. **Feedback to the students in homework has positive relationship with learning achievement**: The students who were provided with regular homework and feedback from the teacher performed better than those who were not provided homework and feedback in both subjects. About 7% community school teachers never provided homework and feedback to the students in Nepali subject.

   Recommendation: Teacher's performance evaluation should be strongly linked with students' performance. Providing homework and feedback in certain subjects (perceived to be difficult) in the form of scaffolding in the schools, continuous assessment service, criterion-based assessment for teaching-learning, formative evaluation system and regular communication with parents about their children's performance, some instructional activities directly influence students' learning achievement.

8. **Students at right age performed better:** Students studying in grade 5 at appropriate age (11 years) performed better than the under or over age students. The gap was found 9 scale score in Mathematics and 11 in Nepali. Similar trend was noticed in the previous study as well. This further means that age wise grade or grade wise age or age appropriate enrollment ensures higher learning achievement.

   Recommendation: Net enrollment practice at the basic level education should be maintained and increased by admitting appropriate age students in the appropriate grades in schools. Age appropriate level of students can learn smoothly with their peers of equal age.

9. **After-school activities determine student learning achievement:** Students' involvement in study during their after-school time has a relation with their performance. The students' who spent about one to two hours' after-school time in the activities like household chores, watching TV, and playing performed better. On the other hand, the students who spent more than two hours of their after-school time in those activities other than study performed relatively lower. The gap was found 17 scale score in Nepali and 16 in Mathematics.

   Recommendation: All schools should make their parents aware of the influence of children's involvement in study with their performance. Besides being involved in entertainment and household chores, the students should be encouraged to study not only the textbooks based on the curriculum but also the literature books,

reference books and other extra-curricular materials as well.

10. **Many students do not have a sense of Mathematics:** In Mathematics, students below minimum learning level have quite limited knowledge and skills in Mathematics. Around 50% of them do not have any sense of reading and writing numbers and number operation. With this limited knowledge and skills, they cannot calculate and solve Mathematical problems. Some of them are able to choose correct answers when options are given in the MCQ items. Mostly, they are unable to perform any Mathematical subjective calculations independently.

    Recommendation: Students experiencing difficulty with learning the numbers and Mathematical calculation should be provided with learning opportunity to developing minimum Mathematical skills and competencies. The teachers should be made accountable to their students' performance. An action needs to be taken if the teachers fail to develop their students' minimum level of competencies. Not only the teachers' performance but also the performance of the Principals of such schools with below minimum learning level of students should be linked up with their career path. The schools with such lower achievement level should be made responsible to develop short term and long-term learning improvement strategies and actions. To monitor the progress of the implementation of the planned strategies, regular standardized tests should be administered and local level monitoring and supervision mechanism should be ensured. Besides, the primary level teachers should be equipped with national level short-term practical, school based professional training to advance their content knowledge and pedagogical skills.

11. **More than half of the students are struggling with the simple Mathematical calculations:** Basic level (level 1) students have superficial knowledge and skills in most of the Mathematical contents. However, they are struggling with calculations. They are able to identify the ordered pair of a point, square pattern in dots, the sum of decimal numbers, the place of a digit in numbers, the relation of kilogram and gram and the numbers, mixed fraction and type of angles. But they also have a limited knowledge of formula for volume and area, they cannot estimate the angle shown in figure and cannot identify the relation between decimal and fraction. They can read the table and bar graphs to take simple information but cannot draw conclusion by comparing the data. They can solve very simple problems of unitary methods, subtract small same denominator fractions, round numbers in the nearest tenth of a decimal number. They can also recognize limited square numbers and express Mathematical sentence in Mathematical language to calculate. They can

also subtract a univariate one-degree Algebraic term from another, find the value of x in one variable equation and subjective, can simplify Algebraic expression (univariate) in Algebra. However, they cannot perform grade level Mathematical calculations independently.

Recommendation: The knowledge of basic level contents is essential to be able to grasp the Mathematical content. Therefore, the teachers should focus on the development of basic level contents among all students. In addition to these skills, the students should be prepared to solve simple grade appropriate Mathematical calculations independently. To improve learning achievement, an emphasis should be laid on the underperforming students instantly through the application of problem-solving method.

12. **In Nepali language, the students below level 1 can read only a few words or sentences but they cannot write the sentences independently.**

    Recommendation: Such underperforming (below level 1) students should be involved more on the activities such as reading and writing words and sentences and describing familiar events independently. Language teaching should focus on meaningful reading and comprehension exercises rather than on reciting the paragraphs in the textbook and rote learning the answers.

13. **Decreasing trend of students' performance in Mathematics:**

    Recommendation: A diagnostic study about the challenges in teaching and learning culture should be carried out. The factors responsible for reduced learning achievement should be identified and disseminated. The involvement of parents and community members should be ensured in making the schools accountable for their student's low performance level.

14. **The achievement and gap related results of NASA 2018 are quite similar to the results of NASA 2012 and 2015 :** The consistent recurring results not only proves the reliability of the NASA study, but also indicates that interventions were not sufficient in improving the quality of learning in the school level.

    Recommendation: MOEST should review the existing plan and policies from the quality concerns.

15. **Overall Recommendation:** As a final step of National Assessment, MOEST should initiate Post-NASA policy review and intervention plan at the national level, sub-national level, and implementing agency level of the education system.

REFERENCES

ACARA. (2017). *NAP Sample Assessment Science Literacy Technical Report:* Sydney: Australian Curriculum Assessment and Reporting Authority (ACARA).

An, Xinming and Yung, Yiu-Fai. (2014). *Paper SAS364-2014 Item Response Theory: What it is and how you can Use the IRT Procedure to Apply it.* Retrieved from https://support.sas.com/resources/papers/proceedings14/SAS364-2014.pdf.

Baker, F.B. (2001). *The Basics of Item Response Theory*. ERIC Clearinghouse on Assessment and Evaluation. Retieved from http://echo.edres.org:8080/irt/baker/final.pdf.

CDC. (2013). *Lower Secondary Curriculum.* Sanothimi: CDC (Curriculum Development Centre).

CDC. (2014). *National Curriculum Framework.* Sanothimi: Author.

Chalmers, P. P. (2012). A Multidimensional Item Response Theory Package for the R Environment. *Journal of Statistical Software*, Volume 48, Issue 6. http://www.jstatsoft.org/ York University.

CMEC (2014). *PCAP 2013: Report on the Pan-Canadian Assessment of Science, Reading, and Mathematics.* Toronto: CMEC.

Cohen, L. and Holliday, M. (1996) *Practical Statistics for Students*. London: Paul Chapman.

ERO. (2013). *Report of National Assessment of Student Achievement 2011, Grade 8.* Sanothimi: Education Review Office.

ERO. (2015). *Report of National Assessment of Student Achievement 2012, Grade 3 and 5*. Sanothimi: Education Review Office.

ERO. (2015a). *Report of National Assessment of Student Achievement 2013, Grade 8.* Sanothimi: Education Review Office.

ERO. (2016). *Report of National Assessment of Student Achievement 2015, Grade 3 and 5.* Sanothimi: Education Review Office.

ERO. (2017). *NASA 2017: Assessment Framework for Grade 5 in Mathematics and Nepali.* Sanothimi: Education Review Office.

Fennema, E., & Sherman, J.A. (1976). Fennema-Sherman Mathematics Attitudes Scales: Instruments designed to measure attitudes toward the learning of mathematics by females and males. *Journal for Research in Mathematics Education, 7,* 324-326. doi: 10.2307/748467.

Greaney. V. and Kellaghan, T. (2007). *Assessing National Achievement Levels in*

Education [edited]. Washington DC: The World Bank.

IEA. (2016). *Methods and Procedures in TIMSS 2015*. Martin, M.O., Mullis, V. S. and Hooper, M. (edited. Boston: International Association for the Evaluation of Educational (IEA).

Kathryn O'Grady, K. and Houme, K. (2015). *PCAP 2013: Technical Report.* Toronto: Council of Ministers of Education, Canada.

Laukaitytė, Inga. (2016). *Statistical Modeling in International Large-scale Assessments.* Department of Statistics Umeå School of Business and Economics.

Le, Dai-Trang. (2013). *Applying Item Response Theory Modeling in Educational Research.* Owa State University.

Lewis, T. (2015). *Replication Techniques for Variance Approximation, Paper 2601-2015.* University of Maryland.

Masters, G. M. and Wright, B. D. (1997). The Partial Credit Model. In W. J. Van der Linden and R. K. Hambleton (edit), Handbook of Modern Item Response Theory, pp 110-122. New York, Springer.

Meinck, S. (2015). *Computing Sampling Weights in Large-scale Assessments in Education. Survey Insights: Methods from the Field, Weighting: Practical Issues and 'How to' Approach.* Retrieved from http://surveyinsights.org/?p=5353.

Murphy, M. & Schulz, W. (2006). *Sampling for National Surveys in Education.* Victoria, Australia: Australian Council for Educational Research (ACER).

Murphy, Paud; Greaney, Vincent; Lockheed, Marlaine E.; Rojas, Carlos [editors] (1996). *National assessments: testing the system (English). EDI learning resources series.* Washington, D.C: The World Bank http://documents.worldbank.org/curated/en/920211468743345819/National-assessments-testing-the-system

OECD. (2009). *PISA Data Analysis Manual*, SPSS, Second Edition. Retrieved from https://www.ipi.lu.lv/fileadmin/_migrated/content_uploads/Datu_analizes_rk.pdf.

OECD. (2016). *PISA 2015: Technical Report*. Retrieved from http://www.oecd.org/pisa/data/2015-technical-report/.

OECD. (2017). PISA 2015 Assessment and Analytical Framework (Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving). Retrieved from www.oecd-ilibrary.org/education/pisa-2015-assessment-and....

Pant L. B., Singh, G. B. and Poudel, L. N. (2016). *NASA 2017: Assessment Framework for Grade 8 in Mathematics, Nepali and Science.* Sanothimi: Education Review Office.

Sudal, L. A. and Sudol, C. (2010). *Analysing Test Items: Using Item Response Theory to Validate Assessments.* Proceedings of the 41st ACM technical symposium on Computer science education, pp 436-440. Retrieved from http://ims.mii.lt/ims/konferenciju_medziaga/SIGCSE%2710/docs/p436.pdf. Verhelst, N.D. &. Verstralen, H.H.F.M. (2008). Some Considerations on the Partial Credit Model. *Psicologica (2008),* 29, 229-254.

Von Davier, M, Gonzalez, E, & Mislevy, R. J. (2009). What are plausible values and why are they useful? In M Von Davier & D Hastedt (Eds.), *IERI Monograph Series: Issues and methodologies in large scale assessments (Vol. 2, pp. 9–36).* Hamburg, Germany: IERI Institute.

Xie, Qingshu, Zhang, MacroSys Ting (2017). *Analyses Using Achievement Levels Based on Plausible Values AIR* - NAEP Working Paper #2017-02 NCES Data R Project Series #01 American Institutes for Research.

Yamamoto, K., & Kulick, E. (2000). Scaling Methodology and Procedures for the TIMSS Mathematics and Science Scales. In M.O. Martin, K.D. Gregory, & S.E. Stemler (Eds.), *TIMSS 1999 technical report.* Chestnut Hill, MA: Boston College.

Yansaneh, I. S. (2003). *Construction and use of sample weights.* United Nations Secretariate.

Zieky, M. & Perie, M. (2006). A Primer on Setting Cut Scores on Tests of Educational Achievement. Educational Testing Service (ETS).

Appendix

Mathematics subject Outputs

1. Mathematics Item Analysis parameters from Joint run (Set 1, 2, and 3 combined)

ConQuest: Generalised Item Response Modelling Software Mon Apr 22 14:12 2019

GENERALISED ITEM ANALYSIS

Group All Students

Item	N	Facility	Item-Rest Cor	Item-Total Cor	Wghtd MNSQ	Avg Delta
item:1 (M1SRQ1)	4484	68.69	0.31	0.36	1.03	-1.16
item:2 (M1SRQ2)	3990	58.55	0.32	0.37	1.05	-0.56
item:3 (M1SRQ3)	4335	68.97	0.37	0.42	0.98	-1.16
item:4 (M1SRQ4)	8892	37.58	0.29	0.35	1.07	0.34
item:5 (M1SRQ5)	4502	48.67	0.34	0.39	1.04	-0.16
item:6 (M1SRQ6)	4502	45.36	0.36	0.41	1.03	0.00
item:7 (M1SRQ7)	4516	80.14	0.33	0.37	0.97	-1.87
item:8 (M1SRQ8)	4497	43.32	0.40	0.45	0.99	0.10
item:9 (M1SRQ9)	8631	46.50	0.37	0.43	0.99	-0.10
item:10 (M1SRQ10)	4480	39.40	0.44	0.49	0.95	0.29
item:11 (M1SRQ11)	4430	36.70	0.35	0.40	1.03	0.42
item:12 (M1SRQ12)	4493	46.90	0.48	0.52	0.91	-0.08
item:13 (M1SRQ13)	4391	49.37	0.42	0.46	0.96	-0.19
item:14 (M1SRQ14)	4455	54.73	0.42	0.47	0.97	-0.45
item:15 (M1SRQ15)	4532	62.69	0.42	0.47	0.95	-0.86
item:16 (M1SRQ16)	13211	39.30	0.35	0.40	1.01	0.26
item:17 (M1SRQ17)	4505	51.94	0.26	0.31	1.12	-0.33
item:18 (M1SRQ18)	8543	49.87	0.32	0.37	1.04	-0.20
item:19 (M1SRQ19)	8896	49.53	0.31	0.37	1.05	-0.20
item:20 (M1SRQ20)	4504	54.88	0.43	0.48	0.95	-0.46
item:21 (M1SRQ21)	8818	55.94	0.41	0.46	0.96	-0.54
item:22 (M1SRQ22)	4219	45.86	0.34	0.39	1.04	0.02
item:23 (M1SRQ23)	4399	49.58	0.35	0.40	1.03	-0.21
item:24 (M1SRQ24)	12904	71.59	0.33	0.38	0.98	-1.34
item:25 (M1CRQ25)	3930	49.62	0.46	0.50	0.93	-0.18
item:26 (M1CRQ26)	3876	32.20	0.40	0.45	0.98	0.71

Item	N	Facility	Item-Rest Cor	Item-Total Cor	Wghtd MNSQ	Avg Delta
item:27 (M1CRQ27)	3880	32.53	0.48	0.52	0.90	0.73
item:28 (M1CRQ28)	3652	45.21	0.35	0.39	1.03	0.11
item:29 (M1CRQ29)	3676	35.32	0.55	0.61	1.04	0.42
item:30 (M1CRQ30)	3592	28.95	0.52	0.55	0.87	0.94
item:31 (M1CRQ31)	3845	55.97	0.50	0.54	0.89	-0.43
item:32 (M1CRQ32)	3278	16.15	0.42	0.47	1.07	1.34
item:33 (M1CRQ33)	6660	45.74	0.53	0.57	0.86	0.08
item:34 (M1CRQ34)	3268	47.83	0.55	0.59	0.86	0.03
item:35 (M1CRQ35)	10246	28.50	0.51	0.58	1.05	0.59
item:36 (M1CRQ36)	3132	51.50	0.46	0.50	0.94	-0.12
item:37 (M1CRQ37)	6753	29.19	0.48	0.55	1.02	0.76
item:38 (M1CRQ38)	3189	15.91	0.43	0.47	1.02	1.50
item:39 (M1CRQ39)	6233	38.60	0.45	0.49	0.93	0.41
item:40 (M1CRQ40)	3212	43.07	0.48	0.55	1.11	0.19
item:41 (M1CRQ41a)	7922	74.64	0.38	0.42	0.94	-1.47
item:42 (M1CRQ41b)	7827	77.51	0.34	0.38	0.96	-1.65
item:43 (M1CRQ42a)	3649	49.99	0.47	0.51	0.93	-0.15
item:44 (M1CRQ42b)	3425	11.53	0.18	0.21	1.07	2.22
item:45 (M1CRQ43)	2778	28.06	0.52	0.57	1.05	0.77
item:46 (M2SRQ1)	4224	40.27	0.33	0.38	1.02	0.24
item:47 (M2SRQ2)	4542	80.34	0.31	0.36	0.98	-1.87
item:50 (M2SRQ5)	4412	59.77	0.31	0.37	1.03	-0.70
item:51 (M2SRQ6)	4429	33.05	0.28	0.33	1.05	0.59
item:52 (M2SRQ7)	4415	55.95	0.28	0.34	1.05	-0.52
item:53 (M2SRQ8)	8788	49.68	0.23	0.29	1.11	-0.26
item:54 (M2SRQ9)	4573	79.55	0.29	0.33	0.99	-1.82
item:55 (M2SRQ10)	4481	37.45	0.35	0.41	1.00	0.35
item:56 (M2SRQ11)	4450	66.72	0.33	0.38	0.99	-1.06
item:57 (M2SRQ12)	4494	40.01	0.26	0.32	1.07	0.22
item:58 (M2SRQ13)	4408	41.81	0.47	0.52	0.90	0.13
item:59 (M2SRQ14)	4495	67.19	0.34	0.39	0.99	-1.09
item:60 (M2SRQ15)	4187	63.89	0.48	0.53	0.88	-0.88

Item	N	Facility	Item-Rest Cor	Item-Total Cor	Wghtd MNSQ	Avg Delta
item:61 (M2SRQ16)	4436	50.45	0.42	0.47	0.94	-0.27
item:62 (M2SRQ17)	4469	46.90	0.47	0.51	0.90	-0.11
item:63 (M2SRQ18)	4454	19.62	0.23	0.27	1.05	1.39
item:64 (M2SRQ19)	8838	35.44	0.31	0.37	1.03	0.43
item:65 (M2CRQ20)	8861	49.99	0.37	0.42	0.98	-0.27
item:66 (M2SRQ23)	4387	48.30	0.32	0.38	1.03	-0.16
item:67 (M2SRQ25)	4438	36.50	0.45	0.50	0.91	0.40
item:68 (M2SRQ26)	4377	42.22	0.15	0.21	1.18	0.13
item:69 (M2CRQ28)	8073	38.85	0.40	0.48	1.15	0.20
item:70 (M2CRQ29)	5360	30.86	0.42	0.46	0.94	0.83
item:71 (M2CRQ30)	3941	59.69	0.31	0.41	1.32	-0.53
item:72 (M2CRQ31)	3646	40.92	0.48	0.53	0.90	0.27
item:73 (M2CRQ32)	3765	24.12	0.41	0.47	1.10	0.88
item:74 (M2SRQ33)	3332	19.69	0.42	0.46	0.91	1.46
item:75 (M2CRQ34)	3140	32.90	0.47	0.51	0.91	0.69
item:76 (M2CRQ35)	3466	20.20	0.45	0.51	0.99	1.16
item:77 (M2CRQ36)	3348	47.07	0.45	0.50	0.93	-0.04
item:78 (M2CRQ38)	3024	19.20	0.48	0.53	0.99	1.10
item:79 (M2CRQ40a)	3856	78.19	0.35	0.39	0.95	-1.69
item:80 (M2CRQ40b)	3613	43.90	0.35	0.40	1.01	0.09
item:81 (M2CRQ42)	7166	60.91	0.47	0.52	0.89	-0.74
item:82 (M2CRQ43)	3661	43.32	0.46	0.55	1.12	0.01
item:83 (M3SRQ2)	3763	51.37	0.34	0.40	1.01	-0.27
item:84 (M3SRQ3)	3580	53.74	0.29	0.35	1.05	-0.42
item:85 (M3SRQ4)	4336	56.20	0.36	0.42	0.99	-0.58
item:86 (M3SRQ5)	4277	60.72	0.35	0.41	0.99	-0.79
item:87 (M3SRQ6)	4204	42.65	0.39	0.45	0.97	0.07
item:88 (M3SRQ11)	4338	36.75	0.29	0.36	1.05	0.35
item:89 (M3SRQ12)	4346	55.48	0.37	0.43	0.99	-0.55
item:90 (M3SRQ13)	4283	65.44	0.34	0.39	0.98	-1.04
item:91 (M3SRQ14)	4363	75.84	0.28	0.33	1.01	-1.63
item:92 (M3SRQ15)	4410	72.47	0.33	0.39	0.98	-1.43

Item	N	Facility	Item-Rest Cor	Item-Total Cor	Wghtd MNSQ	Avg Delta
item:93 (M3SRQ16)	4344	37.06	0.10	0.17	1.22	0.33
item:94 (M3SRQ19)	4281	58.35	0.19	0.26	1.12	-0.69
item:95 (M3SRQ21)	4325	9.83	0.11	0.15	1.08	2.25
item:96 (M3SRQ23)	4233	40.89	0.37	0.43	0.99	0.13
item:97 (M3CRQ27)	3269	31.14	0.38	0.45	1.11	0.73
item:98 (M3CRQ28)	3225	33.02	0.44	0.48	0.94	0.64
item:99 (M3CRQ29)	3261	25.38	0.51	0.57	0.96	0.87
item:100 (M3CRQ30)	2902	40.87	0.50	0.54	0.89	0.28
item:101 (M3CRQ32)	3089	22.39	0.45	0.51	1.01	1.03
item:102 (M3CRQ33)	2940	46.80	0.33	0.39	1.03	-0.03
item:103 (M3CRQ34)	3106	39.83	0.56	0.59	0.84	0.27
item:104 (M3CRQ35)	2820	35.64	0.59	0.62	0.81	0.51
item:105 (M3CRQ37)	3041	41.96	0.53	0.57	0.86	0.19
item:106 (M3CRQ38)	2904	20.11	0.50	0.56	0.99	1.01
item:107 (M3CRQ40)	2971	16.06	0.13	0.17	1.14	1.72
item:108 (M3CRQ41a)	3000	71.60	0.40	0.45	0.93	-1.23
item:109 (M3CRQ41b)	2702	33.31	0.39	0.43	0.98	0.67
item:110 (M3CRQ43)	3084	32.23	0.46	0.51	0.92	0.67

In this analysis 67.00% of the data are missing.

The following results are scaled to assume that a single response was provided for each item.

N 14052
Mean 54.61
Standard Deviation 25.20
Variance 635.18
Skewness 0.32
Kurtosis -0.58
Standard error of mean 0.21

2. The replicate module code of estimating the error of National Mean in Mathematics:

```
Insert                                                    file='C:\Program
Files\IBM\SPSS\Statistics\Addins\Replicates\Macros\mcr_SE_PV.sps'.
OMSEND.
set mprint=yes.

/* Name of macro that was called, followed by the arguments */.
!UNIVAR_PV
      nrep = 350 /
      stat = Mean  /
      dep = MSSPV1 MSSPV2 MSSPV3 MSSPV4 MSSPV5 /
      grp = NOGRP /
      wgt = W_STU /
      rwgt = RWGT /
      cons = 1 /
      PSU = Sch_code  /
    infile = 'd:\nasa\2018\ft\data analysis\report writing 2018\step by step analysis\7.
reporting data\g5math_pl.sav' /.
```

3. A comparative presentation of National Mean and province wise mean and standard error from Replicate module in mathematics

Province	Mean	SE	n_stu	n_sch	Upper	Lower	National mean	CI
National	500	1.293	689836	14174	700.000	494.93	500	5.07
Prov 1	494	2.762	2831	141	504.35	482.70	500	10.83
Prov 2	521	4.017	1739	81	536.50	505.01	500	15.75
Prov 3	505	2.52	3944	193	514.64	494.88	500	9.88
Gandaki	503	4.593	1599	86	521.19	485.19	500	18.00
Prov 5	486	2.603	1943	89	495.81	475.40	500	10.21
Karnali	496	4.847	1071	56	514.65	476.66	500	19.00
Sudur Pashchim	500	3.831	1047	54	515.19	485.16	500	15.02

4. Standard Error of Percentage of students in various levels. in Mathematics

Proficiency level	% of students	SE	N_cases	NU_cases	NU_psu
Level 0: Pre-basic	32.166	1.037	220496.8	4594	567
Level 1: Basic	39.577	0.833	272024.3	5609	677
Level 2: Proficient	24.006	0.837	163676.9	3320	595
Level 3: Advance	4.252	0.473	28300.73	570	187

5. Gender wise National achievement in Mathematics

Gender	Mean	SE	n_stu	CI
Boy	501	1.436	6855	5.627
Girl	499	1.443	6978	5.657
Missing	521	1.982	341	

6. Ethnicity wise National achievement in Mathematics

Ethnicity	Mean	SE	n_stu	N_cases	CI
Brahman/Chhetri	504	1.810	5032	240716.00	7.10
Janjati	496	1.810	5018	250417.00	7.09
Dalit	491	2.152	1892	95288.00	8.44
Others	513	2.398	2232	103415.00	9.40

7. National Mean score by age group *in Mathematics*

Age	Mean	SE	n_stu	CI
9 years or belw	498	3.204	562	12.560
10 years	501	1.782	2682	6.984
11 years	501	1.580	4545	6.192
12 years	500	1.753	3738	6.873
13 years	494	2.513	1395	9.853
14 years or above	492	2.855	604	11.192

8. National mean score by mother's education in Mathematics

Mothers education	Mean	SE	n_stu	CI
Illitrate	493.09	1.774	4095	6.952
Literate	499.43	1.816	3484	7.117
Grade 8	498.68	1.812	2502	7.103
grade 10	512.08	2.077	1653	8.143
Grade 12	516.06	2.432	991	9.533
Bachelors	521.57	4.012	408	15.726
Masters or above	520.076	5.913	194	23.178

9. National mean score by Father's education *in Mathematics*

Fathers education	Mean	SE	n_stu	CI
Illitrate	491.74	1.906	2365	7.471
Literate	495.31	1.925	2993	7.546
Grade 8	497.43	1.784	3015	6.995
grade 10	506.51	1.817	2356	7.122
Grade 12	513.15	2.075	1577	8.133
Bachelors	522.00	2.883	583	11.302
Masters or above	527.19	3.640	412	14.269

10. National mean score by Mother's profession *in Mathematics*

Mothers profession	Mean	SE	n_stu	CI
Agriculture and household work	498	1.473	8374	5.775
Household work only	504	2.193	2420	8.598
Work in other's house	487	2.766	435	10.843
labour	502	4.183	219	16.398
Foreign country	496	3.505	290	13.739
Teaching	525	3.934	353	15.420
Business	513	2.451	742	9.607
Government job	520	4.390	211	17.208
Other	515	3.272	307	12.825

11. National mean score by father's profession *in Mathematics*

Fathers occupation	Mean	SE	n_stu	CI
Agriculture and household work	494.70	1.672	4690	6.554
Household work only	486.99	2.861	409	11.216
Work in other's house	487.72	2.555	309	10.014
labour	498.38	3.385	1041	13.270
Foreign country	504.36	1.805	3672	7.076
teaching	521.96	3.346	483	13.116
Business	509.88	2.281	2202	8.941
Government job	513.99	3.106	1146	12.176
Other	516.15	2.623	1000	10.284
missing	477.64	2.149	530	483.925

12. National mean scores by support after school to students in Mathematics

Support to the students	Mean	SE	n_stu	CI
1 father	501.43	1.8668	342	7.318
2 mother	499.52	2.2038	6390	8.639
3 brother/sister	498.93	1.5214	1733	5.964
4 tuition	511.21	3.0172	3137	11.827
5 friend	495.40	3.3265	172	13.040
6 any other	502.52	6.0789	260	23.829
7 none	504.33	5.4709	1200	21.446

13. National mean score by home language *in Mathematics*

Language	Mean	SE	n_stu	CI
Nepali	499.70	1.510	8975	5.919
Other	502.23	1.865	5199	7.312

14. National mean score by type of schools *in Mathematics*

Type of school	Mean	SE	n_stu	CI
community	492.95	1.441	10126	5.647
Institutional	522.84	2.261	4048	8.864

15. Mean score by involvement in after- school activity *in Mathematics*

Time given	a. Involve in TV, internet, computer	b. play and talk	c. Home chores	d. Homework and study	e. Work for wage	f. Read other books
Not given time	493	495	494	488	505	491
Less than one hour	506	502	503	491	494	504
Up to 3 hours	499	503	503	505	489	507
More than 3 hours	493	489	494	508	489	496

16. Disaggregated mean score of sample districts in Mathematics

Province	Districts	Disaggregated mean
Prov. 1	BHOJPUR	468
	DHANKUTA	487
	OKHALDHUNGA	495
	JHAPA	497
	PANCHTHAR	504
Prov. 2	MAHOTTARI	515
	RAUTAHAT	520
	SIRAHA	523
Prov. 3	SINDHULI	487
	KAVREPALANCHOK	495
	CHITWAN	513
	NUWAKOT	516
	LALITPUR	522
Prov. Gandaki	SYANGJA	497
	LAMJUNG	497
	PARBAT	511
Prov. 5	BARDIYA	472
	PYUTHAN	484
	GULMI	500
Prov. Karnali	KALIKOT	492
	SALYAN	497
	HUMLA	500
Prov. Far-western	BAJURA	490
	BAITADI	503

Nepali subject outputs

1. A comparative presentation of National Mean and province wise mean and standard error from Replicate module in Nepali

MEAN by Province								
	Mean Score	SE	N-population	N-cases	vN-School	Upper	Lower	CI
National	500	0.788	690125	14207	700	501.06	497.97	3.09
Province 1	500	2.634	140972	2795	143	504.95	494.62	10.32
Province 2	489	2.222	79774	1813	81	493.19	484.49	8.71
Province 3	506	1.712	192370	3896	195	509.18	502.47	6.71
Gandaki Province	511	1.191	83713	1617	85	513.20	508.53	4.67
Province 5	496	0.316	84809	1914	86	496.96	495.73	1.24
Karnali Province	488	1.937	54167	1112	55	491.87	484.28	7.59
Sudur Paschim Province	491	0.782	54318	1060	55	492.57	489.50	3.07

2. Standard Error of Percentage of students in various levels. *in Nepali*

Proficiency Level distribution			
proflev	% of Students	SE	N_cases
Level 0 (Pre-basic)	20	0.704	2710
Level 1 (Basic)	35	1.010	4808
Level 2 (Proficient)	30	1.154	4169
Level 3 (Advance)	15	0.608	2198

3. Gender wise National achievement in Nepali

Gender						
	Mean Score	SE	N-cases	Upper	Lower	CI
Boys	499	1.314	6635	501.26	496.11	5.15
Girls	501	0.991	7363	502.56	498.67	3.88

4. Ethnicity wise National achievement in Nepali

Mean by Caste/Ethnicity						
	Mean Score	SE	NU_cases	Upper	Lower	CI
Bhramin/Chettri	501	0.982	6287	503.26	499.41	3.85
Janjati	500	1.376	5363	502.55	497.16	5.39
Dalit	494	1.043	1820	496.50	492.41	4.09
Others	494	3.192	737	500.03	487.52	12.51

5. National Mean score by age group in Nepali

Mean by Student Age						
Age	Mean Score	SE	N-cases	Upper	Lower	CI
Nine and Below	495	1.886	548	498.53	491.13	7.39
10 Years	499	1.748	2427	501.96	495.11	6.85
11 Years	503	1.146	4656	505.18	500.69	4.49
12 Years	502	1.212	4075	504.33	499.58	4.75
13 Years	495	3.403	1409	501.21	487.88	13.34
Fourteen and Above	492	2.141	637	496.16	487.76	8.39
Missing	482	3.313	455	487.99	475.01	12.99

6. National mean score by mother's education in Nepali

Mother Education						
	Mean Score	SE	NU_cases	Upper	Lower	CI
Illiterate	491	0.782	4097	492.17	489.11	3.07
Literate	498	1.462	3278	500.48	494.75	5.73
Grade 8	503	2.703	2545	508.51	497.92	10.60
Grade 10	514	1.367	1601	517.08	511.72	5.36
Grade 12	516	4.264	1100	523.98	507.26	16.72
Bachelor	525	3.010	437	531.37	519.57	11.80
Master	527	4.406	180	535.52	518.25	17.27
Missing	480	1.469	969	483.37	477.62	5.76

7. National mean score by Father's education *in Nepali*

Father Education						
	Mean Score	**SE**	**N-cases**	**Upper**	**Lower**	**CI**
Illiterate	487	1.317	2275	489.92	484.76	5.16
Literate	496	2.652	2942	500.73	490.33	10.40
Grade 8	498	1.369	3081	500.56	495.19	5.37
Grade 10	510	1.132	2248	512.61	508.17	4.44
Grade 12	511	1.187	1632	513.51	508.86	4.65
Bachelor	519	1.891	612	523.18	515.77	7.41
Master	527	2.882	428	532.25	520.95	11.30
Missing	483	1.673	989	486.47	479.91	6.56

8. National mean score by Mother's profession in Nepali

Mother Occupation						
	Mean Score	**Std. Error of Mean**	**N-cases**	**Upper**	**Lower**	**CI**
Household and agriculture	497	0.525	8568	497.76	495.70	2.06
Household work only	506	1.072	2353	508.39	504.19	4.20
Work in other's house	485	2.623	338	490.53	480.25	10.28
Labour	495	3.302	193	501.34	488.39	12.95
Work in foreign country	505	3.043	288	511.29	499.36	11.93
Teaching	528	2.821	376	533.17	522.12	11.06
Business	518	1.901	699	521.77	514.32	7.45
Government job	518	3.445	238	525.09	511.59	13.50
Other	491	1.474	1154	493.75	487.98	5.78

9. National mean score by father's profession **in Nepali**

Father Occupation						
	Mean Score	**Std. Error of Mean**	**N-cases**	**Upper**	**Lower**	**CI**
Household and agriculture	491	0.684	4767	492.61	489.92	2.68
Household work only	494	2.004	549	497.49	489.64	7.85
Work in other's house	484	2.034	483	488.26	480.29	7.97

Father Occupation						
	Mean Score	Std. Error of Mean	N-cases	Upper	Lower	CI
Labour	498	1.766	760	501.21	494.28	6.92
Work in foreign country	505	0.882	3218	506.67	503.22	3.46
Teaching	519	2.701	417	524.13	513.54	10.59
Business	515	1.430	1299	517.89	512.29	5.60
Government job	520	1.782	806	523.94	516.95	6.99
Other	497	1.185	1908	499.11	494.46	4.65

10. National mean scores by support after school to students in Nepali

Support taking						
	Mean Score	SE	N-cases	Upper	Lower	CI
Father	497	1.952	2954	501.17	493.51	7.65
Mother	502	1.679	1877	504.86	498.28	6.58
Brother/Sister	500	0.831	6529	501.99	498.73	3.26
Tuition	510	2.170	985	514.08	505.57	8.51
Friends	497	3.102	367	502.92	490.76	12.16
Any other	503	4.984	160	512.98	493.44	19.54
None	501	3.174	307	507.49	495.05	12.44
Missing	487	2.037	1028	491.23	483.25	7.98

11. National mean score by home language in Nepali

Language						
	Mean Score	SE	N-cases	Upper	Lower	CI
Nepali	504	0.843	9137	505.67	502.37	3.30
Other	492	1.682	4181	495.37	488.77	6.59

12. National mean score by type of schools in Nepali

School Type						
	Mean Score	SE	N-cases	Upper	Lower	CI
Community	491	0.995	10271	493.36	489.46	3.90
Institutional	525	1.112	3936	526.69	522.33	4.36

13. Mean score by involvement in after- school activity in Nepali

Time Given	Time spent on TV, Internet, Mobile, Computer	Time spent on playing and chating with friends	Time spent on household chores	Time spent on studying/ doing homework
Not Given Time	497	501	499	483
Less than One hour	511	506	508	494
One to three hour	505	503	505	509
More thatn three hour	488	493	496	507
Missing	480	479	480	481

14. Disaggregated mean score of sample districts in Nepali

Mean by District							
Province	District	Mean Score	SE	N-cases	Upper	Lower	CI
Province 1	BHOJPUR	485	2.013	446	489.29	481.40	7.89
	DHANKUTA	496	2.190	447	500.01	491.43	8.58
	JHAPA	510	1.546	1062	513.36	507.30	6.06
	OKHALDHUNGA	494	2.805	324	499.06	488.07	10.99
	PANCHTHAR	505	2.174	516	509.31	500.79	8.52
Province 2	MAHOTTARI	485	2.061	530	489.46	481.38	8.08
	RAUTAHAT	487	1.794	706	490.61	483.57	7.03
	SIRAHA	496	2.131	577	500.24	491.88	8.35
Province 3	CHITWAN	522	1.724	874	525.63	518.87	6.76
	KAVREPALANCHOK	501	1.691	873	503.85	497.22	6.63
	LALITPUR	522	1.833	810	525.82	518.63	7.19
	NUWAKOT	499	1.991	568	502.62	494.82	7.80
	SINDHULI	494	1.765	771	497.25	490.33	6.92

Mean by District							
Province	**District**	**Mean Score**	**SE**	**N-cases**	**Upper**	**Lower**	**CI**
Gandaki Proince	LAMJUNG	508	2.206	477	512.51	503.86	8.65
	PARBAT	521	2.655	407	525.88	515.47	10.41
	SYANGJA	503	1.886	733	506.31	498.92	7.39
Province 5	BARDIYA	490	2.044	651	494.11	486.10	8.01
	GULMI	506	1.792	726	509.43	502.40	7.02
	PYUTHAN	487	1.985	537	490.51	482.73	7.78
Karnali Province	HUMLA	491	3.788	171	498.22	483.37	14.85
	KALIKOT	487	2.431	315	491.55	482.02	9.53
	SALYAN	490	1.867	626	493.37	486.06	7.32
Sudur Paschim Province	BAITADI	492	1.691	725	495.09	488.46	6.63
	BAJURA	488	2.518	335	493.43	483.56	9.87

Annex 2: Test Administration Guidelines

विद्यार्थी उपलब्धिको राष्ट्रिय परीक्षण, २०७४ (कक्षा ५) सञ्चालनका लागि प्रधानाध्यापकलाई मार्गदर्शन

विद्यार्थी उपलब्धिको राष्ट्रिय परीक्षण, २०७४ (कक्षा ५) सञ्चालनका लागि प्रधानाध्यापकले निम्नानुसारको व्यवस्थापन तथा सहजीकरण गर्नुपर्नेछ :

१. परीक्षण २०७४ फागुन १९ गते दिनको ११ देखि १:३० बजेसम्म तोकिएको विद्यार्थी सङ्ख्यामा सञ्चालन गर्नुपर्ने छ ।

२. यो परीक्षणले विद्यार्थीलाई पास/फेल गर्ने वा शिक्षक/विद्यालयको मूल्याङ्कन गर्ने नभई विद्यार्थीले दिएका उत्तरका आधारमा सिकाइ उपलब्धिको राष्ट्रिय स्तर निर्धारण र शैक्षिक सुधारका लागि सहयोग गर्दछ । यस परीक्षणबाट विभिन्न आधारमा उपलब्धिको तुलना गर्न, सिकाइका कारक तत्वहरू पहिचान गर्न र सिकाइ उपलब्धि न्यून तथा उच्च भएका विद्यालयका बीचमा भएका भिन्नताहरू पहिचान गर्न सहयोग पुग्दछ । परीक्षणको यस्तो महत्त्वलाई ध्यान दिई यो परीक्षणमा विद्यार्थीले जानेका कुरा सही, स्वच्छ एवम् मर्यादित रूपले लेख्ने वातावरण सिर्जना गर्नु पर्नेछ ।

३. प्रश्नको उत्तर दिनका लागि विद्यार्थीलाई कसैले पनि सिकाउने वा सघाउने गर्नुहुँदैन भने विद्यार्थीहरूले पनि परीक्षण अवधिमा आपसमा कुराकानी तथा छलफल गर्नुहुँदैन ।

४. यस परीक्षणमा प्रश्नावलीको गोपनीयता अत्यन्तै महत्त्वपूर्ण पक्ष हो । कुनै प्रश्नपत्र विद्यालय वा शिक्षकले राख्ने, कुनै प्रश्न वा प्रश्नपत्र कुनै माध्यमले सार्ने वा कपी गर्ने, फोटो खिच्ने वा फोटोकपी गर्ने जस्ता कुनै पनि कार्य गर्नुहुँदैन । शैक्षिक गुणस्तर परीक्षण केन्द्र (भ्च्इ) द्वारा विद्यालयलाई उपलब्ध गराइएका प्रश्नपत्रहरूमध्ये केही प्रयोग नभएमा सुरक्षित साथ प्रश्न प्राप्त भएकै खाममा खामबन्दी गरी पठाउनुपर्नेछ । प्रश्नावली हराएमा, फोटो खिचिएको वा फोटोकपी गरी राखिएको पाइएमा संलग्नलाई नियमानुसार विभागीय कारवाही हुने पक्षमा सचेत गराउने, प्रयोग भएका र नभएका सबै प्रश्नावली अनिवार्य रूपमा गन्ती गरी फिर्ता गर्नुपर्नेछ ।

५. परीक्षण सञ्चालन भएको समयमा सम्बन्धित विषयको विषय शिक्षक परीक्षण सञ्चालन भएको कक्षामा प्रवेश नगरी अन्य शिक्षकहरूबाट परीक्षण सञ्चालन गर्नुपर्नेछ । तर विद्यालयमा जुन विषयको परीक्षण सञ्चालन हुने हो सो विषयको शिक्षकले नै शिक्षक प्रश्नावली भर्नुपर्नेछ ।

६. विद्यालयमा प्राप्त भएका प्रश्नावली सङ्ख्याभन्दा विद्यार्थी सङ्ख्या बढी भएमा च्वलमक क्कउष्लिन विधिबाट तोकिएको सङ्ख्यामा विद्यार्थी छनौट गरी परीक्षण सञ्चालन गर्ने व्यवस्था मिलाउनुपर्नेछ। विद्यार्थी छनोट गर्दा छात्र र छात्राको अनुपात मिलाउनुपर्नेछ ।

७. प्रश्नावलीमा दुई ओटा खण्ड रहेका छन् । पहिलो खण्डमा विद्यार्थीको व्यक्तिगत तथा पारिवारिक विवरणसम्बन्धी प्रश्नावली र दोस्रो खण्डमा विषयगत प्रश्नहरू रहेका छन् । पहिलो खण्डका प्रश्नावलीमा निरीक्षकले विद्यार्थीलाई आवश्यक सहयोग गर्न सक्नेछन् । सबै विद्यार्थीले पहिलो खण्डको प्रश्नावली भरिसकेपछि प्रश्नावली यथास्थानमा राख्न लगाई ५ देखि १० मिनेटसम्म ब्रेक दिई सबैलाई एउटै समयबाट दोस्रो खण्डको विषयगत प्रश्नहरू (गणित, नेपाली र विज्ञानमध्ये कुनै

एक विषयका प्रश्नपत्रहरू हुनेछन्) हल गर्ने गरी परीक्षण सुरु गराउनुपर्नेछ ।

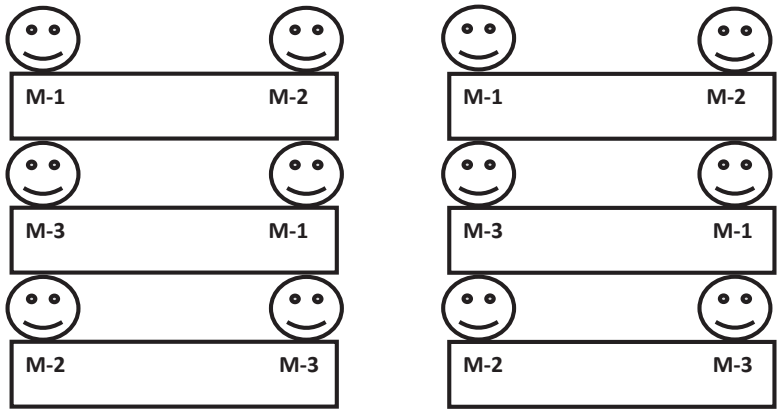
८. परीक्षणमा सहभागी हुने हरेक विद्यार्थीको लागि एउटा कालो मसी भएको एकै किसिमको डटपेन विद्यालयले उपलब्ध गराई (विद्यालयलाई उपलब्ध गराइएको मसलन्दबाट खरिद गर्ने) सोही डटपेन प्रयोग गरी उत्तर लेख्ने व्यवस्था मिलाउनुपर्नेछ ।

९. प्रश्नको उत्तर लेख्नको लागि प्रश्नावलीमा नै खाली ठाउँ दिइएको छ । सामान्यतया विद्यार्थीलाई थप खेस्रा उत्तरपुस्तिका आवश्यक पर्दैन । तर कुनै विद्यार्थीले थप उत्तरपुस्तिका खेस्राका लागि मागेमा विद्यालयले नै उपलब्ध गराउनुपर्नेछ ।

१०. एउटा बेञ्चमा बढीमा २ जना मात्र विद्यार्थी बस्ने गरी बसाइ व्यवस्था मिलाउनुपर्नेछ ।

११. प्रत्येक विषयमा ३ किसिमका प्रश्नपत्रका सेटहरू (ख्मचकष्यलक) उपलब्ध गराइएको छ । गणितमा प्रश्नका किसिमलाई :ज्ञ, :द्द र :घस नेपालीमा ल्ज्ञ, ल्द्द र ल्घ तथा विज्ञानमा क्ज्ञ, क्द्द र क्घ सङ्केत गरिएको छ ।

१२. प्रश्नावली वितरण गर्दा ३ ओटै किसिमका सेटहरू क्रमशः पर्ने गरी वितरण गर्नुपर्नेछ । (तलको चित्रमा गणित विषयमा प्रश्नावली वितरणको एउटा नमूना दिइएको छ ।)



१३. विद्यार्थीले उत्तर लेखिसकेपछि निरीक्षकले प्रत्येक विद्यार्थीका भरिएको प्रश्नावली र उत्तरपुस्तिका दुवैलाई सँगै स्टिच गर्नुपर्नेछ ।

१४. प्रश्नावलीमा कुनै त्रुटी भेटिएमा प्र.अ./निरीक्षकले आफ्नो प्रतिवेदनमा सुझाव संलग्न गरी पठाउनुपर्नेछ तर प्रश्नहरू सच्याउने वा तत्काल कुनै निकायमा सोधखोज वा खबर गर्नु आवश्यक छैन ।

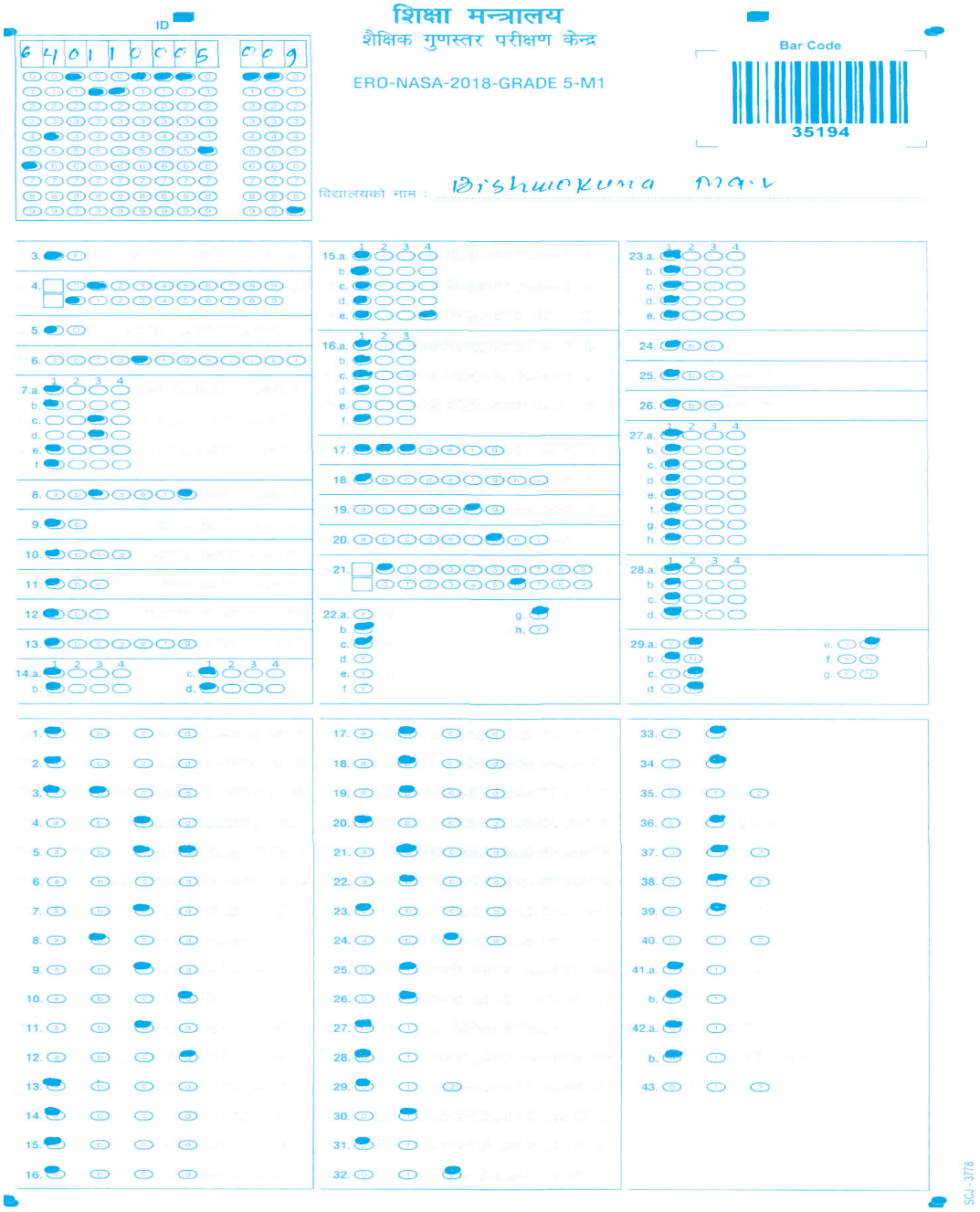
१५. परीक्षण समाप्त भएपछि उत्तरपुस्तिका तत्काल खामबन्दी गरी छिटो र सुरक्षित माध्यमबाट जिल्ला शिक्षा कार्यालयमा बुझाउनुपर्नेछ । उत्तरपुस्तिकाहरू बन्द गरिएको खामको बाहिर प्राप्त भएका, प्रयोग भएका र प्रयोग नभई फिर्ता भएका प्रश्नपत्रको सङ्ख्या तथा विद्यालयको कोड उल्लेख गर्नुपर्नेछ ।

१६. उत्तरपुस्तिका सँगै प्रधानाध्यापक र शिक्षकद्वारा भरिएका प्रश्नावली पनि जिल्ला शिक्षा कार्यालयमा बुझाउनु पर्नेछ ।

परीक्षण सम्बन्धमा थप जानकारी आवश्यक भएमा सम्पर्क:

- शैक्षिक गुणस्तर परीक्षण केन्द्र, सानोठिमी (०१६६३६५१८, ०१६६३२११६), email id: eronasa@gmail.com

Optical Mark Reader sheet designed




Sudur Paschim
Karnali
Gandaki
Province 3
Province 1
Province 5
Province 2
Sample district