नेपाल सरकार
राष्ट्रिय परीक्षा बोर्ड
सानोठिमी, भक्तपुर, नेपाल

# Assessment Framework for

# Basic Level Examination

**Government of Nepal**

**National Examinations Board**

**Sanothimi, Bhaktapur, Nepal**

**March 2022**

**Updated on January 2025**

**Users/Readers are requested to review this draft assessment framework and provide their valuable comments and suggestions.**

**National Examinations Board**

# Foreword

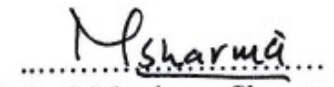
National Curriculum Framework assumes administration of external tests at the end of grade 8 (Basic Level Examination), grade 10 (School Education Examination), and grade 12 (School Level Certificate Examination) in school level education in Nepal. These examinations are within the purview of Local Government, Province Level and Federal Level respectively. Moreover, the National Examinations Board is responsible for providing guidance for conducting, monitoring and evaluating all types of school examinations. The Eighth Amendment of Education Act (2073 BS) underscores the need for examination reforms and standardization of the examinations at various administrative levels. The School Education Sector Plan's result framework also guides local levels for conducting BLE using standardized test items. This reform is also essential to ensure the quality of education and comparable learning opportunities for students across local governments.

The curriculum and the specification grid are the main guidelines not only for developing assessment tools but also for certification. The use of items to measure all cognitive domains, specifically ensuring higher-order thinking skills (HOTS) as intended by the specification grid of the curriculum is essential. Apart from improving the quality of tests and testing, the use of test results to improve classroom teaching-learning is a must. The National Examinations Board has also been working on the capacity development of teachers, schools and the Local Government/Province in these areas with the ultimate aim of enabling the Local and the Provincial Government to take over their role for Basic Level Examination and Secondary Education Examination respectively.

In this context, the assessment framework for Basic Level Examination was prepared in 2022 covering English, Mathematics, and Science and Technology subjects. However, the change and the revision of the curriculum and specification grid of these subjects in the subsequent years made it mandatory to update this framework accordingly.

I express my sincere gratitude to Prof. Dr. Ganesh Bahadur Singh, the lead of the framework updating team and the subject experts who contributed during consultation workshops. I sincerely thank Mr. Janga Bahadur Aryal, Member Secretary, and Mr. Subash Pant, Deputy Controller of NEB for their constructive roles in facilitating the development of this framework.

Last but not the least, I am indeed grateful to the University teachers as well as representative personnel from CDC, OCE and ERO for their contribution in providing constructive feedback in the sharing and discussion sessions during the workshops.



Dr. Mahashram Sharma
Chairperson, NEB

## Abbreviations

| | |
|---|---|
| BLE | Basic Level Education |
| BS | Bikram Sambat |
| CTT | Classical Test Theory |
| CDC | Curriculum Development Center |
| HA | Higher ability |
| HOTS | Higher order thinking skills |
| IRT | Item Response Theory |
| LAQ | Long Answer Question |
| LG | Local Government |
| LO | Learning outcome |
| MCQ | Multiple Choice Question |
| MoEST | Ministry of Education, Science and Technology |
| NEB | National Examinations Board |
| OECD | Organization for Economic Co-operation and Development |
| SAQ | Short Answer Question |
| SESP | School Education Sector Plan |
| SSDP | School Sector Development Plan |
| VSAQ | Very Short Answer Question |

# Table of Contents

**List of Tables**

# Chapter 1

## National Examinations Board and Quality Examinations

The National Curriculum Framework of Curriculum Development Center (CDC) 2076 BS encompasses a two-tier school education system in Nepal: basic and secondary levels. Basic level education includes one year of Early Childhood Development and Education for children, aged four and basic education from grades one to eight for children, aged five to twelve. Secondary level education includes grades from nine to twelve for children, aged thirteen to sixteen.

The Local Government Operational Act 2074 BS and School Sector Education Plan (SESP: 2022/23-2031/32)has made the local government responsible for the operation, monitoring and management of Basic Level Examinations[Ministry of Education, Science and Technology (MoEST), 2022]. The National Examinations Board (NEB) is responsible for providing guidance for conducting, monitoring and evaluating all types of school examinations. The Eighth Amendment of the Education Act underscores the need for examination reforms and standardization of the examinations at various administrative levels. The SESP result framework also guides local levels in conducting Basic Level Examination (BLE) using standardized test items (MoEST, 2022).

Quality education is a shared responsibility, which demands coordinated and combined efforts from respective institutions to ensure the desired learning by the students. Educational efforts should ultimately be reflected in classroom teaching learning and exhibited in students' learning achievement. One of the government's priorities in basic education objectives is to 'ensure student readiness for secondary education by students gaining the required learning competencies' by the end of the basic education level. In order to achieve the improved student assessment and examination system, the government aims to 'develop standardized test items for grades 3, 5 and 8 examinations and standardized examinations at end of Grade 8 as the major intervention. The reform is also emphasized with a special focus on standardization across local governments and the use of the items to measure all cognitive domains as well as skills and attitudes.

This assessment framework for the Grade 8 examination at the end of basic education aims to provide valid, reliable, consistent and transparent guidelines for developing test items in three core subjects: English, Mathematics, and Science and Technology subjects based on the national curriculum and specification grid.

An assessment framework is a descriptive document that briefly outlines an assessment, Grade 8 examination in this case, and explicitly states its characteristics and principles upon which the assessment is built. The framework helps the persons who are interested in and involved with the assessment to understand what the assessment is about. One of the most important purposes of having an assessment framework is to clearly state what is expected from examinations and set a decent ground for the justification of the process of test development. The assessment framework also ensures beneficial backwash on classroom

teaching and learning. It is imperative to bring a transformation in the way that the students are assessed.

## The Need for a Robust Assessment Framework

There are several reasons for developing an assessment framework. One important reason is to provide valid and reliable test scores to certify student's learning achievement at the end of basic education. Validity is the degree to which test scores, by subject or in composite, should reflect what it is supposed to measure and provide information that is relevant to the inferences that are to be made from it. Reliability is the accuracy or precision of a measurement procedure to provide consistent scores (Thorndike Thorndike-Christ, 2014).

Distortions can creep into an assessment in many ways and have a negative effect on validity. For example, having a literacy item that requires a numeracy skill, or an item that gives an advantage to a certain group of students due to background knowledge would not yield a valid assessment outcome. An assessment framework helps in removing such measurement bias as well as distortions from the assessment and increases the degree of validity of the test. Similarly, an assessment framework helps to improve the degree of reliability of the test by guiding the team to develop reliable test items and maintain uniformity in the assessment.

BLE, at the end of grade 8, is a curriculum-based criterion-referenced test. In such a test, quality is assured through developing items in line with the objectives and specifications given in the curriculum. Test development should target a test to the appropriate group by covering suitable content areas. Fairness of the test with respect to equal opportunity to get the item correct among the same ability groups irrespective of gender, location, type of school, etc., is also desired in the assessment. The assessment framework explicitly guides these processes.

Comparability of test scores is essential from one year to another and from one place to another if multiple test sets are used. The assessment framework documents an assessment plan to ensure consistency from one assessment cycle to the next and equivalency of tests while using multiple sets. As a result, the effect of any change in the program in the future cycles can be compared, recorded, and evaluated.

## Key Features of the Assessment Framework

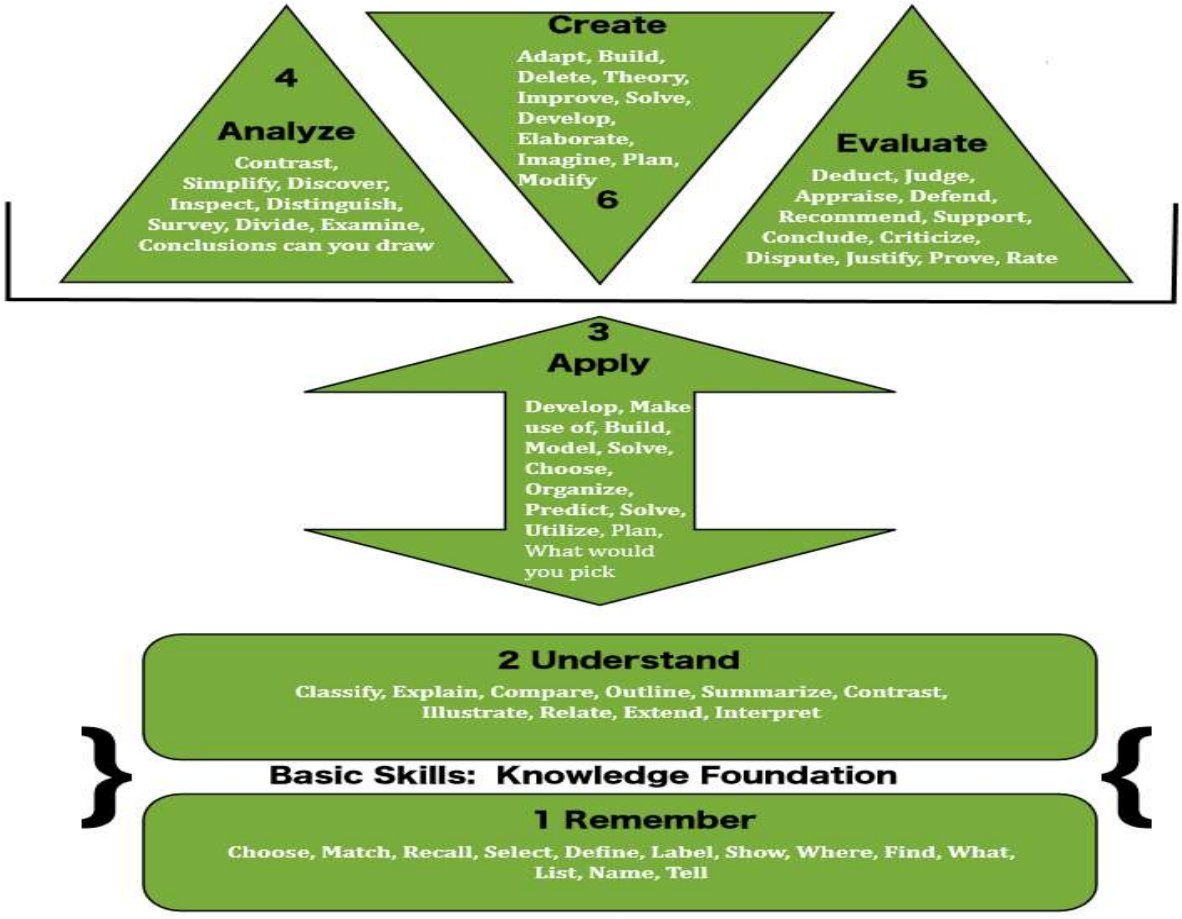
This assessment framework has the following features:

- definitions of the three main test domains - English, Mathematics, and Science and Technology subjects with the sub-domains or strands to be tested;
- a clear linkage between the intended and the implemented curricula, and what is to be tested at the end of grade 8 in the form of summative examination;
- the distribution of strands or sub-domains to be tested;
- the distribution of cognitive behaviour or processes to be assessed;
- the types of question format to be included in the test;
- the distribution of marks within cognitive domains and formats; and
- the marking scheme and/or rubric for each domain.

The information contained in the framework is based on the curriculum and elaborated in the item matrix.

### Review of the Specification Grids

The Basic Education Curriculum (2069 BS), the new grades 6-8 curriculum (2077 BS) and the test specification grids for English, Mathematics and Science and Technology were reviewed. The review highlighted the need for introducing testing of HOTS or higher ability (HA) in the examination to prepare a new generation of learners for the era of technology where creativity and practical application of principles are more critical than rote learning. Tests should also require higher ability thinking by including analysing, evaluating, synthesizing, critical thinking, problem-solving and creating types of items. These are the skills mainly ranked as the upper cognitive processes in the revised Bloom's taxonomy as shown below.




Source: The revised Bloom's Taxonomy https://commons.wikimedia.org/wiki/File:Bloom%27s_Taxonomy_for_Course_Design_

## Enhancing Quality of the Test

One of the main objectives of the examination reform is to enhance the quality of test items and tests. Furthermore, the standardization of the test is desired to determine the quality of items and ensure comparability of test scores over the years and across Local Governments. In order to achieve the standardization of the examination, it is recommended to develop test items following the specification grid laid out in the curriculum and the elaborated test item matrix in the assessment framework. The framework includes a definition of the subject domain. It explains which aspects of the domain are included in the examination. The framework also outlines how the examination is conducted by setting out the proportion of items for each aspect of the learning domain. It describes the response formats that are used and the length and number of items in the examination. Once a pool of items is available and field tested, information on item difficulty, item discrimination, distracter analysis, item correlation, differential item functioning and other essential parameters should be provided for the item selection and test building based on the selected items. The key aspects of item development are described in the chapter 2.

### *Item difficulty*

Item difficulty indicates how difficult or easy an item is for test takers. Item difficulty is also called item facility and the two terms are used interchangeably. P-value is one of the common statistics of item facility. It is the proportion of test takers who answered an item correctly. It is recommended that a test should cover a wide range of item difficulties. It should include some difficult items that can challenge test takers with high ability and some easy items that allow test takers with low ability to show what they know and can do.

### *Item discrimination*

Item discrimination refers to the ability of an item to differentiate test takers who answer the item correctly and who do not. Point-biserial correlation is often used as an index of item discrimination. It is calculated by correlating test takers' answers in dichotomous scores (either right or wrong) with their total scores. As a rule of thumb, it is recommended for test developers to carefully examine any item with a point-biserial correlation lower than 0.2 for any error in the item and apply professional judgment for further actions.

## Grade-level Reporting

In previous years, examination results were reported based on a 101-point scale, ranging from 0 as the zero mark and 100 as the full mark by reflecting the percentage of right answers. However, due to the different difficulty levels of examinations each year, the achievement levels of students could not be compared over the years. For example, the learning level at the test score of 50 in 2015 was not equivalent to that of the same score in 2016. At the same time, when the test was easy, more proportion of students achieved higher marks, while only a small portion of the students could score high marks when the test was difficult. Another pertinent issue is the error in the measurement in a test. If there are more errors in the measurement, there will be more fluctuation in the scores. The error in the measurement can be calculated and the scale used to determine the range of scores can be described in a more

precise manner. In order to address this issue of comparability and error in measurement, a grading system has been introduced since 2015, replacing the 101-point score system. A total of 9 learning levels were in use for grading purposes. The grading system has been revised (NEB, 2021) in which below 35 percent of scores are labelled as 'not graded' with 8 grade levels, viz. A+ Outstanding, A Excellent, B+ Very Good, B Good, C+ Satisfactory, C Acceptable, D Basic, and NG Not Graded are in use. The descriptor of each level is also developed to give information on learning levels. It is essential to bring uniformity in grading throughout grades and subjects and meaningfulness and common understanding in the descriptors. Each score band should have clear descriptors that indicate what a student at that score level knows and can do. These descriptors are based on empirical data for each domain. Thus, student proficiency levels can be clearly defined which helps in improving the standard setting.

Moreover, adopting and defining the grading system is not sufficient for standardization of the test items. Standardization of the test items can only be achieved by conducting empirical testing of the items, analysing the test data with scientific methods, and interpreting the analysis based on the opinions of qualified subject experts. Therefore, the next step after the development of the assessment framework for the examination at grade 8 would be to enhance the quality of test items and tests and then standardize the tests.

### Standardization of the Test

Standardization of the test starts with item development in congruence with learning outcomes and the specification grid of the subject as given in the curriculum. For this purpose, the specification grid is elaborated as a matrix to tap a wide range of possibilities in item development within the premises of learning outcomes and specification grid. Based on the elaborated item matrix, item writer develops item in item card with the required metadata as given below.

---

1) Subject: Science and Technology/Mathematics/English of Grade 8

2) Item cell code:

3) Elaborated item code (For Mathematics and Science and Technology subject)

| Unit | LO | Cognitive skill | Format | Marks |
|---|---|---|---|---|
| | | | | |

3) Elaborated item code (for English):

| Unit | LO | Area (R/W) | If R - skill | Format | Marks | Item type |
|---|---|---|---|---|---|---|
| | | | | | | |

4) Learning outcome (from curriculum):

……..…….....……..…….....

……..……..……..……..……..……..……..……..……..……..……

5) Objective of the item (in line with learning outcome):

……..…….....……..……..……..……

….……..……..……..……..……..……..……..……..……..……..

6) Item (Both in Nepali and English in case of S&T and Mathematics):

……..……..……..……..……..……..……..……..……

……..……..……..……..……..……..……..……..……..…….

7) Key answer/marking scheme (in English or Nepali only):

……..……..……..……..……..……..……..……..……..……..

……..……..……..……..……..……..……..……..……..…….

---

Items submitted by the item writer are then panelled in a team of teachers and experts. Then accepted/finalized items are moderated as per NEB rules. Final draft items are then assembled in a test set as per the specification grid along with the test matrix and marking scheme, answer key and rubric as required. These test sets are pre-tested in the national sample in which both students and items are sampled to pre-test a vast number of items in a short period of time (Redfield, 2001).

The size of the trial population larger the better. However, this should be possible within available resources (Withers, 2005). The sample of students for each trial paper should be 150-250 persons assigned randomly who are similar to those who will attempt the final forms of the test (Izard, 2005a). As there is a chance of loss of data during pre-testing, minimum 200 students should be maintained (Anderson & Morgan, 2008). Pretesting should be conducted under the same conditions as the final test, the length of time allowed for students to take the test should be the same as will be allowed in the final test. Such conceptual deliberation should be clearly explained in the 'Guidelines for Pre-testing, Item Analysis and Item Selection Procedures'.

Izard (2005b) suggested the use of a codebook for the items in pretesting to document vital information such as where an item appears on the test, which area of content and which skills are being assessed, the name assigned to the item (if one is assigned), the number of options, the code used for missing data, any coding values for particular responses, and any notes that provide necessary information about the item. This will be helpful to tracking items from the development to finalization phases. After field testing of items, they are scored, data input/cleaning, and analysis is done. Items are analysed for their difficulty level, discrimination index, and the power of distracters in the case of multiple-choice question (MCQ). During the field testing phase, acceptable and pathological/defective items are detected through item analysis (Izard, 2005b; Metsamuuronen, 2012). The items that meet acceptable quality are banked for later retrieval. Item banking software would be helpful.

Pre-testing helps in the selection of quality test items, but a complete set of items in a final test paper format is not yet assembled and tested. By this time, the test manual should have been prepared and it has to be tested as well. Based on the test specification grid, a complete test paper is assembled from the item bank for the final test. Reliability and validly of the final test are established and reported for the final test in a larger sample.

### Test Administration and Ongoing Test Maintenance

The final test thus is ready to be administered in the regular examination. The item bank needs restocking with fresh calibrated quality items through an ongoing test maintenance activity in an institutionalized way. Restocking of the item bank and the ongoing maintenance of items should be institutionalized. An expert team should be developed and made available to support as and when required.

### Internal Assessment

The internal assessment constitutes fifty percent of the summative assessment scores for grade 8 as provisioned in the curriculum. Student portfolios should be maintained for internal assessment. Along with the skills, the participation of students in classroom activities and their attendance should be included in the evaluation process. It is still not clear how marks obtained by the students in internal and external examinations would be presented in a composite form and meaningfully interpreted. More work is needed in this area.

### Certification

A validated standard-setting procedure informed by statistical analysis of student outcomes should be run to determine the cut score between the sufficient and the insufficient. This score should become part of a table for converting raw scores into grades (A+, A, B+, B, etc.) which should also be decided by the panel doing the standard setting. This panel could also add descriptions in terms of items included in the test to score intervals.

### Result Analysis

Skills in item development are important for developing quality tests, analysing and utilizing students' performance in the test, and planningtheir learning improvement (Khaniya, 2005). Item writing training for teachers also improves classroom teaching learning and students' learning achievement. This will help the teacher understand the nature of items, purposes of the items in relation to the curricular intent for students' learning, and create an appropriate classroom teaching learning environment.

Results can also be analysed at different levels such as school, Local Government, District, Province and National levels and identify poorly performed items. Poorly performed items can be analysed into three aspects: 1) what the item demands, 2) why most of the students were not able to do good, where students might have confusion, and then 3) how to improve classroom teaching learning. The analytical report with specific suggestions can help in gradual improvement in classroom teaching learning and lead to improvement in learning achievement.

# Chapter 2

## Testing in Core Subjects of Basic Level Education

The terminal examination for grade 8 marks the completion of basic education in Nepal. There are six core subjects in basic level education. As per the guidelines of the School Sector Development Plan (SSDP) 2016 and SESP, 2022, quality examination development has been initiated in three subjects: English, Mathematics, and Science and Technology in grade 8. This chapter presents a framework mainly for the summative assessment of these three subjects in the grade 8 examination in line with the curriculum and the specification grid. Competency and learning outcomes desired in the curriculum and requirements sought in the specification grid such as cognitive behaviour, the format of test items, content coverage, and weightage provided to the content domain are the core elements to be considered in a curriculum based test.

## Testing of English Language Subject

English is one of the compulsory subjects taught in schools in Nepal from primary to secondary levels. The curriculum for grades 6-8 (CDC, 2077 BS) reiterates the aim of teaching English as to enable the students to communicate with confidence, expose students to the vast treasures of knowledge and pleasure available in both written and spoken English and develop the linguistic base in English for their further studies. Specifically, the English curriculum expects students to achieve the following competencies by the end of the grade 8:

a) listen in order to understand and respond appropriately in a variety of situations for a variety of purposes;

b) speak appropriately to communicate with different audiences for a variety of purposes in a variety of personal, social and academic contexts;

c) read and understand a variety of literary, informational, and graphic texts, using a range of strategies or cognitive processes to construct meaning;

d) write for an intended purpose and audience by generating, gathering, and organizing ideas and information; and

e) demonstrate good control of vocabulary and syntax to express basic communicative needs.

### Description of Skills for Assessment

The Basic Education Curriculum has organized the curriculum for the English language subject around the four sub-skills: listening, speaking, reading, and writing. While the document acknowledges that the acquisition of these skills by students is by no means linear or mutually exclusive, it enlists separate desired learning outcomes for each skill. The following sections examine the learning outcomes in detail.

**Reading:** Reading literacy is defined as understanding, using, reflecting on, and engaging with written texts, in order to achieve one's goals, develop one's knowledge and potential,

and participate in society (OECD, 2018). The curriculum identifies eleven learning outcomes to evaluate the reading skills of grade 8 students such as understanding the description of events, feelings and wishes, understanding recipes and instructions, finding out the main ideas and supporting details from a text, reading short poems and stories for pleasure and understanding, consulting dictionaries (including e-dictionaries) to learn the different aspects of words (see curriculum for details).

**Writing:** Writing allows students to put their feelings and ideas on paper, organize their knowledge and beliefs into convincing arguments, and convey meaning through a well-constructed text. Spelling, vocabulary, grammar, and organization come together and grow together to help the student demonstrate more advanced writing skills(Michigan State University, 2002). The curriculum identifies ten learning outcomes for writing skills such as writing personal letters and simple official letters, interpreting charts, tables and diagrams, writing short, simple biographies, writing simple stories based on the given pictures or texts or both, writing short simple essays on topics of interest, using punctuation correctly (see curriculum for details).

The learning outcomes for writing comprise writing, interpreting and using tasks, which are largely self-explanatory. The writing skill also includes language functions (19 functions in the list), and 5 marks is allocated for the grammar part in the specification grid.

**Speaking:** Speaking refers to students' ability to employ their communicative skills and engage in discussions in an analytical and creative manner. The curriculum outlines ten learning outcomes for speaking. Speaking is assessed as a part of the internal assessment.

**Listening:** Listening includes students' ability to listen for basic interpersonal, instructional and academic purposes, and use the information to make connections and draw inferences. The curriculum outlines five learning outcomes for listening. Listening is covered in the internal assessment.

### Specification Grid of English Subject

The specification grid provides a framework for developing a test for summative assessment which incorporates reading and writing skills. Reading and writing carry 25 marks weightage each. The structure for each of these skills is also mentioned in the curriculum. The assessment of reading should be done through 'seen' (the passage from the textbook) and 'unseen' (the passage that students may encounter for the first time and is used as a context for answering questions). There will be four questions in reading carrying 25 marks.

1. **Reading 1:** One short reading text from the textbook with one type of comprehension questions (5 marks)
2. **Reading 2:** One short reading text from the textbook with one type of comprehension questions (5 marks)
3. **Reading 3:** One short reading text not given in the textbook with one type of comprehension questions (5 marks)
4. **Reading 4:** One short reading text not given in the textbook with two types of comprehension questions of which one will be of vocabulary test (10 marks).

Text types for reading 1 and 2 will be based on the text type of the textbook. For reading 3 and 4, these text types are suggested as story, dialogue, timetable, menus, charts/graphs, calendar, notice, speech, announcement, instructions, memoirs, diary entry, letter/email, news stories/news report, brochure, leaflet, biography/autobiography, and short essay. The length of the text of readings 3 and 4 should not exceed 250 and 300 words respectively.

The grid has specified the types of questions that can be asked for testing reading comprehension, viz. true/false, fill in the gaps, multiple choice, matching, ordering, and short answer question (SAQ). The item format should be different for Reading 1 and 2 and text types in Reading 1 and 2 should not be repeated in Readings 3 and 4, for example, a letter in Reading 1 and an essay in Reading 2. Furthermore, the specification grid requires that Reading 4 should contain five questions for testing vocabulary.

The testing of students' English language knowledge and skills further demands different types of cognitive processes, including literal comprehension, reorganization, inference, and evaluation and reflection. In order to assess these cognitive processes or levels, 8, 4, 5, and 3 questions for literal comprehension, reorganization, inference, and evaluation will be constructed respectively. As per the weightage given in this section, one element carries one mark.

The assessment of writing skills follows punctuation (5 marks), guided writing (5 marks) and free writing (10 marks). Knowledge in punctuation is assessed with a short and simple paragraph containing ten punctuation errors. Guided writing contains a description, paragraph, news story/story, description of tables/charts and diagrams, announcement, or a set of instructions with some guidelines. Similarly, free writing contains a personal or official letter, an account of events, a diary entry, or a short essay. The specification grid requires that writing task types should not be similar to those of reading texts given in the reading section.

In the case of grammar, two tasks will be given that incorporate grammar items such as article, preposition, tense, connectives, question tag, reported speech, voice, conditional sentences, subject-verb agreement, question, negation, modal verbs, causative verbs, relative clauses, and comparative and superlative forms. Furthermore, the specification grid clarifies that two types of questions should be asked. The reproduction type covers tense, question tag, reported speech, voice, question, and negation (5x0.5=2.5 marks) and the multiple-choice item (individual sentences or a contextual passage) contains any five from the grammar areas, viz. articles, prepositions, connectives, modal verbs, used to, causative verbs, relative clauses, comparative and superlative forms, conditional sentences and subject-verb agreement (5x0.5=2.5 marks).

### Elaboration of Specification Grid for Item Development

Items are developed to assess students' learning outcomes intended by the curriculum. The specification grid provides the explication of the sampling of items to make the test representative in terms of content area, cognitive behaviour, and format of items. Item development is guided by intended learning outcomes, related content area, desired cognitive behaviour, format of the item to be used, and weightage given to each item. Putting these requirements together provides an item development matrix in an elaborated form. The first

step is to understand the specification grid properly. The specification grid for 'Reading' is presented in Table 1.

**Table 1: Specification grid for reading skill**

<table>
<tr><th colspan="2">Item type</th><th>Marks</th><th>Content/context</th><th>Item format</th></tr>
<tr><td>Reading 1: One short reading text from the textbook</td><td>One type of comprehension question</td><td>5 marks</td><td rowspan="2">Given in the textbook</td><td rowspan="4">1. True/False<br>2. Fill in the gaps<br>3. Multiple choice<br>4. Matching<br>5. Ordering<br>6. Short answer questions<br>(Note: Reading 4 should contain the questions for testing of vocabulary.)<br>The distribution of comprehension questions should be:<br>Literal comprehension – 8 questions<br>Reorganization – 4 questions<br>Inference – 5 questions<br>Evaluation and reflection – 3 questions</td></tr>
<tr><td>Reading 2: One short reading text from the textbook</td><td>One type of comprehension question</td><td>5 marks</td></tr>
<tr><td>Reading 3: One short reading text not given in the textbook</td><td>One type of comprehension question</td><td>5 marks</td><td rowspan="2">• From outside the textbook - Any two from story, timetable, menus, charts, calendar, notice, letter, news, brochure, biography/auto biography and short essay<br>• Should not exceed 250 and 300 words respectively.<br>• Text type should be different for Reading 1 and Reading 2. For example, letter in Reading 1 and story in Reading 2.<br>• Type of texts in Reading 1 and 2 should not be repeated in Reading 3 and 4 too.</td></tr>
<tr><td>Reading 4: One short reading text not given in the textbook</td><td>Two types of comprehension questions</td><td>2 x 5 marks</td></tr>
<tr><td>4 texts</td><td></td><td>25 marks</td><td></td><td></td></tr>
</table>

Similarly, Table 2 presents the specification grid for the writing skill.

**Table 2: Specification grid for writing skill**

| Item type | | Marks | Content/context | Item format |
|---|---|---|---|---|
| **Punctuation** | 10 punctuation errors | 5 marks (10 x 0.5) | A short and simple paragraph | Identify and correct – constructed |
| **Guided writing** | One task | 5 marks | description, paragraph, news story/news report, story, description of tables/charts and diagrams, announcement, a set of instructions with some guidelines | Constructed |
| **Free writing** | One task | 10 marks | personal or official letter, an account of events, diary entry or a short essay | Constructed |
| **Grammar** | Reproduction | 2.5 marks | tense, question tag, reported speech, voice, question and negation | Constructed |
| | Multiple choice | 2.5 marks | Individual sentences or a contextual passage - article, preposition, connectives, modal verbs, used to, causative verbs, relative clause, comparative and superlative, conditional sentence, and subject-verb agreement | MCQ – select type |
| **3 questions** | | **25 marks** | | |

In Tables 1 and 2, the items required are formulated in a simplified way with the skills, tasks, weightage, content and context, and item format. The next step is to elaborate the specification grid putting together the learning outcomes and requirements of the specification grid in a matrix form. Such a matrix will form a number of cells for each of the learning outcomes and each cell is coded with a unique number. This code is used to tag the item, trace the individual item, sort the items into different categories, or assemble them in a

set of tests. Item identifier codes are essential for item banking as well as the manual item repository. An example of the elaborated specification grid of the Reading 1 section of English subject is given in Table 3 in which the lesson as per the textbook for the seen text is matched with the possible item format. Cells formed in the matrix when the lesson is cross-matched with possible item format are given unique codes as identifiers for the items.

**Table 3: Example of code for the elaborated specification grid – Reading 1**

| Lesson | Item format | | | | | |
|---|---|---|---|---|---|---|
| | **True/ False** | **Fill in the gaps** | **MCQ** | **Matching** | **Ordering** | **SAQ** |
| 1) A Tour to Central Zoo: Timetable | 1 | 2 | 3 | 4 | 5 | 6 |
| 2) A Request Letter | 7 | 8 | 9 | 10 | 11 | 12 |
| 3) A Father's Letter to his Son | 13 | 14 | 15 | 16 | 17 | 18 |
| **.....** | 19 | 20 | 21 | 22 | 23 | 24 |

Similar coding will be used for Reading 2. In the case of Readings 3 and 4, the following coding list format can be used in each category.

**Table 4: Example of code for the elaborated specification grid – Reading 3**

| Text type | Item format | | | | | |
|---|---|---|---|---|---|---|
| | **True/ False** | **Fill in the gaps** | **MCQ** | **Matching** | **Ordering** | **SAQ** |
| **Story** | 1 | 2 | 3 | 4 | 5 | 6 |
| **Timetable** | 7 | 8 | 9 | 10 | 11 | 12 |
| **Schedule** | 13 | 14 | 15 | 16 | 17 | 18 |
| **.....** | 19 | 20 | 21 | 22 | 23 | 24 |

The item will be given the code number of the cell where it falls into the matrix. This code indicates the content area, the format of the item, etc. as mentioned in the specification grid. For example, the item with the cell code number 9 for Reading 3 is the item from out of the textbook or unseen, for a timetable-type text, in the MCQ format.

### Item Card to Develop Item

The item writer uses item card given in Chapter 1 to develop assigned item/question. If item developers are assigned to prepare a set of tests, they should also provide item metadata information in the test matrix. In both cases, items or test sets submitted go through panelling and moderation at NEB. These processes are essential to ensure that the test developed is

congruent with the curriculum and specification grid. Such congruence is checked through the test matrix.

### Test Matrix for the Assembled Test Set

The panelled and moderated items are assembled in the complete test set meeting the requirement of the specification grid. Items are indicated within the specification grid form where the items of the test set fall. A complete set should fulfil the content weightage area in the column and cognitive level by item format in the row of the test matrix. The format of the test matrix is shown below:

**Table 5: Format of a test matrix for the English test set**

| SN | Skill area | Item Code | Text type | Item format | Marks | Vocabulary (5) | Literal comprehension (8) | Reorganization (4) | Inference (5) | Evaluation and reflection (3) |
|---|---|---|---|---|---|---|---|---|---|---|
| **1)** | **Reading 1** | | | | 5 | | | | | |
| **2)** | **Reading 2** | | | | 5 | | | | | |
| **3)** | **Reading 3** | | | | 5 | | | | | |
| **4)** | **Reading 4A** | | | | 5 | | | | | |
| | **Reading 4B (Vocabulary)** | | | | 5 | 5 | | | | |
| **5)** | **Punctuation** | | | | 5 | | | | | |
| **6)** | **Guided writing** | | | | 5 | | | | | |
| **7)** | **Free writing** | | | | 10 | | | | | |
| **8)** | **Grammar (Reproduction 2.5)** | | | | 2.5 | | | | | |
| **9)** | **Grammar (MCQ 2.5)** | | | | 2.5 | | | | | |
| | | **Total** | | | **50** | **5** | **8** | **4** | **5** | **3** |

## Testing of Mathematics Subject

Mathematics is a vital component of basic education in Nepal. The curriculum organizes the domain of mathematics along three dimensions, namely content, cognitive, and competency. Though the organization is not explicit, the learning expectations for mathematics and the chapters/units for the mathematics domain in the curriculum follow this broad framework. The mathematics curriculum for grade 8 expects students to achieve the following level-wise competencies:

a) solve the practical problems related to sets;

b) solve the practical problems related to the real number system;

c) solve the practical problems related to perimeter, area and volume;

d) solve the problems related to algebraic expressions and equations;

e) construction of plane figures and solid shapes, and verify their properties and facts;

f) solve the problems related to coordinates, Pythagoras theorem and transformation with experiment;

g) classification, presentation and interpretation of data; and

h) develop the ability to relate mathematical knowledge, skills and concepts to interdisciplinary areas and learning areas of other subjects.

### Content Area

The content dimension includes topics that are to be covered in academic years for grades 6-8. The curriculum groups the topics into six areas: sets, arithmetic, mensuration, algebra, geometry, and statistics. The distribution of content areas, their weightage, and marks as provisioned in the Mathematics curriculum for grade 8 are presented in Table 6:

**Table 6: Content areas and their weightage in the Mathematics curriculum grade 8**

| Content Area | Weightage(Teaching Learning Hours) | Marks |
|---|---|---|
| 1) Sets | 10 | 3 |
| 2) Arithmetic | 45 | 14 |
| 3) Mensuration | 15 | 5 |
| 4) Algebra | 30 | 10 |
| 5) Geometry | 50 | 15 |
| 6) Statistics | 10 | 3 |
| **Total** | **160** | **50** |

Assembled test sets should cover the marks designated for each content area as given in the specification grid. The specification grid has further divided each of the areas into units and distributed marks for the units. Item writers should be careful that the assembled test sets

should fulfil the mark distribution assigned to the individual units as well as given in the table 7,

**Table 7: Specification grid of Mathematics for grade 8**

**विशिष्टीकरण तालिका (Test specification Chart)**
**Grade 6 - 8**
**2078**

**Subject : Mathematics** **Full marks : 50** **Time : 2 hrs**

| S.N. | Areas | Total working hours | Knowledge | | Understanding | | Application | | Higher ability | | Total number of items | Total number of questions | Total Marks |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Number of items | Marks | Number of items | Marks | Number of items | Marks | Number of items | Marks | | | |
| 1. | Sets | 10 | 1 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 5 | 2 | 3 |
| 2. | Statistics | 10 | | | | | | | | | | | 3 |
| 3. | Arithmetic | 45 | 2 | 2 | 3 | 4 | 3 | 5 | 2 | 3 | 10 | 3 | 14 |
| 4. | Mensuration | 15 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 4 | 1 | 5 |
| 5. | Algebra | 30 | 2 | 2 | 1 | 2 | 2 | 4 | 1 | 2 | 6 | 3 | 10 |
| 6. | Geometry | 50 | 2 | 2 | 2 | 4 | 2 | 6 | 2 | 3 | 8 | 3 | 15 |
| | | **160** | **8** | **8** | **8** | **12** | **10** | **20** | **7** | **10** | **33** | **12** | **50** |

**द्रष्टव्य**

- प्रश्नपत्र निर्माण गर्दा प्रत्येक क्षेत्रमा र समग्रमा ज्ञान, बोध, प्रयोग र उच्च दक्षताका लागि तोकिएअनुसारको भार मिलेको हुनुपर्दछ। संज्ञानात्मक तहमा २ अङ्कसम्म घटबढ हुन सक्ने छ।
- सन्दर्भ दिएर प्रश्नहरू निर्माण गर्नुपर्ने छ। प्रत्येक प्रश्नमा एकभन्दा बढी संज्ञानात्मक तहका उपप्रश्नहरू समावेश गर्न सकिने छ।
- Application र higher ability तहका प्रश्नहरू निर्माण गर्दा सम्बन्धित क्षेत्रका अलावा अन्य क्षेत्रका विषयवस्तुसँग सम्बन्धित प्रश्नहरू पनि रहन सक्ने छन्।
- हरेक क्षेत्रअन्तर्गत रहेका सबै उपक्षेत्रका विषयवस्तुहरू समानुपातिक रूपमा समावेश हुने गरी प्रश्नहरू निर्माण गर्नुपर्ने छ।



The specification grid also specifies the cognitive levels of the items in each specific item format. These guidelines are to be followed while developing an item with respect to the cognitive behaviour that the item aims to solicit from the student.

## Intended Learning Outcomes in the Curriculum

The items included in the test should assess key objectives of the curriculum. Item writers should be clear about the learning outcomes intended in the curriculum and then plan the development of the items. They should be provided with a matrix of the intended learning outcomes, content area, and the expected cognitive skill level in teaching learning as well as in the assessment of mathematics.

## Cognitive Behaviour

The distribution of cognitive behaviour in the Mathematics curriculum for grade 8 is knowledge 16%, understanding 24%, application 40%, and higher ability 20%. These cognitive aspects are mentioned in the specification grid but not explicitly stated and defined in the curriculum. To develop reliable assessments that assess HOTS, it is necessary to define these skills in the context of each subject domain. These definitions would then form the basis for the development of assessment materials that target specific skills as per the defined

criteria. The cognitive behaviour required by the curriculum and specification grid needs to be clarified to guide item writers in developing test items.

*Knowledge/Knowing* – Retrieving information from short-term or long-term memory such as recalling facts, basic concepts and definitions, identification of figures or diagrams, identification of relationships or characteristics, symbols, and scales.

*Understanding* – Explaining or describing materials from one form to another, and interpreting ideas or concepts such as classifying, discussing, explaining and constructing meaning from instructional messages, including oral, written, and graphic communication.

*Application/Applying* – Applying knowledge, concepts or procedures to solve a task in an unfamiliar or new situation. This may involve demonstrating, solving, interpreting, or concluding.

*Higher Ability*– This skill requires students to analyse or evaluate a task by breaking it into its constituent parts and then determine how the parts relate to one another and an overall structure or purpose; join the elements of a task together to form a coherent or functional whole and reorganize elements into a new pattern or structure; and make informed judgments based on set criteria and standards. The specification grid for the mathematics subject covers analysis, evaluation, and creation as higher ability skills.

### Elaboration of Curriculum and Specification Grid for Item Development

Items are developed to assess students' learning achievement as intended in the curriculum. The specification grid provides an explication of the sampling of items to make the test representative in terms of content area, cognitive behaviour, and format of items. Item writers should pay attention to the following aspects when developing test items:

- Intended learning outcomes specified in the curriculum should be translated as the objectives of items.
- The items should be developed based on one of the elaborated specification grid cells.
- The answer key for the selected type or the marking scheme or the rubric for the constructive response items should be developed along with the first draft of the items and revise/update key answer/marking scheme whenever items are revised/updated.
- The test matrix should correspond with the specification grid, but it should vary with other assembled multiple test sets.

### Item Formats

The specification grid of the mathematics subject does not specify a particular item format. It is open. The grid mentions the number of items and marks they carry. The range of marks is from 1 to 3 for an item. Therefore, there is a possibility of MCQ or very short answer question (VSAQ), each carrying 1 mark and SAQ, each carrying 2 or 3 marks. A reframe of

the specification grid, given in table 7 above, shows several possibilities of item combinations in mathematics as shown in the table below:

**Table 8: Possible combinations of items in Mathematics for grade 8**

| S. N. | Areas | Knowledge | | Understanding | | Application | | Higher ability | | Total number of items | Total no. of questions | Total Marks | QN and marks |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Number of items | Marks | Number of items | Marks | Number of items | Marks | Number of items | Marks | | | | |
| 1. | Sets | 1 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 5 | 2 | 3 | 1) 3[1] |
| | | 1 x 1 | | 1 x 1 | | 1 x 1<br>1x2 | | 1 x 1 | | | | | |
| 2. | Statistics | 1 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | | | 3 | 12) 3 |
| | | 1 x 1 | | 1 x 1 | | 1 x 1<br>1x2 | | 1 x 1 | | | | | |
| 3. | Arithmetic | 2 | 2 | 3 | 4 | 3 | 5 | 2 | 3 | 10 | 3 | 14 | 2)[2]<br>3)<br>4) |
| | | (1 x 1) 2 | | (1 x 1)2<br>1x2 | | 1+1+3 or<br>1+2+2 | | 1 x 1<br>1x2 | | | | | |
| 4. | Mensuration | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 4 | 1 | 5 | 5) abcd<br>1+1+2+1 |
| | | 1 x 1 | | 1 x 1 | | 1 x 2 | | 1 x 1 | | | | | |
| 5. | Algebra | 2 | 2 | 1 | 2 | 2 | 4 | 1 | 2 | 6 | 3 | 10 | 6)[3]<br>7)<br>8) |
| | | 1 x 2 | | 1 x 2 | | (1 x 2) 2<br>(1+3) | | 1 x 2 | | | | | |
| 6. | Geometry | 2 | 2 | 2 | 4 | 2 | 6 | 2 | 3 | 8 | 3 | 15 | 9)[4]<br>10)<br>11) |
| | | (1 x 1) 2 | | (1 x 2) 2 | | (1 x 3) 2 | | 1 x 1<br>1x2 | | | | | |
| | | **8** | **8** | **8** | **12** | **10** | **20** | **7** | **10** | **33** | **12** | **50** | |

[1] 1)abc (1x1)3 & 2)a (1x1) b 1x2 or vice versa
[2] 3) 1+1+1+1 4) 2+1+1 5) 3+1+2 or 3) 1+1 4) 2+1+1+1+1 5) 3+1+2 or many other combinations?
[3] 7) ab 2+2 8) cd 2+2 9) ef 1+1 or 7) abc 2+2+2 8) d 2 9) ef 1+1 or many other combinations?
[4] 10) ab - 1+1 11) cde 2+2+3 12) fgh 3+1+2 or 10) ab - 1+1 +2 11) cde 2+3 12) fgh 3+1+2 or many other combinations?

## Elaboration of the Specification Grid with Cell Codes

Each item should be linked to the learning outcome of the curriculum and the specification grid. In order to trace the individual item, the test writer is expected to sort items into different categories or assemble them in a set of tests. A code is developed and tagged with the item to identify specific characteristics of the item in terms of the content area, cognitive skill level, item format and marks weightage as exemplified in Table 9 for two of the content areas as an example.

**Table 9: Example of item codes in the elaborated specification grid**

| S N | Area/Unit | Learning outcomes | Cognitive level | | | |
|---|---|---|---|---|---|---|
| | | | K | U | A | HA |
| 1 | Sets | Distinguish disjoint and overlapping sets. | 1 | 2 | 3 | 4 |
| | | Identify the proper and improper subsets according to the given sets. | 5 | 6 | 7 | 8 |
| | | Make proper and improper subsets from the given sets. | 9 | 10 | 11 | 12 |
| 2 | Statistics | Receive and provide information from the pie chart and construct pie chart from the data. | 13 | 14 | 15 | 16 |
| | | Find mean, median and mode of individual series of data. | 17 | 18 | 19 | 20 |

Each item will be given a code number of the cell where the item falls into the matrix. This code will indicate the content area, the cognitive skill level, and the format of the item. For example, the item with the cell code number 7 is from unit 1, covering learning outcome 2, and at the application level. An elaborated item code will also be maintained for each of the items to sort it at different indicators.

### Test Matrix for Assembled Test Set

Each of the test sets assembled from a pool of items is expected to differ from one another at least to some degree, but be equivalent in testing. In order to check such variation, the test matrix would be prepared for each of the assembled tests and superimposed with other sets of tests if it is a replica only or varies from others. A sample of a test matrix is given in Table 10.

**Table 10: Sample of a test matrix in the Mathematics subject**

| SN | Areas | Question no. | Cognitive Level | | | | In the set | | Required | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | K (16%) | U (24%) | A (40%) | HA (20%) | Total Marks | No. of items | Total Marks | No. of question (items) |
| **1** | **Sets** | 1 | 1 | 1 | | 1 | 3 | 3 | **3** | **2 (5)** |
| **2** | **Statistics** | 12 | | | 1+2 | | 3 | 2 | **3** | |
| **3** | **Arithmetic** | 2 | 1+1 | | | 1+2 | 5 | 4 | **14** | **3 (10)** |
| | | 3 | | 1+1+2 | | | 4 | 3 | | |
| | | 4 | | | 1+2+2 | | 5 | 3 | | |
| **4** | **Mensuration** | 5 | 1 | 1 | 2 | 1 | 5 | 4 | **5** | **1 (4)** |
| **5** | **Algebra** | 6 | | 2 | | 2 | 4 | 2 | **10** | **3 (6)** |
| | | 7 | 1+1 | | | | 2 | 2 | | |
| | | 8 | | | 2+2 | | 4 | 2 | | |
| **6** | **Geometry** | 9 | 1+1 | | | 1+2 | 5 | 4 | **15** | **3 (8)** |
| | | 10 | | | 3+3 | | 6 | 2 | | |
| | | 11 | | 2+2 | | | 4 | 2 | | |
| | | **Total** | **8** | **12** | **20** | **10** | **50** | **33** | **50** | **12 (33)** |

This is a sample of a test set marks distribution. Other sets of tests should differ at least in some of the cells and marks weightage in the cell while superimposing them. However, same item/question should not be repeated in other sets.

## Testing of Science and Technology Subject

The curriculum revision in 2077 BS changed the name of the subject from 'Science' to 'Science and Technology'. This change also signifies the importance accorded to recent developments in communication and technology. Level-wise competencies intended in the curriculum of Science and Technology are listed below:

a) understand the scientific learning process and use the science process skill;

b) exchange and analyse information and ideas, and proper use of information technology;

c) demonstrate understanding of the relationship between biodiversity and conserve the environment;

d) familiarize oneself with and understand the importance of living beings and life processes;

e) identify and use the basic aspects of the use of force and equipment in daily life;

f) develop basic knowledge of characteristics of various forms of energy used in daily life, their relevant use, and readiness for conservation;

g) acquaint oneself with the properties of substances and use them in daily life;

h) identify and use materials used in daily life appropriately; and

i) demonstrate basic knowledge and understanding of the earth and space.

### Content Areas

The curriculum of Science and Technology includes the distribution of items by content areas/units– Scientific Learning, Information and Communication Technology, Living Beings and their Structure, Bio-diversity and Environment, Life Process, Force and Motion, Energy in Daily Life, Electricity and Magnetism, Matter, Materials of Daily Use, and Earth and Space.

The specification grid has further categorized units into four groups where the total marks should sum up with their assigned group total marks, but ± 2 marks are allowed within the group distribution. Item writers should be careful that the assembled test set should fulfil the distribution of marks assigned by column and row totals as given in the table below.

**Table 11: The specification grid for Science and Technology with item distribution and weightage**

| S. N. | Unit | Working Hour | Cognitive Level | | | | Group wise Weightage | Unit wise Weightage |
|---|---|---|---|---|---|---|---|---|
| | | | K (20%) | U (30%) | A (30%) | HA (20%) | | |
| 1 | Scientific Learning | 10 | MCQ (2×1) | MCQ (3×1) | MCQ (3×1) | MCQ (2×1) | 9 | 3 |
| 2 | ICT | 30 | | | | | | 6 |

| S. N. | Unit | Working Hour | Cognitive Level | | | | Group wise Weightage | Unit wise Weightage |
|---|---|---|---|---|---|---|---|---|
| | | | K (20%) | U (30%) | A (30%) | HA (20%) | | |
| 3 | Organisms and their Structure | 12 | VSAQ (2×1) | VSAQ (2×1) | VSAQ (2×1) | VSAQ (2×1) | 12 | 4 |
| 4 | Biodiversity & Environment | 8 | | | | | | 3 |
| 5 | Life Process | 15 | | | | | | 5 |
| 6 | Force and Motion | 15 | SAQ (1×2) | SAQ (3×2) | SAQ (3×2) | SAQ (1×2) | 18 | 5 |
| 7 | Energy in Daily Life | 20 | | | | | | 7 |
| 8 | Electricity and Magnetism | 10 | | | | | | 3 |
| 9 | Earth and Universe | 10 | LAQ (1×4) | LAQ (1×4) | LAQ (1×4) | LAQ(1×4) | | 3 |
| 10 | Matter | 15 | | | | | 11 | 6 |
| 11 | Matters Used in Daily Life | 15 | | | | | | 5 |
| **Total** | | **160** | **10** | **15** | **15** | **10** | **50** | **50** |

| Detailed Item Plan | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Types of Item | Weightage Per Item | Number of items | | | | Total Questions | Total Weightage |
| 1 | MCQ | 1 mark | 2 | 3 | 3 | 2 | 10 | **10** |
| 2 | VSAQ | 1 mark | 2 | 2 | 2 | 2 | 8 | **8** |
| 3 | SAQ | 2 marks | 1 | 3 | 3 | 1 | 8 | **16** |
| 4 | LAQ | 4 marks | 1 | 1 | 1 | 1 | 4 | **16** |
| **Total** | | | **6** | **9** | **9** | **6** | **30** | **50** |

Note: MCQ = Multiple Choice Question, VSAQ = Very Short Answer Question, SAQ = Short Answer Question, LAQ = Long Answer Question

## Cognitive Behaviour

The cognitive behaviour or processes and their proportions are drawn from the specification grid that divides the domains into knowledge, understanding, application and higher ability. Cognitive behaviour in the specification grid is distributed as: knowledge 20%, understanding 30%, application 30% and higher ability 20%. However, ± 2 marks variation is acceptable for the unit-wise weightage distribution maintaining the overall cognitive levels. It

is also emphasized that no unit can be left uncovered. It is to be noted that the question carrying 2 or 4 marks can be distributed to two or more cognitive behaviours into sub-questions. In such a case, these will be added to their respective cognitive behaviours. A brief description of cognitive behaviours is given below:

*Knowledge/Remembering* – Retrieving information from short-term or long-term memory such as recalling facts and basic concepts, identifying part(s) of organisms or materials, and identifying relationships or characteristics, common scientific instruments, symbols and scales.

*Comprehension/Understanding* – Explaining or describing scientific ideas or concepts such as classifying, comparing, discussing, explaining, interpreting, inferring and constructing meaning from instructional messages, including written or graphic communication.

*Application/Applying* – Applying knowledge, concepts or procedures to solve a new task in an unfamiliar or unique or new situation. This may involve demonstrating, solving, interpreting or concluding.

*Higher Ability* – This skill involves analysing or evaluating a task by breaking it into its constituent parts and then determining how the parts relate to one another and to an overall structure or purpose; creating new ideas and joining the elements of a task together to form a coherent or functional whole and reorganizing elements into a new pattern or structure; and making informed judgments based on set criteria and standards. The specification grid of Science and Technology covers analysis, evaluation and creation as higher ability skills.

### Elaboration of Curriculum and Specification Grid for Item Development

Items are developed to assess students' achievement of the learning outcomes prescribed in the curriculum. The specification grid provides the explication of sampling of items to make tests representative in terms of content area, cognitive behaviour and format of items. Item writers should pay attention to the following aspects when developing test items:

- Intended learning outcomes in the curriculum should be translated as objectives of the items.
- The item should be developed by representing one of the elaborated specification grid cells.
- The answer key for the selected type or the marking scheme or the rubric for the constructed response items should be developed along with the first draft of the items and revise/update the answer key/marking scheme whenever the items are revised/updated.
- The test matrix should correspond with the specification grid, but it should vary with other assembled multiple test sets.

## Item Formats

The specification grid of Science and Technology contains the following item formats:

Multiple-choice questions: Questions requiring the selection of any one of the four options are categorized as MCQ in the specification grid. Each of these items carries 1 mark. Ten MCQs are to be asked as provisioned in the specification grid.

Very short answer questions: Questions requiring the answer in one sentence, phrase or word are categorized as VSAQ in the specification grid. Each of these items carries 1 mark. Eight VSAQs are to be asked as provisioned in the specification grid.

Short answer questions: The specification grid identifies questions carrying 2 marks as SAQ. Eight SAQs are to be asked.

Long answer questions: The specification grid identifies questions carrying 4 marks as LAQ. Four LAQs are to be asked.

## Elaboration of Specification Grid with Cell Codes

For each item developed, the item should be linked to a learning outcome in the curriculum and specification grid. In order to trace individual items and sort items into different categories or assemble them in a set of tests, each item should be tagged with a code. Item code will identify specific characteristics of the items in terms of content area, cognitive skill level, item format, mark weightage as exemplified below.

**Table 12: Example of codes for the elaborate specification grid**

| SN | Area/Unit | K | | | | U | | | | A | | | | HA | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | MCQ | VSAQ | SAQ | LAQ | MCQ | VSAQ | SAQ | LAQ | MCQ | VSAQ | SAQ | LAQ | MCQ | VSAQ | SAQ | LAQ |
| **1** | **Scientific Learning** | | | | | | | | | | | | | | | | |
| 1.1 | Perform a simple survey/ research ..... | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| 1.2 | Apply safety measure ..... | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 |

Each item will be given the code number of the cell where it falls into the matrix. This code indicates the content area, the cognitive skill level, the format of the item and mark it carries. For example, an item with the cell code number 24 is from unit 1; it is a LAQ type and tests the learning outcome 1.2, is of understanding level and carries 4 marks. An item card is used to develop an item as presented in chapter one. This item card presents the learning outcome and essential metadata related to the item so that the item can be identified based on its elements.

## Test Matrix for Assembled Test Set

A test set should be assembled to represent the specification grid and multiple sets of tests can be prepared. These assembled tests are expected to differ from one another at least to some degree. To check such variation, a derivative of the specification grid or test matrix would be prepared for each of the assembled tests and superimposed with other sets of tests if it is the same or varies from others. Table 13 presents a sample of a test matrix with weightage.

**Table 13: Sample of a test matrix**

| Content | Knowledge | | | | Understanding | | | | Application | | | | Higher Ability | | | | Group wise weightage | Total weightage | Group wise weightage (required) | Total weightage required) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MCQ | VSAQ | SAQ | LAQ | MCQ | VSAQ | SAQ | LAQ | MCQ | VSAQ | SAQ | LAQ | MCQ | VSAQ | SAQ | LAQ | | | | |
| **1) Scientific Learning** | | | 2 | | | | | | | | | | 1 | | | | 9 | 3 | 9 | 3 |
| **2) ICT** | | 1 | | | 1 | | | | | | | 4 | | | | | | 6 | | 6 |
| **3) Organisms &Their Structure** | 1 | | | | | | 2 | | | | | | 1 | | | | 12 | 4 | 12 | 4 |
| **4) Bio-diversity & Environment** | | | | | | | | | | 1 | | | | 1 | | | | 2 | | 3 |
| **5) Life Process** | 1 | | | | | | | 4 | | | | | | 1 | | | | 6 | | 5 |
| **6) Force & Motion** | | 1 | | | | 1 | | | 1 | | 2 | | | | | | 18 | 5 | 18 | 5 |
| **7) Energy in Daily Life** | | | | 4 | | | | | 1 | | | | | | 2 | | | 7 | | 7 |
| **8) Electricity and Magnetism** | | | | | | | 2 | | 1 | | | | | | | | | 3 | | 3 |
| **9) Earth & Space** | | | | | | 1 | | | | | 2 | | | | | | | 3 | | 3 |
| **10) Matter** | | | | | 1 | | | | | 1 | | | | | | 4 | 11 | 6 | 11 | 6 |
| **11) Materials Used in Daily Life** | | | | | 1 | | 2 | | | | 2 | | | | | | | 5 | | 5 |
| **Total** | 2 | 2 | 2 | 4 | 3 | 2 | 6 | 4 | 3 | 2 | 6 | 4 | 2 | 2 | 2 | 4 | 50 | 50 | 50 | 50 |

The test set should fulfil the number of items in terms of format as required in the specification grid of the Science and Technology subject. The item format is checked based on the following matrix,

**Table 14: Item format requirement in the Science and Technology subject**

| Types of item | Weightage per item | No of items | | | | Weightage | | | | Total no. of questions | Total weightage |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | K | U | A | HA | K | U | A | HA | | |
| **1) Multiple Choice Question** | 1 | 2 | 3 | 3 | 2 | 2 | 3 | 3 | 2 | 10 | 10 |
| **2) Very Short Answer Question** | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 8 | 8 |
| **3) Short Answer Question** | 2 | 1 | 3 | 3 | 1 | 2 | 6 | 6 | 2 | 8 | 16 |
| **4) Long Answer Question** | 4 | 1 | 1 | 1 | 1 | 4 | 4 | 4 | 4 | 4 | 16 |
| **Total** | | 6 | 9 | 9 | 6 | 10 | 15 | 15 | 10 | 30 | 50 |

The item format required by the specification grid and the item format in the test set are compared to ensure zero differences based on the above matrix.

# Chapter 3

## Standardization of Grade 8 Test – A Way Forward

This framework is intended to guide the test development process and standardization of the test for uniform interpretation and comparability of scores. Test standardization goes beyond item development. The list of the main tasks for the grade 8 test development and test standardization is given below:

**Item development.** Drafting of items for three subjects: English, Mathematics, and Science and Technology will be based on this framework. Item development includes:

- Elaboration of the specification grid with item identifier codes;
- Mapping of Learning Outcomes onto respective cognitive behaviour sand elaboration as per the specification grid;
- Development of an item-writing training manual;

  (https://drive.google.com/drive/folders/1_Lj0X_OnG22iOSh9g8rU2sBpoc1Yyehg)
- Training of item writers – focusing on cognitive behaviours and item formats;
- Item writing as per specification grid requirements;
- Item panelling, moderation, and selection of items for pretesting.

**Pre-testing of items.** The items panelled in the expert team and moderated by the moderation committee of NEB are the final draft items. These items are assembled in a test set form on the basis of the specification grid of the subject. Test sets are then pre-tested in a national representative sample. This process involves the following activities:

- Assembling multiple sets of tests from the item bank/repository
- Sampling of schools/students for pre-testing
- Orientation for field testing
- Field testing
- Answer-book marking
- Data entry and cleaning
- Item analysis and final decision on the items
- Item selection (if needed, item review and pre-testing of revised items)
- Item bank update
- Establishment of an item banking system and training related to item banking, test assembly, item analysis, item selection, etc. Item banking will progress through the following stages:
  - Customization of item bank software
  - Selection of trainees (including Local Government level) for phased training

- Training – item banking, item analysis, reporting, etc. in separate packages and groups as required
- Functionalization of item banking up to the Local Government level
- Maintenance of item banking.

**Final test.** From the item bank, final equivalent test sets will be developed and some of the final sets will be tested in phase 1 LGs. The outcome will be published along with the model sets of tests.

**Test manual:** The final test manual will be developed to be shared with the users of grade 8 standardized test. This will also include descriptors and certification modalities.

**Communication/Dissemination.** During pre-testing and final testing, stakeholders will be familiarized with/orientated to test standardization using the model sets of tests.

**Result analysis:** Students' results/performance will be analysed by item and question group if any. The main purpose of result analysis will be to identify items in which students performed weakly, find out possible confusions/difficulties, and suggest ways to improve learning in specific ways so that students' learning can be improved in a gradual manner.

**Training to Local Governments.** The concerned personnel of the Local Governments will be trained in assembling tests for BLE, marking, data input, and reporting.

**Item bank maintenance.** Item bank maintenance training will be provided to the concerned personnel.

## References


Anderson, P and Morgan, G. (2008). *Developing tests and questionnaires for a national assessment of educational achievement*. Washington: The International Bank for Reconstruction and Development / The World Bank.

CDC. (2069 BS). *Basic education curriculum 2069*. Bhatpara: Curriculum Development Center, Government of Nepal.

CDC. (2077 BS). *Grades 6-8 curriculum 2077*. Bhatpara: Curriculum Development Center, Government of Nepal.

Izard, J. (2005a). *Overview of test construction: Quantitative research methods in educational planning, Module 6, UNESCO 2005*. Unesco.org.iiep Mod6.pdf. Paris: UNESCO.

Izard, J. (2005b). *Trial testing and item analysis in test construction: Quantitative research methods in educational planning, Module 7, UNESCO 2005*. Unesco.org.iiep Mod7.pdf. Paris: UNESCO.

Khaniya, T.R. (2005). *Examination for enhanced learning*. Kathmandu: Author.

Metsamuuronen, J. (2012). *Handbook of basics of research methods in human sciences.* Helsinki: Oy International Methelp Ltd. and Jari Metsamuuronen.

Michigan State University. (2002). Definition, characteristics, and causes of written expression difficulties. https://msu.edu/course/cep/886/Writing/page1.htm

Ministry of Education, Science and Technology. (2016). *School sector development plan, Nepal, 2016–2023*. Kathmandu: Ministry of Education, Government of Nepal.

Ministry of Education, Science and Technology. *National curriculum framework 2069 BS*. Bhaktapur: Education Review Office, Government of Nepal.

Ministry of Education, Science and Technology. (2022). *School sector education plan* (SESP) 2022/23-2031/32. Kathmandu: Author.

NEB. (2021). Letter grading guideline, 2021 provision for grading. https://drive.google.com/file/d/1dr4uTK0u23-oJjB_cIK16Ll-8XaYzq8m/view

NEB. (2075 BS). Implementation guideline for letter grading system at grades 8-12 level 2072 (NEB), 4th amendment (2075/05/01)

OECD. (2018). *"PISA for development reading framework", in PISA for development assessment and analytical framework: Reading, Mathematics and Science*. Paris. DOI: https://doi.org/10.1787/9789264305274-4-en

Redfield, D. (2001). *Critical issues in large-scale assessment: A resource guide*. Washington DC: Council of Chief State School Officers and AEL, Inc.

Thorndike, T. (2014). *Measurement and evaluation in psychology and education*(8th ed.). USA: Pearson New International Edition.

Withers, G. (2005). *Item writing for tests and examinations: Quantitative research methods in educational planning, Module 5, UNESCO 2005*. unesco org iiep Mod5.pdf. Paris: UNESCO.

The revised Bloom's Taxonomy https://commons.wikimedia.org/wiki/File:Bloom%27s_Taxonomy_for_Course_Design_and_Teaching.pdf

# Annex

## Assessment Framework Updating Team (2025)

**Technical Lead**

Prof. Ganesh Bahadur Singh, ADB

**Advisors**

Dr. Mahashram Sharma, NEB

Mr. Jung Bahadur Aryal, NEB

Mr. Subash Pant, NEB

**Language Editing**

Dr. Bal Ram Adhikari, Tribhuvan University

**Contributors in Framework Updating Workshops**

1) Dr. Bal Ram Adhikari, Assoc. Prof. Mahendra Ratna Campus Tahachal, Kathmandu
2) Dr. Mahashram Sharma, Chairman, National Examinations Board Sanothimi, Bhaktapur
3) Dr. Yashodhara Pant, Assoc. Prof., FOE, TU, Kirtipur
4) Mr. Arjun Prasad Dahal, Asst. Computer Operator, National Examinations Board Sanothimi, Bhaktapur
5) Mr. Ashik Singh, Asst. Prof. Mahendra Ratna Campus Tahachal
6) Mr. Bed Nath Adhikari, Deputy Director, National Examinations Board Sanothimi, Bhaktapur
7) Mr. Bil Bahadur Nepali, Section officer, National Examinations Board Sanothimi, Bhaktapur
8) Mr. Deb Raj Rawat, Section Officer, National Examinations Board Sanothimi, Bhaktapur
9) Mr. Gyanendra Prasad Bhatta, Education officer, Madhapur Thimi Municipality, Bhaktapur
10) Mr. Jagannanth Adhikari, Curriculum Officer, Curriculum Development Center, Sanothimi Bhaktapur

11) Mr. Jai Prakash Lal Sirvastav, Secondary Level Teacher, Prabhat Secondary School, Lalitpur
12) Mr. Janga Bahadur Aryal, Member Secretary, National Examinations Board Sanothimi, Bhaktapur
13) Mr. Kabiraj Dahal, Secondary Teacher, Nepal Rastriya Nirman Secondary School
14) Mr. Krishna Prasad Ghimire, Deputy Controller, National Examinations Board Sanothimi, Bhaktapur
15) Mr. Krishna Prasad Niraula, Seconary Level Teacher, Ananda Bhairab Secondary School, Mulpani
16) Mr. Loknath Bastola, Retired Principal
17) Mr. Matrika Subedi, Seconary Level Teacher, Ganesh Secondary School, Kathmandu
18) Mr. Mohan Poudel, Assoc. Prof. Central Department of Education T.U. Kirtipur
19) Mr. Raj Kumar Mathema, Secondary Level Teacher, Sainek Awasiya Secondary School Bhaktapur
20) Mr. Rajan Dhauhadel, Vice-Principal, Juneibe Secondary School
21) Mr. Subash Pant, Deputy Controller, National Examinations Board Sanothimi, Bhaktapur
22) Mr. Sushil Khanal, Resource Teacher, Kathmandu Metropolitan City, Kathmandu
23) Mr. Ubaraj Adhikari, Under Secretary, Education Review Office, Sanothimi Bhaktapur
24) Ms. Anupama Sharma, Secondary Level Teacher, Guhashowari Balsiksha Secondary School Kathmandu
25) Ms. Mina Shrestha, Secondary Level Teacher, Mangala Devi Secondary School, Kirtipur
26) Dr. Rabindra Shiwakoti, Assoc. Prof. Graduate School of Education T.U. Kirtipur
27) Prof. Dr. Ganesh Bahadur Singh, Professor, Tribhuvan University, Kirtipur.

## Assessment framework development team (2022)

**Technical support**

1. Dr. Mee Young Han, ACER, India
2. Prof. Ganesh Bahadur Singh, Ph. D., SSDP TA

**English**

3. Mr. Nabin Khadka, CDC
4. Mr. Shankhar Adhikari, CDC
5. Mr. Nogendra Khanal, NEB
6. Mr. Ramesh Kumar Ghimere, Vishwo Niketan Secondary School, Tripureshwor
7. Mr. Bikash Sharma, FOE, TU

**Mathematics**

8. Mr. Jagannath Adhikari, CDC
9. Mr. Ritu Shrestha, CDC
10. Mr. Bil Bahadur Nepali, NEB
11. Mr. Shyam Acharya, ERO
12. Mr. Sushil Khanal, Koteshwor Saraswati Madhyamik Vidhyalaya

**Science**

13. Mr. Yubraj Adhikari, CDC
14. Mr. Khil Narayan Shrestha, CDC
15. Mr. Mohan Paudel, FOE, TU
16. Mr. Lavadev Bhatta, ERO
17. Mr. Deb Raj Rawat, NEB
18. Mr. Boudha Raj Niraula, Shikharapur Municipality
19. Mr. Achala Thapa, Amarjyoti School, Dholahiti

**Overall**

20. Mr. Durga Prasad Aryal, Member Secretary, NEB
21. Mr. Raj Kumar Thapa, Director, CEHRD
22. Dr. Bhoj Raj Kafle, MoEST