नेपाल सरकार
राष्ट्रिय परीक्षा बोर्ड
सानोठिमी, भक्तपुर, नेपाल

Assessment Framework for Secondary Education Examination (SEE)

Government of Nepal

National Examinations Board

Sanothimi, Bhaktapur

Nepal

April 2025

Foreword

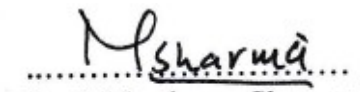
'Quality Education forAll' is considered a guiding principle in Nepalese education system. Quality assessment plays a pivotal role in ensuring quality education. Quality assessment not only communicates uniform meaning but it is also valid, reliable and fair. Secondary Education Examination (SEE), like examinations in other grades in Nepali school education, is based on curriculum and, therefore, the test of each subject in the SEE should directly align with the specification grid of the respective subject.

Reform works on the improvement of SEE started with 'Scoping for Standardized Test Development for SEE' in 2018 with an aim to 'Standardizing Secondary Education Examination with support from World Bank. Based on the 'scoping report', streamlining of curriculum competency, learning outcomes, elaboration of specification grids were undertaken. Item writers were trained, repository of items were pre-tested despite the disturbances during COVID pandemic. Item finalized based on pre-testing result were used on 2021, 2022 and relevant items in 2023 as well in the new curriculum.

Based on the standardization process and documents, this framework has updated SEE assessment framework as per new curriculum in a comprehensive manner covering four compulsory subjects, viz. English, Mathematics, Science and Technology and Social Studies. I hope researchers, planners, implementers and teachers will also benefit from this framework and it will have an impact on the improvement of classroom teaching learning through beneficial backwash effect of improved test items and testing.

I express my sincere gratitude to Prof. Dr. Ganesh Bahadur Singh, lead of the framework development team and the subject experts Dr. Tika Ram Bhatta, Sushil Khanal, Ramesh Prasad Awasthi, Hari Prasad Regmiand Madhav Prasad Dahal for accomplishing this important task in the stipulated time meeting expected quality. I sincerely thank Mr. Janga Bahadur Aryal, Member Secretary, and Mr. Subash Pant, Deputy Controller of NEB for their awesome 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 personnels from CDC, OCE and ERO for their unwavering 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
CBA	Curriculum Based Assessment
CDC	Curriculum Development Center
CERID	Research Centre for Educational Innovation and Development
CRT/M	Criterion Referenced Test / Measurement
ECD	Early Childhood Development
ERO	Education Review Office
GPA	Grade Point Average
HOTS	Higher Order Thinking Skill
HSEB	Higher Secondary Education Board
ICT	Information Communication and Technology
LAQ	Long Answer Question
LG	Local Government
LO	Learning Outcome
MCQ	Multiple Choice Question
MOE	Ministry of Education
MoEST	Ministry of Education, Science and Technology
NASA	National Assessment of Student Achievement
NCF	National Curriculum Framework
NEB	National Examinations Board
NG	Not Gradable
NRT/M	Norm Referenced Test/Measurement
OCE	Office of the Controller of Examination
SAQ	Short Answer Question
SEE	Secondary Education Examination
SESP	School Education Sector Plan
SLC	School Leaving Certificate
SSDP	School Sector Development Plan
S&T	Science and Technology
SSRP	School Sector Reform Plan
UNESCO	United Nations Educational, Scientific, and Cultural Organization
VSAQ	Very Short Answer Question

Table of Contents

List of Tables and Figures

List of Figures

Chapter I

Introduction

The structure of school education in Nepal encompasses three levels of education. Beginning from Early Childhood Development and Education (ECDE), school education includes Basic Level Education (BLE) from Grade 1 to 8 and Secondary Level Education (SLE) from Grade 9 to 12. Secondary Education Examination (SEE) is conducted at the end of Grade 10.This assessment framework has been developed based on the Education Act[1], NEB regulations and Curriculum of Grade 10, 2021 [Curriculum Development Center (CDC)].

NEB Mandate in the Public Examinations

The former Higher Secondary Education Board (HSEB) was responsible for designing curriculum and examination of Grades 11 and 12 whereas Office of Controller of Examination (OCE) used to conduct School Leaving Certificate (SLC) examination at the end of Grade 10. With the 8th amendment in Education Act, 1971 in 2016, HSEB and OCE were both dissolved and merged into National Examinations Board. Thereafter, NEBoperates all the functions related to school level examinations. NEB is made responsible to provide guidance for conducting, monitoring and evaluating all types of school examinations. Specifically, NEB has been given the following two major legal mandates regarding examinations at the school level:

- To devise or cause to devise the policy of the board based on the examination policy approved by the Ministry of Education, Science and Technology (MoEST).
- To evaluate or cause to evaluate whether or not the fairness and quality of Secondary Education Examinations are maintained.

'Fairness' and 'quality' signify vital criteria in any examinations. SEE which has been enacted by the new Education Act, 2016 (2073 BS) and put into effect from the subsequent year, is solely based on curriculum and specification grids developed by CDC. Curriculum and specification grids of all school subjects has been revised and implemented from 2022. Therefore, it is prime time to work on ensuring quality in examination to maintain spirit of new curriculum at Grade 10 as well.

SLC/SEE – Practices and Issues

SLC/SEE at Grade 10 assumes a status of high stake, nation-wide, assessment in school education system in Nepal. SEE assumes status of being a high-stake assessment holding thousands of examinees each year.

A considerable effort has so far been dedicated to improve SLC/SEE. These attempts include development of specification grid, use of marking scheme, training to the item writers, quality improvement of examination tools, capacity development for data processing during Secondary Education Development Project (Wilmut, 2001). Introduction of letter grading with grade descriptors, though rudimentary, is currently in practice. Reports, however, riterate that SLC has never been standardized which is indeed essential for quality assessment. There remain several

[1]https://www.bikashnews.com/wp-content/plugins/pdf-poster/pdfjs/web/viewer.php?file=https://www.bikashnews.com/wp-content/uploads/2023/08/education-act-2080.pdf&download=true&print=false&openfile=false

issues to be tackled and issues/questions to be answered to make SEE a valid, reliable and fair quality test. The subsequent paragraphs briefly deal with these issues and concerns.

Purpose of assessment. SLC examination has fulfilled varied purposes overtime. In a nutshell, the chief purpose of SEE is 'assessment of learning'. It has also been used to inform policy makers about quality of learning based on simplistic/crude analysis of pass percentage in the SLC examination. This analysis often brews intermittent discussion for improving quality of education for some time and it gradually subsidesby the time next SLC result gets published. Nevertheless, Education Review Office (ERO) conducts National Assessment of Student Achievement (NASA) to inform policy makers based onnational sample of large scale assessment in selected subject at grade 3, 5, 8 and 10. SEE and NASA are in practice,but each serving for a different purpose - SEE to assess and certify learning achievement of the students based on national curriculum and NASA to evaluate programme effectiveness and provide feedback to policy.

Establishing validity and reliability. Despite the fact that SEEuses various measures of standardized testing such as use of specification grid, item writing guideline/training, marking scheme, moderation, and scrutiny, actual statistical information is not calculated and published for public. As SLC/SEE does not conduct in-depth analysis of the published result as well as the overall test, it is hard to document, establish and report its validity and reliability.

Basis for interpretation. SLC and SEE have been based on the curriculum approved by the Government of Nepal. SLC had a provision of pass/fail system with a break down of obtained scores in five levels namely First Division with Distinction, First Division, Second Division, Third Division and Fail. With the introduction of SEE, there has been no provision of labelling students as fail. Students were assessed by combining the scores they obtained in theoretical and practical testsin terms ofnine levels of grading system (Please refer to annex A) without a cut score. Nevertheless, eight levels of grade with a separate grades in theoretical as well as practical tests, 'Not-Gradable (NG) label has been introduced from 2024. Accompanying grade descriptors in SEE are general and still rudimentary. Though these practices indicate a curriculum based criterion referenced measurement and absolute grading (Gronlund and Linn, 1990).Basis for interpretation of test scores is not clearly spelled out in the government documents.

Equivalency/Comparability of tests. SLC result which had provision of pass fail system showed pass percentage of students varying from year to year. For example SLC pass percentage of students in SLC jumped up and down from year 1994 to 2015– 31.3% in 1994, 49.2% in 1999, 31.22% in 2002, which leaped to 68.47 in 2009 and slumped down to 47.43% in 2015. Between these years also up and down trend continued which could not be theoreticallyharmonized and explained.

Coverage of learning domains and level of test items. School Sector Development Plan (SSDP)[Ministry of Education (MOE), 2016) pointed out inadequate coverage of learning domains in the test and emphasized that all cognitive domains as well as skills and attitudes should be measured. SSDP also remarked that the test items are commonly at "memorization and

knowledge/recall level. Critical thinking, analysis, and creativity are insufficiently emphasized". A test should not be limited to'what is easy to measure'. A comprehensive study on student performance in SLC by Mathema and Bista (2006) reported that in Science test paper there were 60% of knowledge level questions even though specification grid set knowledge level questions at 20% only. A study by Martin Chautari (2018) also found 'except for Nepali, all other subjects have 60% or more items that measure only lower level abilities'.

Uniformity and Consistency Uniformity in instruction, clarity in wording/phrasing, consistency in administration and reliability in scoring are essential for a good testand appropriate measures should be taken to ensure these qualities in a test. Specifically, answer-sheet markinghas been found problematic causing to reduce the reliability of the test.

Grade descriptors. Grades are described in general rather than subject specific way. Statements in the grade descriptors are described more in a scale having more or less of an attribute. As such it is difficult to have a uniform understanding as well as communicate message in a meaningful way. In other words, the grade descriptors in each subject must convey the competency or the learning outcomes measured. For instance, if a grade descriptor matches the test items that is considered as a proof that abilities stated in the descriptors are tested.

Calibrated item banks and statistical information. SLC/SEE items werenot pre-tested and analyzednor were they stored in the item bank. Unless the quality of test items such as validity and reliability of tests are established before they are used in a realtest, the quality of the test is questionable.There were not any statistical procedures to equate tests or analyze items in the SLC/SEE either. Item quality is also not ensured before its use in the test. Report generation is also not in practice in SLC/SEE beyond analysis of pass and fail. A particular drawback of the data collection has been that student data are entered and reported in an aggregated way, i.e. only the total scores per subject for each student. This makes it impossible to do any regular test and item analysis, be it classical or IRT-based, which would be indispensable for any form of standardization of a test.

In an attempt to improve SEE, NEB developed items for Science, Mathematics and English subjects of Grade 10. Those items were pre-tested and analyzed. These analyzed items were administered in the years 2023and 2024 for the SEE examinations. However, CDC has developed new curriculum and introduced it at Grade 10 from 2024.

Use of ICT. Use of ICT is very limited throughout the assessment/examination process. There has been discussion on use of ICT such as for paperless examination system, test on demand for open and accessible assessment, but ICT use is limited at present.

Above issues indicate that concerted effort is required to improve and enrich SEE for its effective implementation and to derive useful meaning from it. Clarifying concepts for test improvement and underlining associated activities are necessary to improve assessment system in SEE in a systematic and technically correct way.

Purpose of Assessment Framework

One of the main purposes of public examinations is to assess current level of students' learning in the subject and certify their achievement. There is also an ample opportunity to glean the reasons for poor performance in specific content area and improve teaching learning subsequently. Students and their learning is always on the centre of curriculum, teaching learning and assessment. Assessment should contribute to classroom teaching learning whether it is

formative assessment or summative one. This would be central focus of the assessment framework.

SEE is curriculum based criterion referenced test which should fulfill requirements of specification grid to make the test valid. Specification grid is explicit about format of the item, cognitive level to be fulfilled and weightage per item or group of items. Assessment framework should pave ways to ensure requirement of the curriculum and specification grid.

'Teaching for testing rather than learning' is still a dominantphenomenon in Nepali school education despite the fact that backwash effect of test is crucial element in shaping classroom teaching learning. Influence of testing in teaching practices is also seen as an opportunity to improve classroom teaching learning practices by improving quality of test items - less subjugated by recall items, promote thinking and creativity in the students and provide reliable base to determine level of achievement of students for certification. Such improvement in the test will create an environment for achieving beneficial backwash effect – improvement in testing will help improving classroom teaching learning (Hughes, 2003; MoEST, 2016). For this reason School Education Sector Plan (SESP) accounts for making teachers responsible and accountable for "improvement of teaching and learning activities by evaluatingteachers' performance based on, among others, students' result in examination (MoEST, 2022, p. 66). SESP also expresses concern that 'teachers need to be properly trained to analyse the results of school-based assessments and use them in learning improvement in a systematic manner.' (p. 57).

Chapter II

Test Development

Test results should be reliable, fair and adequately represent what it intends to measure. The test must ensure uniformity in terms of instruction given in the test, its administration, scoring, and interpretation. Uniformity in item construction, interpretation and communicationof test results in a reliable and valid manner is essential. A good test should be developed based on sound design, proper test development process, and implemented in a planned way. Development of the test is founded on certain premises such as purpose of the test, coverage of the test, and basis for interpretation of the test results. Based on these premises test is developed; pre-tested, analyzed and revised in several rounds; and final tests are prepared to meet quality of test. Once standardized test is developed and used, rounds of improvement activities continue. Along with the concept of test development, this chapter delineates associated activities to guide through the test developmentprocess in a systematic and technically correct way.

Ascertaining Basis for Test Development

Test development of a test follows specific guidelines and processes. It aspires a team effort and uses step-wise technical process as described in the subsequent paragraphs.

Establishing purpose of the test: Large scale test may serve various purposes such as certification of students' learning achievement; prediction of future learning (selection for higher grade/education); evaluation of school, teacher, programme, curriculum; draw implications for improving curriculum, curricular materials, teaching methods; motivating student, teacher, school. In 'A Nation at Risk', Gardner and others (1983) observed standardized tests of achievement to be administered "at major transition points from one level of schooling to another and particularly from high school to college or work". This report states purposes of such tests to: (a) certify the student's credentials; (b) identify the need for remedial intervention; and (c) identify the opportunity for advanced or accelerated work. In the similar line, UNESCO (2018) identifies purposes of testing as selection for controlling access to different levels of education; accountability through evaluating the effectiveness of instruction and schools; and certification by obtaining knowledge of and reporting on student achievement.

There can be one or more of these purposes in a test. These purposes directly influence the test design and development as well as interpretation of test result. If purpose of the test is multiple it will increase (sometimes confuse) complexity in interpretation and utilization of test result. The purposes of the test are crucial starting point from where the activity of item writing really starts (Ross, 2005).

In case of SEE the main focus is on certification of student achievement in terms of grade and Grade Point Average (GPA). The subject-wise grade and grade points along with cumulative GPA are used by the higher grade institutions for the selection of students for specific subjects. In this regard, SEE also fulfils purpose of student selection. But SEE focuses on curriculum of Grade 10 only, not pre-requisites of higher education. In Nepal, large scale sample based assessment (i.e. NASA) is conducted by ERO to evaluate the education system to support in obtaining results as nuanced in the curricula. The objective of NASA is to create a reliable database on the learning level on selected subjects and grades for benchmarking; compare learning achievement data with earlier achievements in order to monitor the progress over time; and to generate recommendations for policy making to improve learning level of students (ERO,

2015). As evaluation of the education system and input into policy making are the focuses of NASA, SEE can be relieved from these tasks and focus on generating valid and reliable certification of learning achievement at the level of individual student. However, SEE can also provide ample opportunity for item-wise performance of the students, analyse reasons for the poor performance (if any) item by item and make suggestion how classroom teaching learning can be improved, ultimately leading to better performance of the students in subsequent year(s).

Setting standards. 'Standard Setting' in educational assessment usually refers to the procedures used to determine the meaning and value of specific scores: which levels of achievement or competences are reflected by which scores, and at which scores should candidates be given or denied access to further studies. To enrich reporting system, standards should be explicit and closely represented in the curriculum, specification grid and test so that it communicates the standard student can meet (Hill, 2013). Quite often, standard setting takes on the meaning of determining the minimal score that is seen as sufficient achievement, or to pass a subject test.Traditions around pass/fail decisions vary considerably across countries and in some countries it is not even an issue (Bakker, 2010). Grade, level or score, whatever is used to report student's achievement should carry some meaning with it.

SEE uses eight grades including Non-Grade (NG) with a cut-off scorefor measuring student's learning achievement from 2024. These grades are described in general terms for overall performance, not specific to any subjects. However, these grade descriptors raise expectations that may not be realised in SEE tests. The competences in the proposed scheme themselves and the fine distinctions between the levels of mastery of these competences cannot easily be operationalized in test items. And only having a limited number of items it is difficult for grouping students according to the fine distinctions suggested by the grade descriptors. An SEE test is a representative but limited sample of the vast domain of knowledge, understanding and skills students are supposed to have mastered at the end of secondary education. An SEE score is an indication of mastery of the curriculum for a certain subject as such, but brings little information on mastery of sub-component or specific competences. SEE scores may be converted into grades with the help of a simple table as is the current practice. It could be suggested that expert groups, with the help of validated standard setting procedures, should decide on a cut score between what is seen as sufficient and insufficient mastery of the curriculum. To add more meaning, some broader score intervals might be selected and tied to the nature of items which students with a score in that interval are able to answer correctly, e.g.

- Score 80-100 A+, A: *Advanced*; students are able to answer correctly almost all items requiring reproduction of facts and application of procedures and almost all items that require higher order thinking skills.
- Score 60-80: (B+, B) *Fully proficient*; students are able to answer correctly most items requiring reproduction of facts and application of procedures and most items that require higher order thinking skills.
- Score 40-60: (C, C+) *Proficient*; students are able to answer correctly most items requiring reproduction of facts and application of procedures and particularly a few items that require higher order thinking skills.

Pretesting of items and psychometric test and item analysis should guarantee that scores, and especially the cut score, carry the same value over the years, or in other words, that the

examination in one year is not more difficult or easier than the examination in the next year. More is yet to be done on this area.

Coverage of test. A test should reflect objectives and standards being set. The items in the test should manifest the objective(s) of the test and ideally standard set could be backtracked to the items and scores. Test specifications grid, also called test blueprint is an explicit plan to ensure representation of test objectives in the test. Test specification grid should indicate cognitive processes desired and the description of content to be covered (Kubiszyn & Borich, 2016; Thorndike and Thorndike-Chirst, 2011). Specifications should also be provided in terms of the test format, length and item types. These specifications are important for ensuring the content validity of the test: does the test reflect the content, concepts and competences the test is meant to assess?

SLC and now SEE used to cover content and curriculum intents as set in the curriculum which is implemented nationally. Studies have raised several important issues regarding the coverage in the test (Centre for Educational Research and Development [CERID], 1988; Singh, 1996; Wilmut, 2001; Mathema & Bista, 2005; Martin Chautari, 2018) thatthere is dominance of memorization/recall type of questions; textbook is the base for the test development rather than curriculum; and the examination is limited by what can be tested in paper and pencil (written examination). In the existing test specification grid higher ability test items are also mentioned, but in actual test they appear to be lower in number than that of the requirement of the test specification grid. Review of curriculum, specification grid and test items of SEE were undertaken in Science, English and Mathematics subjects by the subject experts[2] in 2018 using test papers of three years. Major findings from these reviews were as follows:

Congruence of specification grid and test items coverage: In terms of weightage of marks in the specification grid and test paper, there was high congruence in Science, Mathematics and English language subjects. In terms of behaviour to be tested as required by the specification grid, there are differences. For example in the chemistry part of Science subject, 20% Knowledge, 30% Understanding, and 50% Higher Abilities items are expected to be constructedin accordance with the requirement of the specification grid. Nevertheless, analysis of test items showed that 50% of Knowledge level items, 27% Understanding level and 23% Higher Abilities items were asked in the actual test. In case of Mathematics, the question paper did not adequately cover application level items nor did they represent need of constructivist teaching approach, which focuses on the importance of previous knowledge while learning a new concept and presents the contents to solve the problems. In case of English language, test item adequately covers the item format as per the specification grid, but in case of cognitivebehavior, several items in the reading section fall mostly at the scanning level intending to measure lower-order skills.

Error in typing, instruction and translations: In the test papers, several mistakes were found in typing and translation of the question/item. Such errorsoften make item incomprehensible and in few cases it has given entirely different meaning in English and Nepali version of the question. For example in SEE question paper 'RE-111''MA'Q. 2 a, English text, '*Liquid column'* is given but correct *is 'liquid column'*. Another example from Science is Q. 8 where question confuses the

[2]Science – Prof. Dr. Rajani Rajbhandari; English Language – Mr. Netra Prasad Sharma; and Mathematics – Mr. Bed Prasad Dhakal.

students as first part asks about meiosis cell division and then ask about changes occurred in anaphase stage of mitosis. This kind of question lacks subjective coherence.

Marking scheme: A good marking scheme explains how student responses to assessment tasks are evaluated. It identifies assessment criteria and articulates qualitative standards of achievement for each criterion.However, faulty marking schemesare mostly indicative of making it open and liable to subjective interpretation. In multi-set of examination papers as well marking scheme differ in the similar test item as well. For example, in order to assess free writing task, the marking scheme prepared for one set of question allocates mark for vocabulary and grammar whereas marking scheme of another set of questions ignores vocabulary and allocates marks for grammar only.

Curriculum reviews were shared in an expert meeting. Experts also suggested that low number of higher ability items might be due to the fact that developing higher level test items are difficult and time consuming. Another important reflection was that there is a lack of training and orientation to the item writers. The reiteration of these issues evidently indicate that test specification grid needs to clearly point out what is to be tested (reflecting objectives, standards, cognitive level) and how it is to be tested (test format, length, item types, weightage). Item writers need training in developing higher level test items, supported by an adequate scheme for classifying items according to cognitive behaviour.

The item format used in SEE is mainly short answer and long answer type. There is a separate question paper and answer sheet. There has been discussion and trialling of using Multiple Choice and long (analytical) answer format as well as having question and answer writing space in the same answer sheet. Dummy SLC examination in 2000 used a version of separate question paper and answer booklet and another version having question and writing spaces in the same booklet were compared (Wilmut, 2001). Single booklet combined with questions of both selected and constructed response format as well as writing space were found to be more practical in several respects.

In case of format of test items, selected response items allow measuring broad range of cognitive skills in a limited amount of time (Lane, Raymond, Haladyna and Downing, 2016). Although constructed-response items, as viewed by Phelps (2007) are of lower reliability and greater cost, they allow students to be creative and demonstrate a more in-depth understanding of content.

Purpose of the test: Interpretation of test results is also related to the purpose of the test. If the purpose of the test is to compare performance of one student against others, comparing against the group can serve the purpose and is termed as Norm Referenced Test/Measurement (NRT/M). As prime purpose in NRT is to compare performance of students, item discrimination is given priority. If the item is too easy or too difficult, its discriminatory index will be low and thus inappropriate for the test even if these would be important learning objectives in the curriculum. Comparison of students' performance can also be against a set standard irrespective of performance of other students in the test – a Criterion Referenced Test/Measurement (CRT/M) (Reynolds, Livingston and Wilson, 2011). In CRT criteria/standard is preset and usually high level of achievement is set. Curriculum is also considered as criteria and where a set curriculum is the basis for test development and administration, and students' performance is compared against the curricular objectives, the test is termed as Curriculum Based Assessment (CBA) or it can be termed as CRT based on curriculum (Thorndike and Thorndike-Christ, 2011). In CBA correct answer by most of the students in some area of curriculum and low achievement in some

other areas does not matter, both conditions provide valuable information. If the purpose is to verify students' mastery over the curriculum, in order for instance to receive a school leaving certificate, the curriculum is the criterion for an achievement test. However, the SEE is also used for student admission to higher studies. For that purpose, scores should also indicate whether a student would do well in higherstudies, or in other words, should have predictive validity. Predictive validity is usually achieved by including items that test skills which are critical for success in academic studies. The specification of a test should be decided in view of the purpose of the test that is on the inferences that are going to be made based on the student outcomes. Sometimes, tests may serve multiple purposes but decisions should be prior to test development process and expressed in the tests specification and other elements that serve to standardize the test.

Action steps to ascertain basis for test improvement

Based on the discussion on the issues and concerns related to SLC and SEE and conceptual deliberation on important aspects for test improvement– some actions were found already in place, some were to be clarified and some require actions as listed below:

1. The purpose of SEE should be on valid and reliable certification of learning achievement at the level of individual student at the end of Grade 10.
2. As the nature of SLC and SEE has been curriculum based assessment, interpretation of score is mainly criterion based on curriculum. However, subject grade and GPA in SEE is used for selection for higher grade entrance test, it is necessary that SEE ensures predictive validity and manifest standard required for higher studies.
3. Grade descriptor should be made specific and tied with the curriculum. Subject-wise grade descriptor should also be developed to tie up with the test items.
4. Specification grid should be improved and updated to clarify required cognitive behaviour, item format, weightage, and number of items. Clear guidelines should be developed to evaluate item meet the standard and intended requirements.

Item Writing, Pre-testing, Analysis and Item Banking of Accepted Items

Once purpose of the test is established, standard is set, interpretation of score is decided, and specification grid is finalized, next important step in test development is item writing. The test should meet its purpose and the items should match a well-designed specification grid focused on one aspect of learning (Anderson and Morgan, 2008). Three qualities should be met by item writer or item development team – adequate content knowledge, proper use of item format and development of item at required cognitive level including higher order thinking skill (HOTS) items. Training to the item writers, rostering of capable item writers and periodic capacity development is essential.

There are some important steps for the preparation of item development: 1) elaboration of specification grid in order to identify possible items by singular item objective as per learning objective and specification grid requirement as given in the curriculum; 2) assigning items to the item writers in which item card is used with necessary metadata and marking scheme[3]; 3) item

[3]The scoring procedures and any assessment criteria should be developed simultaneously with the test development (Withers, 2005). Along with the item writing, marking scheme and rubric should also be developed and tried out.

panelling in an expert team[4]; and 4) moderation of the items and item selection for the item pretesting. Item card is used to write item with required metadata as given below:

---

1) Subject: Science and Technology/Mathematics/English of Grade 10

2) Item cell code:

3) Elaborated item code

Unit	LO	Cognitive Level	Format	Marks

3) Elaborated item code (for English):

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

4) Learning outcome (from curriculum): ……..…….....……..……
……..……..……..……..……..……..……..……..……..……..……..…

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

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

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

Marking schemeis prepared by the item writer along with construction of the item.Together with the items being panelled, moderated, pre-tested and finalized, marking schemes are also revised accordingly. Marking scheme provides guidelines for assigning marks as per weightage to each test item which is mainly based on the elements expected as answer. One mark per element is desirable. The marking scheme should outline all or most of the expected answers with the guideline for awarding marks. This will help to ensure consistent and uniform marking across different examiners. This helps maintain fairness and reliability in the evaluation process, enabling accurate assessment of students' performance.

Draft of items selected by the item writing team should be pre-tested in a sample for which the test is to be used. As there will be vast number of items, not all the items can be tried out with the same sample of the students. Since the number of items for item bank has to be huge in number, testing a sample of items with all the students sampled would not work either. It is, therefore,

[4]The review before the items are tried should ensure that we avoid tasks which are expressed in language too complex for the idea being tested, avoid redundant words, multiple negatives, and distracters which are not plausible. The review should also identify items with no correct (or best) answer and items with multiple correct answers. Such items may be discarded or re-written. (Izard, 2005a).

required to sample both students and items so that some students give answer to some of the items only. Such matrix sampling can be used in pre-testing to try out the vast number of items in short period of time (Redfield, 2001).

The larger the size of the trial population pretesting is much better, but it 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 chance of loss of data during pre-testing, minimum 200 students should be maintained (Anderson and Morgan, 2008)[5]. Anderson and Morgan also suggest 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 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 track an item from development phase to finalization. After trialling items, they are scored, data input/cleaning, and analysis is done. Items are analyzed for their difficulty level, discrimination index and in case of multiple choice items power of distracters, etc. During trialling phase, acceptable and pathological/defective items are detected through logical item characteristics using IRT or other programmes or manually (Metsamuuronen, 2012; Izard, 2005b). Item which meet acceptable quality are banked for later retrieval. Item banking software would be helpful and it is needed to discuss and decide whether to develop a proprietary version in-house or outsource a local provider of IT services, or buy a commercial ready-made one.

Based on the discussion on the item writing, pre-testing, analysis and item banking of accepted item, these action steps are essential: 1) Organize item writing training. Subject-wise item writing teams should be developed and items are developed by the panel of item writers. Building national capacity in item writing is highly preferred. 2) Develop 'Manual for Item Writers'. 3) Prepare 'Guidelines for Pre-testing, Item Analysis and Item Selection Procedures'. 4) Develop national/institutional capacity on item analysis, item banking, and other statistical purposes.

Final Test and Establishing Quality

Pre-testing helps in selection of quality test items, but a complete set of items in a final test paper format are not yet assembled and tested. By this time, test manual[6] 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

[5]Pre-testing of the items (30 sets of tests in English, Mathematics and Science and Technology subjects of grade 8 have been pre-tested by NEB in a national sample in which each of the subjects was tested at the same day by providing number of sample students at the school with alternate subjects and within subject also alternate test set so that at least 30 sets x 3 subjects = 90 students will have none of the sets repeated in any of three subjects in the same examination center.

[6]The test manual explicitly states the purposes and applications for which the test is recommended and identifies special qualifications required to administer the test and interpret it properly. (Redfield, 2001).

established and reported for the final test in a larger sample. For the final test and establishing quality, thesetwo action steps are suggested: 1) Develop 'Test Manual' which should be used for the training and testing to/by the test administrators, examiners and others and 2) Establish and publish reliability and validity of the final test.

Test Administration and Ongoing Test Maintenance

Thus developed test is ready to be administered in the regular examination. Final test developed in short time would require improvement and further development. Similarly, item bank needs restocking with fresh calibrated quality items through ongoing test maintenance activity in institutionalized way. Restocking of item bank and ongoing maintenance of test should be institutionalized.Expert team should be developed and available to support as and when required. For the test administration and ongoing test maintenance, these action steps are suggested: 1) Continued restocking of calibrated items should be made 2) Item writing team should be developed and used. 3) Ongoing maintenance of test item should be undertaken for further development and enrichment.

Certification

A validated standard setting procedure (e.g. Angoff, Bookmark or similar), informed by statistical analyses of student outcomes, should be run to determine the cut-score between sufficient and insufficient. This score should become part of a table for converting raw scores in 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. For certification, these action steps are suggested: 1) Review existing grade descriptors and suggest for refinement. 2) Develop subject-wise grade descriptors in English, Science and Technology, Mathematics and Social Studies to start with subject-wise grade descriptors.

Result Analysis

Skill in item development is not only important for development of a quality test, this skill is also important to anlayze and utilize student's performance in the test and even in an individual item level and plan learning improvement of the student(s) (Khaniya, 2005). Therefore, item writing training should be down streamed to the practicing teacher to improve classroom teaching learning and improve learning achievement of student. This will help teacher understand nature of the item, purpose of the item in relation to the curricular intent for student's learning and designing an appropriate classroom teachinglearning environment to improve student's learning.

Result can also be analyzed at different level such as school, Local Government (LG), District, Province or National level to identify poorly performed items at these level in composite. Poorly performed items can be anlayzed into three aspects, viz. 1) what the item demands, 2) why most of the students were not able to do well, where students might have confusion, then 3) how to improve classroom teaching learning so as to improve learning in such area/content. Such analytical report with specific suggestion can help in gradual improvement in classroom teaching learning and lead to improvement in learning achievement.

Chapter III

English

English, being an international language, forms a basis as a core subject in Nepal. National Curriculum Framework (NCF) of school education establishes English as a compulsory subject to be taught from primary to secondary levels (CDC, 2076, B.S.). It calls for the standardized assessment to produce competing human resource in future. This chapter explicates how reading and writing skills of the students appearing in Secondary Education Examination (SEE) are assessed. It describes the elements of summative assessment of the English language as a subject in the SEE. The chapter also explains how students' performance and metacognitive behaviour in reading and writing practices are measured and reported. This framework also provides important guidelines to the item developers to assess students' levels of learning in the English language.

Level-wise Competencies, Learning Outcomes and Assessment Domain

The curriculum for Grade 10 aims at providing the young adults with a definite linguistic base in English for their future academic and career engagements. The curriculum expects to achieve nine competenciesby the end of Grade 10 (CDC, 2077 B.S.). For instance, the students can communicate with reasonable accuracy and confidence on familiar topics and can read a variety of texts for information and understanding.

Learning outcomes in the English curriculum expect varied skills from the students. All the LOs under each skill may not present the scope to be tested as some of the LOs may need to be assessed in other ways. The curriculum envisions that the students can extract relevant information from the texts as a learning outcome, for instance, to evaluate the reading skills of Grade 10 students. Out of 36 learning outcomes, the curriculum outlines 10 learning outcomes for assessing reading and writing skills each, six for listening and 12 learning outcomes have been outlined for assessing speaking skill.Since grammar and vocabulary are equally significant for developing all four skills of language, the curriculum aspires to teach them in an integrated manner. The same applies for testing as well. They can be assessed via listening, speaking, reading and writing in an incorporated way. (See curriculum for detailedcompetencies and LOs)

The curriculum aspires the assessment of all four skills: Listening, Speaking, Reading and Writing. Listening and speaking skills are assessed in internal assessment whereas Reading and Writing skills are assessed in external assessment.

Assessment of Listening: Assessment of listening is done during the classroom teaching and learning. Each assessment needs to be recorded. (See internal assessment section for details.)

Assessment of Speaking: Assessment of speaking is done during the classroom teaching and learning. Each assessment needs to be recorded. (See internal assessment section for details.)

Assessment of Reading: Reading can be 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 the society (OECD, 2018). Reading refers to an active and complex process of understanding a written text, developing and interpreting meaning and using meaning as appropriate to type of text, purpose and situation. It involves both literary and informational texts. Literary texts cover passages from fiction and science fiction, drama, monologue, and poetry as well as nonfiction texts such as essays, autobiographies and biographies, news stories,

newspaper articles, diary entries, book or film reviews, travelogues and so on. The informational texts include exposition, argumentation and persuasive texts along with procedural texts and documents as well as correspondence texts such as letters, emails, blog posts, manuals, etc. There will be four reading texts: two from textbooks and two from outside but from authentic sources. The reading passages are selected only via expert judgement but not based on readability formulae for passage selection i.e. text analyzer.

The present curriculum mentions broader learning outcomes for testing reading comprehension. However, item developer needs to consider certain abilities of reading comprehension while developing the items. These abilities include: locate explicitly stated information; identify explicitly stated main ideas and supporting details and summarize what is read; determine the main idea of a text by thinking about the details in the text and summarize the information in students' own words; interpret and use information from charts, graphs, diagrams, timelines; explain how authors support ideas by noting the similarities and differences in the point of view they represent; infer the meaning of words from the given context; accurately quote and draw inferences from the given text; explain what a piece of information teaches by referring to details and examples in the text; infer implicit information using context clues and evidence; use information to demonstrate understanding; compare and contrast the most important points of a piece of text; compare and contrast characters, settings or events; understand figurative language, word relationships and nuances in word meanings; describe how a narrator's or speaker's point of view influences a story; refer to details and examples in a text when explaining what the text says explicitly and when drawing inferences from the text; identify a theme of a story, drama or piece of poetry; compare and contrast the point of view from which different stories are narrated; reflect on the theme of a story, drama or a piece of poetry; determine the purpose of the text; interpret multiple levels of meaning such as literal, contextual and intended; appreciate literary texts of grade level; determine the writer's attitude, perspectives, purposes; express opinions and make judgments about ideas, information, experiences presented in the text; arrive at a conclusion; can refer to the referents; can reorganize the ideas and construct students' own meaning; can make predictions about upcoming events in the narrative texts.

Consideration on Text Selection and Text Complexity

The selected text for unseen reading comprehension should be appropriate in content and relevant to the students' age at the target grade level. They should contain materials that students are not likely to have encountered; for example, it should not come from a textbook or other frequently-used teaching resource. The text should be factually accurate; be grammatically correct; be appropriately illustrated; be appropriate in terms of socio-cultural context; be gender-sensitive; be fair, that is, equally accessible for students from all backgrounds likely to take part in the assessment; and be inclusive in the usage of texts and graphics to support and extend meaning. The text should not make students/test-takers emotional. Therefore, test developers should be very careful while selecting the text. The text should be grade level appropriate in terms of difficulty. The following section briefly summarizes the nature of low, moderate and high complexity of the text for reading comprehension.

Assessment of Writing: The ability to write well is essential to productivity and to personal and social advancement. Good writing instruction empowers students to acquire new knowledge and to develop critical thinking skills. The curriculum envisions to encourage the student writers to gradually move from prescriptive to more creative writing approaches. Therefore, both guided and free writing tasks are included in the assessment. Both of these writing tasks measure the students' ability to convey experience, real or imagined, in order to communicate individual and

imagined experience to others. It also measures the students' ability to persuade in order to change the reader's point of view or affect the reader's action. Basically, writing assessment aims to assess the students' ability to inform others of ideas and concepts to readers of information-driven society. There will be four writing tasks: two guided and two free writing tasks.

The present curriculum does not specify the cognitive levels of the writing tasks. However, the item developer needs to focus on giving stimulus or contexts to clarify the writing tasks to measure the expected learning outcomes. The writing task should address real-world, age-appropriate and grade appropriate topics which must be familiar and accessible to students and not controversial in nature. The topics need to encourage the use of effective approaches to thinking and writing. The purpose of the writing task should be clearly stated and be consistent with the audience identified in the writing task. Similarly, the audience should be specified or clearly implied by the context of the task so that students will make choices within parameters provided by the writing task. For example, a writing task intending students to persuade a classmate to read his/her favourite book and the writer would choose the book to write about. The guided writing must provide students with the content while the free writing task must provide students with the stimulus.

The communicative purpose of the writing tasks are to persuade, to explain, and to convey experience. Example audience are principals, parents, peers, community leaders, teachers and government officials, etc. Example approaches to thinking and writing are analyzing, describing, evaluating, narrating, summarizing, synthesizing, reflecting, questioning, and arguing. The writing tasks are measured in terms of depth and complexity, approaches to thinking and writing, details and examples relevant to content while text structure, coherence and focus are measured in organization of ideas. Sentence structure and sentence variety, word choice, voice and tone, grammar, usage and mechanics are other facets of measuring writing responses.

Assessment of Grammar: The curriculum envisions assessment of grammar through reading and writing tasks. However, there will be two types of grammar tasks: reproduction and grammar in context as given in the test specification grid.

Specification Grid

Specification grid is a roadmap of assessment. It mirrors a clear image of the items for the item developers. Without creating a good test, it is almost impossible to obtain valid and reliable result. It is a must for any high-stake test.

Table 1: Specification Grid for Reading

Area	Text types	Content	MultipleChoice	Sequencing	Fill in the gaps(withoutcho	Matching	True/False	Short AnswerQu estion	Vocabulary Question	Marks
Reading	**Seen text1 (100 words)**	Any text of appropriate length from the textbook. (5 items of any one type)								**5**

	Seen text2 (200 words)	Any text of appropriate length from the textbook. (10 items of any two types)								**10**
	Unseen text1 (200 words)	story, notice,advertisement,product guide,letters, newsstory, brochure,instruction/manual, diaryentry,interview,letter,recipe (10 items of any two types)								**10**
	Unseen text2 (300 words)	biography,autobiography,essay,newspaper/magazine article, sciencearticle, book/filmreview,travelogue,editorial,report,BlogSpot (15 items of any three types including one type of items for vocabulary testing)								**15**

N.B. Out of 40 items, reading comprehension questions comprise of 16 literal comprehension, 8 reorganization, 8 inference and 3 evaluation and reflection types of questions and 5 questions for testing vocabulary.

Table 2: specification Grid for Writing

Writing	**Guided Writing1**	Any one task : Paragraph, description of tables/charts/diagrams, a set ofinstructions, recipe, menu, a set of rules and regulations, advertisement and notice **Expected response:** about100 words	**5**
	Guided Writing2	Any one task : News story, skeleton story, message of condolence, message of congratulations, invitation letter, thank you letter, biography **Expected response:** about100 words	**5**
	Free Writing1	Any one task : Paragraph (presenting views, opinion, experience, feeling), leave application, job application, dialogue **Expected response:** about 150 words	**6**
	Free Writing2	Any one task : Personal/official letter, letter to the editor, email, short essay (narrative, argumentative, descriptive), diary, newspaper article, editorial , book review/film review **Expected response:** about 200 words	**8**
Grammar	**Reproduction**	6 tasks representing each of Tense, question tag, reported speech, voice, and interrogation and negation	**6**
	Grammar in context (MCQ in text form)	10 MCQs representing articles, prepositions, tense, question tag, voice, reported speech, connectives, conditional sentence, subject verb agreement, causative verbs, modals, adjectives and adverbs, and relative pronouns	**5**

Assessment Types

The curriculum recommends the inclusion of both internal and external assessments. The internal assessment is conducted to assess listening and speaking skills of the students. It includes 12% for participation and attendance, 32% for listening and speaking each and 24% for score from terminal examinations. The external assessments will be used to assess student performance in reading and writing, and, generate the appropriate grades. The internal assessment will carry 25% of weightage and the external assessment will carry 75%. External assessment covers 40 marks for reading and 24 marks for writing and 11 marks for grammar. The external assessment focuses on reading comprehension and writing along with grammar and vocabulary tasks. This section presents the SG on each assessment area in the external assessment.

Item Format

The present specification grid of English for Grade 10 presents the following item format for assessing reading and writing skills of the students via external examination.

Item Format for Reading

Four reading texts of different genres will be asked. Two short texts are taken from the Grade 10 English Textbook and two short texts are taken from outside source. In order to assess reading comprehension of the students the following item formats will be used.

A. True/False/Not Given items

TRUE/FALSE/NOT GIVEN statements are one of the most difficult items to create. The statements are true only if the given statement gives the same but not similar meaning given in the passage. They are false if the given statements contradicts the information given in the text. Not Given statements are not false statements. They reflect the similar information but this information is not given in the text.

Example: Read the following text and decide whether the given statements are TRUE/FALSE OR NOT GIVEN.

READING TEXT 1

Last month, at an **auction** house in Germany, a single page from a **medieval** manuscripts went up for sale. Among those who were trying to buy it was David Gura, the curator of Ancient and Medieval Manuscripts at the University of Notre Dame. It was another chance for him to save part of a 15th century Book of Hours, which only a few years earlier had been broken up. He is now in a race to find the **remaining** pages of this **manuscript** before they **disappear**.

a. David Gura was trying to buy a book before it gets broken up.
b. David Gura worked for the University of Notre Dame.
c. Only a small number of pages remained of The Book of Hours.

Table 3: Item Explanation

Statement	TRUE/FALSE/NOT GIVEN	Close matching information	EXPLANATION
David Gura was trying to buy a book before it gets broken up.	FALSE	A **single page** from a medieval manuscript went **up for sale.** Among those who were trying	This statement is FALSE because the text clearly states that a single page went up for sale, not a book.

Statement	TRUE/FALSE/NOT GIVEN	Close matching information	EXPLANATION
		to **buy** it was David Gura	
David Gura worked for the University of Notre Dame.	TRUE	**David Gura, the Curator** of Ancient and Medieval Manuscripts at the University of Notre Dame.	This statement is true because David Gura is the curator of ancient and Medieval Manuscripts at the university.
Only a small number of pages remain of The Book of Hours.	Not Given	He is now in a race to find the **remainingpages** of this manuscript before they disappear.	There are remaining pages but it is difficult to conclude how many.

B. Fill in the gaps items

Fill in the gaps items contain questions that contain individual sentences or sometimes a paragraph with missing information with blank space. This space represents a word, a chunk of words or expressions, phrases, number, symbol that an item developer has left out on purpose. The students are expected to fill in the exact word, phrase or information in order to get the answer correct. Fill in the gaps items mostly test the recall power of the students. However, it may also contain higher ability items as well. It is important to note that these questions have only one accurate answer.

Example: Read the given text and fill in the blanks with the correct words highlighted in the text. [**Refer to Reading Text 1**]

a. Our teacher has just submitted the …… of the grammar book to the publisher.

C. Multiple choice questions

Multiple Choice Questions (MCQs) refer to questions in which students/respondents choose a single option as the correct answer from the alternatives given. They consist of a stem i.e. the question and several alternatives i.e. options. A multiple choice item for Grade 10 has a minimum of four alternatives. Out of the given alternatives, one option is the answer i.e. KEY and other three options are called 'distracters'.

Example: Read the text and choose the correct option to answer the question that follows.

READING TEXT 2

One day Kintaro took a hatchet, climbed onto the bear's back, and went off into the mountains with his friends. On the way, they came to a cliff overlooking a big stream and found that there was no bridge across. "I'll knock a tree over and make us a bridge," said the bear. But even though he pushed and pushed, the tree didn't budge. "I'll give it a try," said Kintaro, and he began to push with all his might. When he pushed, the tree began to move, and with a loud crack it fell over and spanned the river. "Hooray!" shouted all of Kintaro's friends.

Which sentence tells us about the most important event in the story?

A. Kintaro climbs a mountain and finds a stream.

B. The bear carries Kintaro on its back.

C. The bear helps Kintaro to build a bridge.

D. Kintaro pushes a tree with all his might and makes a bridge.

D. Matching items

Matching items involve questions with two parts in which students require matching a series of **stems or premises** to a **response** or **principle**. They consist of a set of directions, a column of statements and a column of responses. Matching items in each column must be kept as homogeneous as possible and the list of responses need to be arranged systematically as far as possible either chronologically, alphabetically or numerically.

Example: Read Text 1 above and match the meaning given in column A with the words given in column B.

Column A	Column B
I. a book, document, or piece of music written by hand	A. auction
II. a public sale of goods or property	B. medieval
III. relating to the middle ages	C. manuscript
	D. remaining

E. Ordering of information

With ordering questions, students select the correct sequence of a series of items based on the text. The text based on historical events or a biography of a person has events on chronological order. Item developers can ask ordering questions expecting the correct order of the events or information.

Example: Read Text 2 and write the following sentences in correct order as they appear in the text.

a. Bear wanted to knock the tree down.

b. Kintaro and his friends reach a cliff facing a stream.

c. Kintaro was able to build the bridge of the fallen tree.

d. Kintaro's friends became happy.

e. The bear was made to carry Kintaro on its back.

Answer Key/Marking Scheme: Award 1 mark to each correct answer.

a. E. The bear was made to carry Kintaro on its back.

b. A. Bear wanted to knock the tree down.

c. C. Kintaro was able to build the bridge of the fallen tree.

d. B. Kintaro and his friends reach a cliff facing a stream.

e. D. Kintaro's friends became happy.

F. Short answer questions

Short answer questions require students to construct a concise and focused response that may be factual, interpretive or a combination of the two.

Example: [Refer to Reading Text 2)

a. Who is stronger, Kintaro or the bear? How do you know?

b. Why was the bear not able to make a bridge? Give a reason.

c. Why do you think Kintaro was able to knock the tree?

Cognitive Levels for Reading Tasks

The specification grid clearly mentions four types of comprehension questions viz. literal comprehension, reorganization, inference and evaluation and reflection to assess reading. However, writing assessment is done through writing with clues (Guided Writing) and the students' own writing (Free Writing).

(i) Literal Comprehension: Literal comprehension occurs at the surface level. The details are stated and clear for anyone to identify while reading a particular text. Literal comprehension is often referred to as 'on the page' or 'right there' comprehension. This is the simplest form of comprehension. Items to test literal comprehension includes the following:

- Locating or identifying explicitly presented facts in the text
- Recognizing or recallingdetails and main ideas, sequencing, comparing, examining cause/effect relationships and character traits
- Understanding of information and facts such as dates, times, and locations, etc. which are directly stated in the text

Literal comprehension is recognized as the first and most basic level of comprehension in reading.

Examples:

Questions:

- In which country did the auction take place? (***Refer to Reading Text 1 above.)***
  Answer: in Germany.
- Who is stronger, Kintaro or the bear? (***Refer to Reading Text 2 above.)***
  Answer: Kintaro.

Explanation: The above questions are literal comprehension questions as their answers are not only in the sentence level but the students can locate them easily as they are explicitly stated in the text.

(ii) Reorganization: Reorganization covers the literal understanding of the text by combining information from various parts of the text. Reorganization comprehension requires some curiosity and creativity. Reorganization often requires students to analyze, synthesize, and/ or organize ideas or information explicitly stated in the text. Reorganization items involves the following:

- Rearranging information gained from various parts of the text into new patterns that integrate them into the students' own idea for further understanding.
- Understanding of the coherence of the text which can range from recognizing local coherence between adjacent and/ or non-adjacent sentences to understanding the relationship between several paragraphs across a text.
- Retrieving details and facts from within a text whether it is at the beginning of the text or embedded at the end of the text

- Identifying the information to answer the reorganization questions, test takers ofteninvolve in more than one activity like scanning, skimming, searching for, locating and selecting relevant information.

Overall, the difficulty of the item is determined not only by the complexity of the text but also by whether the information is explicitly stated or whether the student has to categories information, discriminate between two similar pieces of information or go for competing information.

Example: (*Refer to Reading Text 2 above.)*

Question:

- What does the word 'they' in the second line refer to?

  Answer: Kintaro and his friends.

Explanation: This question is a reorganization question because the students have to read the various parts of the text (i.e. two consecutive sentences) and come up with the answer by combining information which are explicitly stated in the text.

(iii) Inference: Inference level covers use of cognitive process to deduce implied, inferred or assumed meaning based on evidence and reasoning. Item developers must always look for the evidence or the clue while creating inference items so that test takers use their prior knowledge based on the evidence or the clue to answer the question. Making inferences involves more than a literal understanding. Inference tasks often involves the following:

- Inferring that one event caused another event
- Comparing and contrasting text information
- Identifying the logical order of the various parts of a text
- Making sensible predictions based on the given text
- Discerning the overall message or theme
- Describing the relationship between two characters

Example: (*Refer to Reading Text 2 above.)*

Questions:

- Why were Kintaro's friends happy?
  Answer: Because Kintaro fell the tree and made the bridge.

Explanation: This question is an inference questions as it seeks clues and evidence such as Kintaro, push, loud crack, fell over, spanned, shouted, bridge to answer this question**.**

(iv) Evaluation and Reflection: Evaluation and Reflection level covers reading beyond lines that requires personal judgment i.e. analyzing, evaluating and creating. Evaluation requires the students to give a global or comprehensive judgment about some aspect of the text. Evaluation and reflection items often expects the following:

- Making generalizations, judgments, recommendations, suggestions and decisions
- Creating alternative ideas based on the text
- Demonstration of a full and detailed understanding and integration of one or more paragraphs to deal with unfamiliar ideas, in the presence of prominent competing information, often requires multiple inferences, comparisons and contrasts
- Evaluatinga complex ideacritically

- Applying sophisticated understandings from beyond the text being based on multiple criteria or perspectives
- Locating and organizing several pieces of deeply embedded information, inferring which information in the text is relevant
- Evaluating,hypothesizing or drawing on specialized knowledge based on the text
- Detailed understanding of a text whose content or form is unfamiliar
- Using formal or public knowledge to hypothesize about or critically evaluating a text

Example: ***Refer to Reading Text 1 above.***

- Do you think David Gura is likely to buy even the remaining pages of the manuscript? Give a reason.
  Answer: Any answer that mentions 'Yes' with the idea'because of the fear of its disappearance/ because it's a holy book'.

  OR

  Any answer that mentions 'No' with the idea 'since the gravity of the book, it's beyond his access.'
- Would you buy the page of the manuscript if you were David Gura? Why, why not? Give a reason.
  Answer: Any answer that mentions 'Yes' with the idea'I wouldin order to preserve the sacred book.'

  OR

  Any answer that mentions 'No' with the idea'I wouldn't because it wouldn't draw my attention as I am an atheist.'

Explanation: These questions are evaluative in nature because they do seek a logical answer from the students' head by relating their views with the text.

The curriculum does not clearly specify the cognitive levels for assessing writing skills of the students. However, since guided writing provides clues to the students, they aim at understanding and application levels of cognitive behavior whereas free writing is considered higher levels of behavior as it is creative in nature.

Elaboration of Specification Grid

The present SG only states the type and the number of items to be constructed but does not present the details of the necessary components an item writer needs to be aware of. These components include information such as how many items on each comprehension levels, which items attempt to test particular learning outcome and the percentage of the weightage. The elaborated SG attempts to address the comprehensive details to help an item writer to develop the items to measure the student's ability to respond.

Development of the standardized tools includes administration guidelines and evaluation criteria such as rubrics to ensure a robust and reliable assessment. Another aspect to be considered would be the time and resources needed for a technically robust assessment of these skills. The present specification grid needs elaboration to clarify some areas.

Item Card

Test developers are required to use the item cards so that learning outcomes are aligned with the course contents and assessment as they are stipulated in the curriculum. A sample item card is given below.

Sample item card (Writing)

1) Subject: English Grade 10

2) Item cell code: XXX

3) Elaborated item code (for English):

Unit	LO	Area (R/W)	If R - skill	Format	Marks	text type
GW 2	5	GW	Chart Interpretation	GW	5	Non continuous

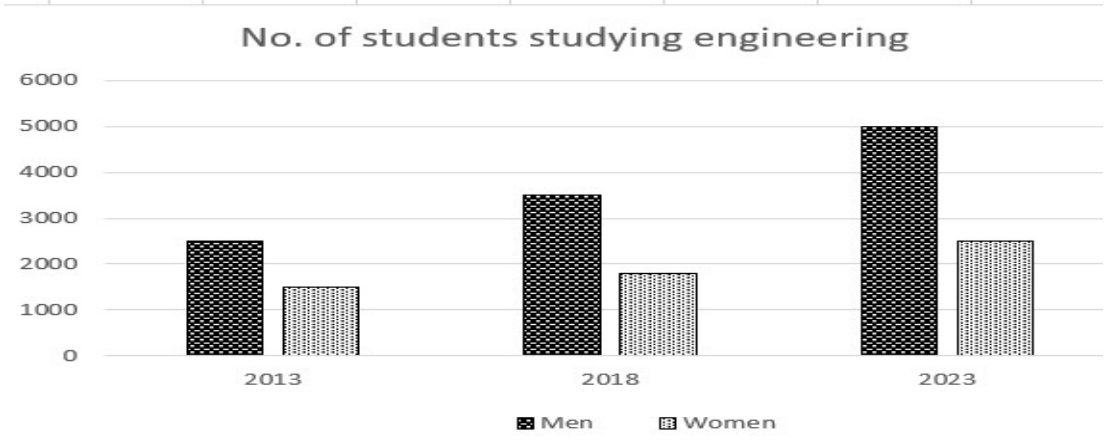
(See Annex for description)

4) Learning outcome (from curriculum): Interpret charts, tables and diagrams.

5) Objective of the item (in line with learning outcome): The objective of the item is to measure whether students will be able to interpret the information given in the charts.

6) Item:

The bar graph below shows the number of men and women studying engineering at different universities in Nepal over the years. Interpret the information in the chart in about 100 words by selecting and reporting the main features. Make comparisons where relevant. You may use words such as illustrate, fall slightly, increase sharply, grow steadily and so on.



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

Answer Key/Marking Scheme: Award marks as per the rubrics attached.

Descriptive Structure: introduction, overview, body paragraphs and conclusion. [e. g. *The chart illustrates/shows the number of male and female students studying…..*]

Key features: key features of the chart, such as highs, lows, changes, and usual data. [e. g. *Both male and female students follow the increasing trend. However, the number of female students is increasing slightly over the years in comparison to the male students*.]

Information Grouping: Information grouped together in terms of similar trends or categories.

Important Details: Important information should not be missing.

Factual: The writing must be avoiding opinions and incorrect information.

Word Limit: The work should be at least 100 words, and 0.5 markswill be deducted if the word limit is not met as far as the writing is praiseworthy.

Marking Scheme (Rubrics)

Content: 2

- Effectively stated and strongly supported content by fully addressing the prompt 2
- Partially stated, often misinterpreted/repeated content with some support 1.5
- Limited content often with broken sentences with no support 1
- Ideas show thought but not tied to the topic 0.5
- No relevant content at all 0

Vocabulary, Layout and Organization: 2

- Appropriate layout, varied vocabulary, coherent sentence structures and linking words 2
- Appropriate layout, varied vocabulary, coherent sentence structures and linking words with occasional lapses 1.5
- Appropriate layout, limited vocabulary, fairly organized sentence structures and linking words 1
- Limited vocabulary, comprehensible but haphazard sentence structures with repetition 0.5
- Incomprehensible vocabulary and sentence structures 0

Grammar and Mechanics: 1

- Negligible grammatical, punctuation and spelling errors 1
- Occasional grammatical, punctuation and spelling errors 0.5
- Significant grammatical, punctuation and spelling errors 0

N.B. Proper training to teachers for capacity building is expected for developing items with appropriate cognitive behaviour and for writing descriptors for marking schemes.

Item assembling for test set generation

Test developers generate test set by assembling the items using the item cards and a test matrix. A number of test matrix can be developed through the specification grid. A sample test matrix is given below.

Table 4: Sample Test Matrix for a single set

Area	Text types	Multiple Choice	Sequencing	Fill in the gaps (without choices)	Matching	True/False	Short Answer Question	Vocabulary	Marks
Reading	**Seen text 1**	**Poem 5**							**5**
	Seen text 2		**Biography 5**				**Biography 5**		**10**
	Unseen text 1			**Job Advertisement 5**			**Job Advertisement 5**		**10**
	Unseen text 2				**Essay 5**	**Essay 5**	**Essay 5**		**15**

Writing	**Guided Writing 1**	**Recipe (with clues)**	5
	Guided Writing 2	**Invitation letter (with clues)**	5
	Free Writing 1	**Views and attitudes (on current/important issues)**	6
	Free Writing 2	**Letter and email (formal/informal)**	8
Grammar	**Reproduction**	6 tasks representing each of tense, question tag, reported speech, voice, and interrogation and negation	6
	Grammar in Context	10 MCQs representing articles, prepositions, tense, question tag, voice, reported speech, connectives, conditional sentence, subject verb agreement, causative verbs, modals, adjectives and adverbs,and relative pronouns	5

Internal Assessment

Internal assessment covers 25 marks. For the purpose of internal assessment, all classroom activities along with student's individual work such as project work, test papers or other proof of the students' performance will be recorded and documented by creating individual student's portfolio. Along with the skills of language, the participation of students in classroom activities as well as their score from terminal exams will be included in the internal assessment component. The table below shows the marks allocated to each assessment area in the internal assessment.

Table 5: Internal Assessment- English

Assessment Area	**Marks**
Participation	3
Listening Test	8
Speaking Test	8
Score from Terminal Examinations	6
Total Marks	**25**

Elaboration of Internal Assessment

a. Participation

Participation of the students is assessed based on their attendance and their performance inteaching learning activities. Out of 3, student attendance carries 1mark and the participation of students in learning carries 2 marks.

b. ListeningTest

As per the curricular provision, teachers needtoprepare two listening tasks themselves so as to assess the students' performance. Each task comprises of 4 questions carrying 1 mark each.The sound files may contain lectures, talks, presentations, radio/TV broadcasts, podcasts/vodcasts,commentaries, interviews, conversations, short discussions, personal accounts (oral anecdotes, pastexperiences) narratives (e.g. radio dramas), instructions and directions, factual accounts (e.g. newsreports, eyewitness accounts) explanations, public

announcements,weather forecast, etc. The sound filesshould be authentic and clearly articulated at normal speed of delivery. Each sound file should be a maximum of three-minute long. The types of questions include multiple choice questions, matching, fill in the blanks and short answer questions.

There is a provision of alternative test method for students with hearing impairment. For the hearing-impaired students, any one of the tasks (paragraph writing on a given topic, writing a letter or writing a description of something) can be given. The allocated time for this is 20 minutes.The record of the task and students' performance should be maintained inportfolios.

c. SpeakingTest

The speaking test will be administered practically in the classroom. It can be done without any formal setting even during the class. However, teacher should have his/her internal advance planning for this. The teacher should prepare the task themselves. The test starts with a greeting and introduction to make the students feel comfortable. The speaking test consists of the following sections and tasks:

1. **Contextual interview or conversation/Describing pictures (4marks)**

The teacher can use any one of the task: Contextual interview/conversation or describing pictures. The students will be asked at least four questions on their personal affairs and immediate contextual situation i.e. situation created to respond using the exponents of language function. (Introduce yourself. Suppose your friend has recently won a speech contest. How do you congratulate him/her for winning the speech contest? You need to be out of the class for sometime. Ask for permission from your teacher, etc). For this type of task, teacher can also make students talk to each other with a partner and assess them. Instead of the above type of task, the students can also be asked to describe the pictures. For this, the students will be given a picture or a set of pictures. They are expected to describe the picture/s in at least 8 sentences. The teacher marks based on their performances.

2. **Speaking on a given topic/presentation of a project work (4marks)**

These two types of tasks are very common and effective in language teaching. These types of tasks are done regularly in course of teaching learning process. Therefore, the teacher can use any one type of the task: speaking on a given topic or presentation of a project work for internal assessment record keeping purposes. For speaking on a given topic, the students will be given a topic like; my school, my hobby, my family, etc. They will get one-minute to think over the topic before speaking on it. This will be done individually. Instead of the task of speaking on a given topic, presentation of a project work can also be used for assessing students' oral presentation skills. The language project the students do in the class in groups or individually can be used for this. Time allocated for this task is 10 to 15 minutes per student.

Alternative test method for students with visual difficulties: For students with visual difficulties, teachers have to ask those students to narrate a sequence of events instead of ‘describing pictures’.

Alternative test method for students with speech and hearing difficulties: For students with speech and hearing difficulties, teachers should give a reading comprehension task worth 8 marks. It will be assessed based on the given rubrics.

It is mandatory for the teachers to maintain the record of the tasks used for assessing students' performance so that it can be presented as portfolio or evidences when asked later.

c. Terminal Test

Terminal test carries 6% marks of marks in the external assessment. At least two terminal tests should be administered and its record should be kept for the external examination. The full marks and weight of written examination of reading and writing (including grammar) will be as per the specification grid developed by Curriculum Development Centre. The marks obtained in the terminal test should be converted to 6% mark and its record should be kept for the final examination. If there are two terminal tests, each terminal test will have the weight of 3 percent. If there are more than two terminal tests, number of any two terminal tests will have to be calculated for the final examination or grade sheet.

N.B. Refer to the internal assessment directives published by CDC for detail.

Result Analysis for Improving Classroom Teaching and Learning

This section presents how result analysis assists classroom teaching and learning.

Area: Reading Comprehension

There are eight planets in the Solar System, and each one is very different. Some planets like Jupiter and Saturn are very large. Others, like Mercury and Mars are smaller. Jupiter has moons that are larger than Mercury. The planets also have different atmospheres. Uranus, Jupiter and Saturn have atmospheres of hydrogen and helium. The atmosphere on Venus is made up of carbon dioxide. Earth has a nitrogen and oxygen atmosphere. Neptune's atmosphere is mostly hydrogen. The planets also have different temperatures. Uranus is the coldest and Venus is the hottest.

Q1. Do the planets have similar atmospheres? Give a reason.

Possible Answers: Students are likely to write the following answers.

A1: No.

A2: No, the planets do not have similar atmosphere because they have different types of gases in their atmosphere.

A3: The planets also have different atmospheres. Uranus, Jupiter and Saturn have atmospheres of hydrogen and helium. The atmosphere on Venus is made up of carbon dioxide. Earth has a nitrogen and oxygen atmosphere. Neptune's atmosphere is mostly hydrogen.

A4: There are eight planets in the Solar System, and each one is very different. Some planets like Jupiter and Saturn are very large. Others, like Mercury and Mars are smaller. Jupiter has moons that are larger than Mercury. The planets also have different atmospheres.

Analysis: Those students who may write A1 have not read the question carefully while those who write A2 have comprehended the text and the question. Those who write A3 and A4 have simply copied the lines without understanding the text.

Classroom Strategies: Teach students the strategies for dealing with the reading comprehension. Some strategies may be teaching vocabulary others may include guessing meaning in the context, relating one sentence with the other, making inferences, etc.

Conclusion: Teaching reading comprehension is the must in the classroom scenario.

Grammar

1. I trust him completely. His conduct is ……..(over/under/below/above) suspicion. (*Choose the correct option to fill in the blanks and complete the sentence*.)

Possible Answers: Students are likely to write the following answers.

A1: I trust him completely. His conduct is over suspicion.

A2: I trust him completely. His conduct is under suspicion.

A3: I trust him completely. His conduct is below suspicion.

A4: I trust him completely. His conduct is above suspicion.

Analysis: Those students who may write A4 have clearly understood the use of over, under, below and above while those who go for A1, A2 and A3 have neither understood the usage of over, under, below and above nor are they able to match the meaning between the first and the second sentence.

Classroom Strategies: Teach students with a clear idea about the distinction between the usage of over, under, below and above with varied examples so that they get a clear concept. Moreover, guessing meaning from context by matching two sentences to come up with a solution is also required to be practiced.

Conclusion: Teaching preposition usage with varied examples is a must.

Chapter IV

Science and Technology

The curriculum of secondary level Science and Technology subject (CDC, 2077 B.S.) aims to promote scientific knowledge, competencies, and the appropriate skills and attitudes necessary at the secondary level. It not only emphasizes the application of skills to address real-world problems but also ensures student's understanding of Science and Technology to enable them to pursue further study in science or related courses. The curriculum developed in accordance with the National Curriculum Framework (NCF) includes level-wise competencies, grade-wise learning outcomes, content, scope and sequence, suggested practical and project activities as well as strategies for facilitating learning and assessing students' progress systematically.

This assessment framework provides guidelines for assessing students' performance through SEE which is conducted at the end of Grade 10. It serves as a reference for designing robust assessments. This chapter outlines test specifications, descriptions of cognitive skills, and sample items aligned with specific competencies as well as deliberates on practical and internal assessment in relation to Science and Technology subject.

Level-wise Competencies and Learning Outcomes

Level-wise competencies

The curriculum of Science and Technology encompasses several essential competencies that the students are expected to demonstrate by the completion of Grades 9 and 10. Some competencies include development of necessary skills, strategies and attitude for scientific research, demonstrating understanding of relationship between components of environment and contribution on conservation of environment, analysis of concept of physical activities and process and use in real life, investigation and use of traditional practices in scientific concepts, etc. Such broad array of competencies are further elaborated into learning outcomes. For example, for the curricular competency, 'analysis of concept of physical activities and processes and use in applied life' is one of the desired learning outcomes mentioned in the curriculum.

Learning Outcomes

Based on the domains, sub-domains, and competencies, a detail learning outcomes (LOs) are given in the present curriculum. There are altogether 84 Los in the 19 units of Science and Technology subject of Grade 10. Some units like 'Life Cycle' and 'Chemical Reaction'comprise of only 2 LOs whereas 'Wave' has a maximum number of 9 LOs. The content-wise number of LOs in the curriculum of Science and Technology subject are as follows:

Table 6: No. of Learning Outcomes - Science

SN	Units	Number of Learning Outcomes
1.	Scientific Studies	4
2.	Classification of living beings	4
3.	Life cycle	2
4.	Heredity	6
5.	Body structure and Life Process	6

6.	Nature and Environment	4
7.	Force and Motion	5
8.	Pressure	3
9.	Heat Energy	4
10.	Wave	9
11.	Electricity and magnetism	6
12.	The universe	4
13.	Information communication Technology	4
14.	Classification of elements	5
15.	Chemical Reaction	2
16.	Gases	4
17.	Metals	3
18.	Carbon and it's compounds	5
19.	Chemicals used in daily life	4
Total		**84**

LOs are essential to determine what to teach and clarify expectations of the learners regarding what they are going to learn. It is also essential to develop a test in order to find out whether the purpose of teaching and learning has been achieved. LOs of Unit 1 'Scientific Study' of grade 10 Science and Technology subjectis given below as an example.

1.1. To identify independent variable, dependent variable and controlled variable in scientific study
1.2. To differentiate between fundamental and derived units.
1.3. To find fundamental units in derived unit.
1.4. To use fundamental units to check homogeneity in physical equation. (CDC, 2077 BS)

Learning outcomes are carefully designed to help students achieve them, enabling them to grasp the fundamental principles and real-world applications of the subject. By dividing the content into manageable units and including designated marks for each unit, the syllabus ensures a comprehensive and fair assessment of students' knowledge and skills. By focusing on the LOs, the syllabus promotes critical thinking, problem-solving abilities, creative thinking and lays a solid foundation in Science and Technology, preparing students for further studies and successful careers in scientific world.

Content Area, Item Format and Cognitive Level

The content area is further elaborated in the curriculum in order to achieve curricular competencies and learning outcomes that encompasses diverse areas of the subject. These domains include Scientific Study, Information and Communication Technology, Biology, Physics and Chemistry. Each domain encompasses a wide range of learning outcomes further covering various aspects of theoretical knowledge, practical skills, problem-solving abilities, creativity and critical thinking in the field. These are fundamental aspects for the assessment of students' learning.

Item formats are multiple choice questions of 1 mark. Each consists four alternatives out of them three are distractors and one is key answer. Very short answer type questions: such items demand only one response in one element or one sentence only and carries one mark only. Short answer questions: these items demand short response in two elements or two major points or two steps. Long answer type questions: these items require a long response with at least in four elements, four individual points or four statements or four steps to solve the problems.

Equally important for the assessment is the cognitive aspect of learning as assumed in the curriculum. Specification grid of Science and Technology subject demands students learning from recalling termed as 'Knowledge' and further demanding 'Understanding', 'Application' and 'Higher Abilities'– the last one incorporates 'Analysis', 'Synthesis' and 'Evaluation' of Bloom's Taxonomy of Educational Objectives. Specification grid of Science and Technology subject divides the full marks into four cognitive levels like 15% for Knowledge, 30% for Understanding, 30% for Application, and 25% for Higher Ability. However, $\pm$ 2 mark variation is allowed within the units of area but marks of content area in the assembled test set be intact.

Aligned to the Bloom's taxonomy, a brief description of different levels of cognitive skill is provided along with some sample items below.

Knowledge Level Items

In Science and Technology, students should respond in fact based items like value of g = $9.8 m/s^2$,G= $6.67 \times 10^{-11} Nm^2/kg^2$ etc. These items can be solved onthe basis of rote learning. Definition and other answers can also be responded with rote learning and recalling. Such types of items are included in knowledge level items. This skill involves recalling, recognizing or retrieving information from short term memory and long-term memory. It is the lowest level of the cognitive process. Test items at this level usually require defining a concept, describing a scientific theory or a scientific phenomenon, stating the statements, recalling the facts, reciting the statement, etc. For example,

Which of the following number of chromosomes is present in the human male? (MCQ)

A) 22+XY
B) 22+ XX
C) 44+ XY
D) 44+ XX

Key Answer: C. 44+ XY

Explanation: This is a knowledge level item since students can solve this by recalling/ memorizing the fact that the human male has 44+ XY chromosome. There is no other alternative way to know/ solve except memorization/recall.

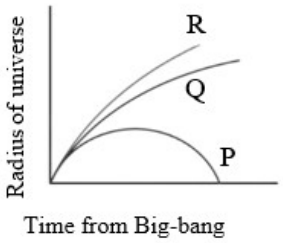
Understanding Level items

Some of the items demand reasons and logics under the fact. What does the fact mean? Students require to comprehend the knowledge under known conditions. These types of items are included in understanding level. It involves explaining, describing, classifying, comparing or constructing meaning from a concept. This often includes comprehension of one or more ideas in order to generate an explanation. It is generally called the comprehension of knowledge. The action verbs that can be used in understanding level items are; explain, interpret,, express, describe,

discuss, report, restate, translate, extrapolate, classify, identify, paraphrase etc. Examples of understanding based item is shown below.

Sample Item 2:(VSAQ)

Types of the universe is shown in the picture. What would be the future of the universe if the critical density is less than the density of the universe?



Answer:

If the critical density is less than the density of the universe then the universe will expand forever,

This scenario is represented by the R curve in the diagram. 1mark

Explanation

This item is of understanding level as it requires an understanding of basic concepts such as density, gravity, and the expansion of the universe. This item also involves interpreting diagrams related to the basic concept mentioned.

Application level items

This cognitive skill involves application of knowledge in anew situation. It assesses s skill of students whether they can use information in a new situation or not. The situation can be familiar as classroom situation or unique situation by presenting the information in unique way. Solving problems or tasks in new situations by applying acquired knowledge, facts, techniques, and rules with different approaches of Science and Technology. This skill requires the student to calculate, predict, solve a problem, apply, or demonstrate an idea, choose, demonstrate, dramatize, employ, illustrate, interpret, operate, schedule, sketch, solve, use, write based on new situation of science and technology.

Sample Item 3 (SAQ)

If primary coil and secondary coil is in the ratio of 2:9 in a transformer and the input voltage is 220 V, calculate the output voltage.

Solution:

Let the number of primary coil (Np)= 2x

Number of secondary coil (Ns) = 9x

Primary voltage (Vp)= 220 V

Secondary voltage (Vs) = ?

We know, $\frac{Vs}{Vp} = \frac{Ns}{Np}$ ……. 1 mark

$\frac{Vs}{220} = \frac{9x}{2x}$

Vs = 990 V ……1 mark

Explanation:

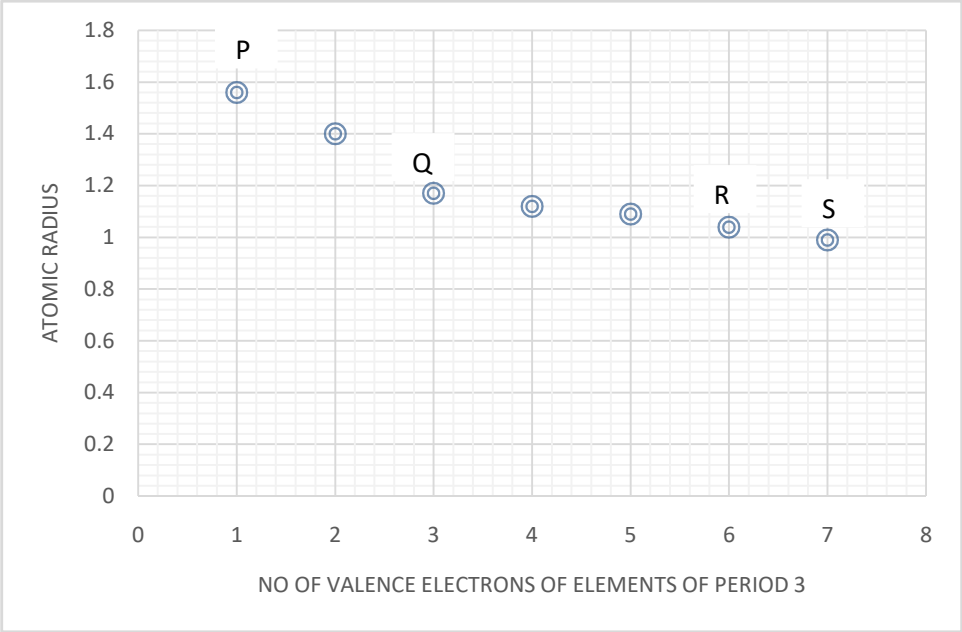
This item is of application level. To solve this item student needs to apply knowledge of relation between number of coils and voltage. Students should choose correct formula to calculate secondary voltage. One should use mathematical skills to solve a real-world problem.

Higher Ability

Higher ability items include a set of higher order thinking skills that enable a student to break down a complex structure into its constituent parts, analyze the relationships between the constituents, make judgment based on standards or scientific evidence available and reorganize the constituent parts to create a new structure in a novel situationin the field of science and technology. Students should have skills to connect Science and Technology with society and are capable of studying society with analytical, evaluation and creating approaches with situation of society and other novel situation. It includes analysis, evaluation and creation skills on the novel situation of daily life and our society. Here novelsituationrefers to the one that is not seen by the respondent and related with daily life or other real life situations.

Sample Item 4: (LAQ)

Study the following graph, in which P, Q, R and S represent the elements of period 3 of modern periodic table.



Based on the graph, answer the following questions.

A) To which block does the element R lie? (1 mark)

B) Among P and Q, which one is more metallic? Explain on the basis of atomic size. (2 marks)

C) Write the molecular formula of the compound formed by the elements Q and S. (1 mark)

Marking scheme:

A) R- lies in p-block (1 mark)

B) P is more metallic in nature than Q (1 mark)
Because element P-lies towards left hand side of the same period than that of the Q. Besides this there is only one valence electron in P but in Q there are 3 valence electrons. It is easier to lose 1 electron than to lose 3 electron. Metallic property is also determined by the tendency of losing electrons. (1 mark)

C) PS_3 (1 mark)

Description of item:

This item is of Higher ability because the situation is novel i.e. students may not have seen such situation earlier. The presentation of data on the basis of atomic radius and number of electrons in valence shell is quite new. Students should study this plot before solving the problem. Additionally, students should require to judge whether P is more metallic or Q on the basis of valence electrons given in x-axis. They need to create molecular formula like PS_3.

Elaboration of Specification Grid for Item Development

In order to develop items in Science and Technology subject, specification grid is further elaborated in line with learning outcomes in the curriculum. This helps to ensure teaching and learning as well as assessment correspond to the learning outcome as designated in the curriculum. In case of Science and Technology subject there are 84 learning outcomes within 19 content units which are to be assessed into 4 cognitive skills and in each cognitive level there would be 4 types of item format. This makes 4x4x84 (1344) possible item cells. One of the content domain of unit 1 scientific study and elaboration of it with required cells are given below as an example. Cell codes and content elaboration of others units can also be prepared by the same way.

Table 7: Sample Elaboration of Specification Grid - Science & Technology

Unit	Area/Unit		Item Cell code															
			K				U				A				HA			
	LO No.	Learning Outcomes	MCQ	VSAQ	SAQ	LAQ	MCQ	VSAQ	SAQ	LAQ	MCQ	VSAQ	SAQ	LAQ	MCQ	VSAQ	SAAQ	LQ
	1.	Scientific Study																
1.	1.1.	1.To identify independent variable, dependent variable and controlled	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16

		variable in scientific study																
	1.2	2.To differentiate between fundamental and derived units.	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
3.	1.3	To find fundamental units in derived unit.	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
4.	1.4	To use fundamental units to check homogeneity in physical equation.	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64

Item Card

It is a sheet in which all the characteristics of item are mentioned. For example; LO, Cognitive levels, format, marks, etc.

Example of Item Card:

Item No 1

1 **Subject:** Science and Technology G 10

2 **Item cell code:** 1

3 **Elaborated item code:**

Area	Unit	LO	Cognitive level	Format	Marks
Scientific Study	1	1.1	1	2	1

Objective of the item (in line with the learning outcome): To identify independent variable, dependent variable and controlled variable in scientific study

4 Item: Define controlled variable.

5 Key answer (marking scheme): the variables which are kept constant throughout the experiment are called controlled variable. 1 mark

In the test, each item will have its own cell code. The cell code of the item indicates the content area, cognitive level and format of the item. For example, item with cell code number 20 represents an item from scientific learning section unit 1 and it is knowledge level and a long question which carries 4 marks.

Test Set Assembly

Items are panelled and moderated along with the key answer in case of MCQs and marking scheme in case of supply items and collected as item repository. Items in the repository are then assembled in a full test set form which should meet requirements of specfication grid. Each of the test set assembled are arranged in a test matrix as shown below as a sample test matrix.

Test matrix

Table 8: Sample Test Matrix - Science and Technology

SN	Units	Teaching Hours	Conginitive Area				Group-wise marks	Unit-wise marks
			K (15%)	U (30%)	A (30%)	HA (25%)		
1.	Scientific Study	5	MCQ		SQ		**8**	**3**
13.	Information and communication technology	10	SAQ	VSAQ	MCQ	VSAQ		**5**
2.	Classification of living beings	9			LAQ		**23**	**4**
3.	Life cycle (Honey bee)	4		VSAQ		MCQ		**2**
4.	Heredity	16	SAQ	LAQ	SAQ			**8**
5.	Human body structure and life process	12	LAQ	SAQ				**6**
6.	Nature and Environment	7	MCQ			SQA		**3**
7.	Force and motion	10	VSAQ			LAQ	**25**	**5**
8.	Pressure	5		VSAQ	MCQ			**2**
9.	Heat Energy	10		MCQ	VSAQ	SAQ		**4**
10.	Wave	15	VSAQ	LAQ	SAQ			**7**
11.	Electricity and magnetism	12		MCQ	LAQ			**5**
12.	Universe	5			SAQ			**2**
14.	Classification of elements	9				LAQ	**19**	**4**
15.	Chemical reactions	6		VSAQ	SAQ			**3**
16.	Gases	8		MCQ	VSAQ	SAQ		**4**
17.	Metals	5		SAQ				**2**
18.	Hydrocarbons and their compounds	6		MCQ		SAQ		**3**
19.	Chemicals used in daily life	6		SAQ		MCQ		**3**

	Total	**160**	**12**	**22**	**22**	**19**	**75**	**75**

This is an example of test matrix. We can make numerous test matrix while developing test paper. All the test sets should fulfil requirement of specification grid in terms of unit-wise and group-wise marks given in left column as well cognitive levels and item formats given the last row of the test matrix. Items in the cells should vary from one set to another in the place in the cell and by question even if placed in the same cell.

Provision of Assessment Types

Internal and External

Both Internal and external assessments are integral part of students' evaluation. Internal assessment in the form of class tests, unit tests, home assignment, oral question answer, peoject work, practical work, regularity, punctuality, participation in learing, etc. provide regular feedback to students, parents, and teachers about the quality of teaching and learning occurring in the classroom. External assessment at the end of the educational program provides insight to students learning achievements based on which students are certified for further enrolment into a course. Year-end summative examination carries greater weightage of marks, and therefore, needs to be fair and reliable. Internal evaluation as the formative evaluation carrying 25 marks out of 100 marks and the remaining marks are allocated to External Examination as the summative evaluation.

Internal Evaluation

Internal evaluation comprises of practical work, marks from two trimester examinations and classroom participation which includes classroom participation and participation in learning.To assess the classroom activities two types of assessments are conducted to achieve learning outcomes of Grade 10. To enhance learning qualities internal assessment is conducted in the various heads.

Internal evaluation consists of five major components. They are:

Criteria	Marks
participation (attendance and activeness and participation in learning activities)	3
experimental work (conduction of experimental work (set up of apparatus, observation, test, tabulation and conclusion) presentation, record keeping of experimental work)	10
project work (conduction of project work [field visit, development of chart, development of models and collection, preparation of apparatus and inquiry (question answer)]	6
result of first and second term	3 marks each

Please see Internal Evaluation Guidelines, CDC-2080, Class 10)

External Evaluation

The external evaluation is a written examination comprising of 75 marks out of total 100 marks. This framework provides guidelines for developing items for the external written examination based on the curriculum and specification gridof CDC. For summative examination following types of items and their weightage are expected by the curriculum as well as the competency levels and their weightage are listed below.

Distribution of test items according to item formats, marks distribution and tentative times.

Table 9: Sample distribution of test items- Science and Technology

Type of item	Score per item	Total item	Total score	Time
Multiple Choice Questions	1	10	10	24 minutes
Very short Answer questions	1	9	9	156 minutes
Short Answer Questions	2	14	28	
Long Answer Questions	4	7	28	
Grand total		**40**	**75**	**3 hours**

Similarly, following requirements for the format and cognitive levelsof items demanded by the items should be fulfilled in each test-set as assumed by curriculum and specification grid.

Table 10: Sample Item Formats mapped to cognitive skills for Science and Technology

Competency levels	MCQ		VSAQ		SAQ		LAQ		Total no. of questions and marks	
	No. of questions	Marks	No. of questions	Marks	No. of questions	Marks	No. of Questions	Marks	Total questions	Total marks
Remembering	2	2	2	2	2	4	1	4	7	12
Understanding	4	4	4	4	3	6	2	8	13	22
Applying	2	2	2	2	5	10	2	8	11	22
Higher Ability	2	2	1	1	4	8	2	8	9	19
Total	**10**	**10**	**9**	**9**	**3**	**28**	**7**	**28**	**40**	**75**

Result analysis/error analysis and classroom teaching learning improvement

Problem 1.

Item: Calculate amount of heat required to raise the temperature from 20^0C to 35^0C of water of 500 g. (Specific heat capacity of water is 4200 J/Kg^{0}C)

Possible Answers:

Correct Answer:

Solution:
Mass of water (m) = 0.5 Kg
Specific heat capacity of water (s) = 4200 J/Kg^{0}C
Difference in temperature (dt) = 35^0C -20^0C =15^0C
Amount of heat (Q) = ?
We know,
Q= m x s x dt
Q = 0.5 x 4200 x 15
= 31,500 J

Students may solve as:

Solution:
Mass of water (m) = 500 g
Specific heat capacity of water (s) = 4200 J/Kg^{0}C
Difference in temperature (dt) = 35^0C -20^0C =15^0C
Amount of heat (Q) = ?
We know,
Q= m x s x dt
Q = 500 x 4200 x 15
= 31,500,000 J

Most of the students do mistakes like, the procedure is correct but answer is incorrect. The unit of specific heat capacity difference in temperature are in SI unit but mass is not. So, 500 g should be in SI unit to make all values in SI i.e. 0.5 Kg. Sometimes, students may not write the unit after calculation as well as write wrong unit. They may make mistake while listing the conditions.

Suggestion for improvement:

The concept of unit homogeneity i.e. uniform units in the same system, CGS, FPS or SI should be made before solving such numerical problems.

Problem 2:

Item: Write any four differences between DNA and RNA.

Possible Answers:

Correct Answer:

DNA	RNA
1. It has double stranded helical.	1. It is single stranded.
2. It transfers genetic information from generation to generation.	2. It helps in protein synthesis
3. It consists de-oxyribose sugar	3. It consists ribose sugar
4. It consists A,T,G,C	4. It consists A,G,U,C

Possible Wrong Difference:

DNA	RNA
1. It has double stranded helical.	1. It helps in protein synthesis
2. It transfers genetic information from generation to generation.	2. It is single stranded.
3. It consists de-oxyribose sugar	3. It consists A,G,U,C
4. It consists A,T,G,C	4. It consists ribose sugar

Suggestion for improvement:

The concept of writing difference with the same feature should be written. Like if definition is written in first part same feature should be written in second part too as given above.

Some of the students do mistakes like information of DNA and RNA are correct but proper difference is not written with proper domains. For example, in first difference structure of DNA has written but function of RNA is written. Sometimes students write about one domain only.

Problem 3

Item: Which one is more reactive between sodium and magnesium?

Possible Answers:

Correct answer: Sodium is more reactive than magnesium.

Most common answer: Magnesium is more reactive than sodium.

This answer is incorrect. In this answer, student's perception might be or they may think larger the atomic number, greater the atomic size and more the reactivity. In fact it is not always true.

Suggestion for improvement:

The concept of reactivity of elements is understood after understanding the relationship between reactivity and following properties.

- Valence electrons
- Atomic size
- Electron gaining or losing tendency of element

Chapter V

Mathematics

Mathematics is considered a core subject in the education system of Nepal. It is taught from basic level to secondary and university levels. Mathematics education in Nepal aims to develop logical reasoning, problem-solving skills, and analytical thinking among students. Thecurriculum developed in accordance with the National Curriculum Framework (NCF) includes level-wise competencies, grade-wise learning outcomes, content scope and sequence, suggested practical and project activities as well as strategies for facilitating learning and assessing students' progress systematically.This chapter attempts to explicate the elements of summative assessment of Mathematics as a subject in Secondary Education Examination (SEE). The core skills, learning outcomes, and the pattern of examination, including the weightage assigned are largely drawn from the curriculum.

Level-wise Competencies and Learning Outcomes

Level-wise Competencies

The level-wise competencies in secondary level curriculum of mathematics in Nepal aim to provide students with a solid foundation in mathematical concepts and skills. The curriculum is designed to prepare students for further education and various career paths by fostering critical thinking, problem-solving abilities, and practical application of mathematical knowledge. The competencies are divided into grades 9-10 and grades 11-12. As such, Mathematics curriculum expects students to achieve various competencies by the end of Grade 10 (CDC, 2077 B.S.). For example, the curriculum explicates altogether 10 competencies at grade 10 such as analyze knowledge, skill and concept related to surface area and volume of solid objects in daily life problem and development of the ability to related mathematical knowledge, skills and concepts to the subject area and the learning areas of other subjects. (SEE curriculum for details).

Learning Outcomes

Learning outcomes in the Mathematics curriculum expects varied skills from the students. The learning outcomes for Grades 9 and 10 mathematics in Nepal aim to provide students with a comprehensive understanding of mathematical concepts and skills with analytical thinking and foster critical thinking, problem-solving abilities, and practical application of mathematical knowledge. There are altogether 15 units in Mathematics subject of Grade 10. In these 15 units, there are altogether 26 LOs. There are quite a few units such as ‘Set’ which has only one LO whereas Unit 12- Circle and Unit 14 - Probability consist of maximum number of LOs. The content wise number of LOs in the Mathematics curriculum are as follows:

Table 11: No. of Learning Outcome - Mathematics

S. No.	Units	**Numberof Learning Outcomes**
1.	Sets	1
2.	Compound Interest	1
3.	Growth and Depreciation	1
4.	Currency and Exchange Rate	1
5.	Area and Volume	3
6.	Sequence and Series	2
7.	Quadratic Equation	1

8.	Algebraic Fraction	1
9.	Indices	1
10.	Triangle and Quadrilaterals	2
11.	Constructions	1
12.	Circle	4
13.	Statistics	1
14.	Probability	4
15.	Trigonometry	2
	Total	26

LOs are essential to determine what to teach and to clarify expectations of the learners regarding what they are going to learn. It is also essential to develop a test to find out whether the purpose of teaching and learning has been achieved. For example, LO of Unit 1 'Sets' of grade 10 Mathematics subjectis to solve the daily life problems related with two or three sets using operation of sets, Venn diagram and cardinality of sets.(CDC, 2077 BS). In orderto achieve these LOs, Grade 10Mathematics covers content areas such as Sets, arithmetic, mensuration, algebra, geometry, statistic and probability, and trigonometry. By dividing the content into manageable units and including designated weightage for each unit, the syllabus ensures a comprehensive and fair assessment of students' knowledge and skills.

Continuous assessments and improvement of teaching learning based on assessment outcome ensures that students achieve expected outcomes effectively. These outcomes guide teachers in planning lessons, assessing student progress, and identifying areas for improvement. Such emphasis is expected to tilt assessment from assessment 'of' learning to assessment 'for' learning and help in better learning achievement of students.

Content Areas, Item Format and Cognitive Levels

Content Area

The curriculum further elucidates LOs associating them to related content domains to achieve curricular competencies and learning outcomes that encompasses diverse areas of the subject. Content domain in Mathematics include seven content areas, viz. Sets, Arithmetic, Mensuration, Algebra, Geometry, Statistics and Probability, and Trigonometry. Each domain encompasses ~~a~~ range of learning outcomes covering various aspects of theoretical knowledge, practical skills, problem-solving abilities, and critical thinking in the field. These are fundamental for the assessment of students' learning.

Item Format

The present specification grid of Mathematics for Grade 10 doesn't present specific item format such as MCQ, very short, short and long. This also means such item formats can be used, but important is to measure expected LOs in proper way. The requirement is 49 sub-questions within 16 questions in a test set. In the test paper of mathematics, level wise items could be arranged in a question with a context.

Example

A square-based pyramid with a base side length of 12 meters and a vertical height of 8 meters are given.

A) Write the formula to find the volume of square-based pyramid when a side (a unit) and vertical height (h unit) are given. **[K] (1 mark)**
B) Find the volume of the pyramid. **[U] (1 mark)**
C) Find the total cost for painting the triangular surfaces of that pyramid at the rate of Rs. 150 per square meter. **[A] (3 mark)**
D) Analyze how the volume changes if the base side length is doubled while keeping the vertical height constant in the square based pyramid. Show your working with showing the suitable example. **[HA](2 mark)**

These cognitive levels are briefly described here with example and description of the items.

Cognitive Levels

The specification grid requires items at four types of comprehension questions viz. knowledge (16%), understanding (24%), application (40%) and higher ability (20%). Above item is used to exemplify different cognitive level.

(i) Knowledge Level

Knowledge level involves recalling or recognizing facts, information, and concepts of Mathematics. It is the lowest level of the cognitive process. It includes tasks such as memorization, identification, and recall of information. It shows memory of previously learned material by recalling facts, terms, basic concepts, and answers. Test items at this level usually require defining a concept, describing a mathematical phenomenon, stating the statements, recalling the facts, reciting the statement etc.

Example

A) Write the formula to find the total surface area of square-based pyramid when a side length of the base (a unit) and slant height (l unit) are given.

Explanation: The above item is related to knowledge level. In this item, students just recall the formula of total surface area of square-based pyramid.

(ii) Understanding

Some of the items require students to give reasons and logics under the fact or comprehend the knowledge under the known conditions. These types of items are included in understanding level. It involves explaining, describing, classifying, comparing or constructing meaning from a concept. The action verbs that can be used in understanding level items are; explain, interpret, express, describe, discuss, report, restate, translate, extrapolate, classify, identify, paraphrase etc. It involves explaining ideas, interpreting data, and summarizing concepts in one's own words. In understanding level, knowledge is necessary but not enough, so use of knowledge is more important in this level.

Example

B) Find the volume of square-based pyramid having a base side length of 12 m and vertical height of 8 m.

Explanation: This item is related to the level of understanding. In this item, the students have to recall the formula, substitute the given value in the well-known formula, and solve the problem.

(iii) Application

This cognitive skill involves application of knowledge in a new situation. It assesses s skill of students whether they can use information in a new situation or not. The situation can be familiar as classroom situation or unique situation by presenting the information in unique way. Solving problems or tasks in new situations by applying acquired knowledge, facts, techniques, and rules with different approaches of mathematics. This skill requires the student to calculate, predict, solve a problem, applying principles to solve problems, and using information in practical situations. In this level, students solve the problem with conceptual development, not only by well-trained procedure alone.

Example

C) A square-based pyramid with a base side length of 12 meters and a height of 8 meters are given. Find the total cost for painting the triangular surfaces of that pyramid at the rate of Rs. 150 per square meter.

Explanation: This item is related to application level. In this item, students find the slant height of the pyramid and recognize the parts of pyramid to colour it, which is not clearly given in the items. Also, this item requires to use two mathematical concepts – area and cost calculation.

(iv) Higher Ability

Higher ability levelisnotlimitedtothesolutionofroutineproblemsbutalsoincludesunfamiliarsituations, complex contexts, and solving multi-step problems using more than one relation andcontexts. It includes a set of higher order thinking skills that enable a student to break down a complex structure into its constituent parts, analyze the relationships between the constituents, make judgment based on standards and reorganize the constituent parts to create a new structure in a novel/new situation in the field of mathematics. Students should have skills to connect mathematics with society and are capable of studying society with analytical, evaluation and creating approaches with situation of society and other noble situation. It includes analysis, evaluation and creation skills on the noble situation of daily life and our society. Here noble situation refers the one that is not seen by the respondent and related with daily life or other real-life situations.It includes tasks such as categorizing, comparing, and contrasting. In evaluation levels, learners make judgments or assessments based on criteria or standards.

Example

D) Analyze how the volume changes if the base side length is doubled while keeping the vertical height constant in the square based pyramid. Show your working with showing the suitable example.

Explanation: This question is evaluative with analyzing and judging in nature because students complete this task by taking any example and decide the decision with logic.

Elaboration of Specification Grid

Specification Grid of Mathematics specifies cognitive content area with cognitive level and weightage of an item as well as it requires 16 questions in total with 49 sub-questions within 75 marks test paper. However, specification grid of Mathematics is not specific on format of the

item and so it is open in the use of item format. Specification grid of Mathematics is given below (CDC, 2078)

विशिष्टीकरण तालिका (Test specification chart)
Grade 9 - 10
2078

Subject : Mathematics **Full marks : 75**

Time : 3 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	12	1	1	1	1	1	3	1	1	4	1	6
2.	Arithmetic	28	2	2	2	3	3	5	2	3	9	3	13
3.	Mensuration	28	2	2	2	3	2	5	2	3	8	3	13
4.	Algebra	32	2	2	2	4	3	7	1	2	8	3	15
5.	Geometry	28	2	2	2	3	2	5	2	3	8	3	13
6.	Statistics and Probability	24	2	2	2	3	2	4	2	2	8	2	11
7	Trigonometry	8	1	1	1	1	1	1	1	1	4	1	4
		160	**12**	**12**	**12**	**18**	**14**	**30**	**11**	**15**	**49**	**16**	**75**

द्रष्टव्य

- प्रश्नपत्र निर्माण गर्दा प्रत्येक क्षेत्रमा र समग्रमा ज्ञान, बोध, प्रयोग र उच्च दक्षताका लागि तोकिएअनुसारको भार मिलेको हुनुपर्दछ। तर 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।
- Applicationर 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।

Figure 1: Test Specification Chart, CDC- Mathematics

In order to develop items in Mathematics subject, specification grid is further elaborated in line with learning outcomes in the curriculum. This helps to ensure teaching and learning as well as assessment correspond to the learning outcome as designated in the curriculum. In Mathematics subject there are 26 learning outcomes within 15 content units which are to assessed into 4 cognitive skills. As specification grid leaves item format open, based on the given weightage there is possibility of different formats of item viz. MCQ, VSAQ (each 1 marks) , SAQ (each 2 marks) and LAQ with more than 2 marks. One of the content domain of unit 5 area and volume with required cells are given below as an example.

Table 12: Sample Cell Code - Mathematics

Unit	Area/Unit		Item Cell code												
			K		U				A				HA		
	LO No	Learning Outcomes													
	5. Area and Volume														
5.	5.1	To solve the daily life problems related with area and volume of a pyramid	MCQ 66	VSAQ 67	MCQ 68	VSAQ 69	SAQ 70	LAQ 71	MCQ 72	VSAQ 73	SAQ 74	LAQ 75	MCQ 76	VSAQ 77	SAQ 78

Cell codes and content elaboration of other units areto be prepared by the same way. Elaboration of specification grid helps in the development of item related to specification grid requirements with the learning outcome expected by the curriculum. Item card is used to develop specific item/question with metadata of the item. A sample of item card is given below

Item Card

1 Subject: Mathematics Grade 10

2 Item cell code: 67

3 Elaborated item code:

Content Area	Unit	LO	Cognitive skill	Format	Marks
Area and Volume	5	5.1	1	2	1

4 Objective of the item (in line with the learning outcome): To recall the formula to calculate total surface area of a square based pyramid.

5 Item: Write the formula to find the total surface area of square-based pyramid when a side length of the base (a unit) and slant height (l unit) are given. **(1 mark)**

6 Key answer/Marking scheme:
The total surface area of a square based pyramid is $a^2 + 2al$. **(1 mark)**

Test set assembling

Items developed by the item writers are panelled in subject expert team and then moderated by moderation committee of NEB. Items accepted by moderation committee as final for item pre-testing are then pre-tested in a representative sample of students. Pre-tested items are calibrated and only those items that meet required parameters are put in the item repository. Then only final test set(s) are assembled which should comply with the specification grid. Test matrix for each of the test set is prepared to ensure that test set meets requirement of the specification grid and items are associated with the LO(s). A sample of the test matrix is given below:

Sample of a test matrix

Table 13: Sample Test Matrix - Mathematics

Q. No.	K	U	A	HA	Total	Area
1	1	1	3	1	6	Set (6)
2		1	2	1	4	Arithmetic (13)
3	1		1	2	4	
4	1	2	2		5	
5	1	1	2		4	Mensuration (13)
6			3	1	4	
7	1	2		2	5	
8	1	2	3		6	Algebra (15)
9		2	3		5	
10	1		1	2	4	
11	1	1	2		4	Geometry (13)
12			3	2	5	
13	1	2		1	4	
14	1	1	1	1	4	Trigonometry (4)
15	1	2	2	1	6	Stats and Probability (11)
16	1	1	2	1	5	
Total	12	18	30	15	75	

In some of the question is to be asked covering all the cells such as question number 1 from set and question number 14 from trigonometry. Though total number of questions, weightage and level of questions are fixed in such question, item writer can ask varieties of questions. While in the question where all the cells are not covered, test matrix will be flexible in terms of cognitive levels to be covered in different test matrix. Such as in arithmetic, mensuration, algebra, geometry and statistics and probability where the number of questions are fixed but level of questions and weightage are different from different lessons in same area.

Provision of Assessment Types

Internal and External

The curriculum recommends the inclusion of both internal and external assessments. The internal assessment is conducted to develop conceptual knowledge, skills with positive attitude of mathematics to the students. The internal assessment will carry 25 marks out of 100 marks. It includes 12% for participation and attendance, 64% for practical and project work and 24% for score from terminal examinations.

Internal assessments are particularly important because they provide a more comprehensive and realistic picture of a student's abilities and learning progress. They help in addressing the diverse learning needs of students, fostering a more inclusive and supportive educational environment.

External Evaluation

The external evaluation is a written examination comprising of 75 marks out of total 100 marks. External assessment covers 40% marks for knowledge and understanding level of items with various areas whereas 60% marks for application and higher ability. The external assessments emphasize skill development, such as critical thinking, problem-solving, and creativity. Also, it ensures students are prepared for real-world challenges. The feedback provided by these assessments is diagnostic, offering detailed information to students, teachers, and parents to identify areas for improvement.

This framework provides guidelines for developing items for the external written examination based on the curriculum and specification gridwhich are developed by CDC. For summative examination following level of items and their weightage are expected by the curriculum as well as the competency levels and their weightage are listed below.

Table 14: Sample distribution of item in terms of cognitive behaviour

Level of item	Score per item	Total item	Total score
Knowledge	1	12	12
Understanding	1 or 2	12	18
Application	1 or 2 or 3 or 4 or 5	14	30
Higher ability	1 or 2	11	15
Grand total		**49**	**75**

Result analysis and classroom teaching learning improvement

SEE is an external summative assessment for grade 10 students and obtain certificate. Purpose of assessment/examination does not end with certification; most important objective still remains there - enabling academically and practically next lot of students based on the shortcoming and weakness at the learning realm so far for grade 10 students. Result analysis at the item level should be done – overall performance of the students, identification of the items in which students relatively perform weak, probable reason for weak performance/confusion, and suggestion teaching learning to improve learning in such areas. Two examples of result analysis and suggestion for improvement are given below:

The Example which is taken from above

Example - 1

Analyze how the volume changes if the base side length is doubled while keeping the height constant in the square based pyramid. Show your working with showing the suitable example.

For this item student may think as follows:

- Students might understand that increasing one dimension changes the volume, but the **nonlinear relationship** (quadrupling effect) when a square side length is doubled can be hard to grasp intuitively.
- For students who are still developing these concepts, the idea that doubling one linear dimension results in more than a doubling of the volume can be counterintuitive.
- The question involves manipulating a formula with **squared terms** and interpreting the impact of squaring in the volume formula. For students who struggle with algebraic manipulation, understanding and applying V= $\frac{1}{3}a^2h$ correctly may be challenging. They need to see how a^2 affects the outcome compared to a linear relationship, which adds complexity.

Suggestions to Make the Question More Accessible

- Showing a 3D model or even 2D cross-sectional views can help students visualize the impact of a change in side length on volume.
- Start by exploring simpler cases (e.g., doubling the height or side length of a rectangular prism) to introduce proportionality before moving to more complex shapes like pyramids.
- Break down the question into smaller steps, such as calculating the base area first, discussing how squaring impacts volume, and then analyzing the volume change step-by-step.

This question could be particularly challenging for students due to its abstract nature, multi-step calculation, and requirement to understand nonlinear relationships. Providing additional support and context would help make it a more favorable question for learning.

Example 2:

Solve $x^2 - 5x + 6 = 0$ by using completing square method.

For this item, students may think as follows:

- Most of the students are more familiar and comfortable to solve it using factor method and applying quadratic formula.
- They may have not seen the completing square method properly in question.
- They may be thought result is important than the process.
- They may not serious about the specific method.

Suggestions to Make the Question More Accessible

- Focus on learning outcomes (To factorize by various method)
- Focus on specified method (given or not) to solve the given item.
- Encourage students to ask questions about why they're doing each step to deepen their understanding, and offer corrections and encouragement along the way.

- Show the solution of same example by factorizing, completing square and quadratic formula method.
- Make sure students practice each step individually before combining them. This scaffolding helps them see how each part fits into the larger process.

Chapter VI

सामाजिक अध्ययन

नेपालमा सामाजिक अध्ययन विषय पठनपाठन हुन थालेको पाँच दशक नाघि सकेको छ । सुरुमा प्राथमिक र निम्न माध्यमिक तहबाट सुरु गरिएको यस विषयलाई हाल कक्षा १२ सम्म समावेश गरिएको छ । माध्यमिक तह कक्षा ९-१० मा अध्यापन गरिने पाँचओटा अनिवार्य विषयहरूमध्ये सामाजिक अध्ययन विषय पनि एक अनिवार्य र महत्तपूर्ण विषय हो । सामाजिक अध्ययन विषयअन्तर्गत विभिन्न सामाजिक विज्ञानका विषयवस्तु समावेश गरिएको छ । यस विषयमा भूगोल, इतिहास, अर्थशास्त्र, नागरिक शास्त्र, समाजशास्त्र, जनसङ्ख्या अध्ययन एवम् समसामयिक विषय लगायतका विषय समेटिएका छन् ।

बदलिँदो विश्व परिवेश, राष्ट्रिय सन्दर्भ, समसामयिक विषयवस्तु र स्थानीय आवश्यकतालाई समेत सम्बोधन गर्ने गरी सामाजिक अध्ययन विषयको पाठ्यक्रम तयार पारिएको छ । यस विषयको अध्ययनबाट विद्यार्थीलाई असल नागरिक भई सफल जीवनयापन गर्न मदत पुग्ने छ । यसका साथै यस विषयमार्फत असल आचरण निर्माण, अनुशासन, जिम्मेवारी, सहयोग, सहकार्य, शान्ति, प्रेम, समानुभूति, विश्वास, भाइचारा र विश्वबन्धुत्व जस्ता मानवीय मूल्यलाई व्यावहारिक रूपमा विकास गराउने उद्देश्य राखेको छ । विविधताको सम्मान, सामाजिक न्यायको आत्मसात् एवम् समावेशिताको मूल्य विकास गराउने लक्ष्य पनि यस विषयले राखेको छ ।

सामाजिक अध्ययन विषयको पाठ्यक्रममा भएको व्यवस्था

राष्ट्रिय पाठ्यक्रम प्रारूप, २०७६ ले मार्गदर्शन गरेअनुरूप सामाजिक अध्ययन विषयको पाठ्यक्रम तयार गरिएको छ । यस पाठ्यक्रममा कक्षा १० का लागि समावेश गरिएका मुख्य विषयवस्तुहरू र प्रत्येक विषय क्षेत्रका लागि निश्चित कार्य घण्टा छुट्याइएको छ । पाठ्यक्रमको पृष्ठभूमि र प्रकृति अध्ययन गर्दा सामाजिक अध्ययन विषयको पाठ्यक्रमका उद्देश्यहरू यस प्रकार रहेको देखिन्छ ।

१. विद्यार्थीमा मानव जीवनसँग सम्बन्धित विभिन्न सामाजिक विज्ञानका विषयवस्तुसँग सम्बन्धित ज्ञान, सिप र अवधारणाको विकास गराई व्यक्तिलाई भविष्यमा आफ्नो समाज, राष्ट्र र अन्तर्राष्ट्रिय परिवेश बुझेर सोअनुसार समायोजन हुन सिकाउनु,

२. बालबालिकामा जीवनोपयोगी सिप र सामाजिक सिपहरूको विकास गर्नु,

३. बालबालिकामा लोकतान्त्रिक आचरण र मानव मूल्यको विकास गर्नु,

४. विद्यार्थीलाई असल नागरिकका रूपमा विकास गर्नु ।

पाठ्यक्रमले सामाजिक अध्ययन कक्षा १० अन्तर्गत चौध वटा सक्षमताहरु हासिल हुने अपेक्षा राखेको छ । उल्लेखित सक्षमताले मूलत विभिन्न पाँच क्षेत्र : **'नेपाली समाज'** (सामाजिकीकरण, मूर्त/अमूर्त सम्पदा आदि), **'नेपाल अध्ययन'** (भौगोलिक अवस्था, ऐतिहासिक घटना),**'नेपालको राजनीति र अन्तर्राष्ट्रिय सम्बन्ध'** (शासन प्रणाली, नागरिक कर्तव्य, दायित्व, अधिकार, अन्तर्राष्ट्रिय सम्बन्ध र प्रभाव, समसामयिक विषय), **'नेपालको अर्थव्यवस्था'** (अर्थतन्त्रका आधार र अवस्था, राजस्व, वित्तीय साक्षरता) तथा **'समसामयिक अध्ययन'** (सक्षमतासित सान्दर्भिक हुने राजनीतिक, आर्थिक, प्राकृतिक पक्ष जस्तै जलवायु परिर्वतन, विपद् व्यवस्थापन जस्ता विषयवस्तु) लाई समेटेको छ ।

सामाजिक अध्ययन विषयको विषय क्षेत्र र सिकाइ उपलब्धि

कक्षा १० को सामाजिक अध्ययनको पाठ्यक्रममा १० विषय क्षेत्रअन्तर्गत ४६ वटा सिकाइ उपलब्धि हासिल हुने अपेक्षा राखिएको छ । विषयवस्तुको अवधारणा बुझ्न, बुझाइलाई व्यवहारमा उतार्न, मूल्यमान्यताको संरक्षण गर्न, सहभागी हुन, ऐतिहासिक विषयवस्तुको खोजी गर्न, आत्मसात् गर्न जस्ता सिकाइ उपलब्धि समावेश गरिएका छन् ।

Table 15: सामाजिक अध्ययनको विषय क्षेत्र र सिकाइ उपलब्धि

एकाइ	विषय क्षेत्र	कार्य घण्टा	सिकाइ उपलब्धि सङ्ख्या	अङ्कभार
१	हामी र हाम्रो समाज	८	३	५
२	विकास र विकासका पूर्वाधार	१०	४	५
३	हाम्रा सामाजिक मूल्यमान्यता	१४	६	८
४	सामाजिक समस्या र समाधान	१२	३	६
५	नागरिक चेतना	१८	६	११
६	हाम्रो पृथ्वी	१८	४	११
७	हाम्रो विगत	१४	४	९
८	हाम्रा आर्थिक क्रियाकलाप	१५	६	९
९	हाम्रो अन्तर्राष्ट्रिय सम्बन्ध र सहयोग	११	६	६
१०	जनसङ्ख्या र यसको व्यवस्थापन	८	४	५
जम्मा		**१२८**	**४६**	**७५**

सामाजिक अध्ययन विषयको सिकाइ उपलब्धिमा राखिएका केही उपलब्धिहरू उदाहरणका रूपमा तल उल्लेख गरिएका छन् :

- सामाजिकीकरणको अवधारणासँग परिचित हुन र व्यवहारमा प्रदर्शन गर्न,
- मानव विकास सूचकाङ्ककको अवधारणा बताउन र नेपालको स्थितिलाई अन्य सार्क राष्ट्रहरूसँग तुलना गर्न,
- हाम्रा लोक तथा शास्त्रीय कलाको महत्त्व बताउन र संरक्षण कार्यमा सहभागी हुन,
- सामाजिक सुधारमा अन्तर्राष्ट्रिय संस्थाहरूले खेलेको भूमिका खोजी गर्न,
- राजनीतिक दलको भूमिका र जिम्मेवारी उल्लेख गर्न,
- नक्सा उतार्ने विभिन्न विधि/तरिकाको अभ्यास गर्न र नक्सासम्बन्धी आधुनिक व्यावहारिक प्रविधिहरूको उपयोग गर्न,
- आफ्नो क्षेत्रका ऐतिहासिक स्थल एवम् स्मारकहरूको खोजी पहिचान, संरक्षण र संवर्धन गर्न,
- आवधिक योजनाको परिचय दिन र चालु आवधिक योजनाका लक्ष्य र कार्यक्रमसँग परिचित हुन,
- स्थानीयकरण र विश्वव्यापीकरणको अवधारणा र महत्त्व प्रस्तुत गर्न,
- विश्वमा जनसङ्ख्याको अवस्था र वृद्धिदरको प्रवृत्ति पहिचान गर्न र वृद्धिदरको नेपालसँग तुलना गर्न ।

प्रश्नहरूको संज्ञानात्मक स्तर (Cognitive Level of the Items)

सामाजिक अध्ययनको पाठ्यक्रम अनुसारको विशिष्टीकरण तालिकामा संज्ञानात्मक तहअन्तर्गतको ज्ञान, बोध र तहको एक समूह, प्रयोग र सिर्जनात्मक सिप समूहको अर्को समूह र उच्च दक्षता (विश्लेषण र मूल्याङ्कन) समूहको अर्को गरी तीन वर्गमा बाँडिएको छ । यी संज्ञानात्मक स्तरहरू यस खण्डमा नमुना प्रश्नहरू सङ्क्षिप्त रूपमा वर्णन गरिएको छ । प्रश्नहरूको विवरणमा विभिन्न पक्षहरु समावेश गरिएको छ । जस्तै विषयवस्तुको क्षेत्र (एकाइ, सिकाइ उपलब्धि), प्रश्नको उद्देश्य, प्रश्नको विवरण, प्रश्नले मूल्याङ्कन गर्ने संज्ञानात्मक तहको विवरण यसरी प्रस्तुत गरिन्छ ।

१.ज्ञान तह (Knowledge level)

ज्ञान तह, ब्लुमको वर्गीकरणमा संज्ञानात्मक डोमेनको सबैभन्दा तल्लो तह हो । यस तहमा विद्यार्थीको सूचनाको पुनःस्मरण गर्ने र सम्झने क्षमताको मूल्याङ्कन गर्ने गरिन्छ । यस तहमा पहिले सिकेका कुराहरूको स्मरण गर्ने, दीर्घकालीन स्मृतिमा रहेको सान्दर्भिक ज्ञानलाई प्रत्याह्वान गर्ने क्षमताका प्रश्नहरू सोध्न सकिन्छ । उदाहरणका लागि

- यस तहमा कोर्सको विषयवस्तुको ज्ञान परीक्षण गर्ने, विशिष्ट तथ्यहरूको जानकारीको लेखाजोखा गर्ने, विधि र प्रक्रियाहरूको जानकारी गराउने, मुख्य घटनाहरूको सूची बनाउने, अवधारणाहरूको जानकारी र सिद्धान्तहरूको जानकारीसम्बन्धी पक्षको मूल्याङकन गर्ने जस्ता पक्षहरूलाई समेटिएको हुन्छ ।
- यस तहमा नयाँ अनुभवलाई पुरानो स्मृतिसँग तुलना गर्ने, दीघकालीन स्मृतिबाट ज्ञान वर्तमानमा ल्याउने प्रक्रियालाई समेटिन्छ । यस तहमा मुख्यतया पहिचान गर्नु (recognize) र प्रत्याह्वान गर्नु (Recall) गरी दुई उपतह हुन्छन् ।

ज्ञान तहलाई सरल माध्यमबाट मूल्याङ्कन गर्न सकिन्छ । यस तहमा को, कहिले, कहाँ, कति, कुन (Who, when, where, how many, which) जस्ता प्रश्न सोधिन्छ । ज्ञान तहका प्रश्नहरूमा परिभाषित गर्न (Define), सूची बनाउन (Enlist), उल्लेख गर्न (State, Mention) आदि क्रियापदहरूको प्रयोग गरिएको हुन्छ ।

नमुना प्रश्न १

प्रश्नको विषय क्षेत्र (Coverage of item) : कक्षा १०, अति छोटो उत्तर आउने प्रश्न, १ अङ्क

सिकाइको उपलब्धि (Learning Outcomes) : सामाजिकीकरणको अवधारणा बुझ्न

एकाइ (Unit) : एक, हामी र हाम्रो समाज

प्रश्नको उद्देश्य (Objective of item-in line with learning outcome): सामाजिकीकरणको अवधारणा र विशेषता बुझ्न

प्रश्न (Item) : सामाजिकीकरणको अर्थ एक वाक्यमा प्रस्ट पार्नुहोस् ।

उतरकुञ्जिका :

समाजमा व्यक्ति–व्यक्तिबिचको अन्तरक्रिया, व्यक्तिको आफ्नो रीतिरिवाज, मूल्यमान्यताप्रतिको सचेतना, अपनत्व, जिज्ञासा वा सामाजिक अन्तर्घुलनसम्बन्धी आशय लेखेमा १ अङ्क

प्रश्नको व्याख्यात्मक टिप्पणी : यो अति छोटो उत्तर आउने प्रश्न हो । यसले विद्यार्थीले सामाजिकीकरणको अवधारणको स्पष्टता मापन गर्छ । यो प्रश्नको उत्तर विद्यार्थीले पाठ्यपुस्तक वा अन्य पाठ्य सामग्रीहरूको अध्ययनबाट सम्झन सक्छ ।

२. बोध तह

संज्ञानात्मक क्षेत्रको दोस्रो तह बोध वा बुझाइ हो । यस तहमा विद्यार्थीहरूले बुझेको कुरालाई अभिव्यक्त गर्ने, निर्देशनमार्फत प्रस्तुत गरिएका सूचनाहरूको अर्थ निर्धारण गर्ने कार्य गर्दछन् । सूचनाको व्याख्या गर्ने, विषयवस्तु तथा सामग्रीको अर्थ बताउने, अनुवाद गर्ने, सारांश बताउने, अनुमान गर्ने तथा भविष्यवाणी गर्ने क्षमतासँग सम्बन्धित पक्षहरूलाई यस तहमा समेटिन्छ । उदाहरणका लागि

- निश्चित शब्दावली, परिभाषालाई आफ्नै शब्दमा पुनर्लेखन गर्ने, पाठ्यपुस्तकबाट सिकेका कुरा बाहेकका उदाहरण दिने,
- शाब्दिक सामग्री चार्ट र ग्राफलाई विश्लेषण गर्ने, सङ्क्षिप्तीकरण गर्ने आदि ।
- यस तहलाई सातओटा उपतहहरूमा वर्गीकरण गरिएको छ । ती हुन (१) अर्थ्याउनु (Interpreting) (२) उदाहरण दिनु (Exemplifying) (३) वर्गीकरण गर्नु (Classifying) (४) सारांश भन्नु (Summarizing) (५) निष्कर्षमा पुग्नु (Inferring) (६) तुलना गर्नु (Comparing) र (७) व्याख्या गर्नु (Explaining)।

बोध तहका प्रश्नहरूमा उदाहरणसहित व्याख्या गर्नु (Exemplify/Illustrate), तुलना गर्नु (compare), व्याख्या गर्नु (Explain) आदि क्रियापदहरूको प्रयोग गरिएको हुन्छ ।

नमुना प्रश्न २

प्रश्नको विषय क्षेत्र (Coverage of item): कक्षा १०, छोटो उत्तर आउने प्रश्न, ४ अङ्क

सिकाइको उपलब्धि (Learning Outcomes): जनसङ्ख्या वितरणको अवस्था बुझ्न

एकाइ (Unit): दश, जनसङ्ख्या र यसको व्यवस्थापन

प्रश्नको उद्देश्य (Objective of item- in line with learning outcome): जनसङ्ख्याको अवस्था र वृद्धिदरको प्रवृत्ति पहिचान गर्न

प्रश्न (Item): तपाईं बसोबास गर्ने प्रदेशको जनसङ्ख्या वितरणको अवस्था र त्यसको सामाजिक प्रभाव तीन बुँदामा स्पष्ट पार्नुहोस् ।

उतरकुञ्जिका :

- आफू बसोबास गर्ने प्रदेशको जनसङ्ख्याको अवस्था (जनघनत्व, जनसङ्ख्या बृद्धि आदि) उल्लेख गरेमा १ अङ्क
- जनसङ्ख्या वितरणको अवस्था उच्च भएमा,
- सामाजिक सुविधा उपलब्धतामा कमी वा पर्याप्तता
- शान्ति सुरक्षा कायम गर्न चुनौती, सामाजिक द्वन्द्व वा यस्तै आशय आउने प्रभाव उल्लेख गरेमा प्रत्येक बुँदाका लागि १ अङ्कका दरले तीन बुँदाका लागि ३ अङ्क (१ × ३ = ३ अङ्क)
- जनसङ्ख्या वितरणको अवस्था पातलो, न्यून भएमा,
    - सामाजिक सुविधा प्रबाहमा कठिनाइ,
    - स्रोत र साधन परिचालनमा कठिनाइ,
    - सामाजिक सद्भाव कायम रहने वा यस्तै आशय आउने उत्तर लेखेमा प्रत्येक बुँदाका लागि १ अङ्कका दरले तीन बुँदाका लागि ३ अङ्क (१ × ३ = ३ अङ्क)

प्रश्नको व्याख्यात्मक टिप्पणी : यो प्रश्नमा विद्यार्थीले नेपालको प्रादेशिक वा भौगोलिक जनसाङ्ख्यिक वितरण बुझे नबुझेको मापन गर्छ । विद्यार्थीहरूले असमान जनङ्ख्या वितरणले कस्तो प्रभाव पर्छ भनी बुझेर आफ्ना विचार लेख्न सक्ने अपेक्षा गरिएको छ तथा उच्च घनत्वले कस्तो प्रभाव पर्छ भनी प्रस्तुत गर्न सक्ने छन् ।

३. प्रयोग सिप

यस तहमा बुझेको कुरालाई भिन्न परिस्थितिमा प्रयोग तथा सामान्यीकरण गर्ने खालका उद्देश्य र प्रश्नहरू पर्दछन् । यस तहमा प्राप्त सूचनालाई नयाँ अवस्था वा परिस्थितिमा प्रयोग गर्ने, सिकेका अनुभवहरूलाई नयाँ र ठोस अवस्थामा प्रयोग गर्ने, नयाँ परिस्थितिमा सिद्धान्तहरूको प्रयोग गर्ने, चार्ट र ग्राफको रचना गर्ने, मुख्य घटनाहरूलाई समय रेखा, ग्राफ, पाइचार्टमा देखाउने, विधि वा तरिकाहरूलाई क्रमबद्ध रूपमा प्रदर्शन गर्ने कार्य गरिन्छ । विद्यार्थीहरूले प्रयोग तहका प्रश्नहरूको उत्तर दिनका लागि आफूले सिकेका ज्ञान र सिपलाई पुनः व्यवस्थित गरी समस्या समाधान गर्नुपर्ने हुन्छ ।

प्रयोग तहका प्रश्नहरूमा उपयोग गर्नु, प्रयोग गर्नु (Apply/Use), सामान्यीकरण गर्नु (Generalize), निमाण गर्नु (Construct) जस्ता क्रियापदहरू प्रयोग गरिन्छ । सिर्जनात्मक तहका प्रश्नहरूमा सिजना गर्नु (Create), ढाँचा तयार गर्नु, विकास गर्नु, योजना तथा प्रस्तावना तयार गर्नु जस्ता क्रियापदहरू प्रयोग गरिन्छ ।

नमूना प्रश्न ३

प्रश्नको विषय क्षेत्र (Coverage of item) : कक्षा १०, छोटो उत्तर आउने प्रश्न, ४ अङ्क

सिकाइको उपलब्धि(Learning Outcomes) : नेपालको पररा‍ष्ट्र नीति बुझ्न

एकाइ (Unit): नौ, हाम्रो अन्तर्राष्ट्रिय सम्बन्ध र सहयोग

प्रश्नको उद्धेश्य (Objective of item-in line with learning outcome): नेपालले लिएको पररा‍ष्ट्र नीति बुझ्न र सोहीअनुसार अन्तर्राष्ट्रिय मञ्चमा नेपालको भूमिका कस्तो हुन पर्ला भनी व्यावहारिक सुझाव दिन सक्नु

प्रश्न (Item) : विश्वमा कतिपय स्वतन्त्र र सार्वभौम मुलुकको आन्तरिक मामलामा बाह्य हस्तक्षेप बढेको छ । असंलग्न पररा‍ष्ट्र नीति अबलम्बन गरेको मुलुक नेपालले संयुक्त राष्ट्रसंघ लगायत विश्व मञ्चमा नेपालले के कस्ता विषय उठाउनु पर्ला ? नेपालले उठाउनुपर्ने विषय समेटी न्युयोर्क स्थित संयुक्त राष्ट्र सङ्घका लागि नेपालको स्थायी नियोगलाई लेख्ने इमेलको नमुना तयार पार्नुहोस् ।

उतरकुञ्जिका :

- इमेलको नमुनाको लागि १ अङ्क
- दुवै राष्ट्रहरूसित आ आफ्ना समस्या अन्तर्राष्ट्रिय मञ्चमा स्पष्ट आह्वान गर्ने,
- संयुक्त राष्ट्र सङ्घ लगायतका शान्ति स्थापनार्थ संस्थालार्य नेतृत्व भूमिका निर्वाह गर्न दबाब दिने,
- द्विपक्षीय द्वन्द्व न्यूनीकरण लागि भएका विगतका प्रयासहरू आधार मानी सोहीअनुसार नेपालले आवाज उठाउने,
- पञ्चशीलको सिद्धान्तको व्यवहारिक अभ्यासका लागि आवज उठाउने वा यस्तै शान्तिपूर्ण गतिविधिमा नेपाललाई सक्रिय सहभागिता गराउने आशयको उत्तर लेखेमा प्रत्येक बुँदाका लागि १ अङ्कका दरले तीन बुँदाका लागि ३ अङ्क (१ × ३ = ३ अङ्क)

प्रश्नको व्याख्यात्मक टिप्पणी : यो प्रश्न प्रयोग र सिर्जनात्मक सिपसँग सम्बन्धित छ । यसले विद्यार्थीको इमेल, वक्तृत्व, प्रश्नावली, अन्तर्वार्ता लगायतका व्यावहारिक लेखन सिपको सिर्जनात्मक क्षमता मापन गर्छ । त्यस्तै नेपालको पररा‍ष्ट्र नीति कस्तो छ ? यसको औचित्य, सान्दर्भिकता, यसलाई प्रभावकारी बनाउन के गर्न पर्ला ? जस्ता विषयमा विद्यार्थीको क्षमता परीक्षण गर्छ ।

४. उच्च दक्षता

विभिन्न परिस्थितिमा ज्ञानको प्रयोग गर्ने, क्षमताको परीक्षण गर्न सक्ने प्रश्नलाई उच्च दक्षतामा वर्गीकृत गरिन्छ । ब्लूमको वर्गीकरणको सन्दर्भमा यस्ता प्रश्नहरू विश्लेषण, संश्लेषण, र मूल्याङ्कन (संशोधित संस्करणमा विश्लेषण, मूल्याङ्कन, र सिर्जना अभिवृत्ति) हुन् । विश्लेषणअन्तर्गत विद्यार्थीले जटिल अवधारणालाई त्यसको आधारभूत भाग वा तत्त्वहरूमा विभाजित गर्दछ । उसले तिनीहरूको सम्बन्ध बुझेको देखाउँछ । जबकि संश्लेषणमा तत्त्व वा भागहरूलाई एक सुसङ्गत पूर्ण ढाँचा बनाउनकोलागि एक साथ राखिन्छ । यसैगरी मूल्याङ्कनमा विद्यार्थीले मापदण्डको आधारमा निर्णय गर्ने क्षमता देखाउँछ । अभिवृत्तिले विद्यार्थीको सोचाइ, धारणा परीक्षण गर्दछ । उदाहरणका लागि

- कुनै कुरालाई यसको घटक/तत्त्वहरूमा अलग गर्न वा सम्बन्धको विश्लेषण गर्न (प्रमाण र परिकल्पना, कारण, प्रभाव, क्रम पत्ता लगाउन, निष्कर्ष निकाल्न) वा संगठनात्मक सिद्धान्तहरूको विश्लेषण गर्न,
- परिकल्पनाबाट तथ्यहरू छुट्याउन, निष्कर्ष र समर्थन कथनको पहिचान गर्न, एक विचारको अर्कोसँग सम्बन्ध पत्ता लगाउन,
- कुनै पनि दुई वा दुईभन्दा बढी अवयव वा घटना जस्तै- परिकल्पना, प्रमाण वा निष्कर्ष, तार्किक भ्रमहरूको पहिचान, कारण र प्रभाव सम्बन्ध तथा क्रमहरूबिचको सम्बन्ध पत्ता लगाउन ।

उच्च तहको विश्लेषणात्मक प्रश्नहरूमा विश्लेषण गर्नु (Analyze), सङ्गठित गर्नु(Organize), गुणदोष, फाइदा बेफाइदा, सबल र दुर्बल पक्ष बताउनु (attribute) आदि जस्ता क्रियापदहरु प्रयोग गरिन्छ । त्यस्तै मूल्याङ्कन तहका प्रश्नहरूमा मूल्याङ्कन, लेखाजाखा गर्नु (Evaluate/judge/appraise/assess), राय, सल्लाह, सुझावहरू दिनु(Suggest/advise/recommend/opinion), तर्कदिनु, प्रतिरक्षा गर्नु, मूल्य दिनु, टिप्पणी गर्नु आदि जस्ता क्रियापदहरू प्रयोग गरिन्छ ।

नमूना प्रश्न ४

प्रश्नको विषय क्षेत्र (Coverage of item) : कक्षा १०, लामो उत्तर आउने प्रश्न, ७ अङ्क ।

सिकाइको उपलब्धि(Learning Outcomes) : राज्यका अङ्ग (कार्यपालिका) हरूको कार्य प्रभावकारिता मूल्याङ्कन, विश्लेषण गर्न ।

एकाइ (Unit) : पाँच, नागरिक चेतना

प्रश्नको उद्देश्य (Objective of item-in line with learning outcome):राज्यका अङ्गहरूमध्ये कार्यपालिकाका कार्यहरु बोध गरी उल्लिखित कार्यहरू विश्लेषण, समसामयिक दायित्वको बोध गर्न ।

प्रश्न (Item) : कार्यपालिका (सरकार) को भूमिका परिवर्तित सन्दर्भमा शान्ति सुव्यवस्था तथा प्रशासनिक गतिविधि प्रभावकारी हुनुपर्ने जन चासो रहेको छ । यस विषयमा सरकारको गतिविधिले कसरी जन चाहना समेटोस् भन्ने अपेक्षा गर्नुपर्छ ? विश्लेषण गर्नुहोस् ।

उतरकुञ्जिका :

परिवर्तित सन्दर्भमा समय सरकारका गतिविधिहरूले समेट्नुपर्ने विषय :

- पारदर्शिता,
- सुशासन,
- प्रविधिमैत्री सेवा प्रवाह,
- मानव अधिकारको प्रत्याभूति,
- सार्वजनिक नीजि क्षेत्र साझेदारीमा आधारित विकासका गतिविधि,
- राष्ट्रिय मूल प्रवाहमा समावेशिता,
- स्रोत साधनको संरक्षण लगाायतका समसामयिक विषय समेटी विश्लेषण वा शान्ति सुव्यवस्था तथा प्रशासनिक प्रभावकारिताको आशय आउने बुँदा लेखेमा प्रत्येक बुँदाका लागि १ अङ्कका दरले पाँच बुँदाका लागि ५ अङ्क (१ × ५ = ५ अङ्क)

प्रश्नको व्याख्यात्मक टिप्पणी : यो प्रश्न उच्च दक्षता तह संग सम्बन्धित छ । यसमा विद्यार्थीहरूले राजयका अङ्गहरुका कार्यहरु पाठ्यपुस्तकबाट बुझ्नेछन् । उल्लिखित कार्यहरू के कति प्रभावकारिता भए नभएको मूल्याङ्कन, विश्लेषण गर्नेछन् । कार्यहरु प्रभावकारी नभएमा वा के कस्ता जन चाहनालाई सम्बोधन गर्नुपर्ला भनी विद्यार्थीहरूको अवधारणा खोज्नु यस्ता उच्च तहका प्रश्नको उद्देश्य हुन्छ ।

सामाजिक अध्ययन विषयको विशिष्टीकरण तालिका

विशिष्टीकरण तालिका मूल्याङ्कनका लागि प्रश्न निर्माण गर्ने मार्ग निर्देशक सामग्री हो । यसले सन्तुलित र व्यवस्थित ढङ्गबाट लिखित परीक्षाका लागि प्रश्न पत्र तयार गर्न सहयोग गर्दछ । कक्षा १० को सामाजिक अध्ययन विषयको प्रश्न निर्माणका लागि पाठ्यक्रम विकास केन्द्रले तपशिल बमोजिमको विशिष्टीकरण तालिका तयार पारेको छ । अहिले विद्यालय शिक्षा परीक्षा (SEE) मा यही विशिष्टीकरण तालिकाअनुसारका प्रश्नहरू तयार पारी प्रयोग गरिँदै आएको छ ।

Table 16:विशिष्टीकरण तालिका-सामाजिक अध्ययन

एकाइ	कार्य घण्टा	अति छोटो उत्तर आउने प्रश्न			छोटो उत्तर आउने प्रश्न			लामो उत्तर आउने प्रश्न			एकाइगत अङ्क विभाजन
		ज्ञान र बोध	प्रयोग र सिर्जनात्मक सिप	उच्च दक्षता	ज्ञान र बोध	प्रयोग र सिर्जनात्मक सिप	उच्च दक्षता	ज्ञान र बोध	प्रयोग र सिर्जनात्मक सिप	उच्च दक्षता	
१. हामी र हाम्रो समाज	८	१(१)					१(४)				५
२. विकासका पूर्वाधार	१०			१(१)	१(४)						५
३. हाम्रो परम्परा एवम् सामाजिक मूल्य र मान्यता	१४					१(४)	१(४)				८
४. सामाजिक समस्या र समाधान	१२	१(१)	१(१)				१(४)				६
५. नागरिक चेतना	१८				१(४)					१(७)	११
६. हाम्रो पृथ्वी	१८				१(४)				१(७)		११
७. हाम्रो विगत	१४	१(१)		१(१)				१(७)			९
८. हाम्रा आर्थिक क्रियाकलाप	१५	१(१)	१(१)						१(७)		९
९. हाम्रो अन्तर्राष्ट्रिय सम्बन्ध र सहयोग	११	१(१)		१(१)		१(४)					६
१०. जनसङ्ख्या र यसको व्यवस्थापन	८		१(१)		१(४)						५
जम्मा	१२८	५	३	३	४	२	३	१	२	१	७५

यस विशिष्टीकरण तालिकाअनुसार अति छोटो, छोटो र लामो गरी तीन प्रकारका प्रश्न सोध्न सकिने व्यवस्था रहेको छ। अति छोटो उत्तर आउने प्रश्नहरू १ अङ्कको ११ ओटा, छोटो उत्तर आउने प्रश्नहरू ४ अङ्कको ९ ओटा र लामो उत्तर आउने प्रश्नहरू ७ अङ्कको ४ ओटा रहने गरी जम्मा ७५ अंकको बाह्य परीक्षाको व्यवस्था गरिएको छ।

तीनओटै प्रकारका प्रश्नमा संज्ञानात्मक तहअन्तर्गत प्रश्नहरूलाई तीन वर्गमा बाँडिएको छ। ज्ञान र बोध क्षेत्रबाट प्रश्न सोध्दा ज्ञान र बोधको सन्तुलन मिलाउनु पर्ने छ। एकाइ ६ बाट छोटो र लामो प्रश्न (प्रयोगात्मक र सिर्जनात्मक तह) सोध्दा रेखा-चित्र (Line Graph), वृत्तचित्र वा कुनै सैद्धान्तिक प्रश्न सोध्न सकिने छ भने लामो उत्तर प्रश्नमा नक्सा कार्य वा परियोजना कार्य सोध्न सकिने छ। प्रश्न १५ मा एकाइ ३ बाट एउटा छोटो प्रश्न सोधिने छ र त्यसको वैकल्पिक प्रश्न एकाइ ७ वा ८ मध्ये कुनै एकबाट सोधिने छ ।

प्रश्न निर्माणका लागि विशिष्टीकरण तालिकाले केही साँघुरो प्रश्नपत्र निर्माण गर्न खोजेको देखिन्छ । जस्तै, 'हामी र हाम्रो समाज' एकाइबाट दुई प्रकारका प्रश्न मात्र बनाउन पाइन्छ र अन्य सम्भाव्य ७ प्रकारका प्रश्न निर्माण गर्न सकिन्छ । हामी हाम्रो समाज एकाइसँग सम्बन्धित ३ वटा सिकाइ उपलब्धि छन् जसले गर्दा यस एकाइबाट २×३=६ प्रकारका प्रश्नहरू बनाउन सकिन्छ ।

प्रश्न निर्माणका लागि विशिष्टीकरण तालिकाको व्याख्या

कुनै सिकाइ उपलब्धिको मापन गर्न पाठबाट कति नम्बरको कुन तहको कुन ढाँचाको प्रश्न सोध्ने भन्ने कुराको निर्क्यौल गर्न विशिष्टीकरण तालिकालाई थप स्पष्ट पार्नुपर्दछ । व्याख्यात्मक विशिष्टीकरण तालिकाको केही अंश तल प्रस्तुत गरिएको छ ।

Table 17: **विशिष्टीकरण तालिकाको व्याख्या**

एकाइ सङ्ख्या	विषय क्षेत्र		प्रश्न सेल कोड								
	सिकाइ उपलब्धि क्रम सङ्ख्या	सिकाइ उपलब्धि	ज्ञान र बोध			प्रयोग र सिर्जनात्मक सिप			उच्च दक्षता		
			अति छोटो उत्तर	छोटो उत्तर	लामो उत्तर	अति छोटो उत्तर	छोटो उत्तर	लामो उत्तर	अति छोटो उत्तर	छोटो उत्तर	लामो उत्तर
१. हामी र हाम्रो समाज											
१	१.१	सामाजिकीकरणको अवधारणासँग परिचित हुन र व्यवहारमा प्रदर्शन गर्न	१							२	
	१.२	हाम्रा सामाजिक /सांस्कृतिक विविधताको पहिचान र सामाजिक एकता संरक्षण गर्न									
	१.३	हाम्रा राष्ट्रिय गौरवको पहिचान र सम्मान गर्न									

व्याख्यात्मक विशिष्टीकरण तालिकाको प्रत्येक कोठाबाट प्रश्न निर्माण गर्नका लागि नम्बर कोडबाट इङ्गित गरिन्छ । व्याख्यात्मक विशिष्टीकरण तालिकाले सम्भाव्य प्रश्नहरु प्रस्तुत गर्दछ र सोअनुसार प्रश्न कार्ड प्रयोग गरी प्रश्न निर्माण गर्न सकिन्छ ।

कुनै पनि विषयमा प्रश्न निर्माण गर्दा पहिले त्यस प्रश्नका आधारहरु स्पष्ट पार्नु आवश्यक हुन्छ, जस्तैः पाठ्य इकाइ, परीक्षण गरिने सिकाइ उपलब्धि, संज्ञानात्मक तह तथा प्रयोग हुने ढाँचा र प्राप्त हुने अधिकतम अङ्क । यी आधारहरु स्पष्ट पारेपछि प्रश्नले मापन गर्ने सिकाइ उपलब्धि स्पष्ट पार्नुपर्छ र प्रश्न निर्माण गरिनु पर्दछ साथै उत्तर कुञ्जिका पनि समावेश गरिनु पर्दछ । प्रश्न कार्डको एउटा नमूना तल दिइएको छ :

प्रश्न कार्ड

१) कक्षा : दश

२) विषयः सामाजिक अध्ययन

३) प्रश्न सेल कोडः २

४) विस्तृत प्रश्न	**एकाइ**	**सिकाइ उपलब्धि नं.**	**संज्ञानात्मक सिप**	**ढाँचा**	**अङ्क**
कोड	१	१.१	उच्च दक्षता	छोटो	४

५) सिकाइ उद्देश्य : सामाजिकीकरणको अवधारणासँग परिचित हुन र व्यवहारमा प्रदर्शन गर्न

६) प्रश्नकोउद्देश्य (सिकाइ उपलब्धिको आधारमा): सामाजिकीरणका तत्त्वले सामाजिकीरणमा पारेको प्रभाव मूल्याङ्कन गर्नु

७) प्रश्नः

पछिल्लो समय सामाजिकीरणमा सामाजिक सञ्जालको प्रयोग बढ्दो छ । हाम्रो सामाजिकीरणमा सामाजिक सञ्जालको प्रभाव कस्तो पाउनुहुन्छ ? चार बुँदामा समीक्षा गर्नुहोस् ।

८) मुख्य उत्तर / उत्तर कुञ्जिकाः

- व्यक्तिको अन्तर्निहित क्षमता विकास भएको,
- छिटो, छरितो ढङ्गले विचार, सूचना प्रवाह भएको,
- जनमत सङ्कलनमा प्रभावकारी भएको
- तथापि व्यक्तिगत सोच, अन्तरक्रियाको अभाव एकलकाँटे प्रवृत्ति बढेको वा यस्तै समसामयिक प्रभाव उल्लेख गरेमा एउटा सही बुँदालाई १ अङ्क, अधिकतम चार बुँदाका लागि ४ अङ्क (१ × ४ = ४ अङ्क)

प्रश्न मेट्रिक्स

पाठ्यक्रमको सिकाइ उपलब्धिको आधारमा विशिष्टिकरण तालिकाको आवश्यकता पूरा हुने गरी प्रश्नहरू निर्माण हुन्छन् । यसरी व्याख्यात्मक विशिष्टीकरण तालिकाको कोठामा निर्माण भएका प्रश्नहरूबाट पाठ्यक्रमले उल्लेख गरेको विशिष्टिकरण तालिकाको आवश्यकता पुरा हुने गरी प्रश्नपत्र तयार गरिन्छ । प्रश्नपत्रमा उक्त आवश्यकता पूरा भएको छ भनी देखाउन प्रश्न मेट्रिक्सको प्रयोग गर्नु पर्नेछ । प्रश्न मेर्ट्रिक्सको नमूना तल दिइएको छ :

Table 18: प्रश्न मेट्रिक्सको नमुना

क्र. स.	एकाइ	कार्य घण्टा	ज्ञान (१०)/ बोध (१८)		प्रयोग/ सिर्जनात्मक सिप (१८)			उच्च दक्षता (१८)			एकाइगत अङ्क विभाजन भएको		एकाइगत अङ्क विभाजन हुनुपर्ने
१	हामी र हाम्रो समाज	८	अति छोटो							छोटो उत्तर			५
२	विकास र विकासका पूर्वाधार	१०		छोटो उत्तर					अति छोटो				५
३	हाम्रो परम्परा एवम् सामाजिक मूल्य र मान्यता	१४					छोटो उत्तर			छोटो उत्तर			८
४	सामाजिक समस्या र समाधान	१२	अति छोटो			अति छोटो				छोटो उत्तर			६
५	नागरिक चेतना	१८		छोटो उत्तर							लामो उत्तर		११
६	हाम्रो पृथ्वी	१८		छोटो उत्तर				लामो उत्तर					११
७	हाम्रो विगत	१४	अति छोटो		लामो उत्तर				अति छोटो				९
८	हाम्रा आर्थिक क्रियाकलाप	१५	अति छोटो			अति छोटो		लामो उत्तर					९
९	हाम्रो अन्तर्राष्ट्रिय सम्बन्ध र सहयोग	११	अति छोटो				छोटो उत्तर		अति छोटो				६
१०	जनसङ्ख्या र यसको व्यवस्थापन	८		छोटो उत्तर		अति छोटो							५
संज्ञानात्मक क्षेत्र जम्मा		१२८	५	४	१	३	२	२	३	३	१		७५
जम्मा प्रश्नगत भार													
प्रश्नको ढाँचा / सङ्ख्या			अति छोटो उत्तर (११)		छोटो उत्तर (९)			लामो उत्तर (४)			जम्मा (२४)		
प्रश्नको भार			११		३६			२८			७५		

माथि उल्लेख गरेअनुसार विशिष्टीकरण तालिकाले प्रश्नपत्र निर्माणमा साँघुरो बाटो दिएकोले प्रश्न पत्र निर्माण पनि साँघुरो भएर

जान्छ । जुन एकाइमा २ वा २ भन्दा बढी सिकाइउपलब्धिछ त्यसबाट फरक प्रकारको प्रश्न निर्माण गरेर प्रश्न पत्र मेट्रिक्समा विविधता ल्याउन सकिन्छ । अझै प्रश्नमा फरक / विविधता ल्याएर हुबहु प्रश्न पत्र निर्माण हुने सम्भावनालाई न्यूनीकरण गर्न सकिन्छ ।

आन्तरिक मूल्याङ्कन

सामाजिक अध्ययनमा आन्तरिक मूल्याङ्कनको निकै महत्त्व रहेको छ । सामाजिक अध्ययन विषयको मूल उद्देश्य विद्यार्थीलाई असल नागरिक बनाउने हो । यसका लागि विषयवस्तुको ज्ञान मात्र भएर पुग्दैन । त्यसका साथसाथै जीवनोपयोगी सिप र सामाजिक व्यवहार पनि सिकाउनु पर्ने हुन्छ । जुन विद्यार्थी केन्द्रित सिकाइ सहजीकरण क्रियाकलापका माध्यमबाट गरिन्छ । सामाजिक व्यवहार र जीवनोपयोगी सिपको विकास भयो कि भएन भनेर लिखित परीक्षाबाट मापन गर्न सकिँदैन । त्यसका लागि अवलोकन लगायतका वैकल्पिक मूल्याङ्कनका साधनहरूको प्रयोग गर्नु पर्ने हुन्छ । जो विद्यालय र विषय शिक्षकले निरन्तर गरिहेका हुन्छन् । शिक्षकले वर्षभरि गरेको मूल्याङ्कन र विद्यार्थीले नियमितरूपमा गरेका सिर्जनात्मक कार्य लगायतका समग्र पक्षको पनि मूल्याङ्कन गर्ने उद्देश्यले आन्तरिक मूल्याङ्कन गर्ने व्यवस्था रहेको छ । आन्तरिक मूल्याङ्कन तर्फको २५ प्रतिशत अङ्कलाई पाठ्यक्रमा निम्नानुसार विभाजन गरिएको छ ।

Table 19: आन्तरिक मूल्याङ्कन - सामाजिक अध्ययन

सि.नं.	मूल्याङ्कन पक्ष,	भार
१	सहभागिता : कक्षामा उपस्थिति र सिकाइ सहभागिता	३
२	प्रयोगात्मक/परियोजना कार्य	१६
३	त्रैमासिक परीक्षाहरू	६

आन्तरिक मूल्याङ्कनका लागि विस्तृत मूल्याङ्कन रमापदण्ड, पाठ्यक्रम विकास केन्द्रले तयार पारेको*विद्यार्थी सिकाइका लागि आन्तरिक मूल्याङ्कन मापदण्ड, २०८०* मा दिए अनुसार पालना गर्नुपर्दछ ।

बाह्य मूल्याङ्कन

सामाजिक अध्ययनमा बाह्य मूल्याङ्कनको भार ७५ प्रतिशत हुने छ ।अन्तिम परीक्षाका प्रश्नपत्र पाठ्यक्रम विकास केन्द्रले तयार पारेको विशिष्टीकरण तालिकालाई आधार मानी निर्माण गरिनु पर्दछ । विशिष्टीकरण तालिकाअनुसार अति छोटो, छोटो र लामोगरी तीन प्रकारका प्रश्न सोध्न सकिनन व्यवस्था रहेको छ ।

Table 20: बाह्य मूल्याङ्कन - सामाजिक अध्ययन

प्रश्नका प्रकार	पूर्णाङ्क प्रति प्रश्न	जम्मा प्रश्न सङ्ख्या	पूर्णाङ्क	समय
अति छोटो उत्तर आउने प्रश्न	१	११	११	२६.४ मिनेट
छोटो उत्तर आउने प्रश्न	४	९	३६	८६.४ मिनेट
लामो उत्तर आउने प्रश्न	७	४	२८	६७.२ मिनेट
जम्मा		**२४**	**७५**	**१८० मिनेट**

त्यस्तै संज्ञानात्मक स्तरअनुरूप प्रश्नको ढाँचा निम्नानुसार तय गरिएको छ ।

Table 21: संज्ञानात्मक प्रश्नहरूको वर्गीकरण

संज्ञानात्मक स्तर	अति छोटो उत्तर आउने प्रश्न		छोटो उत्तर आउने प्रश्न		लामो उत्तर आउने प्रश्न		जम्मा प्रश्न सङ्ख्या र पूर्णाङ्क	
	प्रश्न सङ्ख्या	पूर्णाङ्क	प्रश्न सङ्ख्या	पूर्णाङ्क	प्रश्न सङ्ख्या	पूर्णाङ्क	प्रश्न सङ्ख्या	पूर्णाङ्क
ज्ञान तह	५	५	४	१६	१	७	१०	२८
प्रयोग र सिर्जनात्मक सिप	३	३	२	८	२	१४	७	२५
उच्च तह	३	३	३	१२	१	७	७	२२
जम्मा	११	११	९	३६	४	२८	**२४**	**७५**

नतिजा विश्लेषण तथा शिक्षण सिकाइमा सुधार

सामाजिक अध्ययन (कक्षा १०) को सैद्धान्तिक मूल्याङ्कन ७५ पूर्णाङ्कको हुने छ । सिकाइ उपलब्धि र विशिष्टीकरण तालिकामा उल्लेख भए बमोजिम संज्ञानात्मक तहका आधारमा कुल २४ प्रश्नहरू सोधिनेछन् ।

सामाजिक अध्ययनमा कतिपय त्रुटि प्रश्नपत्र र उत्तरकुञ्जिका निर्माणमा समेत देखिएको छ । उदाहरणका लागि

प्रश्न नं १ *तपाईंको घर, परिवार वा समुदायमा जेष्ठ नागरिक हुनुपर्छ । उहाँहरूको सम्मानित जीवनयापनका लागि तपाईंले के कस्ता कार्यहरु गर्न सक्नुहुन्छ ? आफ्ना योजना चार बुँदामा स्पष्ट पार्नुहोस् ।*

यस प्रश्नमा विद्यार्थीले गर्न सक्ने कार्यहरू उल्लेख गर्नुपर्ने हुन्छ । तर विद्यार्थीले सरकार वा सरोकारवालाले गरेका कार्यहरू मात्रै उल्लेख गरेको पाइन्छ । उल्लिखित प्रश्नमा उत्तरकुञ्जिका निर्माण गरिदा सरकारले गरेका कार्यहरूलाई बुदाँगत रुपमा उल्लेख नगरी विद्यार्थीको प्रस्तुति, भाषा, योजना कस्तो हुनुपर्छ भन्ने विषयलाई विचार गरिनुपर्ने देखिन्छ ।

प्रश्न नं २ *नेपालमा मानव अधिकार संरक्षणमा मानव अधिकारवादी सामाजिक सङ्घसंस्थाको भूमिका महत्त्वपूर्ण हुँदाहुँदै पनि मानव अधिकार हननका घटनाहरु किन बारम्बार दोहोरिन्छन् ? समिक्षा गर्नुहोस् ।*

उल्लिखित प्रश्नमा मानव अधिकारवादी संस्थाको योगदान कस्तो छ ? भन्ने विषय उल्लेख गर्नु पर्नेमा सङ्घसंस्थाको भूमिका कमै विद्यार्थीले उल्लेख गर्ने भएकाले आशातित अङ्क प्राप्त गर्न नसकेको देखिन्छ । यस प्रश्नमा सङ्घसंस्था भूमिका उल्लेख गर्नुपर्ने, कमजोरी पहिचान गर्नुपर्ने र केही सुझाव समेत उल्लेख गरी उत्तर टुक्रयाउनु पर्ने, अङ्क विभाजन सोहीअनुसार हुनुपर्ने जस्ता कुरा विद्यार्थीले नबुझेको, नबुझाइएको र उत्तरकुञ्जिका निर्माण नभएको देखिन्छ ।

यसप्रकार माथि उल्लिखित विषयका अतिरिक्त प्रश्नपत्र उत्तरकुञ्जिका निर्माण गरिदा खासगरी 'उच्च दक्षता', सिर्जनात्मक/प्रयोगात्मक उत्तर प्रस्तुतिमा कमजोरी देखिएकोले

- सिकाइ उपलब्धिलाई आधार मानिनुपर्ने,
- उत्तरकुञ्जिका निर्माण गर्दा प्रश्नको प्रकृतिअनुसार अङ्क विभाजन गरिनुपर्ने, आशय वा आंशिक उत्तर मिलेमा अंक विभाजन गरिनुपर्ने,
- विषयवस्तुको प्रकृतिअनुसार छायाँ पाठ्यक्रम (Hidden Curriculum) को प्रवेश गराउन, खोजमूलक कार्यमूलक समसामयिक विषयलाई पाठ्क्रम सिकाइ उपलब्धिका आधारमा प्रश्नपत्र उत्तरकुञ्जिका निर्माण गरिन पर्ने,
- शिक्षण क्रियाकलाप गरिँदा विषयवस्तुसँग मेल खाने सान्दर्भिक उदाहरण र सिकाइ उपललिब्ध, पाठ्यक्रमको विस्तृतीकरणलाई विचार गरिनुपर्ने ।

Chapter VII

Way Forward

Apart from providing assessment procedures in the SEE in general, this framework is also intended to guide test improvement process – mainly guiding the tasks and processes. Another main intent of assessment is to draw implications for improvement of student's learning. These two main tasks, viz. test improvement of SEE and result analysis for improvement of learning are major thrust for further initiatives.

Item development: Developing items for four subjects, viz. English, Mathematics and Science and Technology, and Social Studies of Grade 10 will be based on this framework. Activities involved in this process will be elaboration of specification grids in order to help singular learning outcome in which an item or sub-question could be developed. These items as well as sub-questions if there are any, would be mapped to their respective LO with respective cell code(s). These are explained in subject-wise chapters above.

Selection of items for item pre-testing: Item (with key answer or marking scheme as applicable) submitted by the item writers will be paneled by the expert team and moderated by the moderation committee. Items selected by this process should be pre-tested to ensure quality. For item pre-testing, test sets will be assembled into multiple sets resembling to the specification grid of the subject. These items will be pre-tested in a representative sample of users. Based on the item analysis, items that meet item selection criteria will be selected and kept in the final repository of items. From final repository of items, equivalent test sets can be prepared for final use. For all the test sets, test matrix will also be prepared individually and checked against requirements of the specification grid.

Final test: Final sets are prepared for implementation. These sets are expected to be comparable.

Establishment of item banking system: Developing IBS and transferring final items in the system is essential. IBS is intended to be used to share model items and test papers for stakeholders use and secured system for accessing final items for final test generation by the concerned authority only. Such IBS would be collaboratively developed and collectively used by the concerned stakeholders.

Result analysis and reporting: Each body responsible for examination will be oriented/trained on result analysis, report generation and prepare learning improvement plan based on the students' performance on the test items. Such report can be generated at different levels – school, LG, province and national.

Item bank maintenance: Item banking maintenance would be as a regular task, as calibrated items need to be restocked as they are used from the IBS. Embedded item pre-testing or IRT based on calibrated items can be adopted for item baking maintenance. Restocking of item bank and ongoing maintenance of test items should be institutionalized under a department/division in NEB or such institutional arrangement. Expert team should be formed and made available to support as and when required for this task.

Certification: Certification of students' achievement is the main purpose of SEE. A meaningful and respectful certification should be maintained by the authorized body for the certification of SEE. Revisiting for a valid standard setting procedure and certification would be useful.

Capacity development: NEB is currently undertaking capacity development of teachers in item development and orientation to the LG and Province personnel responsible for education at their level. Improved test development and result analysis for improving learning of students is also partly covered in these training/workshop. Capacity development in the area of assessment, test development and IBS should be done at various levels –experts team in order to provide expert as and when support needed; School, Province and LG level to enable them to use IBS and contribute in the development and maintenance of IBS; and teachers to provide service as item writer and analyse results to improve learning of students.

References

NEB regulations, 2019 (2076 BS) amended on 7 Feburary 2022 (www.lawcommissions.gov.np).Education Act https://www.bikashnews.com/wp-content/plugins/pdf-poster/pdfjs/web/viewer.php?file= https://www.bikashnews.com/wp-content/uploads/2023/08/educationact2080.pdf & download=true&print= false&openfile=false

Curriculum of Grade 10, 2021 (CDC)

8th amendment in Education Act 1971 in 2016

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. (2016). Letter Grading in School Education: Implementation Manual 2016. Sanothimi: CDC.

CERID. (1988). A Study on Secondary Education in Nepal. Kathmandu: CERID.

Education Review Office. (2016). *National Assessment of Student Achievement 2015 (Grade 3 and 5)*. Sanothimi: Author. www.ero.gov.np

Gardner, David P.; And Others. (1983). A Nation At Risk: The Imperative For Educational Reform. An Open Letter to the American People. A Report to the Nation and the Secretary of Education. National Commission on Excellence in Education (ED), Washington, DC.

Hill, P. (2013). Examination Systems: Asia-Pacific Secondary Education System Review Series No. 1. Bangkok: UNESCO Bangkok Office

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.

Kubiszyn,T and Gary D. B. (2016). Educational Testing and Measurement (11th edition). USA: Wiley.

Lane, S, Raymond, M. R., Haladyna, T. M. and Downing, S. M. (2016). Test Development Process (pp. 3-18). In Handbook of Test Development (Second Edition). Edited by Lane, S, Raymond, M. R., and Haladyna, T. M. New York: Routledge.

Mathema, K. B. and Bista, M. B. (2006). *Study on Student Performance in SLC.* Ministry of Education and Sports, Education Sector Advisory Team.

Martin Chautari. (2018). Measurement of Cognitive Abilities in the SLC Examinations in Nepal Research Brief No. 19, May 2018. Kathmandu: Martin Chautari.

Metsamuuronen, J. (2012). Handbook of Basics of Research Methods in Human Sciences. Researchers' Edition – BETA. Helsinki: Oy International Methelp Ltd. and Jari Metsamuuronen.

MoEST. (2022). School Sector Education Plan (SESP) 2022/23-2031/32. Kathmandu: Author.

MoEST. (2016). School Sector Development Plan (SSDP), Nepal, 2016–2023. Kathmandu: Kathmandu: Author.

Phelps. R. P. (2007). Standardized Testing – Peter Lang primer. New York: Peter Lang Publishing Inc.

Redfield, D. (2001). Critical Issues in Large-Scale Assessment: A Resource Guide. Washington DC: Council of Chief State School Officers and AEL, Inc.

Reynolds, C. R., Livingston. R. B. and Wilson, V. (2011). Measurement and Assessment in Education (Second Edition). New Delhi: PHI Learning Pvt. Ltd.

Singh, G. B. (1996). SLC Examination in Nepal. Kathmandu: CERID.

Singh, G. B. and Bakker, S. (2018). Standardizing Secondary Education Examination: Scoping for Standardized Test Development for SEE. Sanothimi: National Examination Board.

Steven Bakker, Setting Performance Standards, PPT used for training NAEC, Tbilisi, Georgia, 2010.

Thorndike, R. M. and Thorndike-Chirst, T. (2011). Measuremnt and Evaluation in Psychology and Education (Eighth Edition). New Delhi: PHI Learning Pvt. Ltd.

UNESCO. (2018). The Culture of Testing: Socio-cultural Impacts on Learning in Asia and the Pacific. Paris: IIEP UNESCO. http://unesdoc.unesco.org/images/0026/002619/261955E.pdf

Wilmut, J. (2001). Assessment and Examiantion Component Report. Kathmandu: Nepal Second Secondary Education Sector Project Preparation Technical Assistance.

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.

Annex

Annex A: Grade level and descriptor used in SEE

Grade	Descriptor	Explanation
A^+	Outstanding	Has Deep and broad knowledge with highly developed critical insight as well as comprehensive and perceptive appreciation of the theoretical or practical subject matter; an exceptional ability to organize, use, analyze and succinctly present subject matter fluently and clearly with extraordinary performance: a significant capacity for original, creative and logical thinking with superior communication skills.
A	Excellent	Has Deep and broad knowledge with developed critical insight as well as comprehensive and perceptive appreciation of the theoretical or practical subject matter; an exceptional ability to organize, use, analyze and succinctly present subject matter fluently and clearly with exemplary performance: a significant capacity for original, creative and logical thinking with advanced communication skills.
B^+	Very Good	Has broad knowledge with developed insight as well as comprehensive and independent appreciation of the theoretical or practical subject matter; a special ability to organize, use, analyze and succinctly present subject matter clearly with high level of performance: a highly-developed capacity for original, creative and logical thinking with sound communication skills.
B	Good	Has decent knowledge with developed insight as well as comprehensive and independent appreciation of the theoretical or practical subject matter; a special ability to organize, use, analyze and succinctly present subject matter clearly with respectable performance; a develped some capacity for original, creative and logical thinking with reasonable communication skills.
C^+	Satisfactory	Has adequate knowledge with developed insight as well as comprehensive and reasonably good appreciation the theoretical or practical subject matter; an ability to organize, use, analyze and succinctly present subject matter with blameless performance: some capacity for original, creative and logical thinking with sufficient communication skills.
C	Acceptable	Has sufficient knowledge with developed insight as well as comprehensive and acceptable understanding of the theoretical or practical subject matter; an ability to organize, use, analyze and present subject matter with guiltless performance: limited capacity for original, creative and logical thinking with acceptable communication skills.

D	Basic	Has some knowledge with developing insight as well as comprehensive and partial understanding to the theoretical or practical subject matter limited ability to organize, use, analyze and present subject matter with guiltless performance: limited capacity for original, creative and logical thinking with minimal participation in communication skills.
NG	Not Graded	Has less knowledge with incomplete insight as well as comprehensive and negligible understanding of the theoretical or practical subject matter; seriously deficient ability to organize, use, analyze and present subject matter, very limited capacity for original, creative and logical thinking with inadequate communication skills.

Annex B: A sample of grade sheet of SEE

SERIAL NUMBER	SUBJECTS	CREDIT HOUR[1]	OBTAINED GRADE[2]		FINAL GRADE	GRADE POINT	
			TH	PR			
01	COMP. ENGLISH	4	B+	A	B+	3.2	
02	COMP. NEPALI	4	B		B	2.8	
03	COMP. MATHMATICS	4	A+		A+	4.0	
04	COMP. SCIENCE	4	A	A+	A	3.6	
05	COMP. SOCIAL STUDIES	4	C+	A+	B+	3.2	
06	COMP. HEALTH, POP & EDU	4	C+	A+	B+	3.2	
07	OPT. I	4	B+		B+	3.2	
08	OPT. II	4	B	A+	B+	3.2	
GRADE POINT AVERAGE (GPA): 3.25							

Annex: Metadata for the test

Item metadata schema for Science and Technology subject (as per item card)

Metadata field	Description
1) Subject and Grade	S&T G10
2) Item cell code	From elaborated specification grid
3A) Unit	Unit no. 1-19
3B) Learning outcome (LO) no.	From the curriculum
3C) Cognitive skill	KUAHA)
3D) Format	MCQ, VSAQ, SAQ, LAQ
3E) Marks	1 only for MCQ or VSAQ, 2 for SAQ 4 for LAQ
4) Learning outcome	Statement from the curriculum
5) Objective of the item (in line with learning outcome)	Item writer develop
6) Item	Nepali first in Unicode and then in English
7) Key answer/marking scheme/rubric (in English only)	Possible correct answers with allocation of marks

Item metadata schema for Mathematics subject (as per item card)

Metadata field	Description
1) Subject and Grade	Math G10
2) Item cell code	From elaborated specification grid
3A) Unit	Unit no. 1-15
3B) Learning outcome (LO) no.	From the curriculum
3C) Cognitive skill	KUAHA)
3D) Format	MCQ, VSAQ, SAQ, LAQ
3E) Marks	1 only for MCQ or VSAQ, For SAQ and LAQ as assigned in the question
4) Learning outcome	Statement from the curriculum
5) Objective of the item (in line with learning outcome)	Item writer develop
6) Item	Nepali first in Unicode and then in English
7) Key answer/marking scheme/rubric (in English only)	Possible correct answers with allocation of marks

Item metadata schema for English (as per item card)

Field property	Remark
1) Grade	10
2) Subject	Eng
3A) Unit	Unit no.1-18
3B) Learning outcome (LO) no.	From the curriculum per elaborated curriculum matrix: Reading: R1-11 Writing: W1-14
3C) Area	Reading 1) Reading_ 2) Reading_ 3) Reading_ 4) Reading_ Writing 1) Writing _Guided 2) Writing _Free 3) Writing_Punctuation Grammar
3D) If R – Skills	R: literal comprehension, reorganization, inference and evaluation and reflection
3E) Format	1) True False 2) Fill in the gaps 3) Multiple choice question 4) Matching 5) Ordering 6) Short answer question 7) Constructed – essay (LAQ)
3F) Marks	Assinged mark for the item
3G) Text type for Reading	1) Story 2) Essay 3) Schedule/timetable 4) Menus 5) Charts 6) Calendar 7) Notice 8) Letter 9) News 10) Brochure 11) Biography 12) Dialogue 13) Poem 14) …….
3G) Text type for writing (Guided and Free)	1) News Story 2) Story 3) Description of tables

Field property	Remark
	4) Description of charts 5) Description of diagrams 6) Personal letter 7) Official letter 8) A short essay 9) .....
4) Learning outcome	Statement from the curriculum
5) Objective of the item (in line with learning outcome)	Item writer develop
6) Item	In English only
7) Key answer/marking scheme/rubric	Possible correct answers with allocation of marks

Note: Additional meta-data that will be added for each item after pre-testing