

# CHAPTER VI: ANALYSIS OF THE TECHNICAL QUALITY OF TEST MATERIALS USED IN SLC\*

## 1. INTRODUCTION

Educational researchers, despite several decades of constant inquiry, continue to debate as to what factor contributes to school achievement. There is hardly any consensus amongst the researchers. However, there is a common understanding that school achievement is the outcome of a number of school, community, teacher, family, and student related factors. While researchers in both developing and developed countries continue their search for factors that might explain school achievement and/or school effectiveness, testing and assessment experts assert that 'the tests' and the 'marking procedures' – often believed to be neutral and objective - may, in part, explain student and school performance. Continued poor performance of students in SLC and fluctuating SLC results are often believed to be the consequences of poor teaching alone. Many believe that results will improve if classroom teaching is improved. But, if the tests and marking procedure have any link with the outcomes of assessments and examinations, as claimed by the testing and assessment experts, efforts to improve classroom teaching alone will not yield expected results. Therefore, it is imperative to understand how test materials and marking procedure might affect outcomes in examinations. Although a few studies (World Bank 1994) have examined the quality of test materials, available information is only sketchy. How the tests and marking procedures employed in the SLC examinations affect students' outcomes remains largely unknown. In this context, the study 'Analysis of the Technical Quality of Test Materials used in SLC' makes an attempt to examine the quality of test papers and marking procedures and how they might student performance in the SLC examinations.

For details on the objectives and methodology, please refer to 'Analysis of the Technical Quality of Test Materials used in SLC'

## 2. ANALYSIS OF ENGLISH TEST MATERIALS

English is one of six compulsory subjects prescribed for the SLC curriculum. The objective of teaching is to enable students to speak, read, and write in English so that this functional knowledge of an additional (international) language helps them to make their knowledge wider and deeper.

### 2.1 Curriculum and Textbook

The general objectives of the curriculum focus on the understanding of and competence in spoken English followed by reading and writing skills which can lead the learners towards use of a wider variety of English for obtaining knowledge, information, and pleasure. Language learning and teaching activities have been broken down into four language skills. Listening and speaking are emphasized so much so that 20 percent of the total weightage is allotted towards

---

\* This chapter is based on the report 'Analysis of the Technical Quality of Test Materials used in SLC' prepared by Prof. Dibya Man Karmacharya for the SLC Study team.

practical examinations. Two other skills, that is, reading and writing, constitute 80 percent of the theoretical examination in English.

The secondary English curriculum uses two separate textbooks – one for Grade IX, another one for Grade X. Teaching units are divided into functions – 20 for Grade IX and 16 for Grade X. All language skills are incorporated within every unit. The curriculum clearly elaborates the steps of teaching and the methods of assessment too. The curriculum, however, stresses more on the mechanical, reproductive, and market-oriented contents and techniques. Thus the ultimate goals of education, that is cultivation of human values and broad spectrum of human civilization, are lacking in the English curriculum.

The textbooks exactly meet the numbers of functions to be completed. However, there is no separate unit as such for knowledge and comprehension; these levels are not incorporated within the lesson. Most of the reading materials are authored by the textbook writers themselves and only a few of them are adapted. However, all these texts have one serious drawback, that is, more than ten such reading passages are written on first person language and mostly they are about personal or family themes. The poems are original unadapted and their selection is satisfactory. The themes are limited; style is invariant; and sound monotonous.

More than 80 percent of reading texts (Grade IX) have very limited, confined themes. They introduce the world through the eyes of the first person "I". This reduces the variety into stereotype monotonous themes. This "I" element naturally presents a picture of controlled theme, vocabulary, and structure. The themes are duplicated and the world narrowed. Although there are many more 'comprehension passages and texts', a serious lack of variety is a striking feature. More authentic and original passages together with abridged versions and adapted texts must be included in both these textbooks. The dialogues, conversations, and some passages are not very interesting either. Similarly, some lessons do not seem to be appropriate to the age group.

Unlike in Grade IX, the textbook writers themselves created some reading materials in Grade X. Here "I" element is less frequent but then the defects of Textbook IX are duplicated here too. The number of the reading passages in X is less, yet the variety is greater here. Also the passages inform the students about some new or recent concepts, ideas, facts, and issues. The passages appear to be 'artificial'. There is a large body of children's literature, children's classics, and abridged works available in the market today. Such texts should be included there to ensure variety and avoid monotony. They should be introduced to the world of 'fiction' through such readings because the children of this age enjoy entering into the world of imagination and their imagination could be captured by introducing them directly into a more authentic reading materials.

## 2.2 Specification Grid

Principally, the grid should reflect curriculum and test materials should reflect the grid. The final goal of curriculum is to enable candidates to appear in the national examination. The objective of English curriculum is to test candidate knowledge and application of the skills – at both **competence** and **performance** levels. Although Benjamin Bloom's taxonomy of educational objectives has six parts – these are reduced into two in case of English. The grid for English uses only two levels. These are **competence** and **performance**. The grid specifies that competence and performance levels, which mean both knowledge and understanding skills, are integrated into practical abilities and no separate examination is given in this area. On the other hand,

performance includes practical abilities, which classified into language skills such as Listening, Speaking, Reading, and Writing. The grid has perfectly reflected the spirit of the curriculum. The grid has exactly followed marks weightage allotted in the curriculum for each of these four skills, Listening, Speaking, Reading, and Writing, which are 8, 12, 45, and 35 percent.

### **Listening and Speaking**

The present classification grid is alright as it shows a strong linkage with most of the curriculum components. However, the weightage allotted to each, that is, Listening, Speaking, Reading, and Writing needs to be re-adjusted. The main objective of learning English is to be able to use it in spoken form. However, to assess their achievement, only 20 marks are allotted. Furthermore, question types listed for testing listening comprehension are appropriate, but the exercises developed for speaking ability are not sufficient. The question type 'cued situation' needs to be specified in the sense that teachers (examiners) should be instructed to use at least 3 or 4 cued situations to assess students' achievements. To conduct this test type more than one student can be included for role-playing. With these modifications the weightage given to listening and speaking skills could be increased to 30 percent out of which 10 percent should be allotted for the practical assessment to be carried out by the class teacher on a day-to-day basis and 20 percent for the final examination.

### **Reading and Writing**

In case of reading, the types and numbers of vocabulary items are not mentioned anywhere. The reading passages mostly have the contents of common everyday life. It would be better if they were also classified thematically. Besides, all the reading materials, that is, passages are authored, and a few adapted by Nepali textbook writers of English. These books lack the specimen of original writings on native English. These textbooks should include original passages of a wide range of variety as well.

In the case of writing, the curriculum includes various types of writing, but in the classification grid all these types are not listed, although the grid has divided the questions into 3 main types – controlled, guided, and free. The writing activities listed in curriculum are not mentioned under any one of the headings. So these need to be included there too.

### **Vocabulary and Grammar**

To test reading comprehension, vocabulary items are considered as a very effective tool. There are different ways of testing understanding and production of vocabulary, not clarified in the grid. However, in sample questions vocabulary items are included which require students to identify synonyms, antonyms, and definitions only. Instead, a wide range of exercises could be developed for testing vocabulary.

The grid clearly mentions that grammar is not tested separately, but it has to be integrated in the performance particularly in writing skill. Eleven different grammatical aspects are listed for testing; out of them 4 items (articles, question tags, S+V agreements, transformations) are not overtly included in the curriculum and another item, that is voice, is also only partially included.

## **2.3 Test Papers**

In two subjects, Mathematics and English there is a provision for using a parallel set of test papers for each of the five Development Regions. Therefore, English subject has ten sets of

question papers. The distribution of materials and time is exactly reflected in the test papers. While testing the four skills of listening, speaking, reading, and writing, two aspects, that is, grammar and vocabulary, are also tested. The general pattern of the SLC English question is in the specification grid. The distribution of content materials and time in the specification grid is exactly reflected in the test papers.

The grid suggests both seen and unseen materials are suggested for testing reading and writing skills. These components cover, unlike other subjects, only two and a half hours of time and the remaining half an hour is allocated for practical examination that includes listening and speaking components carrying 20 marks arranged at the end of the examination schedule. However, there are some obvious gaps found in the test papers themselves. These are factors like repetition, conceptual and typological errors, and instructional ambiguities that could be corrected with caution and care.

### **Lapses, Errors, and Ambiguities in Test Papers**

Test papers are found to have many errors and lapses, mostly linguistic and inferential, apart from many typological and mechanical ones. Each set has 2 to 4 lapses or correctable points. These errors might have affected the student performance. Analysis reveals that test papers are far from perfect and standard in terms of their internal qualities. There are cases of repetition, vagueness, lack of clarity, confusing instructions, ambiguities, and defective inferences. Such questions have direct impact upon the students and the scorer as well.

### **Repetition Cases in Test Papers**

In total, 40 sets (10 sets per year, that is 2057, 2058, 2059, and 2060) of test papers were checked. There were only eight questions asked in English and many cases of repetition were noticed longitudinally as well as within the same year. In 4 years 40 sets of questions were asked, out of which about one-third of the sets have some proof of repetition. Maximum marks for repeated question is 3.5 and minimum 0.5. This kind of repetition can be ignored as it may justify the need for parallel sets. Regarding the controlled, guided, and free composition questions, three points are worth mentioning.

For some items, clues are supplied, whereas for others no clues are given. Uniformity should be maintained regarding this.

For some items, the number of words is fixed, for others, they are open. The number of words should be limited or fixed for such questions.

Some questions favor more/only a particular group of children who are familiar with these. These have strong bias and pose great problems to especially rural children who are never exposed to such environment.

Topics like Supermarket (2060 Q.5, set Ai), Plane Crash at Thankot (2057, Q.5, set D1), Educational Tour (2058 Q.5 set E1), Twin Otter Crash in Dadeldhura (2057 Q.7 A2), Train Accident (2060 Q. 7 E2), Bus Mishap (2059 Q.7 B2); Applications for the Post of Front Manager to Deluxe Hotel (2057 Q.5 E2), Office Secretary (2057 Q.5 B2), Importance of Television (2057 Q.5, A2), Application for the Post of Deputy Supervisor (2059 Q.5 A2), Assistant Lecturer (2059 Q.5 E2), Agriculture in Nepal (2058 Q.5 A1) are examples of the topics used in the questions that favor only a particular group of children familiar with these situations. Such biases should be rectified, removed and children from deprived classes/places should be kept in mind.

## 2.4 Marking Schemes

To a large extent, the Marking Scheme developed by the OCE avoids confusion and helps maintain uniformity. However, a clear examination of the Marking Scheme sheets supplied for the year 2059 (2003) (for 10 sets of test papers) reveals that there are some obvious lapses, errors, and confusions, even incorrect answers given in the schemes. The weakness indicated in the scheme proves that there is a gap between test papers and marking schemes. There are errors e.g., the marking scheme mentions only one answer as correct whereas there is/ are also other correct answer(s); wrong structure; placing correct answer of one question under another question and not clearly stated. In this sense, the Marking Schemes are far from perfect. This must have influenced the examiners and students must have suffered every year.

## 2.5 Response Patterns of the Examinees

The answer papers were analyzed using eight parameters suggested for analyzing student performance. In most of the cases, judgments were made with the help of facts or evidence available. However, in some cases, such as expression difficulties, originality in writing subjective judgment is inevitable and thus results of analysis may vary from person to person.

### Expression difficulties

Students from all regions have shown expression difficulty – mainly in questions that require the students to write a composition – a letter, application, dialogue, etc. Students seem to have difficulties in questions that cannot be solved simply by filling in the blanks, matching items together, guessing the answers or by rote learning. Such questions require concrete words or sentences and imagination plus accuracy. These skills cannot be developed without a long and rigorous practice. Besides, students' actual competency cannot be tested without eliciting from them answers to such questions. Therefore, those students that lack sufficient practice in developing their vocabulary repertoire and structures find questions which require their own creative or original expressions, naturally difficult. Most of the candidates except from those from the valley of Kathmandu have left these questions unanswered. They usually choose and attempt objective types of questions and those that demand the close ended answers first because these involve a lesser degree of risk of being incorrect in terms of spelling, punctuation, grammar, etc. and even guessing can help them sometimes.

### Originality in writing

Almost all students have the problem of clear, grammatical expression. Lack of adequate vocabulary and correct sentence patterns have posed problems. This is linked with the questions related to comprehension, too. Many of them have shown copying. Overall, originality is lacking in the students' responses.

### Repeated questions

There are only a few cases of questions being repeated. Out of the total the students, only 7 have repeated questions.

### Questions not attempted

This is closely connected with expression difficulty and originality in writing. Only really difficult questions are not attempted.

## Copying

Many of the answers seem copied versions, or reproduced after rote learning. The Mid-Western Region has the highest number of copying and cheating. Thirty-nine out of 40 answerbooks are marked for this. The picture for the remaining regions is like this. Out of the total, 32 percent students from the Eastern Region, 28 percent from Central, 25 percent from Western, and 2 percent each from Mid-Western and Far Western Regions show proof of copying. This observation is based on the experts' judgment.

## Correctness

Copying and correctness are closely related problems. Correctness for the Markers refers to the use of punctuation, correct spelling, proper word order, coherence and appropriate grammar in terms of S-V agreement, tense, including legible handwriting, etc. by the students. Only 7 (9%) candidates wrote correctly in the Eastern Development Region; about half or 50 percent (44) students from Central Development Region districts, Kathamndu, Chitawan, and Bhaktapur have written correctly; and a majority of the students from other regions did not write correctly and accurately to the expected level. In the Far Western Development region, all students lacked correctness in English writing.

## Questions not mentioned or wrongly mentioned

Out of the total, only 18 candidates have failed to mention or have wrongly mentioned questions. This is not a negligible figure as this contributes to about 6 percent. This may cause confusion in the examiners, irritate them and sometimes the candidates may loose marks too, and very often posting error result due to this. This may happen also because of nervousness, hesitation forgetfulness or other unavoidable psychological factors.

## 2.6 Consistency in Markings

The total number of the answer scripts re-examined is 288. Variations are observed in markings. For instance, the percentage of students obtaining pass marks of 26 (out of 80) in the case of the first Marker is 57.6%, but, interestingly, this percentage decreases to 46.8% in the case of Marker 2 and 47.6 in the case of Marker 3. Although the Marker 2 and 3 do not differ much, they do differ from Marker 1. In total, the three Markers vary by 10 percent. This variation is further justified when the range of difference in marking by three markers is calculated. This has been presented in Table 1.

**Table 1. Range of Difference among the Markers in English**

Range of marks	Markers 1, 2, and 3		Markers 2 and 3	
1-5 marks	188	(65%)	256	(89%)
6-10 marks	79	(27%)	30	(10%)
11-15 marks	22	(8%)	2	(1%)
16-20 marks	4	(1%)	No cases	
21	1		No cases	
Total		(100%)		(100%)

Analysis of the marking of 288 answer papers by 3 different markers shows that there is a considerable range of difference among them. Almost two-thirds (65%) of the candidates fall within 1-5 range of difference for Markers 1, 2, and 3. Unlike this, for Markers 2 and 3, the gap further narrows as 89 percent fall within the 1-5 range which proves a high degree of consistency and is also acceptable, if we consider marking a purely

subjective business. Subjectivity may only be a point justifiable for creating such differences. Lack of rigorous training on marking, faulty Marking Scheme, non-use of Marking Scheme, or even negligence may result into cases like ranges of 21 and 35 above.

## 2.7 Scoring Theoretical Versus Practical Examinations

The practical portion in English carries 20 full marks out of 100, the remaining 80 is allocated for written (Reading and Writing) part. The practical portion in English stands for oral test of listening and speaking skills. There is no clear-cut definition of how candidates with speech and hearing impairment can appear such a test. Practical examination is also externally conducted. However, there is no correlation between 'pass' 'fail' in theory and practical examination. For instance, 44 percent of the total candidates (out of 288) have failed the written examination for M1, whereas only 2 percent have failed in the oral or practical. Not only this, practical scoring is quite irrational. One earned 66 percent in written, against 45 percent in oral; another candidate earned 69 and 100 percents in written and oral examinations, respectively; whereas a fourth one got 82 percent in written and 65 percent in oral. This is quite illogical and unjustifiable. An illustration of incomparability between theory and practical modes of examination could be drawn from the available data presented in Table 2.

**Table 2. Average Scores Awarded in Theory and Practical Examinations in English**

Region	Number	Theory %	Practical %
Eastern	135	31	58
Central	121	39	71
Western	48	43	74
Mid Western	31	28	58
Far Western	18	29	72
Total	358	36	65

The data show that there is a great disparity between these two modes of examination and markings. The average marks awarded in practical examination are almost double compared to that of theory (65 vs 36). Among the five regions, the Western region shows the highest performance

both in theory and practical examinations, i.e., 43 and 74 percent. Likewise, the Far Western Region has the lowest rank in both theory and practical examinations. This might be due to that the oral test depends to a large extent on how the examiner perceives, since the answers are orally produced and also on how the students perceive.

## 2.8 Recommendations

First the reading materials in the textbook should include a great variety in terms of theme and content. Originally authored texts carrying contemporary language with literary touch should replace the present ones that are stereotypic, monotonous, and have too common themes.

Reading materials should be compiled according to the interest of thematic variety and more advanced texts from authentic, contemporary writing should be included. Also a few of them should focus on the special needs/underprivileged situations / special needs of Nepal too.

Specifications grid should be made elaborate and explicit and the OCE should ensure that test paper setters strictly follow the specification grid.

The use of multiple sets and parallel sets should be dropped because this practice has no meaning. This has no specific, intended purpose, (e.g. remote areas, disadvantaged children, Dalits, etc). Instead, it leads to disparity and variation (since one set contains dialogue, another

contains essay or letter; one set has simple passage and another contains more difficult or longer one. These depend upon subjective judgment.

Expressive skill (semi-guided and free writing) of the students requires more practice. As this depends upon receptive capacity, that is, Reading, students need to be exposed to a wider variety of reading texts and practices. Extra reading materials can enhance this and for writing, special writing practice is required. This suggestion is made based on the answers to questions 5 and 6 which demand creative and original writing from the students. Open-ended answers and free writing exercises require much more practice. So writing practices should be a major focus of teaching English. Students really face expression difficulty, write answers incorrectly, or leave questions unattempted whenever they have to express originality or creatively. Because of poor vocabulary repertoire, they cannot understand the unseen passages and often write wrong answers. So reading, comprehension, and vocabulary practices should be focused. Extra reading materials can help improve this to a large extent.

Students are weak in vocabulary repertoire and grammatical structure. So extra reading, writing, and grammar practices are essential for improving students' expressive skills. They need to be exposed to a variety of texts and frequent writing followed by correction. This can help them develop originality. Also rote learning, and memorization technique should be discouraged and their self-confidence should be built up. More practice is required in the intensive reading of different texts and passages followed by writing practices so that comprehending unseen passages becomes easier because comprehension power requires development in vocabulary repertoire and structural accuracy too.

Effective measures should be adopted to discourage copying. Copying is a serious problem that should be completely discouraged punishing those that practice this.

Greatest attention should be given to ensure the quality of the Marking Schemes. They should be made elaborate, explicit, and unambiguous and it should be ensured that the examiners strictly follow the same.

Markers need special training on the use of marking scheme, marking techniques, and evaluation process. Every year training or workshops should be organized to recruit trained markers and ensure a greater degree of consistency in marking.

Special measures should be applied in case the SLC wants to continue the practical portion; otherwise, if the present practice persists, there is no point in giving practical, however useful might it be.

### **3. ANALYSIS OF NEPALI TEST MATERIALS**

Nepali is one of the six core subjects prescribed in the SLC curriculum. This section provides the results of assessment of test materials used in the SLC examinations for compulsory Nepali subject.

#### **3.1 Curriculum and Textbooks**

The curriculum consists of six major components, namely: introduction, general objectives, objectives related to language skills, scope of literary forms, teaching techniques, and evaluation. The introduction section is concerned with the purpose and content of the curriculum. Objectives are divided into two categories: general objectives and objectives related to language

skills. Again language skill section of the 'Objectives' is divided into aural-oral, reading, and writing sub- sections. The language aspects are included within the relevant skills concerned; for example, pronunciation within listening and speaking, vocabulary within reading, grammar, and spelling within writing. Thematic contents have been selected from the literary forms like essays, stories, biographies, letters, plays, diaries, and poems and linguistic contents from grammar and vocabulary. These contents have been presented under the section of 'Widhaako krama ra kshetra'. In another section, the curriculum has briefly indicated the 'Teaching technique' for each possible teaching point within the contents mentioned above. Similarly, the 'Evaluation' section of the curriculum includes general guidelines for formative as well as summative evaluation. It has mainly emphasized the nature of questions, area-wise mark distribution, and evaluating lesson-based contents to be presented in literary forms in the textbook. Besides, it also includes free composition, grammar, and vocabulary. However, the provision for oral skill assessment is limited to classroom activities only.

There are two volumes of Nepali textbooks prepared separately for Grades 9 and 10 each. The content is presented in various forms of literature. The scope and sources of such forms are either adapted from the original writings or are prepared by the writers / editors themselves. The literary forms included in the lessons are stories, essays, poetry, plays, biographies, etc. as guided by the curriculum. It has been expected that these forms will be sufficient in achieving curriculum objectives related to various language skills. Each lesson is followed by varieties of language exercises suitable to the literary forms concerned (e.g., reading comprehension from prose lessons, free and guided writing from essays and plays, inferring meaning from poetry, etc.). However, such exercises are centered mainly around reading and writing skills giving the least priority to aural-oral skills. Our contents analysis of curriculum materials shows that there is a proper alignment between curriculum objectives and textbook contents.

### 3.2 Specification Grid

The specification grid has adopted the modified form of usual Bloom's taxonomy. It includes four levels of cognitive domain different from six levels in Bloom's original classification, i.e., knowledge, comprehension, application, analysis, synthesis, and evaluation. This modified classification includes knowledge, comprehension, behavioral skills, and higher levels (application, analysis, synthesis, and evaluation). Each level has been explained in the grid. Thus questions related to memorization of information come under the knowledge level and questions related to explanation and description of the learnt facts come under the comprehension level. Grid further clarifies that questions related to behavior skills will be of rule 'use' and 'adapt' type whereas questions related to higher level performance will include analysis, synthesis, and evaluation. Although this classification represents all the levels specified in Bloom's taxonomy to some extent, the grid does not explain anything about this new classification for being a comfortable option to old classification.

The above levels have been distributed among the questions to be included in the test paper in an ascending order in the grid. According to this distribution, 3 questions fall under knowledge level, 4 questions under comprehension, 8 questions under behavior skills, and 16 questions under the higher-level category (sub-questions are also included in counts). This distribution looks quite logical from the point of view of expected level of language performance on the part of students, but the grid is not found well aware of the level and variety of questions to be included under the comprehension passage in the test paper. Observation of some previous

question sets (e.g., 2058 C, 2058 S, 2059 B, 2059 C, 2060 E etc.) suggests that most of such questions fall under recall types that do not require reasoning as their answers.

### Content coverage in the grid

Generally, a language curriculum consists of language skills (listening, speaking, reading, and writing) and aspects (pronunciation, comprehension, grammar, vocabulary, etc.). These skills and aspects are presented through appropriate thematic contents. Thus, the language skills and aspects are the things to be mastered in language curriculum. In the present discussion as well, the contents have been considered as skills and aspects of Nepali language. The grid is also prepared in the same direction. However, it has totally excluded the part of oral skills. This exclusion might be keeping in view the practicality of administering oral tests, for which it is silent as the curriculum itself. The grid has strictly followed the Evaluation Technique component of the curriculum. As a result the grid is limited to specifying measures of evaluating reading-writing skills and vocabulary-grammar aspects only. The evaluation scheme outlined in the curriculum and the specification grid generally cover all major objectives related to reading, writing, grammar, and vocabulary. But objectives related to listening and speaking have been limited to classroom activities only.

## 3.3 Test Papers

### Clarity in Questions

Some questions are quite long and even seek two or more different tasks to be done on the part of students. For example; questions like (a) 'शिशिरवसन्तको कथा' का प्रतापी राजाकी पहिली रानीको मनमा कस्तो चिन्ता कसरी जाग्यो ? (2060 E 2घ) and (b) 'शहिद' कथाको मुख्य पात्र वीरबहादुर किन र कसरी सहिद भयो ? (2060 B 2ङ) demand two different tasks in their answers. Similarly, problems are ambiguous to the students due for lengthy sentences and their complex formation. Some sentences are even internally deficient in construction. Questions such as (c) जिल्लास्तरीय वक्तृत्व कला प्रतियोगितामा प्रथम भएकी आफ्नी बहिनीको ज्ञान र अभिव्यक्ति कलाको प्रशंसा गर्दै उनको हौसला वृद्धि हुने गरी दाजुले लेख्ने चिठी लेख्नुहोस् (2060 E 6क) and (d) सहरमा पढ्न बसेको भाइलाई विदेशी संस्कृतिको अन्धानुकरण नगर्न र आफ्नो संस्कृतिको संरक्षण गर्ने सल्लाह दिँदै चिठी लेख्नुहोस् (2060 A 6क) are quite long, complex, and loaded with several messages and so they seem quite ambiguous for the average student. In addition to it, in the question (d) above, the two phrases 'विदेशी संस्कृतिको अन्धानुकरण नगर्न' and 'आफ्नो संस्कृतिको संरक्षण गर्ने' are not consistent and belong to two different word class groups (adverbial phrase and adjectival phrase respectively), making the formation of whole sentence internally incorrect from the Nepali common usage point of view. In this case, use of either adverbial or adjectival phrase alone should have been chosen. Questions such as the one in (c) above could have been made more simple and straight if it was framed like 'जिल्लास्तरीय वक्तृत्व कला प्रतियोगितामा प्रथम भएकी आफ्नी बहिनीलाई हौसला मिल्ने गरी दाजुका तर्फबाट चिठी लेख्नुहोस्'।

Many questions do not have verbs that specify the answer to be measured objectively. Moreover, there is no difference between the statement and intention of the questions related to critical answer and short answer. For example, between the two question (e) 'नेल्सन मन्डेलाका स्वभावगत विशेषताहरू के के हुन् ?' (2060 E 2ग) and (f) 'भीमसेन थापाको व्यक्तित्वका प्रमुख विशेषता के के हुन् ?' (2060 D 1क), the former belongs to 'short answer group' carrying 5 marks

whereas the latter belongs to 'critical answer group' carrying 10 marks. From the taxonomy point of view also, the latter type, as the grid suggests, must belong to the higher level category while the former need not. But both of these questions lack a clear-cut indication of how much detail each question should go into in the course of being answered. Besides, they are similar in the performance level of tasks. Moreover, the use of long qualifying phrases has created ambiguity in meaning and such long sentences can be more difficult for L2 learners in identifying the correct response the question demands. Thus the possible reasons for ambiguities in the test papers are mainly language-related, construction-related, and task-related. Language ambiguities include complex sentences; as in (a)-(d) above use of typical Nepali or less frequent tatsam words, as in (e) and (f) above like बखान, छेउ, प्रतिहिंसा, विसङ्गति, राष्ट्रिय अखण्डता, बहुआयामिक; construction ambiguities include similarity in nature and type of questions for various levels, as in (e) and (f) above; whereas the task ambiguities include the lack of limit and specificity of tasks, as in (e) and (f) above as well as in (g) 'भविष्यनिर्माण कथामा कस्तो मनोविज्ञान प्रस्तुत गरिएको छ ? (2058 E 1ख Critical answer).'

Some teachers involved as examiners and head-examiners in the SLC examination have experience that many students are confused in identifying what the question actually intends. For example: विपक्षीका तर्कहरूको खण्डन गर्दै 'समाजमा बाबुको भन्दा गुरुको महत्व ठूलो छ' भन्ने विषयका पक्षमा आफ्ना तर्कहरू प्रस्तुत गर्नुहोस् । (2060 E 7ख). Due to the complex formation of the sentence, the words विपक्षी (wipakshi) and पक्ष (paksha) together might have created confusion in such questions. As a result, the students may answer both in favor of against the topic concerned. The above question would have been clearer if the initial part of the question 'विपक्षीका तर्कहरूको खण्डन गर्दै' had been replaced by the phrase 'वादविवादका लागि...'

### Repetition of Questions

There are quite a few examples of questions with exact word repetition. Examples of such repetitions can be noticed in some short answer questions, as in 2058 A 2ग(पार्वतीले नरेशको भविष्य निर्माण कसरी गरिन् ?), 2059 A 2ड (पार्वतीले आफ्नो छोराको भविष्य निर्माण कसरी गरिन्?), 2059 D 2ग(भविष्य निर्माण कथाकी पार्वतीले नरेशमा उब्जिएको प्रतिहिंसाको भावनालाई कसरी हटाइन्?), 2059 S 2ड (पार्वतीले आफ्नो छोरो नरेशभित्रको प्रतिहिंसाको भावनालाई कसरी हटाइन्?), 2060 A 2छ (पार्वतीले नरेशको मनमा भरिएको प्रतिहिंसाको भावनालाई कसरी हटाइन्?) and in explanation extracts (see 2059 E; 5A(kha) and 2060 D: 5A (kha). In the same way, the repetition of essay topics can also be noticed in different years. Such topics are not exact in word but in content, as in 2058 A, 2058 S, 2059 D, 2060 B, 2060 D and 2060 E sets. Moreover, it has been observed that the repetition of some lessons and their contents was too frequent and these were not excluded in any examinations during the years 2058-2060 BS. Some of the questions were repeated with slight changes of words intending the supply of the same content as their answers. For example, the poem 'Ichchha' was repeated in the year 2058 and 2059 BS. in a critical question and in the year 2060B.S. in short answer questions. Similarly, the stories 'Doshi Chasma' and 'Bhavishiya Nirman' and the play 'Gharko Maya' etc. were included in the question papers of each year in one way or the other. Thus there are lessons which are more 'important' for students from the examination preparation point of view than others in the textbooks. For example, lessons like 'Rashtriya Jhanda', 'Rashtriya Wibhooti Bhimsen Thapa', 'Jaya Bhoonli', 'Sahidharuko Samjhana' etc. were much more frequent whereas lessons like 'Wasant', 'Shishirvasantako kathaa', 'Nyayako pakshya' rarely appeared in the examinations. However,

repetition of the questions or of contents in the parallel sets of the same year should not be taken as seriously as repetition in the following and subsequent years.

### **Relevance of Test Papers in Terms of Taxonomy as Specified in the Present Grid**

Analysis shows that test papers do not deviate sharply from maintaining the taxonomy aspect as instructed in the grid. However, they do not represent exactly the intention of the grid. The question setters are yet to realize differences among levels. As a result, the characteristics of knowledge level questions can be noticed in comprehension questions and the characteristics of comprehension questions in higher level questions. In most of the comprehension passages, the knowledge level questions are in greater number than other levels for which students can just pick up or copy the relevant part of the passage as their answers.

### **Relationship of Test Papers with the Curriculum and Textbooks**

One can observe that the test papers have normally followed the curriculum evaluation guidelines, which are mainly emphatic on textbook lessons. Our analysis shows a close relationship of the test papers with the evaluation section of the Nepali curriculum. Besides, the test papers have strictly followed the specification grid. Thus it can be easily claimed that the test papers in use measure all the curriculum objectives covered by the specification grid.

Although the test papers do not show major deviation from the curriculum guidelines, the nature of questions commonly used for critical answers do not fully reflect the intention of the curriculum, as they usually demand a description of the facts they have memorized from the lessons concerned instead of higher level language performance. The curriculum has clearly guided that such questions should be posed in a way for testing students' analytical and creative level ability. Since textbooks are the main source of teaching learning in our education system, examination cannot be separated from it. As a result, the test papers have a tendency to be too closely associated with the textbook concerned instead of the curriculum objectives. Joshi (2003:101) seems right in claiming that the "test items concentrate mainly on factual recall of textbook information and sometimes encourage repetition of answers based on textbook exercises or commercial guess papers". If we analyze the test papers, we find that some 50% of the marks are allocated for textual knowledge, 25% for free language use whereas the remaining 25% marks are set aside for grammatical items. Thus, the evaluation procedure we practice in SLC level is based on the textbook lessons, not on the curriculum competencies. In such a situation, even a high score in the examination alone may not be much helpful in predicting students' better performance in actual settings.

## **3.4 Physical Quality of Test Papers**

The physical quality of the test papers can be considered in terms of instructions, printing, paper, color, size, length of questions, and the time given for answering. An analysis of the sampled test papers used in the years 2058-60 under the given criteria reveals the following facts:

### **Instructions**

Instructions for questions are given either in group for a number of questions or on an individual basis for a particular question or for a particular type of tasks. Most of these instructions look clear enough for the students concerned. However, in some cases, they seem awkward because of the repetition of the same words in the question and in instruction; for example, as in 'or part' of question no.6.

## Printing

By and large, the test papers possess clarity in printing. They are free from ink smudges, unbalanced inking, unprinted letters, and unusual shift of words and *matras* to the next line. Similarly, type sizes used for various purposes are appropriate for the age level concerned. But typefaces used for instructions, questions, and texts are not justifiable. All instructions should have been printed in bold face and the number of questions to be attempted put in italics or vice versa to attract students, paying due attention towards what they have to do within that question. Besides this, the test papers suffer also from frequent printing mistakes and a few ones, even the distort meaning of the question concerned, e.g., 'कान्छीको (ले?) भने पाँच पाथी धान दिनु' (2060 A2क). Spelling mistakes are too frequent, which may easily misguide students to repeat the same in their answer. Instances of such lapses are prevalent in each question paper here and there. Thus a serious and thorough proofreading is a felt need in the test papers to be used.

## Paper

The printing paper used in the test paper is more or less opaque, yet not exactly appropriate for both side prints. White, glazy, and sufficient thick paper would have been a better preference for the purpose than the present ordinary paper. The color of the paper is also varied from one test paper to another within the same year and year-to-year. At present, white and off-white (brownish) paper has been in use. The recommended thickness of such paper should not be less than 60-gram weight.

## Size

The test papers have maintained a standard size each year. Its consistent size is approximately 6.5 x 9.5 in inches, which is quite handy and portable to look and convenient for use. The size is fixed in a way that the paper covers complete four pages without leaving any space unused.

## Length of Questions

There are no reports as such which claim that the length of questions has a serious effect on examination performance in Nepali subject. Informal queries with the teachers and students also reveal that the length of questions for three hours duration is not that much beyond finishing attempt within the time given.

## 3.5 Marking Schemes

With the purpose of maintaining uniformity in marking the SLC answer-books as far as possible, the system of using marking schemes has been introduced since the year 2000 A.D. These schemes (2058-2060 BS) are guidelines for answerbook examiners and head examiners to help them find the possible right answers of the questions given. However, they do not pose any restrictions to award marks for alternate correct answers. But each examiner/head examiner has to follow them strictly.

The schemes are prepared in a fixed format set by the specification grid 2000. The initial part of schemes is related to general instructions to examiners/head examiners suggesting them how to evaluate correct, partially correct and alternate answers, and when to use their own discretion. But these schemes do not instruct them to consider the problems and difficulties of the students from minorities, and special needs and hardcore groups.

The schemes have fixed evaluation criteria for each subjective question assigning certain marks for each criterion. For the questions related to textbooks (questions 1, 2, and 5), the major points to be included in the answers are also indicated, whereas in the cases of comprehension and summary writing passages (questions 3 and 4), only the scoring technique is suggested. With regard to free composition questions, the aspects to be included in each answer along with the marks distribution have been suggested. But for grammatical questions (questions 8-14 and 16) correct answers have been given with the full marks for each individual item. For such questions, which might have various possible correct answers (i.e., questions 15 and 17), only scoring guidelines are given, leaving the examiners free to use their discretion in allocating marks for every correct answer. Thus effort has been made to give in the marking scheme full coverage of all questions in the test papers in contents and counts. But they do not maintain consistency, especially with regard to the guidelines related to questions 3, 4, 6, and 7. Indicating the main points to be included in answers to these questions would have helped to make marking more objective.

Marking scheme is used not only to maintain uniformity, but also to guide the inexperienced and novice examiners. Usually, it also serves as a reference tool, especially for the examiners from remote areas where library facilities are either rare or nonexistent. This shows that clarity is the first and foremost criterion that a scheme must have. Regarding the marking system so far in use, it has been observed that it needs some improvements to maintain clarity. Instances of printing mistakes can be noticed in most of the sets. Spelling errors like these are common in all of the sets (तीब्र, जमानसिंह, बाविवाद, 2059 A, संरक्षण, राष्ट्रवादी, जन्मभूमिस्च, नेपालीहरु, देश भित्रै, अरू, रुप, श्रब्य, स्वार्थ भन्दा, सर्वोपरी, 2058 A, देहाबसान, तत्कालिन, गजूर, सुरू, शब्दहरु मध्ये, 2058 E). Some of these mistakes even distort the intended meaning, e.g., गरे र (गरेर?) 2059 A, स्तरीय (ता?) 2059 C, शीलता (शीतलता?) 2058 A, भोलाबाट (भोलाबाट) 2060 B etc. Besides, the schemes are not clear as to what to do in the case of partially correct grammatical answers.

### 3.6 Response Patterns of the Examinees

Analysis of the response patterns of students reveals the following:

#### Expression Difficulty in Nepali

It has been observed that many students cannot express themselves with ease in normal Nepali language. Study of 280 answerbooks from different districts has shown that students from Dhankuta, Saptari, Baitadi, Jumla, Ramechhap, Khotang, etc. have noticeable difficulty of expression in Nepali. The reason may be either the absence of sufficient Nepali language environment to be exposed to or inadequate language to minimize the mother tongue interference at previous Grades. Though there is no strong evidence to claim, some examples (15% of the district samples) indicate that even dialectic interference can create expression difficulty in Nepali in the case of Nepali dialects like Jumli and Baitadian.

#### Number of Questions Left Un-attempted

Students are found to leave questions that involve writing long essays, giving critical answers, or explaining certain texts. It indicates that students find themselves at ease to write answers to recapitulatory type questions than to the thought provoking ones. Normally, questions based on textbook lessons are more readily attempted than questions that need thinking and creativity.

Among such questions, questions related to summary writing and dialogue/diary writing were either quite difficult or unfamiliar to them. The un-attempted higher frequencies of these questions also suggest that students have seldom had classroom practices in such writings. In such cases, the teaching skills of such items among teachers can also be put under question. The un-attempted frequencies of questions related to 'explanation', 'comprehension', and essay/letter writing are found standing not far from the above line. However, there is no striking un-attempted figure regarding the grammar questions, although it can be observed that some grammatical questions are more difficult for students than others.

### **Repeated Answers**

Usually, students tend to write answers repeatedly. In this study, 17 (6.07%) students have been observed to attempt questions repeatedly. Although this number is quite negligible, it shows that students lack due awareness during the examination period. Besides this, students write answers repeatedly either because they have lost their confidence in the answer given earlier or are nervous due to exam anxiety. Some students repeat answers to deceive the examiner while others do it for filling answerbook page to show the bulk. However, such tendency somehow has been noticed very frequently.

### **Language Correctness**

Language correctness in writing is a significant problem among the SLC taking students. More than one-third (37.14%) of the sample students could not write Nepali correctly. This problem prevails where the local dominant language is not Nepali, even though it may belong to the same language family or to a different. Language correctness in Nepali writing can be observed even within the Mid-Western and Far Western dialect-speaking areas of the Nepali language itself. The emergence of such tendency among students is not only due to other language/dialect factors but also due to inadequate teaching. Therefore, besides frequent and continuing exposure in Nepali language, its pronunciation and rule practices at the maximum level are a must to avoid such errors in the writings of the students concerned.

### **Questions Not Mentioned or Wrongly Mentioned**

Most of the students (96.36%) have been found to have rightly mentioned the question numbers in their answerbooks. Still, in a few cases (4.64%), students have missed to do this either because of exam anxiety or the of lack of proper orientation.

## **3.7 Consistency in Marking**

To test consistency in marking, a total of 280 answerbooks were re-examined. The marks assigned by the previous examiners (rater 1) and new examiners (rater 2 and rater 3) were computed district/centrewise. Those marks were thoroughly analyzed and were traced out for the range of differences among them assigned by three different examiners. Such differences are shown in Table 3. The table shows that most of the inconsistent cases fall under up to 5 range of difference. Considering the complex nature of language subject, if we ignore the range of the difference of marks up to 5 among the examiners concerned, there is still a significant number of answerbooks (49.29%) suffering from inconsistent markings. It seems quite a serious problem that such gaps are common among the various examiners involved in rating the SLC answerbooks. It has been observed that inconsistency persists even among the examiners with sufficient experience and qualification of head examiner level also. In the present case,

**Table 3. Range of Difference in Marks among Examiners in Compulsory Nepali**

Range of marks	No. of answerbooks rated by all three raters	No. of answerbooks rated by marker 2 and 3	Remarks
Up to 5	142	234	Includes 23 'pass-fail' cases including 12 between marker 2 and 3.
6 – 10	114	43	
11 - 15 <sup>+</sup>	23	3	
Total	280	280	

considering the answerbooks of 11-15<sup>+</sup> range of differences as serious, such examples are presented in Table 4.

The analysis of the above table shows that there are several examples, where the marks given by different raters are inconsistent suggesting lack of uniformity in

rating. Moreover, there are extreme individual cases, which may even affect the result at the extent of 'pass-fail', e.g., the case of one student from Dhanusha in the table above. Similarly, the highest difference was marked 19. Such variations may have occurred due to factors such as lack of specificity in test papers and marking schemes, examiners ignoring or not referring to the marking schemes, poor supervision and scrutiny by head examiners, and tendency among the examiners to value different things (e.g., some giving emphasis on grammar and spelling alone and others on subject matter).

**Table 4. 11 - 15<sup>+</sup> Inconsistent Cases**

S. No.	District	Rater 1	Rater 2	Rater 3	Range of differences	Revised marks
1	Kailali	49	39	37	12	40
2	Kanchanpur	32	45	39	13	40
3	Kanchanpur	44	33	34	11	34
4	Gulmi	45	32	40	13	35
5	Sunsari	40	53	46	13	42
6	Sunsari	42	53	48	11	45
7*	Kapilbastu	54	48	37	17	46
8	Kapilbastu	60	50	46	14	49
9	Kapilbastu	62	49	47	15	45
10*	Kapilbastu	68	62	52	16	49
11	Dhanusa	40	26	33	14	33
12	Dolakha	63	58	51	12	45
13	Dolakha	56	47	45	11	45
14	Dolakha	65	57	54	11	50
15	Sarlahi	45	35	33	12	33
16	Siraha	59	46	44	15	45
17	Siraha	53	45	38	15	40
18	Siraha	45	32	31	14	32
19	Siraha	67	54	49	18	50
20	Siraha	55	44	44	11	44
21*	Rupandehi	62	60	49	13	55
22	Rupandehi	65	51	46	19	48
*23	Tanahu	38	46	50	12	50

\*10<sup>+</sup> differences between R 2 and R 3.

### 3.8 Conclusions and Recommendations

Based on the above analysis, the following conclusions and recommendations have been drawn:

The compulsory Nepali curriculum for Grades 9-10 has not properly addressed problems of other tongue learners as well as remote area dwellers. Their needs in Nepali language are not realised. Further, it does not make any provision for accommodating gender, ethnicity, Dalits, special needs groups and groups, facing difficult circumstances. The scope of content (Widhaako

kshetra) section of the curriculum should include writings in adequate proportion to represent such realities of Nepalese society. Such inclusion is also possible in the curriculum under a new heading of elaboration of contents (Widhaako spashtikaran), if necessary.

The teaching technique (Shikshan prakriya) portion of the curriculum also lacks inclusion of various problems being faced by the children from different sections of the society. It should be improved in a way that all of them feel and find themselves dignified and well-behaved among friends and teachers.

Textbook lessons that are based mainly on literary approach does not seem much favorable to other language students. In fact, lessons should be comprehensible to common L2 learners of Nepali so that they could exploit such lessons for sufficient language practices in various ways.

Textbooks are less representative from the equity point of view. They should be sufficiently equipped with positive references, authorship, success stories, biographies of outstanding personalities belonging to females, underprivileged, ethnic groups, physically impaired groups, Dalits, and so on. One of the ways to minimise the social disharmony is to include letters written to or by members of underprivileged groups and to include references related to family occasions, condolence meetings, festive celebrations, etc. in the lessons and in exercises in adequate number (Niraula 2004; NFD 2004; RDA 2061).

The grid has exactly followed the evaluation guideline outlined in the curriculum. Thus it is more or less representative from the point of view of content coverage. But it still lacks ways of assessing oral communicative competencies. In a real sense, it should lay more emphasis on testing overall communicative competencies than on factual information, which students tend to memorize.

Some of the questions are ambiguous and their ambiguity is language, construction, and task originated. Some questions are quite long, complex, and loaded with several messages and so they seem quite ambiguous for the average students. Some are even internally deficient whereas some seek two or more different tasks to be done (e.g., kina ra kasari). In many cases, instructions are also ambiguous for identifying correct answer. Besides this, many questions are ambiguous because of their unfamiliar contents to certain groups of students. Test papers should be related to student's life situations as far as possible. Various items with unfamiliar contents, especially for students from remote area village settings and from other cultural backgrounds, should be replaced with the common ones. Moreover, the question setters as well as moderators should be aware of using simple, correct, straight and communicable sentences to minimise language ambiguities of all sorts in questions.

Although questions have reasonably followed the taxonomy aspect, the test paper setters are not yet very clear in identifying the exact nature of questions for seeking a level-wise performance of the tasks. Short- term orientation programs and workshops should be organized to make question setters and moderators more acquainted with the taxonomy aspect to be followed for constructing more improved and specific questions belonging to all cognitive levels, as suggested in the grid.

Textual questions have a dominant place in test papers. So, the test papers are not much different from what the curriculum and the textbooks have intended.

There are cases of repetition of questions in the test papers. Some lessons are more frequently asked in the examination and some are rarely. The moderation board should be made aware of avoiding such imbalances.

For the most part, the test papers do not address the problems of L2 learners, remote area dwellers, special needs groups and the socially deprived classes are yet to be considered.

The size and length of the test papers are satisfactory. However, the use of single sober color in one region and a little increment in paper quality of the test papers are preferable. Printing errors have grossly affected the quality of the test papers. So, thorough and serious proofreading is a must.

Efforts have been made to give the marking schemes full coverage of all questions in the test papers in counts. But they do not maintain consistency, especially regarding the guidelines related to some free composition type (nos.3, 4, 6, and 7) and vocabulary (nos.15 and 17) questions. An indication of the main points to be included in answers to these questions also would have helped to make the marking more objective.

The schemes have made no provision for realising the difficulties of L2 background and unprivileged groups of students. Its initial instruction should have elaborated to the extent of addressing such difficulties.

The schemes are not clear as to what to do for partially correct grammatical answers. Spelling errors are enormous even at the level of distorting meaning. Efforts should be made to make them more specific and correct.

Many students have expression difficulty in Nepali. Such difficulty is noticed in areas where non-Nepali dominant language is in use, e.g., Solu, Saptari, etc. or in areas where distinct Nepali dialect is in use, e.g., Baitadi, Jumla, etc.

The pages of answerbooks covered by the students range from 6.7 (Jumla) to 15 pages (Lalitpur) out of the 16 pages in total. The average coverage is 9.72. This is an indication that students commonly either write incomplete answers or leave questions unanswered which has an obvious effect upon their result.

Questions based on textbook lessons are more readily attempted by students than questions of the free composition type. The questions with higher un-attempted frequencies suggest that either they are difficult for students or they seldom have opportunities in classroom practices. In such cases, the teaching skills of such items among teachers can also be put under question in general. Similarly, teaching technique of some grammatical items also seems to be questionable.

Some students have a tendency to attempt questions repeatedly and some others miss mentioning the question numbers concerned. In fact, students should be made fully aware of the exam techniques in their usual situation.

More than one third of the students appearing for SLC cannot write correct standard Nepali. Mother tongue/dialect factors as well as inadequate teaching factors may be responsible behind this situation.

Inconsistency is visible in markings done by various examiners. The range of such inconsistency is 1 to 19 marks. This shows that examiners use their discretion while checking the answerbooks. Appropriate measures should be taken to improve consistency in marking.

#### **4. ANALYSIS OF MATHEMATICS TEST MATERIALS**

Mathematics has long been an integral part of the school curriculum. It is one of the six core subjects, carrying 100 marks each in Grade IX and X. It is considered to be one of the most

difficult subjects. Most of the failure in SLC can be attributed to failure in math. Thus it becomes imperative to look into the curriculum and test materials used in math.

## 4.1 Curriculum and Textbooks

The existing secondary Mathematics curriculum consists of objectives, curriculum structure, contents, teaching methods, evaluation, and scope and sequence. In general, the curriculum comprises computational skills in solving domestic and official arithmetical problems of all kinds; skills in measurement of objects in line, plane, and space of specific configuration; skills of solving problems analytically using algebraic methods; logical establishment of geometric properties and further generalization of the results; and collecting, organizing, represent, and concluding on statistical data for inferences and estimating on the basis of probability theory and also concept building for further study in Mathematics. To achieve these skills, mathematical contents have been presented in the curriculum. However, the teaching activities suggested in the curriculum virtually fail to address the utilitarian and practical aspects of Mathematics. Finally, the students have to memorize almost all the concepts in their effort at remembering for passing the SLC examination. Moreover, textbooks are also designed to fit the contents enumerated in the curriculum with least consideration for the objectives. Eventually, students are led to work like drilling/practicing to memorize the theorems and numerical problems without knowing their utility with the intention to bring high marks in the examination.

## 4.2 Specification Grid

The grid is directed to ensure the validity of test papers and guide the question setter to prepare tests for different cognitive abilities according to their weightage in the grid. The curriculum suggests a grid for developing test papers under the content headings and types of questions with very short, short, and long questions across each content heading. However, the grid is completely silent about the principles of taxonomy, which makes one confused about preparing test items in judging what sorts of cognitive abilities a test paper should contain.

The OCE specification grid reduced Bloom's taxonomy into four categories such as knowledge, understanding, skill, and problem solving in the order of higher ability. In this adaptation, the first two categories go alike whereas skill here means computation and verification and problem solving covers all other higher abilities of cognitive domains. However, this adaptation seems to represent to some extent the philosophy given in Bloom's taxonomy, but the grid does not explain anything about this new classification in order to provide a comfortable and workable option to an already existing classification. The OCE grid has proposed the following pattern and type of questions for the SLC examination test paper: very short, short, and long with 18, 13, and 14 questions respectively. Marks allotted for very short, short, and long questions are 18%, 26%, and 56% respectively.

The knowledge domain commands only very short answer questions. The understanding domain requires both very short and short questions and the skill domain covers very short, short, and long questions, and long type questions fall under the problem solving domain only. This distribution of test items, however, looks logical with respect to testing the ability of students.

The grid is designed according to the curriculum. However, the practical aspect of mathematics teaching is found absent. Some weightage should have been given in the grid to assess student's ability and skills in carrying out life-related problems in Maths practically. The management of

testing practically; is a challenging and cumbersome process nevertheless, it should be considered positively in future. The specification grid does not exactly follow the distribution pattern as presented in the curriculum regarding the types of questions and weightage of the various content units. Analysis suggests that there is a good coverage of the objectives and contents of curriculum in the specification grid.

### **4.3 Test Papers**

The quality of test papers has been assessed using different indicators. A description follows:

#### **Relationship of test papers with the textbooks and curriculum**

For the most part, there is correspondence between test papers, textbooks, and curriculum in terms of the type and number of questions to be used. However, the curriculum does not refer to the taxonomy of educational objectives. This hampers the task of establishing relationship between the test papers, textbook, and curriculum. Test papers are designed to test the memorization of texts that exist in the textbooks instead of the higher level of mathematical competencies. As textbooks are the only source of teaching learning in our education system, examination cannot be separated from it. As a result, test papers have a tendency to be closely associated with the textbook concerned in the absence of desired competencies in terms of the levels of ability according to taxonomy in the Maths curriculum. Analysis of test papers reveals that there is almost non-existence of the test items, which demand students to demonstrate their analytical and creative ability. Thus, the evaluation procedure practiced in SLC is based on the textbook lessons not on the curriculum competencies. Thus a high score in SLC examination alone may not be very helpful in predicting students' better performance in actual settings. Therefore, practical use of Mathematics in research, planning, and decision making and solving their day-to-day problem should be further stressed in the curriculum.

#### **Clarity**

The analysis of test papers reveals that some test items are quite long and require many concepts than very short questions carrying one mark each. Test items carrying one mark under very short type mostly contain word problems, which are not merely of 'knowledge' level but require higher level ability. The above analysis shows that a number of anomalies exist in the test papers. First, the test items do not correspond to a specific ability level to the measured objectively. Second, the test items are often found with discrepancies in units of measurement. Third, the geometric figures in the test papers are not well sketched. Fourth, some sentences are even internally deficient in construction. Fifth, in the test, the items magnitudes and units of measurement are given in separate lines. And, finally, grammatical mistakes, though not common, are found in the test items.

#### **Repetition**

Questions once asked under the very short category have been asked the next year under the short category and questions once asked under the short category are asked under the long category next year. There are many examples of questions with exact word repetition. Most word problems have repeated sentence structures. In arithmetic and algebra, repeated sentence structures vary only in figures (magnitudes) whereas in geometry they vary in naming the geometric figures only. As there are five multi sets of question papers for the five development

regions, the multi sets resemble each other except in magnitudes and the naming of geometric figures.

Parallel sets, in fact, lack their reliability as some sets have a higher difficulty level compared to others of the same year. The analysis of questions of 2002, 2003, and 2004 shows that there are thirty test papers of fifteen parallel and multiple sets. In each content area there are significantly less than thirty different sub-areas, so it is very likely that questions get frequently repeated. Moreover, in geometry there are about ten theorems, which are repeatedly asked in the same year and subsequent years. In Mathematics, repetitions should not be a problem to students' achievement, but the parallel and multiple sets used in the same year sometimes lack equivalent difficulty level.

### **Balance between Grade IX and X textbooks**

Moreover, the analysis of test papers revealed that under very short type questions, 80% of the Arithmetic portion are from Grade IX, 75% of Algebra from Grade IX, and in Geometry 60% questions are from Grade IX whereas 100% questions from Trigonometry, Statistics, and Probability are from Grade IX. This shows that on an average 79% questions under the very the short category are asked from IX content. Similarly, under the short type questions, 66.6% of the Arithmetic portion, 100% of Algebra, and 33.3% of Geometry are from Grade IX. In this category, about 66.6% questions are asked from Grade IX Maths content. In the long type, about 43% weightage is given for Grade IX Maths content. Overall, an SLC test paper covers 55.5% weightage from Grade IX Maths content. The performance of students from Grade IX Maths content is alarming. It shows that the carryover load of Grade IX Maths content could be responsible for the high failure rate in SLC Examination. Despite this scenario, it is mandatory to keep abreast of quite concepts and skills learnt in Grade IX to acquire the concepts and skills of Grade X.

### **Relevance of test papers in terms of grid**

The specification grid (knowledge level, 3 questions; understanding, 9 questions; skills, 23 questions; and problem solving, 10 questions) developed by OCE demands more higher ability questions than the lower ability ones. Review of test papers reveals that the test papers of 2059 did not deviate much from what is suggested in the grid. However, the test papers do not represent exactly the intention of the grid. The question setters seem to have yet to realize the differences among the ability levels of the cognitive domain. The test items should induce the action verbs so as to correspond ability level more precisely. In the test, the number of items of ability level corresponds to some extent to the grid; however, the ability levels of different areas in the curriculum are ignored or unattended.

### **Physical Quality of Test Papers**

The physical quality of the test papers can be considered in terms of instructions, printing, paper, color, size, length of question, and the time given for answering. The analysis of test papers used in the years 2002, 2003, and 2004 under the given criteria reveals the following facts.

#### **Printing**

By and large, the test papers possess clarity in printing. They are free from ink smudges, unbalanced inking, unprinted letters, numbers, signs, and unusual shift in them. Similarly, type sizes used for various purposes are appropriate for the age level concerned. Typefaces used for

instruction and grouping of questions are justifiable. The test papers however, suffer from frequent printing mistakes, few ones even to the extent of distorting the very meaning of the question concerned, e.g., Nepali version of Q.No.9 of 2059 B2. Spelling mistakes are too frequent, which may easily misguide. Instances of such lapses are prevalent in each question paper. The geometric figures often lack proper drawing, which may trouble students to get the right result. Thus a serious and thorough proof reading by the subject expert is a need seriously felt in the test papers to be used.

### **Paper**

The paper used in the test paper is more or less opaque, yet not very appropriate for both-side prints. White, glazy, and sufficient thick paper would have been a better preference. The color of the paper also varies from one test paper to another within the same year and year-to-year. At present, white and off-white (brownish) paper has been in use. The recommended thickness of such paper should not be less than 60-gram weight.

### **Size**

Test papers have maintained a standard size every year. Its consistent size is approximately 6.5x9.5 in inches, which is quiet handy and portable to look and convenient for use. The size is fixed in a way that the paper covers complete six pages without leaving any space unused.

### **Instructions**

Instructions for questions are given in the beginning to cover the whole set of questions. These instructions look clear enough for the students concerned.

### **Length of question**

There are reports, which claim that the length of questions has a serious effect upon examination performance in Maths subject. Informal queries and sharing with teachers and students also reveal that the length of questions for three hours duration is inadequate to complete answering within the time given. For example, the marking schemes of Q5 (a), RE-509B1, B2 show eight and nine steps to get the final result; students may need more time for just two marks.

## **4.4 Marking Schemes**

With the purpose of maintaining uniformity in marking the SLC answerbooks as far as possible, the system of using marking schemes has been in use. These schemes are guidelines for answerbook examiners and head examiners to help them find the possible right answers of the questions given. They do not pose any restrictions over awarding marks for the alternate correct answers. But each examiner/head examiner has to follow them strictly. The schemes are prepared in a fixed format set by the specification grid 2000. The initial part of the schemes is related to general instructions to the examiners/head examiners suggesting them how to evaluate correct, partially correct and alternate answers and when to use their own discretion. But these schemes are silent regarding the problems and difficulties of the students from minorities, special needs, and hardcore groups.

The general instructions in the marking schemes since 2000 AD stand still and unchanged word by word each year. The instructions given in question papers are hardly realized in the marking scheme. The instructions in the question paper say that the examiner will encourage answers

written in the examinee's own words in a creative way rather than answers memorized or as a carbon copy of the text books. But the marking schemes do not instruct at all. The marking schemes have fixed evaluation criteria for each question assigning certain marks for each criterion.

There are no uniform criteria suggesting award for the marks. Some schemes enumerate each and every step of the answer and specify criteria for the award of marks, but others simply specify criteria showing the desired steps for the award of the mark. Marking schemes are not free from mistakes either. In the marking scheme (RE-509 A1) of Q 1(c) 2003,  $3/y$  is written for  $3/4$  and in Q 14, the question demands two numbers but the marking scheme stresses two natural numbers, which may confuse the examiner. Similarly, Q.1(c) of the marking scheme has a mistake (RE-509 A2). In another case, the marking schemes of RE-509 B1, B2- Q.21 contain wrong notations. Q.7 (f) requires a tree diagram but the marking scheme is silent about this. The marking schemes for the same type of questions, 5(a) and 8(c) of RE-509 A1, A2 carry different criteria.

The marking scheme exists not only for uniformity purposes but also for guiding the inexperienced and amateur examiners. Usually, it also serves as a reference tool, especially for the remote area examiners where additional logistics besides the textbook are almost non-existent. The language used for instructions in the marking scheme may not be very appropriate for the targeted examiners. The instruction at the outset in geometry that deducts marks for badly drawn figures may confuse the examiner while awarding marks of very short and short geometric problems which do not require figures in answering at all. This shows that clarity is the first and foremost criterion that a scheme must have. Therefore, the marking scheme intended for compulsory Mathematics subject needs improvements to maintain clarity.

#### **4.5 Response Patterns of the Examinees**

As a part of our assessment of test materials, we did examine the response of students. The results are given below for Math.

##### **Medium of language used**

In the answerbooks sampled, about 85 percent students were found using Nepali as their medium of expression and 15 percent using English as the medium of expression. Students of public schools mostly used Nepali and those of private schools English. This trend is consistent with the overall practice of using Nepali medium in textbooks and teaching in the Government/public schools and using English medium in textbooks and teaching in the private schools. Regionally, the proportion of English medium user was as high as 39 percent in the central region, 14.3 percent in Eastern, and 10 percent in Far Western but none was found from the Western and Mid- Western regions. The average marks obtained by students using English medium was 46.1 percent, which is significantly higher than the 21.4 percent obtained by students using Nepali as their medium of writing.

##### **Expression Difficulty**

Language proficiency, understanding of the mathematical concept, and practice in the school are accounted for creating expression difficulty. About 50 percent of the students were found to have expression difficulty, due possibly to the length of the test papers that can create induce

haste among the examiners resulting in expression difficulty. The use of difficult words in test papers with low weightage, which requires more time to comprehend the question, is also responsible for expression difficulty.

### Questions not attempted or partially attempted

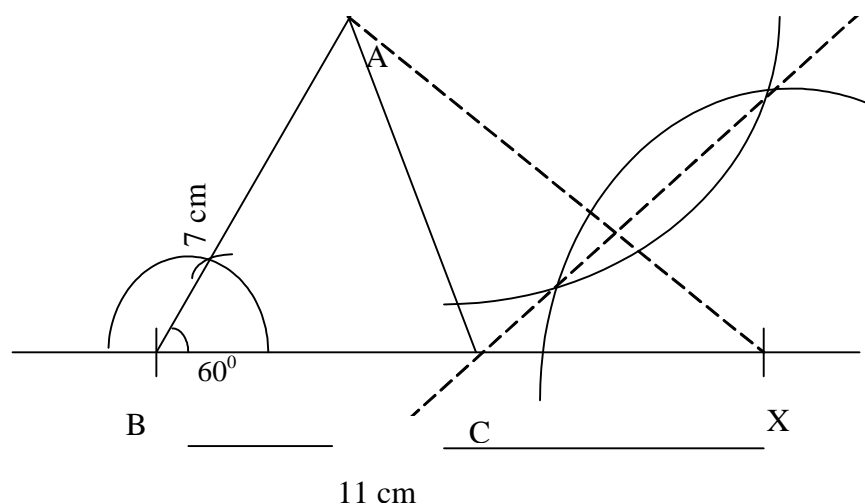
The incidence of failure to attempt questions or attempting partially is very high. Analysis of sampled answerbooks revealed discouraging results. Some 67% of the students did not attempt or attempted partially very short answer questions; and 77% did the same with the short answer questions; and 34% with the long-answer questions. Students' performance on the short and very short type questions was found to be poor. Students' performance on the long type questions was also far from satisfactory contrary to the conviction that Mathematics is high a scoring subject.

### Repeated answers of the same questions

The repeated answer means answering the same question more than once. It can happen either when the students cannot solve the problem in a single attempt or they are unable to respond to other problems. Some cases of answering more than once were found.

### Originality in writing

Originality/creativity in writing is a very important quality of a student. Such quality can be inculcated in them through proper teaching. It is only the original/creative thinking that can contribute to original/creative writing. Current classroom practices do not encourage students to be original and creative. Expert judgment of students' answers revealed that only 15 percent of the sample answerbooks have originality/creativity in writing. Ironically, examiners do punish when students demonstrate originality. A correct construction in geometry of a student was awarded zero mark by two markers. The following figure is one example of such a case.



The marking scheme instructs the student to draw  $AB = 7$  cm horizontally and then complete the construction. The student did differently and correctly for which he was punished. This shows that markers blindly try to follow the marking scheme or the textbooks contrary to the instructions given in the marking scheme. Finally originality/creativity in writing demands that

students be not confined to the textbooks and classroom only, but they should be given an opportunity to think creatively.

### Copying

Here copying means copying from other students' answerbooks or copying to cheat. To ascertain the instances of such practice, it is sometimes unfair from a simply go through the answerbooks of a single student. Such judgment can be made from one-to-one similarity of the steps, correctness, or wrongness in the steps, etc. through comparison of two or more answerbooks of the students. Such acts of copying can be assessed also from the inconsistent steps detected, e.g., in the alternation of correct and wrong answers noticeable in the answerbooks. Screening of sample answerbooks revealed as many as 31 percent of the students appearing in the examinations could have copied the answer, with the highest frequency in the Eastern and Mid-Western regions and lowest in the Far Western region.

### Mention of question number

Analysis reveals that failure to mention the question number is quite frequent among students. About 65 percent of the students were found omitting such mention or wrongly referring the number of the question they attempted.

## 4.6 Consistency in Marking

There is no consistency in the allocation of marks across the different markers (Table 5). This inconsistency, too, resulted in pass/fail cases also which cannot be ignored. There is a significant number of cases of difference of scores among the markers in the range of 1-5, which is higher in the case of M2 and M3. The score difference of six and more among the markers is nearly one-fourth of the total students, and between M2 and M3 this is nearly one-tenth. Moreover, there are extreme cases, which may even affect the result to the extent of passing and failing.

**Table 5. Marks Awarded by Different Markers in Maths**

Range of Score	Markers 1, 2, and 3		Markers 2 and 3	
1-5	190	(76.3%)	199	(91.3%)
6-10	46	(18.5%)	18	(8.3%)
11-15+	13	(5.2%)	1	(0.4%)
Total	249	(92.9%)	218	(80.7%)

Such cases account for 7.8 percent of the total students, a very serious issue that demands immediate attention. Such gaps are common among those involved in marking the SLC answerbooks. It is observed from a sample of 35 answerbooks with the difference of more than seven marks among three

examiners. That inconsistency persists among the examiners, with sufficient qualification and experience even at the head examiner's level. In all the 35 answerbooks, a lot of inconsistency is obvious, which may be due either to following the marking scheme blindly or to ascertain lack of seriousness in observing the process followed or just to verifying the result of the question which simply shows that the markers do not realize the gravity and sensitivity of the task at all. Some examples of anomalies in the marking of thirty-five answerbooks reveal that answerbooks considerably suffer from inconsistent marking. It seems quite a serious problem that examiners generally ignore the marking scheme or do not understand its intention. The analysis reveals that majority of the examiners, even head examiners use their discretion very loosely without keeping in mind the true intention of marking scheme by and large.

## 4.7 Conclusions and Recommendations.

The analysis, leads to the following conclusions and recommendations:

The curriculum should focus on the learning competencies directed towards learning practically applicable to daily life using mathematical contents. Henceforth the teaching technique should be improved in a way that children from all sections of the society feel and find mathematical learning comfortable and useful.

The curriculum should guide through its components to achieve the objectives such as 'to expose the inner ability and talent', to bring hardcore people in society in the national stream' in general, and 'to prepare healthy and capable or ideal citizens able to develop livelihood-oriented mathematical knowledge and skills, and enable them to enjoy Mathematics' in particular.

Textbooks should be made more inclusive from the point of view of equity perspective and alternate measures to cope with the disability of different group (physical or social).

The specification grid should correspond with the curriculum objective. Besides, it should lay more emphasis on testing practical and creative abilities than factual information which students need to memorize.

Test papers should relate to real life situation as far as possible. Various items with unfamiliar contents for students from remote areas and other ethnic groups should be replaced with familiar contents to them in order to realize localization in students.

For the development of originality/creativity in students, the curriculum of Mathematics subject should make the provision of theory and practical part. This provision will let students study and work beyond the confinement of the class room and textbook alone, which will eventually create a positive thinking in them about the uses and value of Mathematics as an important component of life skills. Finally, this will also help students in realizing the underpinning beauty of Mathematics creating more interest in its study.

## 5. ANALYSIS OF SCIENCE TEST MATERIALS

Science is also a compulsory subject. It is taught right from Grade one. The SLC measures both theoretical and practical understanding of science as a subject, with 75 marks for theory and 25 for practical. Students have to secure pass marks in both the theory paper and in practical examination. Science is also known to be one of the 'killer' subjects, in which a large majority of the students fail annually in the SLC examinations. The following sections provide results of analysis of curriculum and testing materials of science as a subject.

### 5.1 Curriculum and Textbooks

The main objective of teaching science is to develop the basic knowledge of scientific concepts, principles, and laws; impart the skills of observation and inquiry, and develop competence in applying knowledge and skills for the solution of problems in daily life. The science curriculum consists of four prominent areas of Science: Physics, Chemistry, Biology, and Astronomy and Geology. The curriculum has given 35 percent weightage to Physics, followed by Biology (30%), Chemistry (25%), and Astronomy and Geology (10%). Such an unequal distribution of weightage is hard to justify. Each area is further divided into a number of units. The unit contains specific objectives, showing the development of expected outcomes among students.

The given specific objectives can be classified into knowledge and comprehension type and higher ability type. The analysis of these objectives reveals that, overall, knowledge and understanding type of objectives overwhelmingly dominate: Chemistry (79.4%), Biology (94.7%), Astronomy and Geology (95.24%), and Physics (50.8%). The objectives related to higher abilities have obviously been given marginal space.

For the most part, textbooks are written according to the prescribed curriculum. Textbooks include illustrations, examples, and exercises. Each chapter contains a summary and activity-orientated problems. However, there are some shortcomings which can be removed in order to make the textbook more useful to the students. While writing school science textbooks attention should be given in coining Nepali words for scientific and technical words. Often, the lack of standardized Nepali words for scientific and technical words creates confusion among students and teachers.

Errors: There are some errors in printing which should be eliminated. For example, derived units are called **तत्जन्य एकाइ** in Nepali. But in the textbook of Grade IX it is written as **तज्जन्य एकाइ** which is wrong.

Although the textbook contains many figures and sketches, there are instances where figures and sketches are necessary but not mentioned. For example, figures should be given to illustrate terms such as **आयताकार, त्रिकोणाकार, वृत्ताकार, गोलाकार, वेलनाकार**.

For example, **displace** is denoted as **स्थानान्तर** in the science textbook. However, **displace** should be denoted by **विस्थापन** and not **स्थानान्तर**.

The content in the textbook is presented in simple Nepali language. However, the content presentation in some cases is not clear and is difficult to understand. For example, Grade X page 65 mentions:

"तत्वहरूका गुणहरू तिनीहरूका पारमाणविक सङ्ख्याका पेरियोडिक कार्यस्वरूप हुन्छन्।" The words and language used in this sentence are not clear enough to communicate the meaning properly to the students.

The textbook of Grade IX (page 184) mentions:

"मस्तिष्कको धेरै भाग ठूलो मस्तिष्कले ओगटेको हुन्छ।" This sentence does not give clear meaning.

The textbooks are well illustrated with examples. However, in some cases examples are missing. There are explanations about **asteroids**, but **examples are not given**. In many instances, textbooks do not explain the terms clearly. In the Universe chapter (Grade X, pp. 184,185) words like **plasma, hydrogen-helium** reaction in the sun are mentioned, but the terms are not explained. These terms need to be explained properly. Lack of adequate explanation leaves both the students and teachers handicapped.

## 5.2 Specification Grid

Specification grid in science published in 2058 has classified the expected abilities into three categories as knowledge, understanding, and higher abilities. The higher abilities include application, analysis, synthesis, and evaluation. The specification grid allots around 50 percent weightage to the development of higher abilities in Physics, Chemistry, Biology, and Astronomy and Geology. There are some discrepancies in the weightage given in the curriculum and the grid. The curriculum weightage in Physics is 35 percent but the weightage in specification grid is

40 percent. Similarly, in Chemistry, the curricular weightage is 25 percent, but the weightage in specification grid is 20 percent. The grid has moved 5 percent weightage from Chemistry to Physics whereas Physics has already got the highest weightage in curriculum. There is no perfect alignment between curricular weightage and marks allotted. For instance, the curricular weightage in Physics is 35%, but the number of marks allotted is 30. On the other hand, the curricular weightage given in Chemistry is 25% and the mark allotted is just 15. Such an unjustified distribution of marks is one main reason of the low content coverage in Chemistry. The distribution of curricular weightage and marks thus needs to be revised and reallocated uniformly.

### 5.3 Test Papers

In Science, different sets of question papers are administered in different regions. Each set of question paper contains ten main questions, which are again sub-divided. Again, in some cases, each sub-item is further divided into two or three small items. Thus every SLC question paper contains 37-43 items of questions, which adds a heavy burden on the students.

#### Test papers and Specification Grid

The specification grid has fixed the number of question items along with the weightage of expected abilities in each area of science test papers. However, the test papers do not seem to have followed the suggestions given in the grid. For instance, the specification grid demands 20% of the question items only from knowledge level. But in the year 2058BS/2001AD, the A, B, C, D, E sets of test papers included 60, 55, 48, 44, and 56 percent of knowledge level question items respectively. Thus there is a big gap between the specification grid and test papers as far as knowledge level question items are concerned. Moreover, the weightage of the understanding level of question items in different sets even in the same year was found to be quite different. Thus, in 2058, the E set included 29 percent of understanding level items whereas the D set of paper included 41 percent, with a variation of 13 points. A, B, and C sets had a tolerable range. Again, the discrepancy in the specification grid and test papers is very much noticeable in the question items of higher abilities. The specification grid demands 49 percent of question items from higher abilities whereas the test papers of 2058 BS gave 8 to 22 percent weightage only on the different sets. Similar variations could be noticed in different years.

#### Errors

In many instances, both the Nepali and English versions of the question items do not denote the same meaning.

For example

#### Set B-2058 Q.1. ka is

Q.1. क= निम्न संख्यालाई छोटो तरिकाले लेख्नुहोस:

Write the following numbers in **Shorthand form**:

The word "**shorthand**" denotes rapid writing in a system using signs or shorter forms for letters, words, etc. which is not applicable in this case.

The English version of this question item should be "**write the following numbers in short form**"

Similarly, set-B, 2058 BS contains an item in Biology which does not convey the same meaning in Nepali and English versions.

Q.7. ka. दिइएका चित्रहरू वीजाणुबाट प्रजनन हुने विरुवाहरू ढाड नभएका जनावरहरूको जीवन चक्रको विभिन्न अवस्थाहरू हुन् । ती विरुवा वा जनावर र अवस्थाहरूको नाम लेख्नुहोस् ।

- In English, "The given diagrams show the different stages of plants propagated with spores or invertebrate animals. **Find** the plants or animals and their stages."
- Here "Find the plants or animals and their stages" is not appropriate.
- The sentence should be like, "**Identify the plants and animals and their stages**".

### Set-B (2058 BS)

The B set of paper in 2058 BS contained the following abnormalities in the English and Nepali versions of test items.

Q.2. Ga. उर्जा संकट भनेको के हो ? **उर्जा संकट दुर गर्ने** एउटा उपाय लेख्नुहोस् ।

What is energy crisis? Give a measure **to push further the energy crisis**.

Now "to push further the energy crisis" denotes deepening the energy crisis. Hence the appropriate sentence would be "What is energy crisis? Give a measure to remove the energy crisis".

Q.7. ka (i) उन्चुको प्रजनन क्रियालाई कति चरणमा विभाजन गरिएको छ र ती के के हुन् ?

- Life cycle of a fern plant is divided into how many generations and what are they called?
- Here the word generation is wrong. It should be replaced by stages.
- The English version should be "Write the different stages of the life cycle of a fern plant".

Q.8. ka कुन प्रकारको कोष विभाजनबाट तन्तुहरूको वृद्धि हुन्छ ?

Tissues grow because of which type of cell division?

- The correct sentence should be "Name the type of cell division which helps tissues to grow".

### Set-E (2058 BS)

Q.2. Ga। शक्तिको रूपान्तर भन्नाले के बुझ्नुहन्छ?

What is law of conservation of energy?

- The English version is totally wrong.
- Law of conservation is not the same as शक्तिको रूपान्तर । शक्तिको रूपान्तर means transformation of energy.
- The correct sentence should be "What do you understand by transformation of energy?"

### Set-C (2058 BS)

Q.8. Ga। जीव भू-रासायनिक चक्र भनेको के हो, एउटा उदाहरणसहित लेख्नुहोस् ।

- What is a **bio-terrestrial-chemical cycle**? Explain with an example.
- Here, the terrestrial denotes only land. The correct sentence should be **"What is bio-geo-chemical cycle? Explain with an example."**

### Set-C (2059 BS)

Q.6.ka तलको चित्रमा प्रयोगशालामा एमोनिया ग्यास बनाउने विधि देखाइएको छ । ग्यास जार भरिए, नभरिएको कसरी पत्ता लगाउने? यो ग्यासलाई किन पानीले तलतिरको विस्थापन विधि वा हावाको माथितिरको विस्थापन विधिबाट ग्यास जारमा जम्मा गर्न सकिदैन, व्याख्या गर्नुहोस् ।

- How do you find out whether the gas jar is full or not? Also explain why is it that this gas cannot be collected in the gas jar by the downward displacement of water nor by the upward displacement of air.
- The English version item is not complete as in the Nepali version. The English version of तलको चित्रमा प्रयोगशालामा एमोनिया ग्यास बनाउने विधि देखाइएको छ is not mentioned. The item should be read as **"The figure given below shows the method of preparing ammonia gas in the laboratory. How would you find out whether or not the gas jar is full? Also explain why this gas cannot be collected by the downward displacement of water and upward displacement of air."**

Q.6 kha ब्युटेनको एउटा प्रमुख प्रयोग पनि लेख्नुहोस् ।

State a common use of butane.

- Common use does not mean " प्रमुख प्रयोग । The sentence should be **"Write one main use of Butane"**.

### Set B (2060 BS)

Q. 3 Ga. पानीको विशिष्ट तापधारणा शक्ति  $4^{\circ}$  से. मा सबैभन्दा बढी हुन्छ किन ?

Why is the **density of water** highest at  $4^{\circ}\text{C}$ ?

- Density does not mean विशिष्ट तापधारणा । विशिष्ट तापधारणा is **specific heat** and **density** is **घनत्व** The sentence should be पानीको घनत्व  $4^{\circ}$ से. मा सबैभन्दा बढी हुन्छ किन ?

Q. 6 Ga. नाइट्रोजन, फस्फोरस र पोट्यासियमको कमीले विरुवामा के के असर पर्दछ ? तामालाई अक्सिजनसँग तताउँदा के बन्दछ ? समीकरणसहित लेख्नुहोस् ।

What effects on plant will be caused by the deficiency of nitrogen, phosphorus and potassium?  
Write with chemical equation.

- Here the English version of तामालाई अक्सिजनसँग तताउँदा के बन्दछ ? is not mentioned.
- **"What product is formed when copper is heated with oxygen? should be added"**.

Q.8.ka.दिइएको चित्र कुन कोष विभाजनको कुन अवस्थाको हो ? यस अवस्थाका दुइवटा लक्षणहरू लेख्नुहोस् ।

What type of cell division and what stage is shown in the diagram?

- The correct sentence should be **"Name the stage and type of cell division shown in the diagram"**.

### Set-D(2060 BS)

Q.5.kha (i) अम्लको परिभाषा लेख्नुहोस् । किन पानीलाई अम्ल तथा क्षार दुवै मानिन्छ?

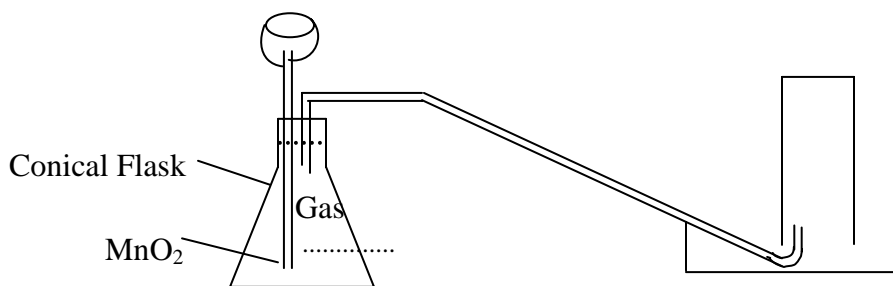
**Give the definition of acid and base.** Explain why water can be considered as an acid as well as a base.

- Here the Nepali version is incomplete. The Nepali version of base is not mentioned. It should be अम्ल र क्षारको परिभाषा लेख्नुहोस् । पानीलाई अम्ल तथा क्षार दुवै मानिन्छ, किन ?

### Set-E (2060 BS)

Q.5.ka तलको चित्रको आधारमा सोधिएका प्रश्नहरूको उत्तर लेख्नुहोस् ?

Answer the following questions with the help of the given figure:



The materials mentioned in this figure are not sufficient to prepare any gas. So the questions asked in item Q.5 ka cannot be answered.

Question items repeated regularly in SLC examination. Dozens of examples of repetition have been noted. It is found that there are few question items, which appear every year.

## 5.4 Marking Scheme

The marking scheme has been developed in science to reduce the possible deviation in the scoring of answerbooks. However, there are some shortcomings in the marking scheme. There are few instances of wrong answers given in the marking scheme. For example:

### Set-A (2058 BS)

**Q.8.Ga:** समुदाय भनेको के हो? समुदायमा च्याउ कुन प्रकारको जीव हो?

एउटै प्राकृतिक वातावरणमा मिलीजुली बाँच्ने वनस्पति र प्राणीहरूको समूहलाई समुदाय भनिन्छ ।

- Direction given in the marking scheme च्याउ विच्छेदक हो ।
- च्याउ विच्छेदक हो । is the wrong answer given in the marking scheme.

### (2059 BS)

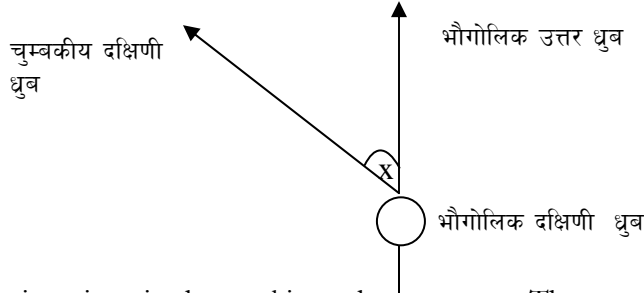
**Q.3. ka II:** कन्भेक्स लेन्सले घामलाई कागजमा केन्द्रित गर्दा बल्छ । कारण सहित व्याख्या गर्नुस् । is the question

**Direction given in the marking scheme** घाममा भएको इन्फ्रारेड पनि कागजमा केन्द्रित हुन्छ । यसले गर्दा तापक्रम बढ्न गई कागज बल्छ ।

The correct answer should be - कन्भेक्स लेन्सले प्रकाशका किरणहरू केन्द्रित गर्ने भएकोले कागज बल्दछ ।

(2059 BS)

Q.4 kha II चित्रमा देखाइएको ह कोणको नाम लेख्नुहोस् ।



Direction given in the marking scheme **अवपात** । The meaning of **अवपात** is **angle of dip**. The angle of dip is a wrong answer. The correct answer is **angle of inclination**. The angle of inclination is called **दिव्पात** in Nepali. **अवपात** is wrong. The right answer is **दिव्पात** .

(2059 BS )

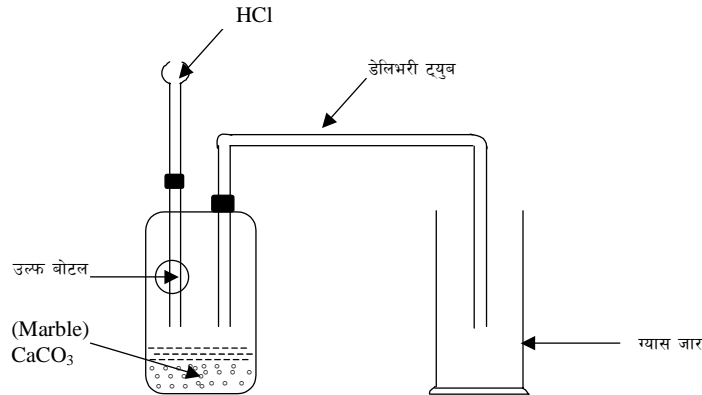
Q.8.ga खाद्य शृंखला भनेको के हो?

The answer given in the marking scheme is सजीवहरूमा खाद्य पदार्थको एक जीवबाट अर्कोमा सर्दै जाने क्रियालाई खाद्य शृंखला भनिन्छ ।

The correct answer should be - उत्पादक, उपभोक्ता र विच्छेदकको आपसी अन्तरसम्बन्धबाट निरन्तर रूपमा शक्ति आदानप्रदान प्रक्रियालाई खाद्य शृंखला भनिन्छ ।

(2058 BS)

Q.6.ka Wrong Labeling of Apparatus Parts



The labeling of Woulfe's bottle here is wrong: the **arrow shows the Thistle funnel as Woulfe's bottle**.

## 5.5 Response Pattern of the Examinees.

The results of analysis of students' answers are given below:

### Language and Expression Difficulty

In the 300 samples of answer books, almost 96 percent students have been found to use Nepali as their medium of expression. However, 20 percent of the students showed difficulty in their

expression in Nepali language. That may be due to factors such as the use of English word in scientific and technical terms.

### **Attempting Same Questions Repeatedly**

Sometimes students tend to write answers repeatedly in the answerbooks during examination to deceive the examiner. Their number was small (9.3%). Such repetition is a wastage of time and effort on the part of the students. Therefore students should be well orientated during teaching learning process that repeated writing of answers in the examination yields a negative impact on themselves because it deducts the time they can devote to thinking and attempting other questions. It also produces a negative impact upon the examiner.

### **Originality in writing**

Almost 81 percent of the students have been found to possess originality in their writing. This is a positive aspect on the part of the students. Students should be encouraged to understand the concept and then express their understanding by writing in their own language.

### **Copying**

The data above show that more than 17 percent (52) of the students tend to copy in the examination. It is a serious problem. Students should be properly oriented about the negative aspects of copying. Examination management should be made more efficient to deal with such a tendency.

### **Correctness in Writing**

The data show that 81 percent of the students have written their answers in correct language. Language is one of the main factors in writing answers. If students cannot use correct language, answers cannot be given properly even if students understand the answers to the questions. Allowing students to practice regular writing and checking by the teachers can easily improve this weakness.

### **Mentioning of the Numbers of Questions**

More than a quarter (29%) of the students either did not mention the numbers of the questions or mentioned them wrongly. Such a practice increases the chances of losing scores.

### **Questions not Attempted or Partially Attempted**

Overall, one-fourth of the total questions were either not attempted or only partially attempted. This could be one reason for the high rate of failure. Students from the Central Region performed better compared with the students from other regions. Students from the Far Western Region did poorly compared to others. Most of them could not attempt science questions of higher abilities. This may be due to many factors such as rote memorization by students, classroom teaching with emphasis on lecture, and giving notes to students and students not habituated to higher abilities questions.

## **5.6 Consistency in Marking**

Some 300 answerbooks were re-examined to find out the extent of consistency in marking by different examiners. Three markers examined the same answerbook, including the original

examiner appointed by the OCE. Thus the patterns of inconsistency found in marking are reported in Table 6.

**Table 6. Inconsistency in Markings**

Range of Scores	Answer books examined by M-1, M-2 & M-3		Answer books examined by M-2 & M-3	
1-5	184	(62.8%)	218	(72.7%)
6-10	78	(25.0%)	60	(20.0%)
11-above	38	(12.2%)	22	(7.3%)
Total	300	(100.0%)	300	(100.0%)

The table clearly shows that inconsistency persists extensively in the marking of science answerbooks. The inconsistencies among the markers were very high (63 % among the three markers and 73 % between the two markers with marking range of 1-5

followed by the range of 6-10. However, inconsistent marking with marking range 11 - above was found to be rarest among the markers (7% and 12.7%). Inconsistency in marking was distinct even between the two selected examiners (M2 and M3) who were quite experienced and senior science teachers. Such inconsistency in marking distinctly affects the SLC results not only in its pass/fail rates but also in the percentage score of an individual student, which is detrimental to both career and scope for pursuit of the higher studies. The inconsistency it self may be due to factors such as the subjectivity of the markers, defective marking scheme, and incompetence of the examiner and head examiners. Generally, a science teacher with B.Sc. or M.Sc. degree will not be able to teach school science unless s/he studies or gets trained in all the areas of science such as Physics, Chemistry, Biology, astronomy and geology.

One main factor in inconsistency in marking is the negligence on the part of the answerbook markers who often do not follow the marking scheme. For example, question 6 kha in the C-set of year 2059 BS SLC examination for the Western region reads as:

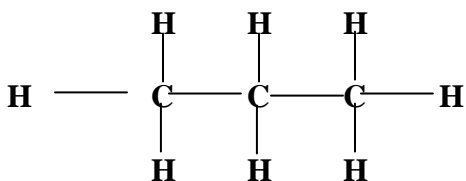
Write the structural and molecular formula of butane. State common use of butane. (Mark 1.5)

**Direction given in marking scheme**

- $C_4H_{10}$  ..... 0.5
- $$\begin{array}{ccccccc} & H & H & H & H & & \\ & | & | & | & | & & \\ H & - & C & - & C & - & C & - & C & - & H \\ & | & | & | & | & & \\ & H & H & H & H & & \end{array}$$
 ..... 0.5
- इन्धनको रूपमा प्रयोग गरिन्छ ।..... 0.5

One of the answerbooks shows:

Q.No. 6 (ख) ब्यूटेनको संरचनात्मक सूत्र :



ब्यूटेनको अणु सूत्र :  $C_4H_{10}$

The answerbook clearly shows a wrong answer. However, the first marker has happily given 1.0 mark to this answer out of 1.5, which shows the examiner's negligence.

**Q.No.9ka in 2059 BS (2003) set B Central Region asks:**

What is meant by asexual reproduction? Write two differences between male gamete and female gamete. 1 + 2 = 3

The direction given in the marking scheme is:

भाले र पोथी लैंगिक कोषहरूको समागम नभई हुने प्रजनन क्रियालाई अमैथुनिक प्रजनन भनिन्छ। ..... 1.0

**भाले ग्यामेट**  
यसलाई शुक्रकित भनिन्छ।  
यो सानो हुन्छ।  
यो सक्रिय हुन्छ।

**पोथी ग्यामेट** ..... 1+1=2  
यसलाई ओभम वा अण्डा भनिन्छ।  
यो भाले ग्यामेटभन्दा ठूलो हुन्छ।  
यो निस्क्रिय हुन्छ।

### Answerbook and marking

(क) अमैथुनिक प्रजनन भनेको भाले र थमिट र पोथी ग्यामेन्ट समागम नभई सन्तान उत्पान्न गर्नेलाई अमैथुनिक प्रजनन भनिन्छ।

**भाले ग्यामेट**  
यसमा डिम्ब उत्पादन हुन्छ।  
यसले डिम्बलाई जोगाउने काम गर्दछ।

**पोथी ग्यामेट**  
यसमा शुक्रकीट उत्पादन हुन्छ।  
यसले शुक्रकीटलाई सुरक्षित राख्दछ।

## 5.7 Scoring Practical versus Theoretical Examinations

The average marks obtained by students in the theoretical and practical components of science do not correspond. The average marks in practical lie in the range of 79 to 83%, while the same in theory range from 21 to 36%. Apparently, there is no consistency in awarding marks in theory and practical papers. It may be due to student pressure because practical marks are awarded by schoolteacher her/himself. Many schools do not engage students in practical activities. Students do not get any lab experience. Strange enough, their marks in the practical papers are often higher than those obtained in the theory paper. Thus students remain handicapped in practical knowledge and skills in school science.

## 5.8 Conclusions and Recommendations.

On the basis of the analysis presented above, the following conclusions and recommendations are made:

- The curricular weightage and the weightage of specific objectives given in the secondary school science do not coincide. The discrepancy in curricular weightage and specific objective weightage is very sharp in Physics, Biology, and Astronomy and Geology areas. In Chemistry, such a discrepancy is nominal. Further, the discrepancy in the expected development of higher abilities in science curriculum and specification grid is very large, especially in Chemistry, Biology, Astronomy, and Geology areas. This discrepancy in the

curricular weightage and the weightage in specification of grid should be avoided. Moreover, school science teachers should be given training in conducting science classes to develop higher abilities among students. Secondary School Science Curriculum should be revised to remove the discrepancy in curricular weightage and specification grid.

- Science textbooks should be written in a simple and clear language so that students understand the concepts in science easily. Scientific and technical terms should be carefully translated in Nepali so that both words (English as well as Nepali) give the same meaning. Each science concept should be carefully explained, with clear illustrations. Figures and sketches should be given wherever possible and should be well labeled. If possible, multicolored figures and sketches should be given in the science textbooks. The presentation of the content materials needs simplification. The use of science activities in each chapter is a welcome approach. If the science lesson is taught by the simple lecture method, which is very probable in our country, one has to think about the impact of this kind of activities orientated science textbooks on the student's learning outcomes. In simple terms, this kind of science textbooks will be useful only in such schools where students' activities dominate science class. The inclusion of summary in each chapter is also a welcome approach. However, the summary parts of each chapter need to be organized well with the help of basic facts and principles. Moreover, the inclusion of exercises "Do, observe, and learn" parts in each chapter are very useful for students. While presenting text contents, the materials, and examples chosen should be simple and familiar to the students. Moreover, science textbooks should include a glossary of scientific and technical terms to enhance the quality of the book. Secondary School Science textbooks should be revised to make the topics more simple, clear, and free of errors.
- There are many mistakes in the language of question items. Both English and Nepali versions of the question items must convey the same thing with same meaning. Moreover, mistakes with technical terms must also be eliminated. In order to remove the mistakes in scientific and technical terms, the setting of and moderation of science question papers should be given to experienced and well known science educators. Repetition of questions should be avoided or minimized. Also, the SLC test papers do not cover the specification grid with expected abilities. Large variation is found between the specification grid and the test papers in the expected abilities.
- In the marking scheme many mistakes are found where mentions the expected correct answers are not indicated. For example, in the marking scheme, in the year 2058 BS, set A of eastern region, question 2'Ga' does not mention the conditions for nuclear fissions. Similarly in that very set, question 3 'Ka' does not mention the complete ray diagram and images formed. Q.8'Ga' mentions mushroom as विच्छेदक which is **wrong**. Thus while developing marking scheme, care should be taken so that all the marking schemes are correct and to the point. The points of the expected answers should be clearly mentioned. The expected definitions should be clearly written. In some cases, definitions are clearly mentioned in the marking schemes while in other cases they are not. Such discrepancy have to be avoided.
- There is expression difficulty among the SLC science students. Although a majority of the students use Nepali language, the level of difficulty of expression is significant. More

than 20% of the students suffer from such difficulty. It may be reduced with improvement in Nepali language teaching. Language of the SLC test items should be simple and clear. Generally, rote memorized answers are written correctly. But answers which need reasoning are found to be expressed with difficulty. This may be due to the nature of teaching-learning activities. Students, who are generally asked to listen to teachers' lecture, try to memorize the lecture, class notes, and textbooks.

- There is no consistency in the marking of the science answerbooks. The range of such inconsistency lies between 1 and 26 clearly showing that the examiners do not use marking scheme and use their own discretion loosely. Such inconsistency in marking can be removed by appointing experienced and trained science teachers as examiners. The marking scheme must include the expected main points in all question items. The distribution of marks must be done according to the expected answers.
- Many students either did not attempt or attempted only partially most of the questions. For instance, more than 65 percent of the students from the MWDR could not attempt Chemistry questions. From the FWDR, more than 58 percent of the students could not attempt Physics questions. More than 95 percent of the students from FWDR could not attempt questions from Astronomy & Geology. This may be due to poor teaching in schools, too many question items, and overloaded science curriculum or the inclusion of higher abilities questions to which students are not used or else, the inclusion of question items from Grade IX which they learned a year before. Overloading of science curriculum with too many concepts should be adjusted to the age level of students. Research findings also show school science teachers to be of the opinion that secondary school science curriculum is overloaded. Well trained science education graduates and science graduates should be appointed as school science teachers. Only those persons must be appointed as school science teachers who have graduated in science education. In the case of science graduates, they must be provided intensive training in science education with Physics, Chemistry, Biology, Astronomy, Geology, and Pedagogical courses. School administration in general and science teachers in particular should be held responsible and punished accordingly for not completing science courses in time. However, reward should also be given to science teachers who completed the course in time.
- The difference in the range of marks in theoretical science and practical science is wide enough, that is, 20.8% in theory and 83.2% in practical. In fact, the practical classes and examinations in science are often not done in schools. Still, full marks in science practicals are awarded to all students irrespective of their abilities. Thus, practical science has become a big joke in our school system. This must be taken seriously and improved appropriately. In order to improve the situation, schools should be compelled to run practical classes in science with examination held by external examiners. All schools should be made responsible to run practical classes and keep records and assess regularly as prescribed in the curriculum. Science educators, senior science teachers, and school science supervisors should be allowed to evaluate students' performance in practical science.
- Too many question items in science (37-43 items in 2059) should be avoided. The number of question items should be about 15 to 20. Questions must be set on the basis of the science curriculum and specification grid. All sets of science questions must

contain an equal number of question items of similar abilities. Questions in science must be included to test the organization ability of students.

## **6. ANALYSIS OF HEALTH, POPULATION, AND ENVIRONMENT EDUCATION (HPE) TEST MATERIALS**

Health, Population, and Environment (HPE) education is one of the core courses of secondary education introduced since 2001. The subject carries 100 marks in SLC, where 25 marks are practically assessed by the subject teacher and 75 marks assessed theoretically in SLC Examination. The examinee must pass in both the examinations.

### **6.1 Curriculum and Textbooks.**

Health Education, Population Education, and Environment Education are three separate subjects, which have been integrated to form a core course entitled Health, Population and Environment (HPE). The curriculum of Health Education is designed to bring changes in health knowledge, attitude, and behavior of the individuals. Similarly, the courses of Population Education are designed to develop awareness about and understanding of population situation as well as rational attitude and behavior of the students toward those situations by realizing their own responsibilities so that they can better deal with the population problems for the attainment of quality life. Likewise, the courses of Environment Education are designed to make students recognize the values of environment and clarify concepts to understand and appreciate the interrelatedness among man, his culture, and his biological surroundings. There are altogether 19 units in HPE at the secondary level.

Most of the contents included in the curriculum are basically focused on the development of knowledge, attitude, and behavior towards HPE education. However, some specific units are also designed to address adolescent girls and boys, elderly people, adult women, and responsible parents. Contents like sex determination, gender issues, role of mothers, sexual and reproductive rights, women empowerment can solve gender discrimination in the society. Adolescent sexual and reproductive health can specifically address the problems of adolescent boys and girls. It also, puts emphasis on the adolescents' need to develop their life skills to tackle the problems. Similarly, the contents on the composition and cultures of geographical locations and ethnicity can address the status and activities of various ethnic groups. However, they cannot specifically address the specific problems and activities of indigenous/ethnic minorities. The curriculum is silent in dealing with special needs children and the children living in difficult circumstances.

In some areas, conceptual clarity is required. For example, the purpose of including cultural heritage, environmental health, and family life education should clearly be spelt out. Similarly, some texts in population education and environmental education are overlapping. For example, waste management, cultural heritage, ecosystem, biodiversity, and management of environment are found in more than one unit. Likewise, population management and family planning contents are mentioned in two units.

This subject is basically concerned with the development of knowledge, attitude, and practice of students towards HPE. Therefore there is no question of equity raised. However, among some ethnic groups, the question may have to address the socio-cultural factors, affecting HPE. For example, occupations, dresses, and traditional values differ according to the ethnicity and geographical locations.

The curriculum and textbooks of HPE are successful in addressing the growing needs of the school children. However, they can not fully address the situation and status of poor households, ethnic minorities, and special needs students. The previous studies also pointed out that the curriculum and textbooks could not address the behavior of students in many ways (Curriculum Analysis 2004 AD).

In totality, an attempt was also made to integrate three subjects into one, but the demarcation line of each subject is clearly seen in the course except in three units. Also, the Environment Education is more focused on Environment Science, which ultimately is unable to address the nature of the subject. Moreover, the subject matter included in Environment Education is already included in Science.

Recently the curriculum and textbooks have been revised according to the demands of different teachers. The contents on reproductive health, gender issues, and common diseases are incorporated into the revised curriculum and textbooks. Attempt was also made to introduce life skills approach in presenting materials in the textbooks. However, from the perspectives of disabled, and Dalit children, some contents need to be added while dealing with the composition of population.

## 6.2 Specification Grid

According to the specification grid, provision is made to assess the students internally and externally. Out of the total 100 marks, 25 marks are allotted to internal assessment, which should be based on totally practical activities. Similarly, 75 marks are allotted to external assessment or final examination using multiple choices, short answer, and long answer type questions.

The provision of internal assessment has been made in the curriculum in terms of practical activities for assessing the development of attitude and behaviors of the students. Unfortunately, most of the subject teachers award full marks in internal assessment without engaging students in the activities suggested by the curriculum. Theoretical examination emphasizes only the assessment of cognitive behavior of students.

Specification Grid is silent in addressing the levels and taxonomy of behaviors in multiple-choice items and the question setters are free to construct questions at any level and domains in this category. However, most of the questions asked in this section are of recall type.

For setting 12 short answer type questions, the question setters have been given little option. Among the twelve questions three from knowledge, four from understanding, four from application, and one from synthesis and evaluation level of the cognitive domain are asked whereas for the long answer type questions, the question setters are provided the options to construct one question each from understanding, application, and analysis level.

As far as the content coverage from the curriculum is concerned, the specification grid could not address all the given units equally though the weightage of the units was similar to each other. As a result, only one multiple choice question each was asked from the units of (i) Primary Health Care; (ii) Population, Environment, and Development; (iii) Consumer Health, and (iv) Safe Motherhood. This provision allowed the teachers to skip certain units while teaching. As a result, the learning outcomes of the subject could not meet the expectation of the curriculum.

Though long answer type questions are asked each from the three areas, the specification grid is environment-biased in the case of short answer type question. The highest number of short answer type questions (5) was asked from the Environment part, whereas only three questions

from Population and two questions from Health Education asked. Likewise, two questions were asked from the integrated units.

The discussion above clearly indicates the need for revision of the Specification Grid and the course so that questions can be asked equally from all the units according to their weightage. Also, it needs a little flexibility in constructing questions from different units. Considering the problem of content coverage, the specification grid has been revised recently to incorporate questions from all the units. According to the revised grid, very short answer type questions have replaced the multiple-choice items. Likewise, 17 short answer type questions are to be asked from different units carrying 3 marks each. The candidates have to attempt any 15 questions from this group. This provision induces the question setters to set questions from all the units. As a result, teachers are compelled to teach all the units of this subject.

### **6.3 Test Papers**

Out of 100, 75 percent marks are assessed for theoretical examination in SLC with the remaining allotted for internal assessment for practical examinations within the concerned schools. The examinee is required to pass in both examinations. Altogether 15 sets of questions introduced from 2057 BS to 2059 BS are reviewed (three years x 5 sets) to analyze them. A brief description of the analysis of the questions is given below.

#### **Objective Questions**

There is a provision for asking 10 multiple-choice questions carrying one mark each in the objective section, which requires the candidate to complete answering within 15 minutes. The examiners are strictly instructed to construct the multiple-choice questions from particular areas; therefore, it was observed that all the questions are set according to the grid.

#### **Repetition**

Some questions were repeatedly asked more than others. For instance, topics of descriptive subject matter, e.g., Family Life Education, Population, Environment, and Development, were repeated more often. There is a tendency among the question setters to copy questions from the previous years. As a result, students are compelled to collect old questions to study rather than thinking critically on the subject matter. Hence, the learning outcomes will not be achieved even if the candidates score more marks in the paper. This tendency hampers the development of attitude and behaviors of the students towards HPE Education.

#### **Clarity**

In total, about 16 percent of the multiple-choice questions in three years were found correct in terms of their structure. Question papers lack language editing. Still, the meaning of the question is somehow communicated to the candidates. Also, question setters constructed multiple choice items comparatively better in terms of content coverage, use of terminologies, and time allotment.

#### **Short Answer Type Questions**

There is a provision for asking twelve questions in this category; out of them candidates have to attempt any nine. Each question carries 5 marks; therefore, the total marks allotted for this section is 45. However, the question setters are allowed to exercise flexibility in asking sub-

questions within the framework of the aforesaid number. No ambiguities were found in the questions of the above three years. Similarly, questions were obviously not so difficult because candidates had alternative questions to attempt. The pass percentages were above 97 % in all three years' examinations, suggesting that the questions were not too difficult. The reasons may also lie in the nature of the subject matter where candidates could use their own experience drawing from their day-to-day life. In long answer type questions, candidates have to attempt any two questions out of three carrying 10 marks each. It was observed that questions in this section were clearer and avoided duplication. However, some questions on population measures may be difficult for students with poor Mathematics background. Repetition of questions was observed in very few sets. However, the question setter had changed the figure and data in the questions. It is easier to set questions from the concerned units because the units have a large coverage of contents as compared to other units.

### Relevancy in Terms of the Grid

Questions were also assessed to find out whether or not they were set as per the instructions outlined in the taxonomy and the given grid. Analysis shows that most of the short answer type questions were asked from the knowledge level. According to the grid, compared to three questions required from the knowledge level, the question setters had constructed up to 10 questions and questions on analysis level were also asked in all the sets though not required. But the questions on understanding and application were rarely asked and no questions were asked on evaluation and synthesis. Clearly the question setters were not following the specification grid as prescribed. As regards the long answer type questions, 45 percent were asked from the knowledge level against 13 percent from application and analysis although the specification grid clearly prescribes the percentages as 33.33 percent from understanding, application, and analysis. The question setters were thus not following the specification grid according to the suggested levels and taxonomy. Since the schools lacked teachers with HPE background, the questions were asked randomly from different levels. The moderators also did not look properly into this anomaly, analysis of the 15 sets of question papers suggests that a region-wise question paper is not required. Moreover, the questions asked in the different regions do not address the respective regions. Instead of introducing five sets of papers in separate regions, parallel sets of papers would work better. But attention should be given to make them equal in terms of the level of difficulty.

### 6.4 Marking Schemes.

Altogether seven sets of marking schemes (2058 BS B, C and E; 2059 BS B, C and D; and 2060 BS B) were reviewed. As far as the marking schemes of the subjective questions were concerned, they were found to have developed in two ways. The first type emphasized only the distribution of marks. For example, the question asked was: **“What is meant by physical and mental health?”** (2 marks). The marking scheme instructed to give one mark for introducing physical health and another one for introducing mental health without any specific points. In this case, the marking scheme is not clear about the specific points to be written. As a result, the examiner may have to consult text materials to find the specific answer. Thus, the marking may vary in this type of scheme. Another type of marking scheme was found prepared specifically, with possible answers. This type of marking scheme is comparatively better, especially for HPE Education teachers, because most of the teachers are from another subject background. While reviewing the contents of the aforesaid schemes, almost fifty per cent schemes were developed

specifically. This clearly indicates that additional efforts are needed to prepare specific marking scheme.

## **6.5 Response Patterns of the Examinees**

The two examiners assigned by the SLC study checked the answerbooks separately and made separate reports on the checking. For preparing the report, they were given some criteria to follow. The reports prepared were specifically studied to obtain the necessary information. The findings needed and the subject specialist's views are presented below:

### **Expression Difficulty**

About 90 percent of the examinees used Nepali as their script of writing. Most of the examinees from private schools used English. The markers had different impressions about the difficulty of expression. In either case, it was found that more than 20 percent of the examinees felt difficulty in expression.

### **Number of Questions Left**

Questions from the Environmental Education portion have not been attempted by more than 40 percent of the examinees indicating that Environment Education was felt either to be a difficult subject or the teaching of this subject did not take place in the schools.

### **Repeated questions**

About 25 percent of the students attempted more questions than what was required. The examinees show a general tendency to tempt the examiner to award marks by answering more questions than asked. About 3 percent of the examinees repeated the answers reflecting the same tendency of the examinees to tempt the examiners.

### **Originality in Writing**

Originality was usually found lacking in the students' responses. They were good at reproducing texts from the textbooks.

### **Copying**

About 10 percent of the answerbooks gave evidence of copying noticed in the appearance of the mistakes of similar nature in the copies of examinees who sat next.

### **Correctness of Language**

Only about 20 percent of the examinees showed correctness in writing.

### **Question not mentioned**

About 10 percent of the examinees either did not write the numbers of the question or wrote it wrongly due possibly to the nervousness or carelessness.

## 6.6 Consistency in Marking

A review of the marks given for 280 papers by three different raters shows, more consistency in the answerbooks marked by the Markers 2 and 3. The results of the consistency analysis are displayed in Table 7.

**Table 7. Range of Difference among the Markers in HPE**

Range of Scores	Markers, 1, 2, and 3	Percent	Markers 1 & 2	Percent
0	2	0.71	9	3.21
1-5	76	27.14	119	42.50
6-10	120	42.86	129	46.07
11+	82	29.29	23	8.22
Total	280	100.00	280	100.00

The table shows that the range difference of marking was high between the Marker 1 and Marker 2/3. If the range of difference of 5 marks is ignored, more than 70 percent of the answerbooks gave evidence of inconsistent marking. This implies the examinees suffered from inconsistent marking. The grand total of the Marker1 was found to be

comparatively lower than that of the Markers 2 and 3. As noted earlier, Marker1 did not even add up the marks secured in the objective questions to the total, which increased the range.

Likewise, the range of the difference in marking between Marker 2 and 3, assigned by the study team, was found to be comparatively low. However, more than 50 percent of the answerbooks featured a range of 6 or more marks. The possible reason for this difference was that Marker 3 was more lenient. Inconsistency in marking could raise questions on the validity and reliability of the test papers. Therefore, the subject specialist also tried to scrutinize some of the answerbooks, which had a range difference of more than 10.

The average marks obtained by the examinees in the same answerbooks were different among the three different markers. The first and the second markers showed consistency in checking (44% vs. 47%), but the third one was found somewhat lenient. There was a difference of 10 marks between the first and third Markers and of 7 marks between the second and the third on an average.

Analysis of the marks given by the first marker shows the first marker did not rate some of the answers. S/he did not add the scores given for the objective items in 12 out of the 27 answerbooks scrutinized in the process of reviewing the marks offered by the three different markers. This may be and large explain the low average marks given, an error found also in the process of retotaling the answerbooks at OCE.

In totality, the first and second markers and the subject expert marked the answerbooks almost similarly. But the third marker gave the marks leniently in all the answerbooks. This is simply a problem of attitude of the examiner and suggests the answerbook markers require proper orientation.

## 6.7 Scoring Theoretical vs Practical Examinations

HPE is a subject where students get practical and theoretical marks separately. The students have to secure at least pass marks separately in each to get through. In practice, students get comparatively better marks in the practical whereas in theory paper they often fail. Therefore, an attempt was made to see the difference between these two parts in the available answerbooks. Altogether 276 cases out of 280 were examined for a comparison of their practical and theory

**Table 8. Scores Awarded in Theory and Practical Examinations in HPE**

Type of Examination	Total Examinees	Full marks	Average marks
Theory	276	75	32.5 (43.3 %)
Practical	276	25	20.5 (81.9 %)

marks. In the case of four examinees, the practical marks were missing. Table 8 presents the actual deviation between these two marks.

The table shows that out of 75 marks, in theory paper, the candidates scored 32.5 (43.3%) on an average. The theory marks

are based on the marks given by the OCE. On the other hand, out of 25 marks in practical, the candidates scored 20.5 (81.9%) mark on an average. The practical marks are based on internal assessment, given by the school (subject teacher) itself. The results show that there is a big gap between the scores awarded in theory examination (final) and practical examination (in the school).

## 6.8 Conclusions and Recommendations

Based on the foregoing analysis, some conclusions and recommendations are drawn:

- Most of the questions asked are recall type in multiple-choice items, where students can directly copy from their textbooks. Orientation program is needed for the question setters especially in the construction of multiple-choice questions and of questions from different domains.
- There is a scope for copying answers from the neighboring candidates because multiple sets are not available in this paper within the regions. Parallel sets of question papers should be introduced instead of giving five sets of papers separately in rogue at present.
- The nature of the question papers clearly shows that all the setters are not from the subject background which determines the level of the question. Appropriate persons must be identified for setting questions in HPE.
- All the question setters have not followed the specification grid, especially in setting up the subjective type questions. Question setters should strictly be instructed to follow the specification grid.
- Candidates can take maximum advantage from the paper because of the repeatedly asked questions even if however it is very difficult to secure very high marks. Repetition must be avoided.
- More than 40 percent of the examinees did not attempt questions from Environment Education. The subject matters included in Environment Education is basically from Environment Science. This indicates that the area of Environment Education was either not taught properly or the area is difficult for the students to understand. Appropriate actions need to be taken in this regard.
- Different examiners rated differently though they were provided with marking schemes. This shows that either the markers did not follow the marking scheme strictly or the marking schemes were not specific enough. However, there is no significant difference in the marking of the answerbooks.
- The pass percentage of this subject was found to be very high (97.25); however, the average mark that the candidate secured in the paper is comparatively low, which shows

that it is easy to get pass mark even with less effort. This might be due to general nature of the subject that is related to the life situation.

- The content load is found to be heavy for the students in this subject if the curriculum is followed properly. Separate examinations for Grade nine and ten should be conducted to assess the students' performance so that their load can be relieved.

## 7. ANALYSIS OF TEST MATERIALS RELATED TO SOCIAL STUDIES

Social studies, one of the eight subjects taught in the secondary Grades, plays an important role in preparing children to take on the duties of citizenship. It is an integrated field of knowledge, which is imparted to develop a balanced personality. This section summarizes the findings of the analysis of curriculum and test materials of social studies.

### 7.1 Curriculum and Textbooks

Social studies draws contents from a number of disciplines within the field of social sciences such as history, geography, economics, civics, sociology, anthropology, social psychology, and philosophy. Of these, history, geography, and civics are the core areas. Geography occupies the top priority in the social studies curriculum with 25% weightage, followed by historical studies (20%). Other fields such as economics (15%), environmental studies (15%), and civic studies (10%) are also part of the curriculum. The curriculum also incorporates topics related to community, nation, development, social values, tradition, international understanding, peace, and cooperation.

It should be mentioned that the social studies textbooks lack in rich/wealthy information. The quality, extent, length, and standard of the subject matter in the textbooks are not satisfactory. There are many lessons but they lack the subject matter according to the need of learner and nature of exercises. Maps, pictures, sketches, charts, and diagrams are essential illustrative materials that make textbooks more meaningful, attractive, and informative. Evaluation of the nature of placement of illustrations shows they are not appropriately done: their size, signs, symbols, and letters used are not clear and distinct. They are given in black and white. A map of Nepal in Grade IX textbook is associated with rivers and streams of Nepal. The line indicating the major river is smaller than that of its tributaries, which only creates confusion among the readers. Thus, a number of the illustrations given in the books are not properly designed and incorporated.

Exercises are essential to evaluate the performance of the students. They are also important tools for developing knowledge and skills in the students. In the social studies textbooks some exercises have been included at the end of each lesson. They are of varied nature—subjective and objective. Most of the questions are of subjective type—long and short. In the objective type, matching, true and false, and fill in the gap type items are included, but multiple-choice questions have not been considered. The arrangement of exercises, too, is not systematic in all the cases.

The exercises in the secondary textbooks are activity-oriented. However, all of these exercises are not practicable because a lot of time and reference materials are required for solving them. It is not possible to collect the materials required in most of the areas of Nepal.

## 7.2 Specification Grid

Test items are constructed according to the specification grid in order to maintain uniformity and consistency in marking. Specification grid also provides guidelines to cover the course. The grid has identified four areas of learning: knowledge, comprehension, higher abilities, and practical exercises as mentioned in the curriculum. Knowledge and comprehension are associated with the cognitive domain. These areas carry 40 percent marks that include one long answer question of 10 marks with the rest left for the short answer (too short 6 and short 8 questions) type questions. The items dealing with higher abilities carry 50 percent marks and are related to knowledge, attitude, and skills. Under this, one long question is associated with practical abilities with reference to map work in geographical studies. It is the area of psychomotor domain. Thus the grid has made an attempt to cover each area of grid, a modified version of Bloom's taxonomy.

## 7.3 Test Papers

The test papers are based on the specification grid. It is also obvious that the grid and test papers have made an attempt to follow the guidelines provided by the curriculum. However, the grid had deviated 1-3 points from the weightage suggested in the curriculum. The test paper has followed the grid while constructing the questions. Test papers for the last three years were analyzed and its results are given below:

### Clarity

The language used in the test paper is simple and clear in most of the cases. However, there were some errors.

### Repetition

There is no problem of exact repetition in the test papers. Some of the topics like map work and field visit have been repeated. However, the styles of interpretation are different and they are also modified and specified. Within the various sets used in the same year and different years no exact repetition could be noticed.

### Test items in terms of curriculum and grid

Test items are prepared according to the curriculum and grid constructed. They have followed the weightage given in the curriculum and grid. Geographical and historical studies are given the top priority in the curriculum. The grid has also attempted to consider this aspect. The test items have considered each of the aspects mentioned in the curriculum and grid.

### Physical Quality

Under this component, the instructions, printing, quality of the paper, color, size, and length of the questions are considered.

### Printing

Test papers possess clarity in printing. Since the paper is double-folded, it may be uncomfortable for the examinees to stretch in a limited space and turn back and forth.

### **Quality of Paper**

The quality of paper is not very good. In some of the pages, the facts and figures look opaque. The physical aspects of the question papers are ordinary. Since the paper is thin and opaque, the illustrations disturb the visibility of the text on the opposite side. It would be better if thicker paper were used to make question papers more distinct and clear.

### **7.4 Marking Schemes**

Marking scheme is based on the nature of the question papers as also on the specification grid which provides specific criteria for setting questions. Two sets of marking schemes reviewed and are found to be good and enough to deliver the message. However, a number of problems such as lack of clarity and poor sentence formulation were noted.

### **7.5 Response Patterns of the Examinees**

Our analysis of the responses revealed the following

#### **Difficulty in expression:**

Most examinees were found to have used Nepali as their medium of writing. These students showed problems in using the correct language. Altogether 138 (45%) of the examinees had expression problem indicating expression practice in correct language is essential at the school level. Most of the answers looked unsatisfactory due to the lack of study, proper understanding of the questions, and confusion created due to nervousness. Thus, understandability, flow of language, consistency, formation of sentences, and correctness of language were the areas associated with difficulty in expression.

#### **Questions not attempted**

There were altogether 9 major questions with 33 sub-questions. Students seem to have difficulties with the items related to geographical studies. Out of 309 students 52 percent attempted questions related to geographical studies, of which 114 (71%) left one or another sub-question unanswered. Similarly, 80 percent of the students attempted questions related to history only partially. The same problem appeared in the area of international understanding, peace, and cooperation. It is obvious from the response patterns of the examinees that questions need to be made more specific so that the examinees could understand them clearly. Furthermore, it is clear that the students were not well prepared in the subject matter.

#### **Questions Repeated**

Cases of repetitions were observed in the answerbooks, here 22 (7%). The major causes may be students' carelessness and poor preparation in the subject matter. The students might have expected marks even if they repeated the questions, expecting they could deceive the examiners. The most repeated questions were found to be 4, 6, and 7. More cases of repetitions were found in geography and history sections.

### Correctness in Writing

The answerbooks of Kathmandu, Dang and Dhankuta maintained standard to some extent. The students in the other districts showed low levels of correctness in writing. The examinees from Pyuthan, Rukum, and Siraha districts, appeared very poor in this aspect.

### Question Numbers Mentioned Wrongly

Some of the examinees were found to write the numbers of some questions wrongly. The frequency of such cases was 84 (27%). In a few cases, the number was given in both Nepali and English, creating confusion. Some were written in English while others were in Nepali. Non of the answerbooks from Dhankuta and Sindhuli showed such errors. The reason for writing the wrong number may be confusion in using Nepali and English numerals in the answerbooks or just nervousness.

## 7.6 Consistency in Marking

The range of differences marks is displayed in Table 9 shows the range of marks among the examiners varies from 1 (lowest) to 19 (highest). The number of cases under the group with a discrepancy of up to 5 marks was 16 percent higher in the case of Markers 2 and 3 compared to Markers 1, 2, and 3. In the case of 6 to 10 range, the number was found to be higher among the three markers (38%) compared to markers 2 and 3 (27%). The number of cases decreased as the range of discrepancy increased. Although there was no greater difference in the marks awarded, discrepancy was observed in most of the cases attributable to subjectivity and lack of seriousness

**Table 9. Discrepancy in Marks**

Range	M1, M2, and M3		M2 and M3	
1-05	148	(48%)	198	(64%)
6- 10	118	(38%)	82	(27%)
11-15	38	(12%)	27	(9%)
6-20	5	(2%)	2	(1%)
Total	309	(100%)	309	(100%)

of examiners in following the marking schemes, as mentioned in an earlier part of this analysis.

Overall, student performance has remained poor, as represented by the average marks (44). The possible reasons for such poor performance of students and discrepancy among the markers are: lack of specificity of the test papers, lack of clarity in marking

schemes, lack of uniformity in the use of terms, poor translation from Nepali to English, and lack of required understanding on the part of students.

## 7.7 Conclusions and Recommendations

The analysis leads to the following recommendations for further improvement:

- Social studies is an integrated subject comprising various aspects of human life, but teachers are oriented only on single subjects such as history, geography, economics, and political science making it difficult to find teachers who can handle the social studies curriculum properly and adequately. This should be addressed through pre-service and in-service teacher education courses.
- Test materials need further improvement. The language and key terms used in the test papers should not be made ambiguous. Questions should be made more specific so that the examinees know exactly what has been expected. Special care should be taken during the process of moderation and editing of the test papers to make it error free.

## 8. MAJOR CONCLUSIONS AND RECOMMENDATIONS

This section summarizes the major findings and conclusions of the study 'Analysis of the Technical Quality of Test Materials used in SLC' and offers a set of recommendations.

### 8.1 Curriculum and Textbooks

No major flaws were reported in the curriculum materials. However, the analysts have noticed a number of problems. First, the use of mechanical, reproductive, and market-oriented contents and techniques instead of materials likely to contribute to the cultivation of human values and broad spectrum of human civilization in English. The textbooks contain materials that are 'artificial' rather than genuine, original, and authentic. Second, the texts and activities chosen in Nepali are centered mainly on reading and writing giving least priority to aural-oral skills. The curriculum and textbooks do not favor the non-Nepali speaking students because of their over-emphasis on the literary aspect of learning rather than language skills. Third, in Mathematics, Science, and HPE curricula are content-heavy. The provision for testing of both Grade IX and X materials in SLC add heavy burden on students who are already strained due to the contents-driven curricula and textbooks. Fourth, there is sufficient overlapping in contents across the different subjects. The following actions are therefore suggested:

- As in English, a provision should be made to introduce of practical test in Nepal. Similarly, the Math and Social Studies curricula should emphasize practical orientation.
- The overlapping of contents from one subject to another should be minimized through joint curriculum workshops of subject experts.
- More authentic, original, and genuine materials should be included in the textbooks.
- SLC should only include texts set aside for Grade X. The existing provision of testing both Grade IX and X materials has added unnecessary burden, leading to high failure rates. This is consistent with many other education systems that focus on the achievement of single year in their public examinations.

### 8.2 Specification Grid

The specification grid that provides a plan for testing the learning achievement of children is believed to promote the quality of test papers in terms of maintaining curriculum coverage, testing students' higher-order and lower-order abilities, and developing valid and reliable test items and serves as a guide to the test developers that has been in use since 2001. The grids of six core subjects were reviewed by the experts. The review suggested that the grids are consistent with the curriculum – one essential feature of the grid. Since what is tested in SLC is determined by the grid, which eventually determines what will be taught in schools, the OCE must ensure that the grid is technically sound.

### 8.3 Test Papers

In general, the SLC test papers have more or less followed the direction given in the specification grids. In some subjects, however, a large variation was recorded between the specification grid and the test papers (e.g., Science). The test papers used in SLC suffer from a number of deficiencies. First, the parallel sets of test papers were not comparable in terms of their difficulty level. Most importantly, the multiple sets of papers were not serving any real

purpose. Second, tests were poorly constructed in terms of their language, structure, accuracy, clarity, and purpose. Third, test items demanded memorization of texts rather than problem-solving and creative, analytical, and critical thinking on the part of students. There was evidence of test developers not following the grid. Fourth, there was also a tendency to repeat question items of the previous years. Fifth, in Science and Math there were simply too many questions and there was no match between the tasks and the time allotted. Sixth, in some cases, the two versions of test papers (English and Nepali) did not match due to the poor quality of translation work done. Finally, there were also problems with the physical quality of the papers. The actions suggested to improve the quality of test papers are:

- The tests developers and moderators should take their work seriously about making the tests error free, unambiguous, impartial, and judicious. They should be made accountable for any errors committed by them.
- The test developers should follow the specification grids while developing the tests.
- The test developers should be aware of using a simple language to keep the questions free from ambiguity and confusion.
- Short-term workshops and training programs on test development and moderation should be organized within certain intervals to refresh the old test developers and prepare the new ones.
- A question bank should be established at the OCE to store quality questions in each subject.
- Parallel sets of tests instead of multi or alternate sets should be developed.
- The test paper should be produced in the form of a booklet with two or three stitches so that it is easy to handle and use.
- Better quality 80 gm map litho paper should be used for printing test papers. For a better result, color printing with different letter faces would be worth while and welcome from the students.
- Twelve (12)-point letter fonts instead of 10 in practice, as at present, should be used for printing test papers that are visible even in rooms with poor light.
- Instructions should be printed in bold faces and the number of questions to be attempted should be put in italics or vice versa to attract students' attention towards what they are supposed to do.
- Too many and too lengthy questions combined with inadequate time do not permit students to perform well. Therefore, the time given should be commensurate with the magnitude of work assigned to the examinees.
- Many of the problems reported earlier result from the lack of standardization of test items. The OCE should take actions towards the standardization of test items.

## 8.4 Marking Schemes

Marking schemes have helped in many ways to ensure consistency in marking. However, in some cases the marking schemes themselves were responsible for inconsistency in marking.

Vagueness, lack of clarity, poor structure, choice of confusing words and terminologies, etc. were observed in the marking schemes. In some cases, the schemes suggested the desirable response, while in others no such suggestion was found. In such context, the OCE must consider improving the quality of the marking schemes.

### 8.5 Response Patterns of the Examinees

- Analysis of the response patterns of students shows that some examinees have inadequate language ability resulting in difficulty in expression. The students coming from Non-Nepali speaking background suffer from inadequacy of language. They are face difficulty in answering questions that demand long answers and creative or imaginative responses. One's language ability is found to determine the chances of obtaining high marks in almost every subject. Non-Nepali students will benefit if measures are taken to help them improve their language ability in Nepali. The Government should consider teaching Nepali as a second language to the non-Nepali speaking student population.
- Students' responses lack originality and creativity. It is not surprising given the fact that the questions used in SLC themselves call for reproduction of texts. At times, those who try to demonstrate their originality and creativity are penalized. There was one case in Science where the student's creative response did not yield the mark the student deserved. Examiners should be instructed clearly to value students' original and creative thinking, a message that should get across the education system.
- The study shows that quite often questions in almost every subject are either not attempted or only partially attempted. Such a situation arises due to a number of reasons such as poor and inadequate teaching, content overload, lack of preparation on the part of students, etc. The best bet would be to improve teaching and giving opportunity to the students to learn .
- It was common among the question-setters to use question items used in the previous years. Such repetition promotes selective teaching. Teachers and students leave out certain portions of the curriculum and over-emphasize others that are likely to repeat. It results in inadequate coverage of the learning outcome. The incidence of repetition can be avoided through adequate and careful moderation. A provision for item-banks can also be made a priority.
- Students do not seem to be familiar with the norms and expectations of the examinations. Attempting the same question more than once, writing answer on the back of the cover page of the answerbook, failing to mention the question number or mentioning the wrong number, inability to follow the given instructions, treating long-answer questions as short-answer ones and vice versa, inability to write legibly are a few examples. Schools should prepare their students adequately for the examinations.

### 8.6 Consistency in Marking

Re-examination of a sample of SLC answerbooks revealed inconsistency in marking which is true for all the six subjects. While the marking scheme has played a major role in ensuring uniformity in marking, there were cases where these schemes were not followed. The lack of uniformity may have resulted from a number of reasons: vagueness in marking scheme, failure

of the examiners to consult the marking scheme, carelessness of the examiner, poor formulation of questions, overloading the examiner with too many answerbooks, lack of seriousness on the part of examiners, poor supervision by the head examiner, unclear instructions, lack of standardization of terminologies, etc. These problems have to be addressed to ensure consistency in marking.

### **8.7 Theoretical vs. Practical Examinations**

In three subjects, the practical aspects of learning are also assessed - English, Science, and HPE. While English (speaking and listening) is tested externally, marks in HPE and Science are given by the schools themselves. In English, the oral test conducted in the examination centers is just a ritual. Tests lack objectivity and seriousness from both sides— examiners and examinees. Students appear in the test without adequate practice. In the case of Science and HPE, there are unacceptable differences between the marks obtained by the students in the theory and practical aspects: the average in the practicals stand for higher than the scores obtained in theory papers. A student securing about 80% marks in practical and 30 to 40% in theory raises questions and doubts. Often, students are not engaged in practical activities in schools. Here two measures can be suggested. First, the MOES must ensure that schools engage students in practical activities as envisaged in the curriculum through proper monitoring and supervision. Second, some kind of adjustment in the internal marks would be necessary to reduce the disparity between the theory and practical marks.