CHAPTER – ONE
INTRODUCTION

1.1   General Background

Language is a voluntary vocal system of human communication. It is the most widely used means of communication for human beings. We express our feelings, thoughts, emotions, desires and experiences through language. Through language we store knowledge and transmit message, and experiences from one person to another and one generation to another. Language can also be taken as a social phenomenon which is used in our society to establish the relationship among the human beings. Through the interaction among the members of community by means of language a society is alive. Language is not just the juxtaposition of words or noises of sounds but it is a systematic arrangement of language components. Richards et al. (1999:196) define "language as 'The system' of human communication which consists of the structured arrangement of sounds or written representation in to larger units, e.g. morphemes, words, sentences, utterances. Language has enabled man to establish great civilizations on this earth".

There are nonlinguistic symbols which are used for communication, yet they are not known as flexible, comprehensive, extensive and perfect as language is. Language is 'species-specific' and species –uniform possession of human being. It is common to all human beings. Man differs from the other species because he / she possesses unique faculty of species. Mankind is the only species gifted with the power of speech. Human beings perceive the whole picture of world in terms of language. We pray and think, read and write, mediate and dream in terms of language. It is the universal medium for conveying the

common facts. No language is superior or inferior to other language in terms of community values. The major function of language is to communicate.

It is used to send and receive the message. We use language for various purposes. It is language which expresses past, present and future. Various linguists and scholars have cited the definition of language. According to Robinson (1965: 14) "Languages are symbol systems. … based on pure/ arbitrary convention … infinitely extendable modifiable according to the need of the speakers". Similarly, Wardhaugh (1972:3) says language is a system of arbitrary vocal symbols used for human communication".

In the same way Chomsky (1957:13) says "A language to be a set of (finite and infinite) sentences, each infinite in length and constructed out of finite set of elements."

From the above definition, we can say that language is universal medium for conveying the facts including complex thoughts, emotions and feelings of everyday life.

1.2   Importance of the English Language

Education is the basic need of human beings and one of the main characteristics to be different from other living beings. It broadens our capacities and makes us a sensible creature. It is the main gate through which we can see the modern civilization of the world.

There are more than 6000 languages spoken in the world. Among them no language can be thought to be superior or inferior to other language in terms of communicative values. However, some languages

have more dominant role in a particular situation. For example, English is the most widely used language in the world and it is accepted as a lingua franca. It has made international communication possible. It is the most dominant language of the world. English vocabulary is about two million words and it is claimed that one in every seven people can speak English. Pahuj (1995) states, "if we look at the media we find that over 50% of the world's newspapers, over 50% of the world's scientific and technical periodicals and more than 60% of the radio stations use English as a medium of communication" as quoted in Dawadi, 2004. Undoubtedly, it is the passport to travel the whole world. It is a global language which at present is most widely taught as a foreign language over 100 countries. So it is principal language of the international communication. Speaking in English is a fashion, a necessity and sometimes also a passport to success in ones career. In hotels, hospitals, airlines, courts, public school, scientific, technical management institutes and collage, speaking English is much in fashion. It is necessary for the students who are supposed to study abroad. It is also one of the official languages offered by the UN.

1.3 ELT in Nepal

English has got the standard of international language because it is used in the international affairs, like formal meeting, seminars, trades, politics, communications, science and technology. So in our context also the importance of the English language has been realized.

In the context of Nepal, the English language has been taught in schools since 1853 with the establishment of Durbar High School by the first Rana Prime-minister Janga Bahaduar Rana after his short visit to Britain. He was impressed with the English language and therefore, he introduced English in Nepal. But the nature of course, teaching and

testing system is still unknown publicly. After the establishment of Tri-Chanda collage (1918), English was formally taught as a compulsory subject at the higher level. Since the implementation of the NESP in 1971, the English language was introduced from class four as a compulsory subject up to Bachelor level and in 2003, from class one as well. In the beginning literature was taught as a subject matter. Reading and writing were the language skills which were given focus in testing language skills. English literature was thought to be an end for its own sake rather than being a means of communication. The medium of instruction and testing was English. Two compulsory papers of English were taught and tested. This course was taught until 2010 (Pandey, 1999:6).

More than five decades before education was extremely limited to the people of higher class. No opportunities were given to the unprivileged people. The political changes of 1951 opened up a favorable condition in this regard and created a new era of development and extreme demands for educational activities.

The most notable affairs in these changes were the curriculum changes proposed be Nepal National Education Commission (1956), All Round National Committee 1961, Natinal Education system Plan 1971, Curriculum Implementation Plan 1981 and the latest one National Education Commission 1992 and High Level National Education Commission 1998. The stated commissions had significant contribution in the development of the present curriculum. English curriculum has also undergone changes and reforms in the light of changing needs of the world accordingly.

After the establishment of democracy in 1950 A.D. some reforms could be seem in the field of education. Despite, several researches regarding education and several efforts made for the development. But there was not seen any drastic change in English education. Then after the restoration of democracy in 1989 A.D. several efforts have been made in the field of education. Furthermore, the National Education Commission (NEC) 1992 report and several meaningful efforts have laid great emphasis on introducing English as a compulsory subject in all schools of Nepal from the very beginning of school Education.

1.4 Teaching and Testing

Testing in a broad sense has always been an inherent part of teaching. Assessment of learning is as old as education itself. From the time when teaching began, the teacher has always been keen to know the extent to which his teaching has been effective in making the learner understand what has been taught. Testing is used as a process of scrutinizing how far learners have learned what the teacher wishes them to learn.

"Both testing and teaching are so closely interrelated that it is virtually impossible to work in either field without being constantly concerned with the other" (Heaton 1988:5). Teaching is imparting knowledge and skills to a person, where as testing as measuring the persons ability to retrieve and use that knowledge and skills.

In some situation test is geared to teaching i.e. to motivate the students and to reinforce learning and in some situation teaching is geared to the test i.e. to assess the student's performance in language. With the

help of testing we can find the areas that need more attention and we can do the remedial works in teaching.

Further more, Heaton (1988:6) "Tests may be constructed primarily as devices to reinforcement learning and to motivate the student or primarily as means of assessing the student's performance in the language. A language test which seeks to find out what candidates can do with language provides a focus for purposeful, everyday communication activities. Such a test will have a more useful effect on the learning of particular language then a mechanical test of structure. In the past, even good tests of grammar, translation or language manipulation had a negative even harmful effect on teaching. A good communicative test of language however should have a much more positive effect on learning and teaching and should generally result in improved in habits."

According to Richard et al. (1999:377) "Testing is the use of tests or the study of theory and practice, their use, development and evaluation etc. Test is any procedure for measuring ability, knowledge or performance."

Similarly, Davies (1968:5) says "The good test is an obedient servant since it follows and apes the teaching."

In conclusion, teaching and testing are closely related. Testing is an inseparable part of teaching. In the absence of testing teaching remains lame or the student's performance can't be seen.

1.5   Qualities of a Good Test

The evaluation of student's performance for the purpose of comparison or selection is only one of the functions of the test. The

classroom test belongs to enabling the teachers to teach more effectively and the final examination is concerned with the selection of students or grading them. A language teacher should conduct both types of tests. A good test has the following characteristics;

(a) Validity

Validity is one of the main qualities of good test. Validity in general refers to the appropriateness of a given test. It is a relative term and can be defined as the degree of accuracy of a test. The test must measure what it is supposed to measure and nothing else. Heaton (1988:159) defines validity as, "The validity of test is the extent of which it measure what it is supposed to measure and nothing else." Every test whether it is informal class test or a public formal examination should be as valid as the construct can make or any its component parts as a measure of what it is purposed to measure. A test is said to be valid to the extent that it measures of what it is supposed to measure. It follows that term valid when used to describe a test should be accompanied by the preposition for any test then may be valid for some purpose but not for others.

Harrison (1991:11) defines validity as, "The extent to which the test measures what it is intended to measure". This is the most acceptable definition by a number of scholars so validity is the degree to which the test actually accurately measures what it is supposed to measure.

(b) Reliability

Reliability of test means the quality of being sameness of the result even if it is conducted many times by different examiners to the same group. If the scores of the student is same after testing by different testers in the interval of different time, the test paper seems to be reliable.

There are different ways of measuring reliability. Spilt halves, Kuder Richardson and test retest are the major ways of measuring reliability. Among them test- retest is the most popular method of testing reliability. It is an important quality of a good test.

In this process reliability is measured by applying the same question papers to the same group of the students. The same paper is administered to the same group after a certain time interval. There may be different examiners but the scores must be the same. If the score is the same then the test is called reliable. However, there are some factors that effect the reliability for e.g. time interval, classroom environment etc.

According to Heaton (1988:163) "Test-retest method of reliability is the method of measuring reliability of test is to re administer the same test after a lapse of time. It is assumed that all candidates have been treated in the same way in the interval that they have either all been taught or that none of them have provided that such assumptions (which are frequently hard to justify) can be made, comparison of the two results would then show how reliable the test has proved."

In other words, the same test is administered to the same group of learners twice, keeping the considerable interval between the two administrations. The term 'considerable' refers to not too long because learning or forgetting will have been taken place. After administering the same group of learners, the two sets of scores are compared. Thus, in this method, the reliability of a test is determined by the extent to which a test gives the same results if it is administered at two different times in a considerable interval of time. It is estimated from the coefficient of correlation which is obtained from the two administration of the test. The specific objectives of the test – retest methods are:

) To develop the techniques to remove the factors which affect the reliability.

⟩ To find out the consistency with the sets of tests scores and tests.

⟩ To provide the interesting materials which are needed to the students for information, knowledge and pleasure.

⟩ To be able to allocate the time to the test.

(c) Discrimination

Discrimination is the degree to which a test or an item in a test distinguishes among better and weaker students who take the test. It is important to know how it distinguishes between students at different levels of ability. This quality of test has the capacity to discriminate among the different students and to reflect the differences to the performances of the individual in the group.

(d) Practicability

A test must be practicable. The length of time, the collection of question papers, the reading time of the instruction should be managed well. The stationery materials should be available on time. The paper should be neatly printed or typed.

(e) Economy

Any test to be good one, it should be economical in all senses. That is to say any good test should be as economical as possible in time and money without dissolving the other qualities of a good test. Economy of a test refers to the construction, administration and scoring of test at minimum cost and labor.

(f) Clear Rubric

The instruction of a test must be clear. The instruction should not puzzle the students. The time allocated should be clear to the examinees. What question and how many of them should be solved or left, should be clearly instructed.

(g) Wash Back Effect

Effect of test on teaching and learning is called wash back effect. There are good and bad or positive and negative wash back effects. Any test to be good one, it should have a good wash back effect. A good test should leave influence on teaching and language planning. After the out come is determined by scoring the answer sheets, the result must provide the teachers, course designers and the textbook writers with some guidelines to reform their respective areas.

1.6 Testing Language Skills and Aspects

Language skills refer to listening, speaking, reading and writing. These all language skills are taught in a language class. What a teacher actually teaches in the classroom can be judged only from testing. So testing is the mirror of teaching. The language skills should be tested according to curriculum to find out the student's proficiency. To test listening live or recorded materials can be used and objective type questions are preferable to test student's ability to manipulate language. Interviews, question answer, spelling the word, naming the things, debate, contest and extempore speech can be given to the students to test speaking skill.

According to Heaton (1988:8) "A relatively difficult or complex text should be shorter than a straight forward one. Generally, the length of the reading extract might vary from 50 to 100 word at elementary level, 200 to 300 at the intermediate level and 400 to 600 at the advanced level. The items to be employ for testing reading should be true / false, Multiple choice (both shorter and longer texts), completion type, rearrangement items, cloze items and open-ended questions etc. For testing writing, the test should test the skills related to language use, mechanical skills,

treatment of content, judgment skills and stylistic skills. Punctuation (general) and spelling should be evaluated in elementary level. The basic level writing test should include the items like writing short personal letters, post cards, dairy entry and forms filling etc."

While testing writing all the types of writing exercises are better means to evaluate students writing ability. Controlled writing, free writing, skeleton controlled writing, and cloze test are some of the useful items to test the writing ability of the students.

Heaton (1988:8) further mentions that "it is the test constructor's task to assess the relative importance of these skills at the various levels and to devise an accurate means of measuring the student's success in developing these skills. Several test writers still consider that their purpose can best be achieved if each separate skill can be measured on its own. But it is usually extremely difficult to separate one skill from another for the very division of the four skills is an artificial one and the concept itself constitutes a vast oversimplification of the issues involved in communication."

Some of the test items that could be included in a test of English are presented in the table from Ghimire (2006).

S.N.	Language Skills	Types of items
1	Listening comprehension	Short utterances, dialogues, talk, lectures etc.
2	Speaking	Interviews, Picture description, role play, pair work, group work, problem solving etc.
3	Reading	Questions related to understanding the gist of the text, extracting, specific information (subjective type, objective type question , cloze test etc.
4	Writing	Writing different types of letters according to the level of the pupils, writing reports, memos, messages, instructions and accounts of past events etc.

1.7 Review of Literature

Review in general refers to the thorough study of previously done work related to the researcher's topic. Review of literature gives an insight to the researchers to conduct his/ her study very effectively. So it is an integral part of the step wise study of something. The researcher's topic is related to the lower secondary teacher's efficiency in constructing the question papers of English for lower secondary level. The researcher's topic is a new one. Only a researcher has conducted a research on the topic that is related to primary level. The researcher studied some of the books related to testing. The researcher studied some of the thesis in the CRC section of Department of Education, Kritipur, Kathmandu.

Khanal (1997) in his study an unpublished M.Ed. thesis on "Effectiveness of Cloze Test against Conventional Test Items" concluded that cloze tests are better tools to evaluate language skills in an integrative way. Therefore, they should be included in the English language tests.

Baral (1999) had done a study on 'Measurement of Readability of Course in General English.' He found the difficulty level of 46 passages in general English and rank them in terms of their readability level.

Shah (2003) studied on the "Effectiveness of Objective Items and Subjective Items for class Ten Students" and found that objective items were better tools to test reading skills of the tenth graders. He recommended that both subjective and objective questions should be asked to test reading skills of the students.

Bhattarai (2004) carried out a study on an "Overview of the Test Materials of SLC English." He found that multiple sets or parallel sets do not serve any specific purpose except for trying to control copying,

cheating etc. As these do not address to the needs of the candidates with special needs, region language or whatever, this practice is costing the board much.

Ghimire (2006) Conducted research on "Ability of Teachers to Construct Test Items." The objective of the study was to find out the ability of primary level English teachers to construct a question paper of English. It was found that neither of the respondents has set the test items for listening and speaking skills however, the primary curriculum has allocated 6 marks for listening and 6 marks for speaking in the annual examination.

Till now no one has conducted research on this topic.

1.8   Objectives of the Study

The objectives of this study were:

(a)   To find out the lower secondary level English teachers' proficiency in setting a good test.

(b)   To analyze the question paper in terms of:

    (I)   Instructional clarity (Rubrics)

    (II)   Reliability

    (III)   Validity

    (IV)   Economy

    (V)   Practicability

    (VI)   Errors

        a. Grammatical Errors

        b. Punctuation

1.9 Significance of the Study

This study is significant because it gives insight to educationist and ELT teachers as to design a good test. The study is significant to the teachers, curriculum designers and other concerned authorities. Examiners and teacher trainers will also benefit from the study.

CHAPTER – TWO
METHODOLOGY

The researcher had adopted the following methodology during the study in order to achieve the objectives.

2.1 Sources of Data

In this process of study the research data were collected from both sources, primary and secondary.

2.1.1 Primary Sources

The primary sources of the data were question papers prepared by the teachers and answer sheets written by the students.

2.1.2 Secondary Sources

The lower secondary level English curriculum, specification grid, model question paper, previously used question papers of English, journals and other theses were taken for the theoretical background of the study.

2.2 Population of the Study

The population for the purpose of the study includes all the lower secondary level English teachers of Kanchanpur District teaching at grade VII and all the lower secondary level students studying at grade VIII in Kanchanpur district.

2.3   Sample Population

For this research the researcher randomly selected 50 lower secondary level English teachers of Kanchanpur District. In the same way, the researcher selected 30 students of grade VIII from Siddhanath Secondary School, Badaipur, Kanchanpur.

2.4   Tools for Data Collections

The test papers prepared by the selected teachers and the responses made by the students were the tools for collecting the data.

2.5   Process of Data Collection

The researcher visited each of the selected schools and selected two teachers from each. After the selection of the teachers she distributed the request forms to them and requested them to design a question paper that covered the reading and writing skills of the present course of English for Grade VII. Then she collected the question papers from them. Then she randomly selected 30 students of grade VIII from Siddhanath Secondary School, Badaipur, Kanchanpur and administered the test paper prepared by the teacher. The same question paper was administered four times in equal interval (i.e. each after ten days). Then she collected all the answer sheets and checked the answers sheets. In this way, she collected the data for the study.

2.6   Limitations of Study

The study had the following limitations:

a) This study was limited to 50 Lower Secondary level English teachers and 30 students of grade VIII of Kanchanpur district.

b) The primary data for this study was collected only from the written test items.

c) Same test was administered for four times within forty days in the same class.

d) Question papers were concerned only with the reading and writing skills.

e) Wash back effect and discrimination were not included in this research work.

CHAPTER – THREE
ANALYSIS AND INTERPRETATION

This chapter is concerned with the analysis and interpretation of the data obtained from the teachers and students. For this research work, the analysis is done in the following way:

  a. Rubrics
  b. Reliability
  c. Validity
  d. Economy
  e. Practicability
  f. Errors

The researcher has used different statistical tools to analyze and interpret the data. She has used mean, standard deviation and reliability to analyze the data.

To analyze the data the researcher has used the following formula:

A. $$r_{xy} = \frac{N\sum xy - \sum x \sum y}{\sqrt{\left(N\sum x^2 - (\sum x)^2\right)\left(N\sum y^2 - (\sum y)^2\right)}}$$

Where,

  r_{xy} = correlation between x and y

  N = Number of given scores

  x and y = Individual scores

B. Correlation coefficient

  $r_{xy} < 0.3$ | poor correlatives

  $0.3 < r_{xy} < 0.7$ | moderate/Satisfactory

  $0.7 < r_{xy}$ | highly correlated

3.1   Rubrics

Most of the respondents were able to give clear rubrics except for some grammatical lapses. The following table shows the mistakes found in rubrics in the sample question papers.

Table No. 1

Types of Errors

S.No.	Types of error	Frequency of error in each paper	%	Remarks
1	Omission Article Preposition	 5 7	 1.54 2.15	Total instructions 324
2	Wrong Spelling	8	2.48	
3	Wrong Lexical item	4	1.23	
4	Addition words / morpheme	6	1.85	
5	Substitution Letters Preposition	 9 10	 2.77 3.86	
6	Punctuation marks	38	11.98	
7	Ambiguity	5	1.54	
Total		85	$X = 85/9$ $=9.44$	

The above table shows all the error areas that were found in the question papers. The types of errors have been categorized into nine categories. The total number of instructions used in all question papers by all the teachers is 324. The highest number of mistakes (11.48) was found

in punctuation marks. The major errors were with capital and small letters full stops, question marks and sometimes with comas.

The teachers have used either full stop or nothing at the end of instruction although they should have used either colons (:) or dash (-).

The second highest error was in the use of prepositions which carries 3.86% of total instructional error. The teachers have used incorrect prepositions for e.g. 'That evening Pasang, Pemba and Sonam slept on their tents near the south summit.' Some other places, the teachers have inserted prepositions.

3.2 Reliability

There are various ways of measuring reliability. Some of the major ways are spilt-halves, Kuder-Richardson and test-retest method. Test-retest method is the most used method that is used to measure reliability. In this reason, test re-test method was used to measure reliability.

To analyze data, the researcher has used the following formula:

A. $\quad r_{xy} = \dfrac{N \sum xy - \sum x \sum y}{\sqrt{\left[N \sum x^2 - (\sum x)^2\right]\left[N \sum y^2 - (\sum y)^2\right]}}$

Where,

rxy = correlation between x and y

N = Number of given scores

x and y = Individual scores

B. Correlation coefficient

$\quad$ rxy < 0.3 | poor correlatives

$\quad$ 0.3 < rxy < 0.7 | moderate/Satisfactory

$\quad$ 0.7 < rxy | highly correlated

Table No. 2

Comparison of Scores between Test A and Test B

S. No.	Name of the students	Test A (X)	Test B (Y)	XY	X2	Y2
1	Dhana Thapa	12	19	228	144	361
2	Dropati Joshi	11	17	187	121	289
3	Ranjeet Rawat	23	15	345	529	225
4	Sarita Mahara	16	14	224	256	196
5	Arjun Bist	17	21	357	289	441
6	Nawa Raj Bhatta	11	9	99	121	81
7	Gita Singh	31	34	1054	961	1156
8	Dhiraj Kunwar	30	25	750	900	625
9	Chet Raj Pant	13	20	260	169	400
10	Yasoda Joshi	14	15	210	196	225
11	Nanda Kala B.K.	21	20	420	441	400
12	Naresh Raj Pant	24	20	480	576	400
13	Abhilash Saud	25	22	550	625	484
14	Meen Bdr.Bohara	25	21	525	625	441
15	Laxmi Thagunna	34	30	1020	1156	900
16	Kamala Thagunna	26	24	624	676	576
17	Kalawati B.K.	8	10	80	64	100
18	Kosh Raj Bhatta	30	24	720	900	576
19	Deepak Mahara	16	20	320	256	400
20	Liladhar Bhatta	17	27	459	289	729
21	Gita Bogati	10	13	130	100	169
22	Naresh Bhatta	17	27	459	289	729
23	Dipesh Nepali	32	34	1088	1024	1156
24	Suman Thapa	20	15	300	400	225
25	Dammara Joshi	32	34	1088	1024	1156
26	Pawan Bhatta	15	16	240	225	256
27	Dibya Raj Joshi	21	29	609	441	841
28	Kamala Lohar	17	18	306	289	324
29	Tilak Nath	21	15	315	441	225
30	Dhana Magar	17	11	187	289	121
		X =606	Y= 619	XY =13634	X2 = 13816	Y2 = 14207

Mean Test   A = 20.2
          B = 20.3
Standard Deviation A = 7.24
          B = 5.66
Reliability = 0.58

The above table shows that the mean of test 'A' is 20.2, test 'B' is 20.3. Standard deviation of test 'A' is 7.24, test 'B' is 5.66 and reliability of test 'A' and 'B' is 0.58

Table No. 3

Comparison of Scores between Test A and Test C

S. No.	Name of the students	Test A (X)	Test C (Y)	XY	X2	Y2
1	Dhana Thapa	12	15	180	144	225
2	Dropati Joshi	11	16	176	121	256
3	Ranjeet Rawat	23	17	391	529	289
4	Sarita Mahara	16	22	352	256	484
5	Arjun Bist	17	29	493	289	841
6	Nawa Raj Bhatta	11	13	143	121	169
7	Gita Singh	31	33	1023	961	1089
8	Dhiraj Kunwar	30	32	960	900	1024
9	Chet Raj Pant	13	17	221	169	289
10	Yasoda Joshi	14	17	238	196	289
11	Nanda Kala B.K.	21	19	399	441	361
12	Naresh Raj Pant	24	17	408	576	289
13	Abhilash Saud	25	21	525	625	441
14	Meen Bdr.Bohara	25	22	550	625	484
15	Laxmi Thagunna	34	28	952	1156	784
16	Kamala Thagunna	26	18	468	676	324
17	Kalawati B.K.	8	9	72	64	81
18	Kosh Raj Bhatta	30	29	870	900	841
19	Deepak Mahara	16	19	304	256	361
20	Liladhar Bhatta	17	22	374	289	484
21	Gita Bogati	10	15	150	100	225
22	Naresh Bhatta	17	22	374	289	484
23	Dipesh Nepali	32	27	864	1024	729
24	Suman Thapa	20	11	220	400	121
25	Dammara Joshi	32	27	864	1024	729
26	Pawan Bhatta	15	19	285	225	361
27	Dibya Raj Joshi	21	20	420	441	400
28	Kamala Lohar	17	20	340	289	400
29	Tilak Nath	21	20	420	441	400
30	Dhana Magar	17	14	238	289	196
		X =606	Y= 610	XY =13274	X2 = 13816	Y2 = 13450

Mean Test   A = 20.2
            C = 20.33
Standard Deviation A = 7.24
            C = 5.90
Reliability = 0.74

The above table shows that the mean of test 'A' is 20.2, test 'c' is 20.33, standard deviation of test 'A' is 7.24, test 'C' is 5.90 and reliability of test 'A' and 'C' is 0.74. This is highly reliable test.

Table No. 4

Comparison of Scores between Test A and Test D

S. No.	Name of the students	Test A (X)	Test D (Y)	XY	X2	Y2
1	Dhana Thapa	12	18	216	144	324
2	Dropati Joshi	11	20	220	121	400
3	Ranjeet Rawat	23	16	368	529	256
4	Sarita Mahara	16	23	368	256	529
5	Arjun Bist	17	25	425	289	625
6	Nawa Raj Bhatta	11	10	110	121	100
7	Gita Singh	31	21	651	961	441
8	Dhiraj Kunwar	30	35	1050	900	1225
9	Chet Raj Pant	13	19	247	169	361
10	Yasoda Joshi	14	20	280	196	400
11	Nanda Kala B.K.	21	22	462	441	484
12	Naresh Raj Pant	24	21	504	576	441
13	Abhilash Saud	25	26	650	625	676
14	Meen Bdr.Bohara	25	27	675	625	729
15	Laxmi Thagunna	34	25	850	1156	625
16	Kamala Thagunna	26	20	520	676	400
17	Kalawati B.K.	8	11	88	64	121
18	Kosh Raj Bhatta	30	27	810	900	729
19	Deepak Mahara	16	21	336	256	441
20	Liladhar Bhatta	17	20	340	289	400
21	Gita Bogati	10	20	200	100	400
22	Naresh Bhatta	17	20	340	289	400
23	Dipesh Nepali	32	30	960	1024	900
24	Suman Thapa	20	21	420	400	441
25	Dammara Joshi	32	30	960	1024	900
26	Pawan Bhatta	15	20	300	225	400
27	Dibya Raj Joshi	21	15	315	441	225
28	Kamala Lohar	17	15	255	289	225
29	Tilak Nath	21	19	399	441	361
30	Dhana Magar	17	19	323	289	361
		X =606	Y= 636	XY =13642	X2 = 13816	Y2 = 14320

Mean Test   A = 20.2

D = 21.2

Standard Deviation A = 7.24

D = 5.28

Reliability = 0.69

The above table shows that the mean of test 'A' is 20.2, test 'D' is 21.2. Standard deviation of test 'A' is 7.24, test 'D' is 5.28 and reliability of test 'A' and 'D' is 0.69.

Table No. 5

Comparison of Scores between Test B and Test C

S. No.	Testes	Test B (X)	Test C (Y)	XY	X2	Y2
1	Dhana Thapa	19	15	285	361	225
2	Dropati Joshi	17	16	272	289	256
3	Ranjeet Rawat	15	17	255	225	289
4	Sarita Mahara	14	22	308	196	484
5	Arjun Bist	21	29	609	441	841
6	Nawa Raj Bhatta	9	13	117	81	169
7	Gita Singh	34	33	1122	1156	1089
8	Dhiraj Kunwar	25	32	800	625	1024
9	Chet Raj Pant	20	17	340	400	289
10	Yasoda Joshi	15	17	255	225	289
11	Nanda Kala B.K.	20	19	380	400	361
12	Naresh Raj Pant	20	17	340	400	289
13	Abhilash Saud	22	21	462	484	441
14	Meen Bdr.Bohara	21	22	462	441	484
15	Laxmi Thagunna	30	28	840	900	784
16	Kamala Thagunna	24	18	432	576	324
17	Kalawati B.K.	10	9	90	100	81
18	Kosh Raj Bhatta	24	29	696	576	841
19	Deepak Mahara	20	19	380	400	361
20	Liladhar Bhatta	27	22	594	729	484
21	Gita Bogati	13	15	195	169	225
22	Naresh Bhatta	27	22	594	729	484
23	Dipesh Nepali	34	27	918	1156	729
24	Suman Thapa	15	11	165	225	121
25	Dammara Joshi	34	27	918	1156	729
26	Pawan Bhatta	16	19	304	256	361
27	Dibya Raj Joshi	19	20	380	361	400
28	Kamala Lohar	18	20	360	324	400
29	Tilak Nath	15	20	300	225	400
30	Dhana Magar	11	14	154	121	196
		X =609	Y= 610	XY =13327	X2 = 13727	Y2 = 13450

Mean Test   B = 20.3
            C = 20.33
Standard Deviation B = 5.66
              C = 5.90
Reliability = 0.29

The above table shows that the mean of test 'B' is 20.3 test 'C' is 20.33. Standard deviation of test 'B' is 5.66 'C' is 5.90 and reliability of test 'B' and 'C' is 0.29.

Table No. 6

Comparison of Scores between Test B and Test D

S. No.	The Name of the students	Test B (X)	Test D (Y)	XY	X2	Y2
1	Dhana Thapa	19	18	342	361	324
2	Dropati Joshi	17	20	340	289	400
3	Ranjeet Rawat	15	16	240	225	256
4	Sarita Mahara	14	23	322	196	529
5	Arjun Bist	21	25	525	441	625
6	Nawa Raj Bhatta	9	10	90	81	100
7	Gita Sing	34	21	714	1156	441
8	Dhiraj Kunwar	25	35	875	625	1225
9	Chet Raj Pant	20	19	380	400	361
10	Yasoda Joshi	15	20	300	225	400
11	Nanda Kala B.K.	20	22	440	400	484
12	Naresh Raj Pant	20	21	420	400	441
13	Abhilash Saud	22	26	572	484	676
14	Meen Bdr.Bohara	21	27	567	441	729
15	Laxmi Thagunna	30	25	750	900	625
16	Kamala Thagunna	24	20	480	576	400
17	Kalawati B.K.	10	11	110	100	121
18	Kosh Raj Bhatta	24	27	648	576	729
19	Deepak Mahara	20	21	420	400	441
20	Liladhar Bhatta	27	20	540	729	400
21	Gita Bogati	13	20	260	169	400
22	Naresh Bhatta	27	20	540	729	400
23	Dipesh Nepali	34	30	1020	1156	900
24	Suman Thapa	15	21	315	225	441
25	Dammara Joshi	34	30	1020	1156	900
26	Pawan Bhatta	16	20	320	256	400
27	Dibya Raj Joshi	29	15	435	841	225
28	Kamala Lohar	18	15	270	324	225
29	Tilak Nath	15	19	285	225	361
30	Dhana Magar	11	19	209	121	361
Total		X = 619	Y = 636	XY= 13749	X2= 14207	Y2= 14320

Mean Test   B = 20.3
          D = 21.2
Standard Deviation B = 5.66
          D = 5.2
Reliability = 0.57

The above table shows that the mean of test 'B' is 20.3, test 'D' is 21.2, standard deviation of test 'B' is 5.66, test 'D' is 5.2 and reliability of test 'B' and 'D' is 0.57.

Table No. 7

Comparison of Scores between Test C and Test D

S.No.	Name of the Students	Test C (X)	Test D (Y)	XY	X2	Y2
1	Dhana Thapa	15	18	270	225	324
2	Dropati Joshi	16	20	320	256	400
3	Ranjeet Rawat	17	16	272	289	256
4	Sarita Mahara	22	23	506	484	529
5	Arjun Bist	29	25	725	841	625
6	Nawa Raj Bhatta	13	10	130	169	100
7	Gita Sing	33	21	693	1089	441
8	Dhiraj Kunwar	32	35	1120	1024	1225
9	Chet Raj Pant	17	19	323	289	361
10	Yasoda Joshi	17	20	340	289	400
11	Nanda Kala B.K.	19	22	418	361	484
12	Naresh Raj Pant	17	21	357	289	441
13	Abhilash Saud	21	26	546	441	676
14	Meen Bdr.Bohara	22	27	594	484	729
15	Laxmi Thagunna	28	25	700	784	625
16	Kamala Thagunna	18	20	360	324	400
17	Kalawati B.K.	9	11	99	81	121
18	Kosh Raj Bhatta	29	27	783	841	729
19	Deepak Mahara	19	21	399	361	441
20	Liladhar Bhatta	22	20	440	484	400
21	Gita Bogati	15	20	300	225	400
22	Naresh Bhatta	22	20	440	484	400
23	Dipesh Nepali	27	30	810	729	900
24	Suman Thapa	11	21	231	121	441
25	Dammara Joshi	27	30	810	729	900
26	Pawan Bhatta	19	20	380	361	400
27	Dibya Raj Joshi	20	15	300	400	225
28	Kamala Lohar	20	15	300	400	225
29	Tilak Nath	20	19	380	400	361
30	Dhana Magar	14	19	266	196	361
Total		X = 610	Y = 636	XY= 13612	X2= 13450	Y2= 14320

Mean Test   C = 20.33
          D = 21.2
Standard Deviation C = 5.90
          D = 5.28
Reliability = 0.72

The above table shows that the mean of test 'C' is 20.33, test 'D' is

21.2. Standard deviation of test 'C' is 5.90, test 'D' is 5.28 and reliability

of test 'C' and 'D' is 0.72.

The above tables show that the test 'A' and test 'C' are highly reliable. Its reliability is 0.74. The Reliability of test 'A' and 'B' is 0.58, test 'A' and test 'B' is 0.69, and test 'B' and test 'C' is 0.29. Similarly, the reliability of test 'B' and test 'D' is 0.57 and test 'C' and test 'D' is 0.72.

3.3 Validity

While analyzing the test papers prepared by lower secondary level English teachers, it can be said that the test papers are not valid. The teachers were asked to make test papers only on reading and writing skills. Majority of the teachers did not include the items mentioned in the curriculum of grade seven. Most of the test papers had not covered the objectives set in the curriculum of grade Seven. Some of the teachers asked "Reading Passage" for testing writing skill, which is not given in the grid. This shows that the teachers are not familiar with the curriculum.

Table No. 8

Types of Questions Asked by Teachers

Mode	Types of Question	No. of teachers	% of teachers
Objections	Fill in the gapes	15	30%
	Matching Items	9	18%
	Multiple choice	5	10%
	Ordering / Rearranging	20	40%
	True/ False	15	30%
	Vocabulary Testing	13	26%
Subjections	Knowledge testing	26	52%
	Free writing	18	36%
	Guided writing	13	26%
	Describing picture	4	8%
	Sentence making	40	80%
	Punctuation	12	24%
	Grammar	7	14%

The above table shows that most of the teachers were not familiar with the specification grid. Though they were asked to prepare question papers only for reading and writing skills. Some of them set the question papers that included grammatical items as well. Fourteen percent of the total number of teachers used grammatical items in their question papers. It shows that they prepared question papers without having knowledge of specification grid. Only 24% teachers used punctuation item but a question that carries 4 marks is compulsory according to the grid. Similarly, there must be two questions for free and guided writing, one for each item. Very few teachers (36%) asked questions from free and guided writing. The other fact that can be understood from the table is that a few (25%) of teachers asked questions on vocabulary.

Table No. 9

Numbers of Questions and Average Mark Allocation.

No. Of Questions in each papers(X)	No. of teachers (F)	F X	Mark Allocation
2	6	2X6 =12	50%
3	4	3X4 = 12	33.33%
4	7	4X7 = 28	25%
5	22	5X 22 = 110	20%
6	5	6X5 = 30	16.66%
7	2	7X2 = 14	14.28%
8	1	8 X1 = 8	12.5%
10	1	10 X 1 = 10	10 %
Total N = 50		**FX =324** $X = FX/N$ $= 324/50$ $= 6.48$	

The above table shows that the number of questions asked in the final examination differs from teacher to teacher. Some teachers have included less than 3 questions and some others have included 10 questions. The specification grid mentions four questions for testing reading and four questions for testing writing skill. This shows that the teachers did not use specification grid to allocate marks for the questions.

3.4   Economy

Economy of the test refers to the construction, administration and scoring of the test at minimum cost and labor. For test to be good one, it

should be economical in all senses. The question paper which was administered to the students was economical. There were two passages; they were short and most of the students finished them in proper time. From the view point of examiner, it did not take too much time to construct and score the question paper.

3.5 Practicability

The length of time, the collection of question paper, the reading time of instruction should be well managed. The question paper was easy for the students to understand. Some of the students had problem because they had not brought stationery materials. Later then they could solve the problem. The question paper was well printed.

3.6 Errors

a. Grammatical Error

Grammar is the basics of language, which gives certain shape and meaning of linguistic unit. Without grammar the intention of a sentence or word can not be understood. To analyze the question papers of the teachers, 6 different areas were selected. The areas were use of articles, prepositions, number, S-V agreement, conjunction and selection of lexical items. The grammatical categories were studied on omission, substitution and insertion. The grammatical errors found in the respondents question papers are listed below:

Table No. 10

Grammatical Errors

S.N.	Grammatical areas	Omission		Substitution		Insertion		Total	
1	Article	6	37.5%	3	18%	7	43.75%	16	21.85%
1	Preposition	10	29.41%	19	55.88%	5	14.70%	34	44.73%
2	Number	2	40%	1	20%	2	40%	5	6.57%
4	S-V agreement	13						13	17.10%
5	Selection of lexical item	3	60%	1	20%	1	20%	5	6.57%
6	conjunction	3						3	3.94%
Total		37	48.68%	21	27.3%	15	19.73%	76	

The above table shows that the majority of errors (44.74) are committed under preposition. The second highest category of error (21.45) is in articles. Similarly, the least error category is conjunction. Only 3.94% of the total errors are committed in the use of conjunction.

Similarly, errors were studied under three categories viz., omission, substitution and insertion. The highest number of errors are committed due to the omission (48.68%). The second highest error category is substitution. 27.3% errors are committed due to substitution. The least error category is insertion. .

b. Punctuation

Punctuation in general refers to the act of using punctuation marks accurately. The punctuation marks play a crucial and meaningful role in

conveying a message in a piece of a text. Using full stop, commas, question marks, colon, capital letters and small letters in proper place is the art of punctuation. The same piece of writing conveys different messages due to the use of different signs. The signs that were found errors in question papers are presented in the table below:

Table No. 11

Types of Errors

Punctuation Title	Omission	|Substitution	Insertion	Total	%
Question Mark (?)	4	3	2	9	2.77
Full Stops (.)	15	2	-	17	5.24
Capital Letter		8		8	2.46
Small letter		12		12	3.70
Spelling	7	-	-	7	2.16
Colon	23	94	-	117	38.58
Commas	5	-	3	8	2.24
Total	54	111	5	178	
	31:76%	65.29%	2.94%		

As the above table shows, the teachers committed the highest number of errors in the use of colon. It represents 38.50% of total errors in setting the question papers. The teachers have mostly used either full stops or left empty instead of colon (:) in giving instruction.

The types of errors were studied under three categories viz, omission, substitution and insertion. The highest number of error was committed due to substitution of different elements for example, capital letters were used instead of small letters or vice – versa, one punctuation mark is used instead of another one. Even in the use of spelling errors were found. For e.g. 'hornbills' was spelt as 'honbill.' It is the least error area which carries only 2.16% of the total errors.

CHAPTER – FOUR
FINDINGS AND RECOMMENDATIONS

4.1   Findings

After the tabulation, interpretation and analysis of the data, the following facts were found from the study:

1.   The teachers were proficient to set a good test.

2.   Regarding the test papers prepared by them the following points can be made:

(I)   The question papers lack clear rubrics. There were lots of grammatical lapses.

(II)   Most of the test papers was reliable. The test scores among the four test items seem highly reliable. The reliability of test 'A' and 'B' is 0.58, 'A' and 'C' is 0.74, 'A' and 'D' is 0.69, 'B' and 'C' is 0.29 (poor), 'B' and 'D' is 0.57 and 'C' and 'D' is 0.72.

(III)   36% of the test papers lack validity. Some of the teachers asked an extra question which is not mentioned in the grid and some others left the item which is mentioned in the grid. 'Reading Passage' were asked for writing skill. 14% of teachers asked grammatical item in writing skill which is not given in the grid.

(IV)   The question paper was economical as it was neither too long nor too short.

(V)   The question paper was practical in the sense that there was no problem to administer it to the students and to collect answer sheets. Most of the students had finished most of the questions on time.

(VI)   Lots of grammatical and punctuation errors were found in test papers. The major problems were with selections of wrong lexical items and subject verb agreement, wrong use of colon, semi colons, and commas and even in small and capital letters.

4.2   Recommendations

(1)   The study shows that the ability of teachers to set a test paper was satisfactory. So the teachers are recommended to follow the same trend while designing the test papers.

(2)   The test papers lack clear rubrics. Majority of the teachers have used either full stop or nothing at the end of instruction. Therefore, they are recommended to use either colon (:) or des (-) at the end of the instruction.

(3)   The test paper was reliable. The score differences between two tests seem satisfactory. All the teacher should try to follow this while designing a test paper.

(4)   The test papers lack validity. Majority of the teachers did not include items mentioned in the curriculum of grade VII and some of the teachers asked extra items so they are strongly recommended to follow the curriculum and specification grid of grade VII while designing question papers.

(5)   The test paper was practical and economical. It was neither too long nor too short. So teachers are recommended to follow this trend while designing a test paper.

(6)   The test papers had lots of errors in the use of grammar and punctuation. The teachers should make question paper error free.

REFERENCES

Baral, C. 2004. *Measurement of Readability of Course in General English*. An Unpublished M.Ed. Thesis; Kathmandu: T.U.

Bhattarai, A. 2001. *A Brief Survey of School Level English Language Teaching Situation in Nepal*. Ratna Pustak Bhandar, Nepal.

Bhattarai, G.R. 2004. *An Overview of the test Materials of S.L.C. English*. Journal of NELTA Vol. 9 No.1-2.

Chomsky, N. 1957. *Syntactic Structures*. The Hange: Mounton.

Davies, A. (ed.) 1968. *Language Testing Symposium: A Psycholinguistic Approach*. Oxford: OUP.

Dawadi, Saraswati. 2004. *Analysis of Grade Seven English Textbook*. An Unpublished M.Ed. Thesis; Kathmandu: T.U.

Ghimire, N.R. 2006. *Ability of Teachers to Construct Test Item*. An unpublished M.Ed. Thesis; Kathmandu: T.U.

Harrison, A. 1983. *A Language Testing Handbook*. London, Modern English Publication.

Harrison, A. 1991. *A Language Testing Handbook*. Longman: ELTS.

Heaton, J.B. 1988, *Writing English Language Tests*. Longman: English Language Book Society.

Khanal, K.P. 1997. *Effectiveness of Cloze Test Against Conventional Test Items*. Unpublished M. Ed. Thesis; Kathmandu: T.U.

Khaniya, T.R. 2005. *Examination for Enhanced Learning*. Kathmandu: Millennium.

Kumar R. *Research Methodology*. London, SAGE Publication Ltd.

Lower Secondary Level Curriculum 2003. Bhaktapur; Curriculum Development Center.

Nunan, D. 1993. *Research Method in Language Learning*. Cambridge CUP

Richards et al. 1995. *Longman Dictionary of Applied Linguistics*. Essex: Longman.

Robinson, W.P. 1997. *Language and social Behaviour*. Harmonds Wort: Penguin.

Shah, Rajendra 2003. *Effectiveness of Objective Items and Subjective Item for Grade 10*. An Unpublished M.Ed. Thesis; Kathmandu: T.U.

Specification Grid of Lower Secondary Level 2004. Bhaktapur CDC.

Wardhaugh, R. 2000. *An Introduction to Sociolinguistics*. New York: Basil Black Well.

Appendix - I

REQUEST FORM

School:

Teacher's Name:

Dear Sir/ Madam:

     You are requested to prepare a question paper of English for grade seven. Please, follow the following instruction.

1. See the specification grid carefully that I have given to you.
2. Include only reading skill and aspects mentioned in the grid.
3. Balance the mark allocation and questions according to grid.
4. Write instructions clearly and accurately.

REQUEST FORM

School:

Teacher's Name:

Dear Sir/ Madam:

You are requested to prepare a question paper of English for grade seven. Please, follow the following instruction.

1. See the specification grid carefully that I have given to you.
2. Include only writing skill and aspects mentioned in the grid.
3. Balance the mark allocation and questions according to grid.
4. Write instructions clearly and accurately.