

# CHAPTER ONE

## INTRODUCTION

### 1.1. GENERAL BACKGROUND

Language is defined as a means of human communication through which we express our feelings, thoughts, wants, desires and so on. Richards, et al. (1995 : 2) define language as, “The System of human communication which consists of the structured arrangement of sounds into larger units, e.g. morphemes, words, sentences, utterances.”

Regarding the English Language, it is an international language. According to Otto Jespersen states, “Language is not an end in itself ... it is a way of connection between souls, a means of communication (1904).” There is no doubt that English is widely spoken by majority of the world population as a means of communication. Most of the books of the world are written in English. It has gained the status of being the most dominant language in almost all the areas like Commerce, Trade, Mass Media, International Diplomacy, Science and Technology.

As we know the language is species-specific to mankind i.e. only human beings can have the capability to speak language and their mind is genetically equipped with it. Language consists of the vocal noises made by human beings. Vocal sound such as sneezing, coughing, snoring etc. can communicate some meanings but cannot be considered as language because these sounds are not deliberately used for the purpose of communication. These are involuntary sounds. Thus, "Language as the voluntary vocal system of human communications" can be considered as the more adequate definition.

In the context of Nepal, the English language entered here in 1910 B.S. with the establishment of Durbar High School. Since then, it has been taught in schools and colleges as a foreign language. This is

taught as a compulsory subject from class three to Bachelor's level. By the established of the Durbar School in 1983 A.D. formal teaching and learning of English curriculum was developed. At University level most of the books and reference books are found in the English medium.

The importance of the English language in the present day world not to be overemphasized. It is a principal language for international communication and a gateway to the world body of knowledge. In view of all these facts the English language is given great importance in the education system of Nepal.

### **1.1.1. Introduction of a test and its types**

A test is defined as a means to measure the learner's knowledge. Tests are either predictive or indicative of achievement of current level of knowledge. A test is a process of evaluating the ability or knowledge of a learner. It is a measuring instrument which is applied to learners, not to teaching materials or teachers. It is a device used to show what a learner knows. Tests are such devices to reinforce learning and to motivate students or to assess their performance. When we read a book, novel, listen to the news, music on T.V. or prepare a meal, we are testing hypothesis and making judgments.

Language learners are testing their newly acquired forms of language almost every time they speak.

A test, in ordinary words, is a method of measuring a person's ability or knowledge in a given area. The definition captures the essential components of a test. A test samples performance but it also infers certain competence. A driving tests for a driver's license is a test requiring a sample of performance.

Generally, the test and exam are taken as synonymous in the 'testing literature' (Davies, 1977 : 49). "A test may be defined as an activity the main purpose of which is to convey how well the testee knows or can do something. This is contrast to practice, the main

purpose of which is sheer learning. Any procedure for measuring ability, knowledge and performance is called test,” (Richards et al.). Harrison (1983:1) says, “It is a natural extension of the classroom work providing teaching, teachers and students that can serve as a basic for improvement.” A test is a set of questions or problems for determining a person’s knowledge or ability. It is an attempt to see whether the things taught have been learned.

A test is defined as a standardized procedure to measure quantitatively or qualitatively one or more than one aspect of language by means of a sample of verbal or non-verbal behaviours. A language test samples language behaviour and infers general ability in language. A test of reading comprehension, may consist of some questions following one or two paragraphs, a tiny sample of second language learners and total reading behaviour. By the result of that test the examiner infers a certain level of general reading ability of the learner in question.

Thus, a test is a measuring device, which we use when we want to compare an individual with other individuals who belong to the same group.

There are different types of tests. Mainly, they are of five types which are given as follows: Proficiency test, achievement test, diagnostic test, aptitude test and placement test. These are described as follows:

**(a) Proficiency test**

A Proficiency test is defined as “a measure of how much of language someone has learned.” Davies et al. (1999 : 76). It is designed to measure people’s language ability without considering the course they have been taught and in which situation they have learned. The Proficiency test is a way of measuring the language ability of students in anticipation of some tasks that they are expected to perform. For example, one would ask: Does a candidate know English enough to cope with an English Medium University Course?

Does he has adequate command over English to work as an English Teacher? The tests are based rather on anticipation of the skills and abilities that are needed for the testees to cope with the course of study or the situation that they are going to get involved in a proficient manner. Proficiency test are said to have no relation with any course of instruction. However, Davies (1985 : 76) argues, “the proper accolade for a good proficiency test is that it allows itself to be outdated – it becomes the achievement test for a teaching syllabus...”

The Proficiency test is a forward looking test in the sense that it defines the ability of a student to use a language with reference to a particular task which the learner is expected to perform in the future. This is a test about how much of what he needs to know does a candidate actually know. Proficiency test is based on analysis of a job or a situation or a course of study. The purpose of administering the proficiency test is to investigate whether or not a candidate has adequate proficiency in the language in question in order to cope with the anticipated course or job.

There are certain proficiency tests which are not targeted to anticipated tasks but just at findings out whether certain standards with reference to specified abilities are reached or not. In this case, proficiency is more a general concept. There are several well known proficiency tests such as the tests – TOEFL, IELTS; entrance examinations at different Colleges or Universities, Cambridge Examination (First Certificate Examination and Proficiency Examination), and the Oxford EFL examinations are some examples of the proficiency test.

**(b) Achievement test**

Davies et al. (1999 : 77) describe, "an achievement test as an instrument designed to measure what a person has learned within or up to a given time. It is based on a clear and public indication of the instruction that had been given." Unlike proficiency tests, achievement /attainment tests are directly based on predetermined courses. That means, the achievement test has to measure the extent to which the learners have achieved what they are supposed to achieve in relation to the contents and objectives of the course.

Some well-known examples of achievement tests are: the SLC Examinations, Higher Secondary Examinations, the examinations administered by the office of the controller of Examinations, Tribhuvan University, etc. There are certain presumptions that the students are expected to learn based on a prescribed course of study. They may also reflect how the courses should have been taught. However, they do not necessarily reflect what has actually been taught. The reason is that the achievement test should be based on the course objectives.

Mainly achievement tests are of two types: class progress test and final achievement test. A class progress test is a teacher-designed test to be used to examine the extent to which the students have learned from the instructions presented in the classroom. It is used to see how students are progressing towards achieving the objectives of the course which can take a form of unit test or can be broken into a well-defined short objectives.

Final achievement test is administered at the end of the academic session. This test deals with an estimate of the present achievement of the learner, e.g. what he knows and what his learning ability is. This test serves both purposes – forward looking and backward looking purposes. The relevance of the test tasks depends upon the relevance of the course of study. It is a norm-referenced test in the sense that it shows how a learner has achieved in comparison to others.

**(c) Diagnostic test**

A diagnostic test is like a diagnosis of a medical doctor. As the doctor makes a judgment on an illness after examining the person in order to pursue further treatment, a teacher administers a diagnostic test to ascertain the strengths and weaknesses of the students so as to determine the kind of further remedial action needed for a particular group of students. The main purpose of it is to find out student's strengths and weakness. The another purpose of it is always remedial. There is no use of information unless we improve the performance for better learning.

The main areas that the diagnostic test works better are discrimination tests (pair discrimination), grammar and usage, controlled writing, etc. This test gives both quantitative and qualitative information about the problem. The diagnostic test is good for instruction and self-instruction. Learners themselves can find out where they have problems. As an achievement test tells us how much a learner knows, a diagnostic test tells us what he knows.

**(d) Prognostic test (Aptitude test)**

An aptitude test is defined as a measuring instrument to find out “the extent to which an individual possesses specific language learning abilities.” (Davies et al. 1999 : 82). It is also described as an instrument to assess learning potentiality of the learners. This test is used to find out the probable performance in a foreign language, which is not yet started. This test involves intelligence, motivation, memory and phonological sensitivity. This test generally predicts the students’ probable strengths and weaknesses in learning a foreign language. This is done by measuring performance in an artificial language. That is why it is sometimes called prognostic test.

**(e) Placement test**

As its name suggests, it is useful and intended to provide information which will help to place students at the appropriate stage of the teaching programme. The placement test is administered to assign students to different classes and levels according to their proficiency level. Obviously, it should reflect the features of the teaching context such as the level of the students, methodology, syllabus, etc. The tests may vary from institution to institution.

### **1.1.2. Language Teaching and Language Testing**

Language teaching and language testing both are two sides of a coin. One would be meaningless in the absence of another. "In as much as language teaching and testing concerned, both teaching and testing are so closely interrelated that it is virtually impossible to work in either field without being constantly concerned with other" (Heaton : 1985 : 5).

Language testing includes both linguistics and psychology because it is concerned with language and with learning. Testing is an integral and essential part of education. Education involves both teaching and testing. And, therefore, we can say that there is whole-part relationship between education and testing. Language teaching and testing can function like the combination of a pick and a shovel to dig deep into the language education.

Testing is used as a process of scrutinizing how far learners have learned what the teacher wishes them to learn. Testing is used in the classroom or after the classroom teaching. In many classroom situations, thus, teaching and testing are hardly separated. A large number of examinations in the past have encouraged a tendency to separate testing from teaching. But it has been proved that language teaching can't go as smoothly as possible in the lack of language testing.

The objective of teaching anything is to help the learners to learn it. Teaching, therefore, should be geared to facilitating learning on the part of the learners. This is true of language teaching as well. Hence, the objective of teaching a language is to facilitate learners in learning it. The errand of language teaching activities therefore begins with questions, who are the learners? Why do they learn the language in question?

Testing plays an important role in teaching from the educational point of view, so testing is an essential part of it. One of the functions of an examination is to reflect the course of objectives. Without testing, no teaching will be done in a good way. So testing is as important as teaching. Similarly, Richards et al. (1985 : 377) define, testing as “the use of test or the study of theory and practice their use, development and evaluation. Test is any procedure for measuring ability, knowledge or performance.”

### **1.1.3. Requirements of a Good Test**

Examinations are administered for some purpose. In order to serve the purposes for which exams are conducted, they must be of good quality. In

another sense, the quality of an exam is examined in light of the extent to which it serves the purposes for which it is administered.

Any test to be a good test should have common characteristics. Although the characteristics of test differ from author to author, some common characteristics of good tests are mentioned below. These characteristics should be taken into consideration while writing the tests otherwise it becomes “...just as it is impossible to play chess without knowing how a knight moves across the board, so it is pointless to write tests without a basic understanding of the principles behind them” (Harrison, 1991 : 10).

There are various views on what makes a test as a good test. Some experts say that there are three constituents of exam efficiency, which are: Validity, reliability and practicality. In this respect, validity, reliability and practicality should be seen as relativistic concepts. The whole idea of considering the three constituents of exam efficiency is to build up a framework for designing a good test. Those issues which come under the quality of a good tests are described below:

- (i) Reliability
- (ii) Practicality and
- (iii) Validity

**(i) Reliability**

Reliability is one of the two most important characteristics of a good test, the other being validity. It is a prerequisite for a valid test. That is to say, if a test is not reliable then no talk of its validity. In other words, any test to be valid, it should be reliable but a reliable test is not necessarily a valid test. It is another essential quality of a test which refers to the consistency of scores or performance of the same or similar test administered within a reasonable time. It is a matter of the extent to which we can believe that performance is true, how likely it is that the performance will be repeated next time. Reliability can be dealt with at two levels: test and retest of students and marking and



remarking of the examiners. It is a statistical concept. Reliability is related with examining consistency in the performance of the examinee.

The underlying concept of reliability is that whatever is measured is not only a property of the exam, it is also a property of the examinees' performance. It implies that reliability is not like a permanent quality of an exam, which remains the same, if it is administered again and again, reliability of an exam may change because of variations in performance. It seems that testing conditions can also have an effect on the reliability of an exam.

Basically, there are three methods of determining reliability of the exam. They are test retest method, parallel tests method and internal consistency methods. A test retest method refers to a test that is administered two times within a short time gap. And, in parallel tests method, two similar tests are administered to the students. The assumption is that the students will have similar type of performance on the two tests. The correlation between the two parallel tests will indicate the reliability of the test in question. In case of having difficulty in administering two tests, reliability of a test can be established by splitting the test into two halves. The two halves will be considered two equivalent tests. This method is called "internal consistency method."

## **(ii) Practicality**

Practicality is different from other qualities of a test. Absence of this quality in a test will lead the test to be of no use. It has been shown that reliability and validity are the most important aspects of exam efficiency. Another aspect of it, though non-technical, is practicality in the absence of which even a valid and reliable exam can be of no use.

Basically, the aspect of exam efficiency is important because failing to achieve the practicalities would lead to have problems to convince the authority concerned. If any exam is likely to create some other problems because it did not fit into the given situation, it would be difficult to convince the people who would actually implement the exam.

In order to develop a test with a reasonable degree of practicality, it is necessary to pay attention to the following issues: human resources, materials resources and time. Here, human resource refers to test writers, markers, test administrators and clerical support. Material resource refers to space, equipment and materials etc. And, time refers to the time available for the development, implementation time given for students to perform the given tasks, scoring and analyzing, etc.

### **(iii) Validity**

Validity is one of the main qualities of a good test. Validity refers to the degree to which a test, actually measures what it is designed to measure. It means a test which is designed to measure the function actually measure the same, then it is known as valid test. If a test doesn't measure what it ought to be measure then it is not the valid test, validity is the complex criteria of finding a good test.

Validity is a very important quality of a test. "A measure is valid if it does what it is intended to do..." (Davies et al. 1999 : 94). The validity of a test is measured on the basis of how far the information it provides is accurate, concrete and representative in light of the purposes for which it is administered. Validity is a relative term and can be defined as the degree of a test. The most complex criterion of a good test is validity in the sense that the concept of validity reveals a number of aspects.

To be clear further, we can deal with some concepts of validity which are specified as follows:

Harrison (1991 : 11) defines, validity as "the extent to which the test measures what it is intended to measure."

Heaton (1995 : 153) mentions that "the validity of a test is the extent to which it measures what it is supposed to measure and nothing else." He further views about the terms 'validity' that every test, whether it be as short, informal classroom test or a public examination should be as valid as the constructor can make it. The test must aim to provide a true measure of the particular skill

which it is included to measure to the extent that it measures external knowledge and other skills at the same time: it will not be a valid test.

Lastly, Richards et al. (1993 : 396), define the term validity as, “the degree to which a test measures what it is supposed to measure or can be used successfully for the purpose for which it is intended.”

Validity, as a quality of language test can be divided into the following types:

- (i) Criterion-related Validity
- (ii) Construct Validity
- (iii) Face Validity
- (iv) Washback Validity
- (v) Content Validity

**(i) Criterion-related Validity**

Criterion-related validity is also termed as empirical validity. Criterion-related validity is a form of the validity of a test arrived at by comparing the test with one or more criterion measures. Criterion measure is a standard against which a test can be compared as a measure of its validity.

Criterion-related validity and predictive validity are statistical concepts which are also called “empirical validity”. This type of validity is established employing a process of comparing the results of a test with the results of some criteria already set or the subsequent performance of the students. The validity of a test established by comparing with a set criterion measure is called Criterion-related. It is important that a test must also be empirically validated to ensure that it has elicited the information it was supposed to elicit.

Criterion-related validity can be established by giving the students an established test with similar nature which has proved to be valid. The test can be administered at the same time or in a short gap ensuring that no additional learning opportunity is given. Criterion-related validation procedures determine

the efficacy of an examination in predicting the examinee's future performance in a pre-specified situation (Anastasi 1982 : 137).

Criterion-related validity is divided into two types they are: Concurrent Validity and Predictive Validity. The main difference between the two types of validation procedure is time interval: if the exam scores are validated against the criterion approximately at the same time, it is concurrent validity and if the exam scores are validated against the criterion after a stated time interval, it is Predictive Validity.

By given name, the Concurrent Validity of a test means that the process of determining the validity against the set criterion at the same time. Test developers tend to establish the validity of the new test by comparing the performance of the students on this test against their performance on a test of similar kind already established. The correlation between the two tests is said to be the concurrent validity of the new test.

Predictive Validity of a test can be defined as the degree to which a test can predict candidates' future performance. In other words, any test is said to have its predictive validity if the scores of a group of learners on the test correlate with their scores on a valid test administered at a later time in future. This type of validity is established by comparing test results with another criterion such as success in a particular job or in higher education. Predictive validity is important in the sense that a test is supposed to predict the future performance of the candidates which can be established against the external criteria.

## **(ii) Construct Validity**

Construct validity can be defined as a form of validity which is based on the degree to which the items in a test reflect the essential aspects of the theory on which the test is based. The term 'Construct' refers to any underlying ability which is hypothesized in a theory of language ability. In other words, in language testing, it is defined as the extent to which a test represents an

underlying theory of language learning. A test can be said to have construct validity if it measures just the ability which it is expected to measure and nothing else.

In the literature, construct validity is discussed as the center of all methods of validating a test. Accepted the inclusiveness of the construct validity, it can be argued that if a test provides evidence that it has content and criterion-related validity, the exam also provides some evidence that it has construct validity as well.

The concept of Construct Validity is thus associated with the validation procedures based on systematic examination of the abilities we want to measure and evidence of their measurements. Cronbach (1971 : 443) further says about construct validity: construction of a test itself starts from a theory about behaviour or mental organization derived from prior research that suggests the ground work for the test.

Lastly, Messick (1975 : 103) discusses construct validity in a simple way as, “A measure estimates how much of something an individual displays or possesses. The basic question (of construct validation) is, what is the nature of that something?”

### **(iii) Face Validity**

Generally, any test is said to have its Face Validity if it looks like valid in surface. Face Validity of a test is not sufficient but essential quality of a valid test. In such a test, the test items look right to other testers, teacher and testes in surface. Face Validity is the validity of a test in layman’s eye. Face Validity can be defined as the degree to which a test appears to measure the knowledge or abilities it claims to measure. In other words, Face Validity is defined as “What it appears superficially to measure” (Anastasi 1982 : 136).

The importance of face validity cannot be underestimated in the sense that if the examinee does not consider it a valid exam. Anastasi’s (1982 : 136) argument that “Face Validity should never be regarded as a substitute for objectively determined validity” can be accepted. The concept of face validity

is far from new in language testing but the emphasis now placed on it is relatively new. In the past, Face Validity was regarded by many tests written simply as a public relations exercise.

**(iv) Washback Validity**

Washback is inherent in an exam: an exam is bound to influence teaching and learning. The possible reason for why the exam is influential on education should be seen in terms of its functions. The functions of the exams can be summarized under its forward looking and backward looking purpose though they are not exclusive. The backward looking functions of the exam have to do with a sense of achievement and evaluation of the effectiveness of teaching. The forward looking functions are related with making decisions about the examinees.

Morrow (1986 : 6) argues that one of the principal responsibilities of testing or examining boards and institutions, which operate in the public domain is to provide a “package” with a powerful and positive Washback effect into the classroom. Morrow’s Concern is specifically with the Royal Society of Arts’ Communicative use of English as a Foreign Language (CUEFL) exam, which is a test taken, to all intents and purposes, voluntarily. We all recognize that Washback validity is important – we should be grateful to the Criterion-referenced testing movement for highlighting this point. The real significance of Morrow’s (1986 : 6) remarks about Washback validity lie not so much in their emphasis on validity as such, but on the highlighting of the purpose of validation.

**(v) Content Validity**

Content Validity is one kind of validity which depends on a careful analysis of the language being tested and of the particular course objective. Content validity is the degree to which the test items represent or cover the whole text or course. To achieve this validity the language or material being

tested and the particular course objectives should be carefully analyzed by the expert, who may be the language teacher.

Content Validity is defined as “Whether the items composing the test do, in fact, constitute representative sample of the content domain of concern” (Brown 1976 : 122 – 123). In case of a final achievement test, it is said to have content validity, if its test items are considered to be a representative sample of the tasks as can be seen in the course objectives. Anastasi (1982:132) describes the nature of content validity as, “Content validity involves essentially the systematic examination of the test content to determine whether it covers a representative sample of the behaviour domain to be measured.”

For further, Brown (1976 : 123) says, Content validity is “a measure of the adequacy of sampling.”

In the same way, Richards et al. (1999 : 81) describe, content validity is “a form of validity which is based on the degree to which a test adequately and sufficiently measures the particular skills or behaviour it sets out to measure. For instance, a test of pronunciation skills in language would have low content validity if it tested only some of the skills which required for accurate pronunciation.” By above mentioned definitions, we can conclude that content validity is one of the most important qualities of validity. Mainly, it covers two things (components), they are content relevance and content coverage.

### **Content Relevance**

"Validity is essentially a matter of relevance. Content relevance refers to whether or not the test tasks included in an exam are relevant to the language activities that are expected to be exercised under the given course. For a test to be valid we expect the content and conditions to be relevant and that there will be no irrelevant problems which are more difficult than the problems being tested” (Lado, 1999). Content relevance involves the specification of the ability domain and the test method facts. The amount or the way that subject matters are closely related with the course is content relevance.

### **Content Coverage:**

Content coverage is the degree to which the tasks required in the test adequately represent the behavioural domain in the questions. Hughes (1989:15) views that Content validity depends on the how many of the functions are tested in the component and how representative they are of the complete set of functions included in the objectives. Similarly, Brown (1976:124), says, it is clear that in a two or three hours test one cannot use all the contents from the Syllabus or Course. The selection of tests, therefore, to be included in the test is indispensable. The basic question is therefore, whether the test items that compose an exam constitute an appropriate representatives sample of behavior domain under Consideration.

In the words of Harrison (1991 : 11), he claims that Content validity is established by considering the purposes of the assessment and then drawing up a content lists. Lastly, we can say that Content coverage is amount or the way that the subject matter of the Course covers an area.

#### **1.1.4. Introduction to Major English at B.ED. Level**

The National Education Commision (NEC, 1992) of Nepal recommended phasing out of proficiency certificate level programmes by adopting the Higher Secondary Level and three year Bachelor's (10 + 2 + 3) programme. In conformity with these recommendations and in view of the trends and practices taking place at different universities in other SAARC countries, Tribhuvan University (T.U.) introduced a three years Bachelor's Programme from the year 1996 in all faculties and institutes.

In the case of specialization subjects, mainly there are three (3) major subjects in B.Ed second year for the students of 'Three-year Bachelor's Degree' in education (B.Ed.) which are specified as follows:

- (i) Communicative English III paper (Eng. Ed. 303)
- (ii) ELT Theories and Methods (Eng. 390)
- (iii) ELT Materials and Practices (Eng. Ed. 391)



The presented thesis is related to Major English Subject i.e., ELT Materials and Practice (Eng. Ed. 391). This is a specialization course of the English language for the students' of 'three-year bachelor's degree in education especially for second year students. This course book has been designed by different writers in different ways and has been divided into two parts. One is theoretical and another is practical. Both parts contain 50 and 50 marks and the total full marks of it is 100 and 35 is its pass mark. Moreover, this course, the English Language Teaching Materials and Practices, has been designed with a view to acquainting teacher-trainees with different aspects of ELT. The course book has eight units namely, syllabus, textbook, teaching aids and materials, Lesson Planning Testing and Evaluation, Correction and Remediation, Micro-teaching and Practical error analysis. Each unit offers a theoretical foundation on an ELT Topic – followed by practical exercise on it. The first three units deal with the material aspects and remaining ones deal with the pedagogical aspects of ELT. This course book is useful for B.Ed. II<sup>nd</sup> year (T.U.), one year B.Ed. (T.U.), one year and 3 year B.Ed. (M.S.U.), 3 year B.Ed. Purbanchal University and Teacher Training Programme.

Basically, in testing system of our context, most of schools, colleges and universities follow the only annual examination system. For this purpose written, oral and practical examinations are administered according to the nature of the course. This is the same case for this presented book because the annual written examination is administered to measure the student's ability. And practical examination is also done after finishing the written examination. The distribution of marks and time allotment in form of table can be shown as follows:

**Table No: 1**

**Unit Wise Weightage and Time Allotment of the Whole Course**

Unit		Theory	Marks	Practical I	Allotment of time
			Internal	External	
I	The Syllabus	8	4	Field material 10	15%
II	The Text Book	8	4	Examinations 15	15%

III	Teaching Aids and Materials	8	4		15%
IV	Lesson Planning	8	4		15%
V	Testing and Evaluation	4	2		10%
VI	Correction and Remediation	4	2		10%
VII	Micro-teaching	5	2.5		10%
VIII	Practical error Analysis	5	2.5		10%
Total		50	25	25	100%

### 1.1.5. Course Structure of the Bachelor of English Education

There are altogether eight (8) papers, each paper carries 100 marks and are divided into two major groups. Group A for specialization and it carries six (6) papers; and Group B for Elective and it carries two papers. In Elective group there are many papers but any two papers will be offered as prescribed by the concerned subject committee in the campuses for teaching learning processes.

**Table No. 2**  
**Specialization**

S.N.	Course Title No.	Subjects	Marks
1.	Eng. Ed. 301	Fundamentals of Language And Linguistics	100
2.	Eng. Ed. 302	English Sounds and Structures	100
3.	Eng. Ed. 303	Communicative English	100
4.	Eng. Ed. 390	ELT Theories and Methods	100
5.	Eng. Ed. 305	Writing Skills in English	100
6.	Eng. Ed. 391	ELT Materials and Practices	100

**Table No. 3**

**Electives**

<b>S.N.</b>	<b>Course Titles No.</b>	<b>Subjects</b>	<b>Marks</b>
1.	Eng. Ed. 311	English For Mass Media	100
2.	Eng. Ed. 392	Teaching English Literature	100
3.	Eng. Ed. 393	English Language Teaching Methods	100
4.	Eng. Ed. 313	Fundamentals of Language and Linguistics	100
5.	Eng. Ed. 314	Writing Skills	100
6.	Eng. Ed. 315	English for Business	100
7.	Eng. Ed. 316	English for Science and Technology	100

*Source: CDC, T.U.*

From the above data and the description we know that at B.Ed. first year two specialization papers: Eng. Ed. 301 (Fundamentals of Language and Linguistics) and Eng. Ed. 302 (English Sounds and Structures) are adopted. In second year, 3 papers from the specialization group; Eng. Ed. 303 (Communicative English), Eng. Ed. 390 (ELT Theories and Methods) and Eng. Ed. 391 (ELT Materials and Practices) are adopted. Similarly, in the third year, one paper from specialization group, i.e. Eng. Ed. 305 (Writing Skills in English) and any two (2) papers from the elective groups are adopted.

The students who specialize in the Major English at B.Ed. level will study the above mentioned subjects, compulsory English in First year, do the teaching practice in third year, read other compulsory and elective subjects in different years as well. Each paper has prescribed different textbooks, reference materials to cater the objectives of the course.

### **1.1.6. The Syllabus of "ELT Materials and Practices"**

This course, the English Language Teaching (ELT) Materials and Practices, is designed with a view to acquainting teacher – trainees with different aspects of ELT.

### **General Objectives**

- The course aims at providing the teacher – trainees with sound insight into various aspects of ELT and then equipping them with necessary pedagogic skills for carrying out the task of English Language Teaching effectively and efficiently.

### **Specific Objectives**

- To develop in the teacher – trainees an ability to analyze and design school level English language Syllabuses.
- To train them how to construct and use different kinds of teaching aids.
- To develop in them an ability to critically evaluate the school level English Language textbooks.
- To help them plan their everyday lessons for effective and efficient teaching in the existing and make them able to design and use different kinds of testing materials.
- To enable them to better analyze and explain errors committed by the learners of English as a second or foreign language.

The contents of this course book can be divided into eight units as follows:

- Unit I : The Syllabus.
- Unit II : The Textbook.
- Unit III : Teaching Aids and Materials.
- Unit IV : Testing and Evaluation.
- Unit V : Lesson Planning.
- Unit VI : Micro – teaching.
- Unit VII : Practical – error analyses
- Unit VIII : Correction and Remediation.

**Source : CDC, T.U.**

Note: The more detail of this course contents is given in the appendix – 1

## 1.2. LITERATURE REVIEW

Generally, many researchers have carried out their research in the field of language testing and about the content validity in different subject matters but no specific researches have been carried out on this presented topic. However, very near ones which are related to this topic or in the field of language testing are reviewed below:

Aryal (2005), conducted a research entitled, “a study on content validity of grade twelve compulsory English examination 2061.” He has analyzed the question paper from different angles (by rubrics, length, difficulty level and content coverage). The study concluded that the rubric of all questions were simple, scientific expect few items; the length of the question matches to the allotment of time for the examinees; the question of English are moderate difficulty level; the questions related to "Heritage of words" lack content validity and was unscientific.

Batalia (2004), carried out a research on “the validation of the S.L.C. English examination.” His study concluded that content validity of the S.L.C. exam English question paper has been found having high content validity as the test items have represented 102 units out of 140 units of the whole S.L.C. English course, i.e. 72.86% (percent)

Bhattarai (2005), concluded a research entitled, “The Content Validity of Compulsory English textbook for Grade Eight.” She has found that the textbook has less content validity in terms of interest and level, listening text or objectives, speaking, reading and writing exercise and communicative functions of language.

Dhakal (2002), studied on, “Effectiveness of Discrete point Test and Integrative Test as measure of English Language Proficiency.” The study was an attempt to compare the out comes of two types of test discrete point test and integrative test. This was a cross – sectional study. It concluded that the score in the textual (seen) material was naturally higher than in the unseen text and

girls' performance was better than boys' performance. The study had the weakness that it did not show which of the test was better.

Khaniya (1990), has conducted a research on "Examination as Instruments for Educational Change: Investigating the Washback effect of Nepalese English Exams" and comes to the conclusion that SLC exam fails to assess the language skills that the SLC English course intends to develop in students ... because of its textbook and previous exam paper oriented nature, it does not encourage students and teachers to focus on language skills entailed in the objectives (Khaniya 1990 : 245). Finally he has concluded that: (i) Washback is an inherent quality of exam, (ii) Ingredients of exam determine Whether the Washback is negative or positive and (iii) Teaching for final exam is inevitable.

Khaniya (2002), has written an article on, "Washback: Emerging Validity" in the journal of NELTA (Nepal English Language Teacher Association) August – 2000, Vol.: 5 : 31 and concluded that "Washback" is an inherent attribute of an examination. To be a good examination, an exam shouldn't only exert a negative influence, but it must also have the potential to exert a beneficial influence on teaching, where necessary.

Similarly, Neupane (2005), carried a research on, "Washback Effect of Examinations: A case of communicative English". She found that the most favourable technique for the preparation as well as the theoretical examination was memorization of exponents. She also found that examination had failed to follow the genuine spirit of the course objective.

Although, the studies above mentioned are related to testing and few are related to content validity also but this research work entitled "content validity of English test paper on ELT Materials and Practices at B.Ed. level" is different from the above presented research work. So the researcher is interested in this topic to find out whether the ELT test papers have content validity or not. This study will be the first kind in the Department of English.

### **1.3. OBJECTIVES OF THE STUDY**

The following are the objectives of the proposed research:

- (i) To examine the content validity of English tests on “ELT Material and Practices” of B.Ed. second year in terms of Content Coverage and Content Weightage.
- (ii) To suggest some pedagogical implications and insights for the improvement of the test papers.

### **1.4. SIGNIFICANCE OF THE STUDY**

This study provides information on whether the administered test papers especially for “ELT Materials and Practices” had content validity or not. Furthermore, this presented study will be more beneficial to all those who are concerned and involved in language teaching and learning process. And it will be particularly useful and helpful for textbook writers, test designers, curriculum designers, subject experts, examiners, teachers and language learners etc. It will help to set valid question (test) papers having content coverage and relevance. If the teachers and designers are familiar with construction of the test papers having content validity, they will be aware of the variation and mistakes hidden in the sets and try to minimize them. So, it will help to make any examination or test more valid.

# CHAPTER TWO

## METHODOLOGY

The researcher followed the following methodology during the study in order to achieve the specified objectives.

### **2.1. SOURCES OF DATA**

For the collection of data, the researcher used only the secondary sources of data.

### **2.2. SECONDARY SOURCES OF DATA**

As secondary sources of data, the researcher has used the five years question papers from 2059 to 2063 years B.S. of the course entitled "ELT Materials and Practices" of B.Ed. Second year. Apart from this, the researcher used the various books, journals, reports and articles available related to the topic. For instance, Hughes (1989), Lado (1999), etc.

### **2.3. TOOLS FOR DATA COLLECTION**

The researcher used five years question papers (from 2059 to 2063) of the course entitled "ELT Materials and Practices" both subjective and objective questions. Besides this, he analyzed the textbook entitled "A course in ELT Materials and Practices" written by Dr. Bal Mukunda Bhandari and Thaneshwor Gnyawali of B.Ed. second year based on its objectives.

### **2.4. PROCESS OF DATA COLLECTION**

For the process of data collection, the researcher collected the five years test papers (from 2059 to 2063) asked in the previous annual examinations of B.Ed. second year, Majoring English, in course entitled "ELT Materials and Practices" containing both subjective and objective questions. He analyzed and



judged the test papers whether they have content validity or not in terms of coverage and weightage.

## **2.5. LIMITATIONS OF THE STUDY**

This study has the following limitations.

- The study is limited to the content validity of the test papers of ELT Materials and Practices.
- On the basis of five years test papers from 2059 to 2063 B.S. asked in T.U. examination their content validity has been analyzed.
- This study has been done on the theoretical basis only.
- The statistical tools used for analyzing the data were tabulation and percentage.

# CHAPTER THREE

## ANALYSIS AND INTERPRETATION OF DATA

### 3.1. ANALYSIS OF THE CONTENT VALIDITY

This Chapter deals with analysis and interpretation of the raw data used in the study. The main concern of this research work was to find out the content validity of test papers of ELT Materials and Practices at B.Ed. Second year. For this research study, this chapter has been divided into two parts. The first part deals with the analysis of content validity of the ELT Materials and Practices tests in terms of coverage, and second part deals with the content validity in terms of weightage.

Here, the question papers of the ELT Materials and Practices of B.Ed. Second year which were administered in Tribhuvan University (T.U.) examinations during 5 years period (2059 to 2063) were analyzed in terms of coverage and weightage. The researcher tried to find out whether the test papers had content validity or not.

There are altogether eight (8) units of the whole course entitled "ELT Materials and Practices" taught at B.Ed. second year. It is both theoretical and practical subject and carries 100 full marks and requires 35 to pass it. Course contents and its weightage is clearly specified in the syllables, which is given in Appendix I. Table no. 4, 5, 6, 7, 8, 9, 10 and 11 shows the representation of units 1, 2, 3, 4, 5, 6, 7 and 8 respectively in the examination from 2059 to 2063. The researcher analyzed subjective as well as objective questions of those academic years. And the question papers of those 5 years are given in appendix II.

As we have different views on content validity as, Hughes, A (1995 : 22 - 3) and Bachman (1998 : 244 - 45) have emphasized two things for content validity "content representation" and "content relevance" but here, the researcher used mainly the Heaton's (1998: 160-61) view on content validity.

In Heaton's view there are two components through which we can judge whether or not they have content validity; they are "representative sample of the course" and "weightage of the course."

Thus, using the above two principles or guidelines as "representative sample of the course", and "percentage weightage" the researcher tried to find out whether the administered tests have content validity or not. Whether the tests had strictly followed the weightage of the course or not.

### 3.1.1. Representative Sample of Content Coverage: Comparison between course contents and test contents.

For the purpose of examining content validity of ELT Materials and Practices test at B.Ed. second year during 5 years (2059 to 2063 B.S.), the researcher compared the test contents in relation to course contents. In other words, he examined whether the test contained a representative sample of the whole course or not. If the coverage of content is above 50% it is believed that the test papers are nearer to content validity. To the contrary, if it is below the 50%, it is supposed to lack content validity. The more test papers are constructed the more chances of having content validity. If more than 60% of course contents are covered in a test, then it is supposed to have high content validity.

Therefore, to find out the content validity of a test in terms of course representativeness, the researcher examined and analyzed the ELT Materials and Practices test papers of the years 2059 to 2063 of its subjective as well as objective test papers. The researcher used only descriptive method and tabulation method to analyze the data.

### 3.1.1.1 Examining Course Representativeness in Unit: One

Table No. 4

Representation of test contents in terms of course content in Unit: One

S.No.	Course Contents	Test Contents				
	Course Items	Test Items Represented				
Unit: 1	The Syllabus	2059	2060	2061	2062	2063
1.1.	Some views on Syllabus					
1.2.	Introduction	S <sup>1</sup>				
1.3.	Types of Syllabus (grammatical)					
1.3.a.	Structural Syllabus			O <sup>1</sup> , O <sup>2</sup> , S.S <sup>1</sup>	O <sup>8</sup>	
1.3.b.	Communicative /Functional Syllabus				S <sup>5</sup>	O <sup>1</sup>
1.3.c.	Linear Syllabus	O <sup>1</sup>				[ S <sup>5</sup> ]
1.3.d.	The Spiral Syllabus				O <sup>1</sup>	
1.4.	Principles and Processes of Syllabus Designing				O <sup>9</sup>	O <sup>4</sup>
1.5.	Criteria of Selecting Language Items	O <sup>2b</sup>				
1.5.1.	Frequency					
1.5.2.	Range					
1.5.3.	Coverage					
1.5.4.	Learnability					
1.6.	Model Course Formal					
1.7.	An ELT Syllabus for the beginners					
1.8.	Analysis and Evaluation of existing					
1.9.	Secondary English Curriculum					
	Total Number of asked questions in Unit: 1	2Obj, 1S.S		2Obj, 1S.S	3Obj, 1S.S	2Obj, 1S.S
	<b>Note:</b> O <sup>1</sup> = Objective Question: Question No. 1					
	O <sup>2</sup> = Objective Question: Question No. 2					
	S.S <sup>1</sup> = Subjective Short Question No.1					
	O <sup>2b</sup> = Objective Question No.2b					

In 2059, 1 subjective short question and 2 objective questions were asked in the examination.

The above table indicates that in unit: one, there are 17 language items from 1 to 1.9. If we see diachronically, among these 17 language items, there is no any most representative language item repeating in each and every year from 2059 to 2063. The language items 1.3.a. (structural syllabus), 1.3.b. (communicative syllabus), 1.4 (principles and processes of syllabus designing) have been represented only in two years questions out of five years. 1.2 (introduction), 1.3.c. (linear syllabus), 1.3.d. (the spiral syllabus) have been represented only in one year out of five years. Similarly, the language items both 1.3.c (linear syllabus) and 1.3.d. (spiral syllabus) have been represented as combined question in one year out of five years. The language items which were not represented at all are 1.1. (some views on syllabus), 1.3 (types of syllabus), 1.5 (criteria of selecting language items), 1.5.1. (frequency), 1.5.2. (range), 1.5.3. (coverage), 1.5.4. (learnability), 1.6. (modal course format), 1.7. (an ELT syllabus for the beginners), 1.8. (analysis and evaluation of existing), 1.9. (secondary english curriculum).

If we see synchronically, 3 questions (2 objectives and 1 subjective) were asked in 2059. In the case of subjective short questions, that was asked from 1.2. (introduction of syllabus) and 2 objective questions were asked from 1.3.c. (linear syllabus) and 1.5. (criteria of selecting language items). In 2060, no objective and subjective questions were asked from any unit. Similarly, in 2061, 3 questions (2 objective and 1 subjective questions) were asked. One (1) objective question was asked from 1.3.a. (structural syllabus) and another objective question was asked from the same unit also and 1 subjective short question was asked from 1.3.a. (structural syllabus). In 2062, 4 questions (3 objective questions and 1 subjective short question) were asked from this unit. 1 subjective short question was asked from 1.3.b. (communicative syllabus) and 3 objective questions were asked from 1.3.a. (structural syllabus), 1.3.d. (the spiral syllabus) and 1.4. (principles and processes of syllabus designing)

respectively. Lastly, in 2063, 3 questions (2 objective questions and 1 subjective question) were asked from this unit. The objective questions were asked from 1.3.b. (communicative syllabus) and 1.4. (principles and processes of syllabus designing) and 1 subjective short question was asked from both 1.3.c. (linear syllabus) and 1.3.d. (spiral syllabus).

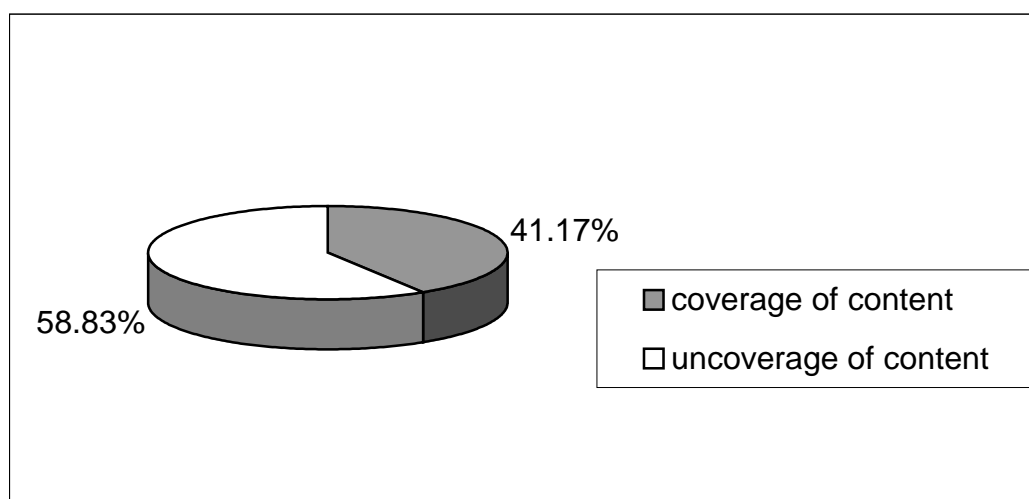
The researcher also found that there was no any representation of long questions from this unit during 5 years (2059 – 2063) question papers.

### **Repeated Items**

The researcher also found that from the unit : one, within five years, 1 objective question was repeated or same. 2062 objective question no.9 and question no.4 (Objective) of 2063 were exactly same. They were from 1.4. (principles and processes of syllabus designing).

In conclusion, the above presented table and the description can be shown in the pie chart as follows:

**Figure No.1**



From the above table and the description it is clear that there are 17 language items in unit one from 1 to 1.9 according to course contents but the representation of the test items is only 7 language items. 10 language items

were neglected while constructing test papers. It means the coverage of course contents in test contents in unit one is 41.17%. Fifty eight (58.83%) percent contents were not covered in the test papers. Therefore, ELT test papers have low content validity according to unit one because it (test papers) covers less than 50 percent course contents.

### 3.1.1.2 Examining Course Representativeness in Unit: Two

**Table No. 5**

Representation of test contents in terms of course content in Unit: Two

S.No.	Course Contents	Test Contents				
	Course Items	Test Items Represented				
		2059	2060	2061	2062	2063
Unit: 2	The Textbook			O <sup>5</sup>		O <sup>8</sup>
2.1.	Introduction and important					
2.2.	Traditional Vs Communicative textbooks			O <sup>3</sup>		S <sup>1</sup>
2.3.	Correlation between syllabus and textbooks					
2.4.	Characteristics of a good textbook.	S.S <sup>2</sup>		S.L <sup>6</sup>		
2.5	Evaluating of a textbook					O <sup>10</sup>
2.5.1	Appropriateness of target group					
2.5.2.	Organization of items (selection and gradation)		O <sup>10</sup> , S <sup>1</sup>		O <sup>2</sup>	
2.5.3.	Peripheral features		O <sup>3</sup>			
2.5.4.	Supplementary materials				S <sup>1</sup>	
2.6.	A note on English grade 9 textbook					
2.7.	A note on English grade 10 textbook					
	Total Number of asked questions in Unit: 2	1S.S, No Obj	2Obj, 1S.S	2Obj, 1S.L	1Obj, 1S.S	2Obj, 1S.S
	<b>Note:</b> O <sup>3</sup> = Objective Question: Question No. 3					
	S <sup>1</sup> = Subjective Question No.1					
	S.L <sup>6</sup> = Subjective Long Question No.6					

The above table shows that in unit: two there are 12 language items from 2 to 2.7. If we see diachronically among these 12 language items, there is no any most representative language item repeating in each and every year from (2059 to 2063). The language items 2. (the textbook), 2.2. (traditional Vs communicative textbooks), 2.4. (characteristics of a good text book), 2.5.2. (organization of items) have been represented in two years questions out of five years. The language items 2.5. (evaluating of a textbook), 2.5.3. (peripheral features) and 2.5.4. (supplementary materials) have been represented only in one year question out of five years. Besides them, the language items which were not represented at all are 2.1. (introduction and importance), 2.3. (appropriateness of target group), 2.6. (a note on English grade 9 textbook) and 2.7. (a note on English grade 10 textbook).

If we see synchronically, 1 question as subjective short question was asked in 2059 year. There was not asked any objective question in the same year. In 2060, 3 questions (2 objectives and 1 subjective short question) were asked from unit: two. In case of objective questions these questions were asked from 2.5.2. (organization of items) and 2.5.3. (peripheral features). In case of subjective question, that was asked from 2.5.2. (organization of items). In the year 2061, 3 questions (2 objectives and 1 subjective questions) were asked. In case of objective questions they were asked from 2. (the textbook) and 2.2. (traditional Vs communicative textbook). One subjective long question was asked from 2.4. (characteristics of a good textbook). Similarly, in the year 2062, 2 questions were asked. 1 objective question was asked from 2.5.2. (organization of items) and 1 subjective short question was asked from 2.5.4. (supplementary materials) and lastly, in 2063, 3 questions (2 objectives and 1 subjective short questions) were asked. In the case of objective, they were asked from

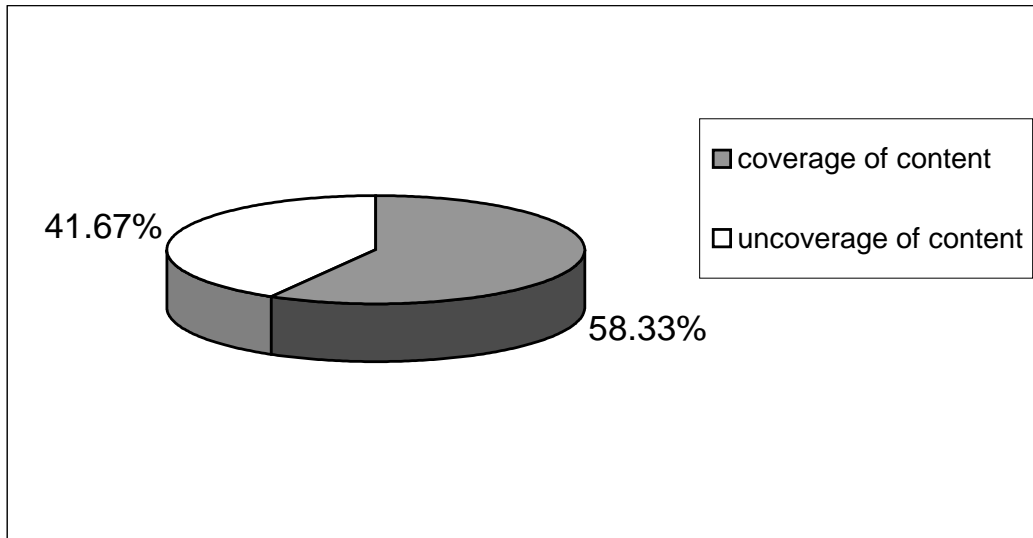


unit: 2. (the textbook) and 2.5. (evaluating of a textbook). 1 subjective short question was asked from 2.2. (traditional Vs communicative textbooks).

The researcher also found that there was one representation of subjective long question from this unit during the past 5 years.

In conclusion, the above presented table and the description can be shown in the pie chart as follows:

**Figure No.2**



From the above table and the description it is clear that there are only 12 language items in unit: two according to course content but the representation of the test items is only seven (7) language items. Five (5) language items were neglected while constructing test papers. It means the coverage of course in test contents, in unit two is 58.33(%) percent. Similarly, 41.67 (%) percent contents were not covered in the question papers. According to the unit two, ELT test papers have content validity because it (test papers) covers over than 50 percent course contents.

### 3.1.1.3 Examining Course Representativeness in Unit: Three

**Table No. 6**

Representation of test contents in terms of course content in Unit: Three

S.No.	Course Contents	Test Contents				
	Course Items	Test Items Represented				
		2059	2060	2061	2062	2063
Unit: 3	Teaching Aid and Materials					
3.1.	Introduction					
3.1.1.	Display devices		[ S <sup>1</sup> ]		O <sup>3</sup>	
3.1.2.	Visual Materials	O <sup>4</sup>				
3.1.3.	Supplementary Materials	S <sup>5b</sup> , O <sup>3</sup> , S.S <sup>5b</sup>				
3.1.4.	Electronic Aids					
3.2.	Value of teaching aids	S <sup>3</sup>			S <sup>2</sup>	S <sup>2</sup>
3.3.	Limitations of teaching aids					
3.4.	Suggestion on the use of teaching aids					
3.5.	Construction and use of teaching aids.					
3.5.1.	Blackboard /Chalkboard		O <sup>4</sup>	S <sup>3</sup>		
3.5.2.	The flannel board					
3.5.3.	Magnet board					
3.5.4.	Cork board					
3.5.5.	Pocket chart (board) Pictures		O <sup>1</sup>			
3.5.6.	Realia					
3.5.7.	Magazine Cutouts					O <sup>2</sup>
3.5.8.	Match Stick figures					
3.6.	Language games and puzzles					
3.7.	Songs and rhymes			O <sup>7</sup>		
	Total number of asked questions in Unit: 3	2Obj, 3S.S.	2Obj, 1S.S	1Obj, 1S.S	1Obj, 1S.S	1Obj, 1S.S
	<b>Note:</b> [S <sup>1</sup> ] = for two items i.e., display and visual materials					
	S.S <sup>5b</sup> = Subjective Short Question No.5b					

The above table represents that in unit three (3) there are 20 language items from 3 to 3.7. If we see diachronically, among these 20 language items,

there is no any most representative language item repeating in each and every year from 2059 to 2063. The language item 3.2. (value of teaching aids) has been only represented item in three years questions out of five years. 3.5.1. (blackboard /chalkboard) has been only represented in two years questions out of five years. Similarly, the language items 3.1.1. (display devices), 3.1.2. (visual materials), 3.1.3. (supplementary materials), 3.5.5. (pocket chart /board), 3.5.7. (magazine cutouts) and 3.7. (songs and rhymes) have been represented in one year out of five years. And the language items which were represented as combined questions are 3.1.1. (display devices) and 3.1.2. (visual materials) in 2060 year. The language items which were not represented at all are 3 (teaching aid and materials), 3.1. (introduction), 3.1.4. (electronic aids), 3.3. (limitations of teaching aids), 3.4. (suggestion on the use of teaching aids), 3.5. (construction and use of teaching aids), 3.5.2. (the flannel board), 3.5.3. (magnet board), 3.5.4. (cork board), 3.5.6. (realia), 3.5.8. (match stick figures) and 3.6. (language games and puzzles).

If we see synchronically, 5 questions (2 objectives and 3 subjective short questions) were asked from this unit in 2059. In case of objective questions, 2 questions were asked 3.1.2. (visual materials) and 3.1.3. (supplementary materials) and 2 subjective short questions were asked from 3.1.3. (supplementary materials) and 1 subjective short question was asked from 3.2. (value of teaching aids). In 2060, 3 questions (2 objectives and 1 subjective question) were asked from this unit. In the case of objective questions, 2 objectives were asked from 3.5.1. (Blackboard) and 3.5.5. (pocket chart) and 1 subjective short question was asked from two sub units, 3.1.1. (display devices) and 3.1.2. (visual materials). In 2061, 2 questions (1 objective and 1 subjective questions) were asked from this unit. 1 objective question was asked from 3.7. (songs and rhymes) and 1 subjective short question was asked from 3.5.1. (black board /chalkboard).

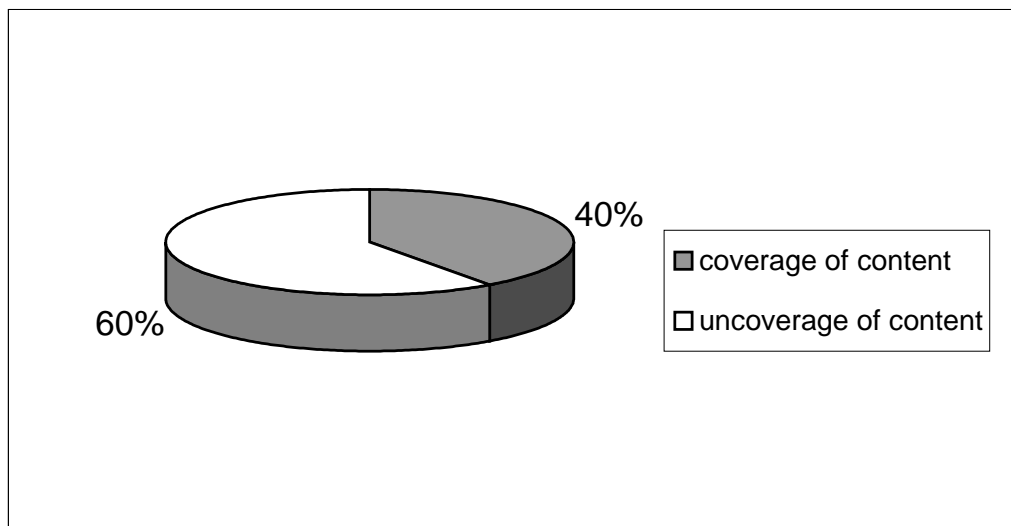
Similarly, in 2062, there were asked 2 questions as 1 objective and 1 subjective short question. 1 objective question was asked from 3.1.1. (display devices) and 1 subjective question was asked from 3.2. (value of teaching aids).

Lastly, in 2063, there were also asked 2 questions (1 objective and 1 subjective question) from this unit. The objective question was asked from 3.5.7. (magazine cutouts) and subjective short question was asked from 3.2. (value of teaching aids).

There was not asked any long subjective question during 5 years from this unit.

In conclusion, the above presented table and the description can be shown in the pie chart as follows:

**Figure No. 3**



For the above presented table and the description it is clear that there are 20 language items in unit three from 3 to 3.7. according to course content but the representation of the test items is only 8 language items. Twelve (12) language items were neglected while constructing test papers. It means the coverage of course contents in test contents in unit three is 40%. Similarly, sixty (60%) percent contents were not covered in the test papers. According to the unit three, ELT test papers have less content validity because test papers cover less than 50 (%) percent course contents.

### 3.1.1.4 Examining Course Representativeness in Unit: Four

**Table No. 7**

Representation of test contents in terms of course content in Unit: Four

S.No.	Course Contents	Test Contents				
	Course Items	Test Items Represented				
		2059	2060	2061	2062	2063
Unit: 4	Lesson Planning	S.S <sup>5a</sup>	O <sup>5</sup>	S <sup>5c</sup>		O <sup>5</sup>
4.1.	Meaning and Importance		S <sup>5</sup>			
4.2.	The Structure of lesson plan			O <sup>4</sup>		S.L <sup>6</sup>
4.2.1.	Teaching Unit					
4.2.2.	Teaching item or topic					
4.2.3.	General Objectives					
4.2.4.	Specific Objectives	O <sup>10</sup>			S <sup>5a</sup>	
4.2.5.	Lesson timing					
4.2.6.	Lesson methodology					
4.3.	Model lesson plan					
4.3.1.	Evaluation					
4.3.2.	Post reading activities					
4.3.3.	Pre-listening activities					
4.3.4.	While-listening stage					
4.3.5.	Post-listening stage					
4.4.	Annual work plan					
4.4.1.	Annual Instructional work plan					
4.4.2.	Work plan for a unit					
	Total number of asked questions in Unit: 4	1Obj, 1S.S	1Obj, 1S.S	1Obj, 1S.S	1S.S	1Obj, 1S.L.

The above table indicates that in unit four (4), there are 18 language items from 4 to 4.4.2. If we see diachronically among these 18 language items, there is no any most representative language item repeating each and every years from 2059 to 2063. The language item 4 (lesson planning) has been only represented in four years out of five years. The language items 4.2. (the structure of lesson plan) and 4.2.4. (specific objectives) have been represented in two years questions out of five years.

Similarly, the language item 4.1. (meaning and importance) has been only represented in one year's questions out of five years. And, the language items, besides them, which are mentioned in the above table have not been represented in any year.

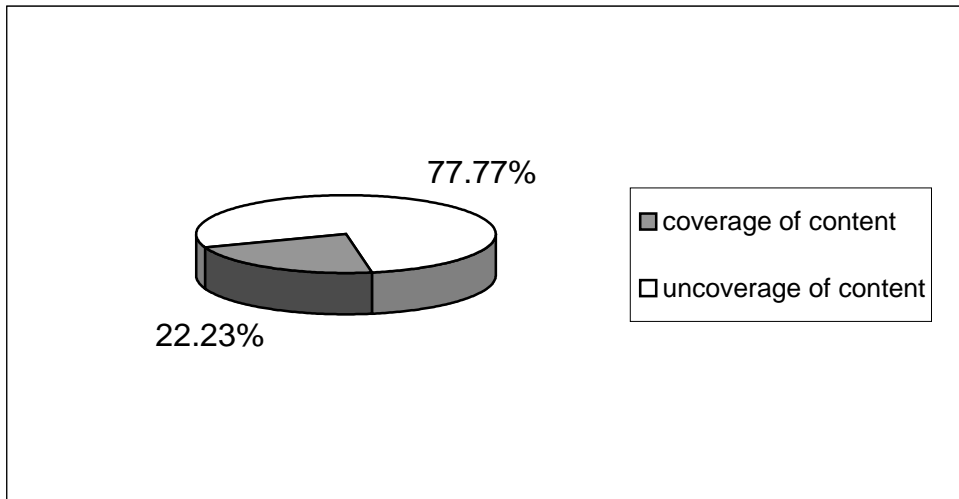
If we see synchronically, 2 questions (1 objective and 1 subjective short question) were asked in 2059. In the case of objective, these question was asked from 4.2.4. (specific objectives) and subjective short question was asked from 4. (lesson planning). In the year 2060, there were also asked 2 questions (1 objective and 1 subjective short question). 1 objective question was asked from 4. (lesson planning) and 1 subjective short question was asked from 4.1. (meaning and importance). In the same way, in 2061, 2 questions (1 subjective and 1 objective) were asked from this unit. In the case of subjective, the question was asked from 4. (lesson planning) and the objective question was asked from 4.2. (the structure of lesson plan). In 2062, there was asked only one subjective short question form 4.2.4. (specific objectives) and no any objective question was asked in this same year from this unit. Lastly, 2 questions (1 objective and 1 subjective long questions) were asked in the year of 2063. In the case of subjective, the long question was asked from 4.2. (The structure of lesson plan) and 1 objective question was asked from 4. (lesson planning).

### **Repeated Items**

The researcher also found that from the unit four, within five years, 1 objective question was repeated or same. 2060's objective question no.5 and question no.5 of 2063 were exactly same. They were from unit 4. (lesson planning).

In conclusion, the above presented table and the description can be shown in the pie chart as follows:

**Figure No. 4**



From the above table and the description it is clear that there are 18 language items in unit four from 4 to 4.4.2. according to course contents but the representation of the test items is only 4. language items. And, 14 language items were not represented while constructing test papers. It means the coverage of course contents in test contents in unit four is 22.23 (%) percent and 77.77(%) percent contents were not covered in the test papers. Therefore, ELT test papers have more less content validity according to unit four because it (test papers) covers less than 50 (%) percent course contents.



### 3.1.1.5 Examining Course Representativeness in Unit: Five

**Table No. 8**

Representation of test contents in terms of course content in Unit: Five

S.No.	Course Contents	Test Contents				
	Course Items	Test Items Represented				
		2059	2060	2061	2062	2063
Unit: 5	Testing and Evaluation					
5.1.	Formative assessment					
5.2.	Test, examination and evaluation	O <sup>7b</sup>	O <sup>7</sup>		O <sup>10</sup>	
5.3.	The purpose of testing					
5.4.	Types of test (Goal based test)		S.S <sup>2</sup>		O <sup>4</sup>	
5.4.1.	Placement test					S.S <sup>5b</sup>
5.4.2.	Diagnostic test		O <sup>2</sup>			
5.4.3.	Achievement test	[ S.S <sup>4</sup> ]			O <sup>6</sup>	
5.4.4.	Proficiency test					
5.4.5.	The close test	O <sup>5</sup>		S.S <sup>5a</sup>		
5.4.6.	Mode based test					
5.4.6.a	Subjective test					
5.4.6.b	Objective test			S.S <sup>4</sup>		
	(i) Fill in the blank test					
	(ii) Matching test					
	(iii) True /False test					
	(iv) Multiple choice test	S.S <sup>4</sup> (opt.)				
	(v) Jumbled word or sentence test					
	(vi) Close answer questions					
5.4.7.	Approach based tests					
	(a) discrete point test	O <sup>6c</sup> , S.S <sup>4</sup> (opt.)				
	(b) integrative test			O <sup>8</sup>		
5.5.	Test-design					
5.6.	Marking the tests					O <sup>6</sup> , SS <sup>5c</sup>
5.6.a.	Marking subjective tests		S.S <sup>4</sup>			
5.6.b.	Marking objective tests					
5.6.c.	Marking close tests					
5.6.d.	Marking oral tests					

5.7.	Evaluation of test (qualities of a test)		S.S <sup>4</sup> (opt.)			
5.7.1.	Validity			O <sup>6</sup>	O <sup>7</sup>	
5.7.1.	(a) Face Validity					
	(b) Content Validity					
	(c) Predictive Validity					
	(d) Construct Validity					
5.7.2.	Reliability					
	Methods of determining reliability					
5.7.3.	Scorability					
5.7.4.	Practicality /administrability					
5.7.5.	Objectivity					
5.7.6.	Economy					
	Total number of asked question in unit: 5	3Obj, 3S.S.	2Obj. 3S.S.	2Obj, 2.S.S	4Obj.	1Obj, 2S.S.

The above table indicates that in unit five, there are 33 language items from unit 5 to 5.7.6.

If we see diachronically, among these 33 language items, there is no any most representative language item repeating in each and every year from 2059 to 2063. The language items 5.2. (test, examination and evaluation) has been only represented in three years questions out of five years. The language items 5.4. (types of goal based test), 5.4.5. (the close test) and 5.7.1. (validity) have been represented in two years questions out of five years. Similarly, the language items 5.4.1. (placement test), 5.4.2. (diagnostic test), 5.4.3. (achievement test), 5.4.6.b. (objective test), 5.4.6.b. (iv) (multiple choice test), 5.4.7. (a) (discrete point test), 5.4.7.(b) (integrative test), 5.6. (marking the tests), 5.6.a. (marking subjective tests) and 5.7. (evaluation of test) have been represented in one year's questions out of five years. The language items 5.4.3. (achievement test) and 5.4.4. (proficiency test) have been represented as combined subjective short question in 2059 year from this unit. Besides them, the language items which are mentioned in the above table have not been represented in any year.

If we see synchronically, 6 questions (3 objectives and 3 subjective short questions) were asked in 2059 year. In the case of objectives, that questions were asked from 5.2. (test, examination and evaluation), 5.4.5. (the close test) and 5.4.7.(a) (discrete point test) and 3 subjective short questions were asked from 5.4.3. (achievement test) and 5.4.4. (proficiency test), 5.4.6. (b) (iv) (multiple choice test) and 5.4.7. (a) (discrete point test) respectively. Two subjective short questions were asked as an optional questions from this unit in 2059 year which are from 5.4.6. (b) (iv) (multiple choice test) and 5.4.7. (a) (discrete point test). In 2060, 5 questions (2 objective and 3 subjective short questions) were asked. 2 objective questions were asked from 5.2. (test, examination and evaluation) and 5.4.2. (diagnostic test). In the case of subjective, 3 subjective short questions were asked from 5.4. (types of goal based test), 5.6.a. (marking subjective tests) and 5.7. (evaluation of test) from this unit and one subjective short question was asked as optional question from 5.7. (evaluation of test).

Similarly, in 2061, 4 questions (2 objective and 2 subjective short questions) were asked from this unit. In the case of objectives, two questions were asked from 5.4.7. (b) (integrative test) and 5.7.1. (validity). And, for subjective case, 2 subjective short questions were asked from 5.4.5. (the close test) and 5.4.6. (b) objective test.

In the same way, in 2062, only four objective questions were asked from this unit and no any subjective question was asked. 4 objective questions were asked from 5.2. (test, examination and evaluation), 5.4. (types of goal based test), 5.4.3. (achievement test) and 5.7.1. (validity). Lastly, in 2063, 3 questions (1 objective and 2 subjective short questions) were asked. In the case of subjective, 2 short questions were asked from 5.4.1. (placement test) and 5.6. (marking the tests) and 1 objective question was asked from 5.6. (marking the tests).

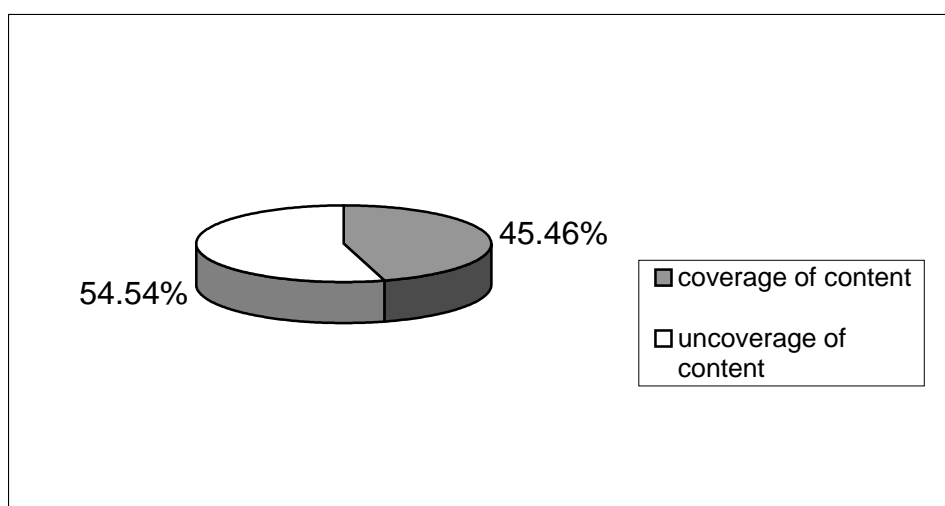
The researcher also found that there was no any representation of long subjective question from this unit.

### Repeated items

The researcher also found that from unit : five, within five years, 1 objective question was repeated or same. 2061's objective question no.6 and objective question no.7 of 2062 were exactly same. They were from 5.7.1. (validity).

In conclusion, the above presented table and the description can be shown in the pie chart as follows.

**Figure No. 5**



From the above table and the description it is clear that there are 33 language items from unit 5 to 5.7.6. according to course content but the representation of the test items is only 15 language items and 18 language items were neglected while constructing test papers. It means the coverage of course in test contents in unit five is 45.46 (%) percent. Similarly, 54.54 (%) percent contents were not covered in the question papers. Therefore, ELT test papers have less content validity according to unit five. Because it (test papers) covers less than 50 (%) percent course contents.

### 3.1.1.6 Examining Course Representativeness in Unit: Six

**Table No. 9**

Representation of test contents in terms of course content in Unit: Six

S.No.	Course Contents	Test Contents				
	Course Items	Test Items Represented				
		2059	2060	2061	2062	2063
Unit: 6	Correction and Remediation	S.S <sup>5c</sup>				
6.1.	Attitude towards errors					
6.2.	Explicit and implicit correction					
6.3.	Remedial teaching				S.S <sup>5c</sup>	
6.4.	Correction techniques			S.S <sup>5b</sup>		
6.4.1.	Self-correction					
6.4.2.	Peer-correction					
6.4.3.	Teacher correction					
6.5.	Mixed correction					
	Total number of asked question in unit: 6	1S.S No. Obj.	-	1S.S.	1S.S.	-

The above table represents that in unit six, there are 9 language items from 6 to 6.5. If we see diachronically, among these 9 language items, there is no any most representative language item repeating in each and every year from 2059 to 2063. The language items 6. (correction and remediation), 6.3. (remedial teaching), 6.4. (correction techniques) have been represented only in one year's questions out of five years. The language items which were not represented at all are 6.1. (attitude towards errors), 6.2. (explicit and implicit correction), 6.4.1. (self correction), 6.4.2. (peer correction), 6.4.3. (teacher correction) and 6.5. (mixed correction).

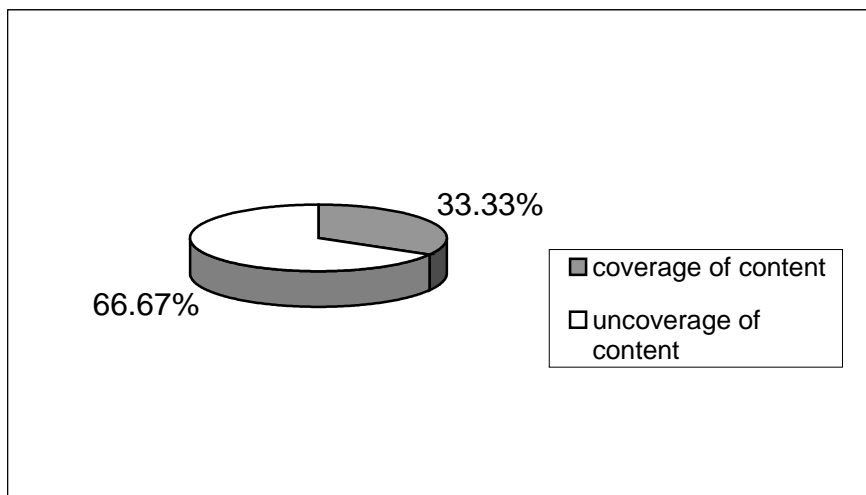
If we see synchronically, only one subjective short question was asked and no any objective question was asked in 2059 year. For 2060 year, there was not asked

any subjective and objective question from this unit. And, in 2061, there was asked only one subjective short question and no objective question was asked. Similarly, in 2062, there was also asked only one subjective short question. Lastly, in 2063, there was not asked any subjective and objective question.

The researcher also found that there was not asked any subjective long question from this unit during the passed five (2059 to 2063) years.

In conclusion, the above table and the description can be shown in the pie chart as follows:

**Figure No.6**



From the above table and the description it is clear that there are only 9 language items in this unit according to course content but the representation of the test items is only 3 language items. 6 language items were neglected while

constructing test papers. It means the coverage of course in test content is 33.33% (percent) and 66.67 % (percent) contents were not covered in the test papers.

According to the unit six, ELT test papers have less content validity because test papers cover less than 50% (percent) course contents.

### 3.1.1.7 Examining Course Representativeness in Unit: Seven

**Table No. 10**

Representation of test contents in terms of course content in Unit: Seven

S.No.	Course Contents	Test Contents				
	Course Items	Test Items Represented				
		2059	2060	2061	2062	2063
Unit: 7	Micro-teaching			S.S <sup>2</sup>		O <sup>7</sup>
7.1.	Objectives of micro-teaching					
7.2.	Some features of micro-teaching				S <sup>3</sup>	
7.3.	Stages of micro-teaching	O <sup>8c</sup>		O <sup>10</sup>		
7.3.a.	Briefing					
7.3.b.	Teaching					
7.3.c.	Commenting					
7.3.d.	Re-teaching					
7.4.	Limitations of micro-teaching					
7.5.	Lesson plans for micro-teaching					
	Total number of asked question in unit: 7	1Obj.		1Obj, 1S.S.	1S.S	1Obj.

The above table reveals that in unit seven, there are only 10 language items from unit 7 to 7.5. If we see diachronically, among these 10 language items, there is no any most representative language item repeating in each and every year from 2059 to 2063. The language items 7. (micro-teaching), 7.3. (stages of micro-teaching) have been represented in two years out of five years. The language item 7.2. (some features of micro-teaching) has been represented only in one year out of five years. Similarly, the language items which were not represented in any year are 7.1. (objectives of micro-teaching), 7.3.(a) (briefing), 7.3.(b) (teaching), 7.3.(c) (commenting), 7.3.(d) (re-teaching), 7.4. (limitations of micro-teaching) and 7.5. (lesson plans for micro-teaching).

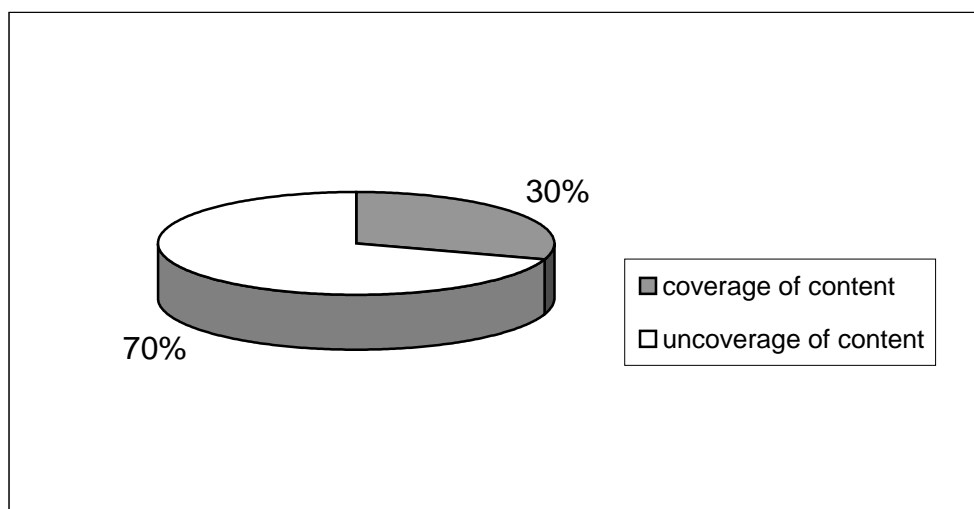
If we see synchronically, there is only 1 objective question was asked in 2059 year and no subjective question was asked from this unit in the same year. In 2060, there were not asked any subjective and objective questions from this unit. In the same way, in the year 2061, 2 questions (1 objective and 1 subjective question) were asked from this unit. For objective, the question was asked from 7.3. (stages of micro-teaching) and the subjective short question was asked from 7 (micro-teaching). In 2062, there was asked only one subjective short question and no objective question. At last, in 2063, there was asked only 1 objective question and no any subjective question was there.

The researcher also found that there was no any subjective long question asked from this unit during 5 years (2059 to 2063).

In conclusion, the above presented table and the description can be shown in the pie chart as follows.



**Figure No. 7**



From the above table and the description it is clear that there are 10 language items in unit seven according to course content but the representation of the test items is only 3 language items and 7 language items were neglected while constructing test papers. It means the coverage of course in test contents in unit seven is 30% (percent). 70% (percent) contents were not covered in the test papers. Therefore, according to the unit seven, ELT test papers have less content validity because the test papers cover less than 50 percent course contents.

### **3.1.1.8 Examining Course Representativeness in Unit: Eight**

**Table No. 11**

Representation of test contents in terms of course content in Unit: Eight

S.No.	Course Contents	Test Contents				
	Course Items	Test Items Represented				
		2059	2060	2061	2062	2063
Unit: 8	Practical error analysis					
8.1.	Definition and aim of error analysis		O <sup>6</sup>			
8.2.a.	Stages in error analysis				S.S <sup>4 (opt.)</sup>	
	- Recognition					
	- Description					
	- Explanation					
8.2.	Identification of errors					O <sup>9</sup>
8.2.1.	Overt Vs Covert errors					
8.2.2.	Group errors Vs individual errors					
8.2.3.	Systematic Vs unsystematic errors					
8.3.	Description of errors				S.S <sup>4</sup> , S.L <sup>6</sup>	
8.3.1.	Omission					
8.3.2.	Addition					
8.3.3.	Selection					
8.3.4.	Ordering					
8.4.	Types of errors at different level of language (committed errors)		O <sup>9</sup>			S.S <sup>4</sup>
8.4.1.	Phonological error					
8.4.2.	Graphological error					
8.4.3.	Lexical error					
8.4.4.	Grammatical errors					
8.4.5.	Stylistic errors					
8.5.	Explanation of errors (causes of errors)					
8.5.1.	Interlingual error or language transfer	O <sup>9a</sup>			O <sup>5</sup>	
8.5.2.	Intralingual analytical errors inherent in the T.L.					
8.5.2.a	Over generalization					O <sup>3</sup>
8.5.2.b	Under generalization					
8.5.2.c	Hyper correction			O <sup>9</sup>		
8.6.	Steps of error analysis (stages)	S.L <sup>6</sup>	S.L <sup>6</sup>			S.S <sup>3</sup>
8.6.1.	Collection of Data					
8.6.2.	Identification					

8.6.3.	Description and classification					
8.6.4.	Explanation					
8.7.	Practical work					
	Total number of asked question in unit: 8	1 Obj, 1S.L.	2Obj, 1S.L.	1Obj,	1Obj, 2S.S. 1S.L.	2Obj, 2S.S.

The above table indicates that in unit eight, there are only 30 language items from 8 to 8.7. If we see diachronically, among these 30 language items, there is no any most representative language item repeating in each and every year from 2059 to 2063. The language item 8.6. (steps of error analysis) has been represented in three years questions out of five years. The language items 8.4. (types of errors at different levels of language) and 8.5.1. (interlingual errors /language transfer) have been represented in two years question out of five years. Similarly, the language items 8.1. (definition and aim of error analysis), 8.2. (Identification of errors), 8.3. (description of errors), 8.5.2.(a) (over generalization), 8.5.2.(c) (hyper correction) have been represented in one year's question out of five years. And, one subjective short question has been represented as an optional question from 8.1.(a) (stages in error analysis) also. Besides them, the remained language items which are mentioned in the table have not been represented in any year.

If we see synchronically, in 2059, 2 questions were asked (1 objective and 1 subjective short questions) from this unit. In the case of objective, the question was asked from 8.5.1. (interlingual error) and the subjective long question was asked from 8.6. (steps of error analysis).

In 2060, 3 questions (2 objective and 1 subjective long questions) were asked. In the case of objective, 2 objective questions were asked from 8.1. (definition and aim of error analysis) and 8.4. (types of error at different levels' language) and 1 subjective long question was asked from 8.6. (steps of error analysis). Similarly, in 2061, only one objective question was asked from 8.5.2.(c) (hyper correction). In the same way, in 2062, 4 questions (1 objective, 2 subjective short and 1 subjective long questions) were asked. 1 objective question was asked from 8.5.1. (interlingual error) and in the case of subjective,

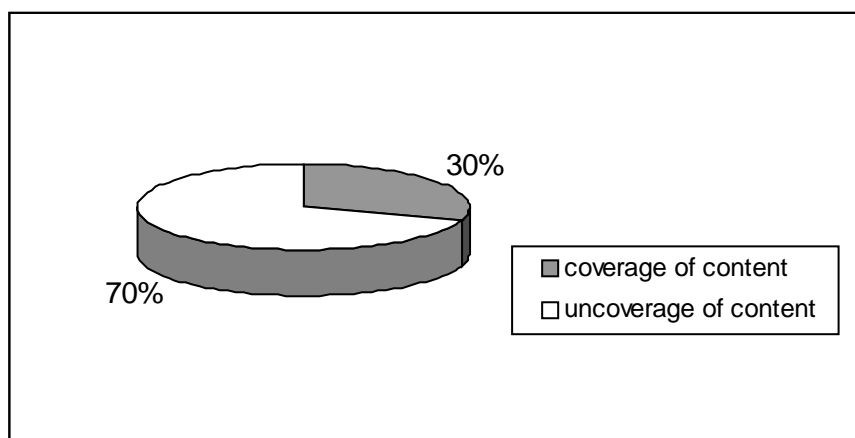
2 subjective short and 1 subjective long questions were asked. 1 objective question was asked from 8.5.1. (interlingual error) and in the case of subjective, 2 subjective short questions were asked from 8.1.(a) stages in error analysis as an optional question and from 8.3. (description of errors). And, one subjective long was asked from 8.3. (description of errors). Lastly, in 2063, 4 questions (2 objective and 2 subjective questions) were asked from this unit. In the case of objective, 2 questions were asked from 8.2. (identification of errors) and 8.5.2.(a) (over generalization). In the case of subjective, 2 subjective short questions were asked from 8.4. (types of errors at different language levels) and 8.6. (steps of error analysis).

### **Repeated Items**

The researcher also found that from unit: eight, within five years, 1 objective and 1 subjective long questions were same. The objective question no. 9a of 2059 and question no. 5 of 2062 were exactly same. They were from 8.5. (interlingual error). And another were subjective long question number 6 from 2059 and question number 6 of 2060 exactly same. They were from 8.6. (steps of error analysis).

In conclusion, the foregoing table and the description can be shown in the pie chart as follows:

**Figure No. 8**



From the above table and the description it is clear that there are 30 language items in unit 8 according to course contents but the representation of the test items is only 9 language items and 21 language items were neglected while constructing test

papers. It means the coverage of course contents in test contents in unit eight is 30% (percent). 70% (percent) contents were not covered in the question papers. Therefore, ELT test papers have less content validity according to the unit: eight because the test papers cover less than 50% (percent) course contents.

3.1.2 Examining Content Validity of the test papers on the whole in terms of coverage:

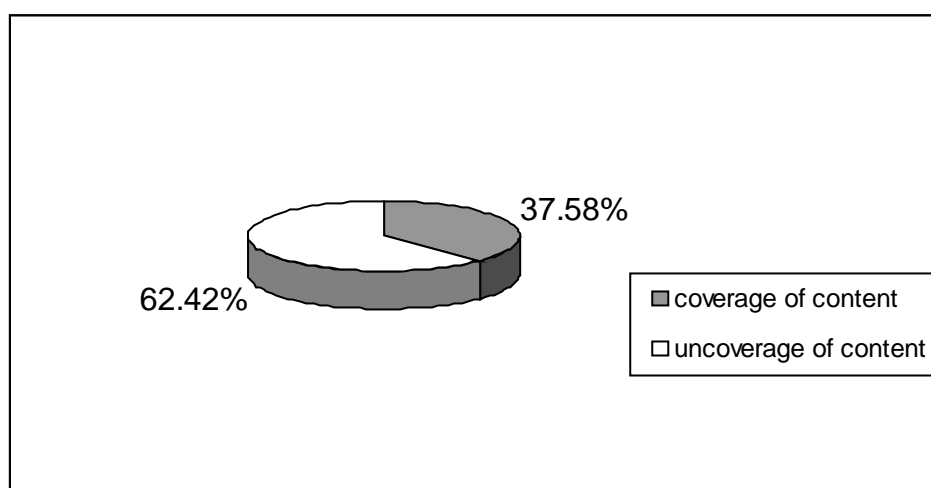
**Table No.12**

<b>S.No.</b>	<b>Units</b>	<b>Course Contents Language Items</b>	<b>Test Contents Language Items</b>	<b>Test Coverage in Percentage</b>
1.	1	17	7	41.17
2.	2	12	7	58.33
3.	3	20	8	40.00
4.	4	18	4	22.23
5.	5	33	15	45.46
6.	6	9	3	33.33
7.	7	10	3	30.00
8.	8	30	9	30.00
<b>Total</b>		149	56	37.58

The above table reveals that on the whole in the eight units, there are 149 language items in the whole ELT Materials and Practices course at B.Ed. Second Year. Out of 149 course contents language items, the representation of the test contents language items are 56. Ninety three (93) language items were neglected while constructing the test papers. It means the coverage of contents in whole units is 37.58 percent. Sixty two (62.42) percent contents were not covered in the test papers.

In conclusion, the above presented table and the description can be shown in the pie-chart as follows.

**Figure No.9**



Therefore, the content validity of the test papers as a whole is not satisfactory because it covers less than 50 percent course contents. When we see in individual sense, out of eight units, only one unit has content validity i.e., unit : 2, others remained seven units have low content validity. We can also say that most of the units have not represented the course contents and they have not really tested what they have claimed to test in the testees. Thus, the ELT test papers have not good content validity in terms of coverage. It has low content validity.

### **3.2. WEIGHTAGE PERCENTAGE:-**

Comparison between weightage of the course contents and weightage of the test contents:

This is the second part of this chapter which deals with an analysis of the proportionality of weightage of the ELT Materials and Practices test papers. In other words, for the purpose of examining content validity of the ELT Materials and Practices test, the researcher compared /examined whether or not the marks weightage in the course contents according to syllabus is proportional with the weightage of the test contents for this purpose, the researcher examined ELT Materials and Practices 5 years (2059 to 2063) test papers as a whole.

Examining the content validity of the test papers on the whole weightage

**Table No.13**

Weightage of the course contents			Weightage of the test papers				
S.No.	Units	Full Weightage	2059	2060	2061	2062	2063
1.	1	8	2 + 6	-	2 + 6	3 + 6	2 + 6
2.	2	8	6	2 + 6	2 + 10	1 + 6	2 + 6
3.	3	8	2 + 6 + 3	2 + 6	1 + 6	1 + 6	1 + 6
4.	4	4	1 + 3	1 + 6	1 + 3	3	1 + 10
5.	5	8	3 + 6 + (12)	2 + 12 + (6)	2 + 9	4	1 + 6
6.	6	5	3	-	3	3	-
7.	7	4	1	-	1 + 6	6	1
8.	8	5	1 + 10	2 + 10	1	1 + 16 + (6)	2 + 12
Total 50 marks			53 + (12)	49 + (6)	53	56 + (6)	56
<b>Total</b>			50	50	50	50	50

Note:- In 2059,  $2 + 6 = 8$  marks from unit 1 has been asked as compulsory questions marks (2 objective questions contain 2 marks and 1 subjective short question carries 6 marks).

The above table indicates that out of 50 marks of the whole ELT Materials and Practices theoretical course at B.Ed. level, unit one has carried 8 marks according to syllabus. In comparison between weightage of the course contents to weightage of the test papers, the researcher found that though **unit one** is said to be carrying 8 marks, it is found to be followed in test papers because, in 2059, there asked 8 marks as a compulsory questions marks (2 marks for 2 objective questions and 6 marks for 1 subjective short question). In 2060, there were not asked any subjective and objective questions. So it seems that there was not followed the syllabus weightage. In 2061, there were asked 2

marks for 2 objective questions and 6 marks for 1 subjective question. So, it is found that the test papers were followed by the syllabus weightage. Similarly, in 2062, the weightage of the test papers from this unit was 9 marks as compulsory questions marks (3 marks for 3 objective questions and 6 marks for 1 subjective short question). In 2063, the weightage of the test papers was 8 marks as compulsory questions marks (2 marks for 2 objective questions and 6 marks 1 subjective short question). In 2063, the weightage of the test papers seems satisfactory. Thus, as a whole in unit one, the researcher found that in most of the years the marks weightage in the test papers were proportional to the weightage of its course contents. And, this shows that the ELT Materials and Practices test papers have content validity.

**Unit two** carries 8 marks weightage according to syllabus but the test papers have been carried different weightage as follows: In 2059, there was only 6 marks as compulsory question marks for subjective short question. In 2060, the weightage of the test papers was 8 marks as compulsory questions marks (2 marks for 2 objective questions and 6 marks for 1 subjective short question). It seems that the test papers weightage is satisfactory. In 2061, the weightage of the test papers was 12 marks (2 marks for 2 objective questions and 10 marks for 1 subjective long question). So, it seems rather odd instead of weightage of course contents 8 marks from this unit in this year, its weightage is 12 marks. 4 marks was over weightage for this unit on the test papers. In 2062, the weightage of the test papers was 7 marks (1 mark for 1 objective question and 6 marks for 1 subjective short question).

Similarly, the weightage of the test papers was 8 marks in 2063. It seems that the test papers weightage is satisfactory in 2063. Finally, as a whole in unit two, the weightage of the test papers was rather satisfactory. The researcher found that in some of the years the marks weightage in the test papers were not proportional to the weightage of its course contents. Thus, this shows that ELT Materials and Practices test papers lack content validity.

In the case of **unit three**, it carries 8 marks weightage according to the syllabus but the test papers have been carried different weightage as follows: in



2059, 11 marks as compulsory question marks (2 marks for 2 objective questions and 9 marks for subjective short questions). In 2060, 8 marks (2 marks for 2 objective and 6 marks for 1 subjective question) has carried in test papers. Similarly, in 2061, 7 marks as compulsory questions (1 mark for objective and 6 marks for 1 subjective questions). In 2061, similarly, 7 marks (1 mark as objective and 6 marks as 1 subjective short questions) as compulsory marks weightage. In 2062, similarly, 7 marks (1 mark as objective and 6 marks as 1 subjective short questions) as compulsory marks. Lastly, in 2063, there was also 7 marks as compulsory marks. Thus, as a whole in unit three, the researcher found that out of all years in some of the years the mark weightage in the test papers were not proportional to the weightage of its course contents because of it's a little over and under weightage. Therefore, it shows that ELT Materials and Practices test papers lack content validity.

In the case of **unit four**, it carries 4 marks weightage according to the syllabus but the test papers have been carried different weightage as follows: in 2059, 4 marks as compulsory marks for 1 objective and 1 subjective short question. In 2060, there was 7 marks (1 mark for 1 objective question and 6 marks for 1 subjective short question). Similarly, in 2061, the weightage of test paper was 4 marks. It seems that the weightage of test paper is satisfactory in this year. In 2062, the weightage of test papers was 3 marks only. So, the weightage of test papers was rather satisfactory. Lastly, in 2063, the weightage of test papers was 11 marks (1 mark for 1 objective and 10 marks for 1 subjective long questions). It has carried 7 marks over than scheduled of the syllabus. Thus, as a whole in unit four, the researcher found that in some years the marks weightage in test papers was satisfactory and in some years there was not any norm or tendency in the distribution of marks in the test papers. This shows that ELT Materials and Practices test papers lack content validity.

Similarly, in the case of **unit five**, it carries 8 marks weightage according to syllabus but the test papers have been carried different weightage as follows: in 2059, there was asked 9 marks as compulsory marks weightage and 12 marks as an optional marks for 2 subjective short questions. In 2060, the

weightage of test papers was 14 marks (2 marks for 2 objective questions and 12 marks for 2 subjective short questions) and 6 marks for 1 optional subjective short question marks. It has carried 10 marks over than scheduled of the syllabus. Similarly, in 2061, the weightage of test papers was 11 marks (2 marks as 2 objectives and 9 marks as 2 subjective questions). There was also 3 marks over than scheduled of the syllabus. The weightage of test papers has been carried only as 4 marks in 2062. This marks were asked for 4 objective questions. It shows that there was not followed the syllabus weightage. Lastly, in 2063, 7 marks (1 mark for 1 objective and 6 marks for 1 subjective questions) was asked as compulsory question marks. Thus, as a whole in unit five, the researcher found that there was not any norm or tendency in the distribution of marks in the test papers because in some of the years it was under weightage and in some of the years it was over weightage. This shows that ELT Materials and Practices test papers lack content validity.

In the case of **unit six**, it carries 5 marks weightage according to syllabus but the test papers have been carried different weightage as follows. In 2059, 2061 and 2062 there were asked only 3 – 3 marks as compulsory marks weightage. But in 2060, there was not asked any mark's question and similarly, the weightage of test papers was not also asked in 2063. It seems that there were not followed the syllabus weightage in two years i.e., 2060 and 2063. Thus, as a whole in unit six, there was also not any norm or tendency in the distribution of marks in the test papers because of its under weightage and in some of the years it was no weightage of test papers. This shows that ELT Materials and Practices test papers lack content validity.

Similarly, **unit seven** carries 7 marks weightage according to the syllabus but the test papers have been carried different weightage as follows: in 2059, only 1 mark for 1 objective question and in 2060, there was not asked any marks test papers as objective and subjective questions. It seems that the syllabus weightage was not followed by 2060 year. The weightage of test papers was 7 marks (1 mark as 1 objective and 6 marks as 1 subjective questions) in 2061. Similarly, in 2062, 6 marks was asked for 1 subjective

short question. Lastly, for 2063, there was asked only 1 mark as objective question from this unit. Thus, as a whole in unit seven, the marks distribution for each year's test papers was not satisfactory. The researcher found that there was not any norm or tendency in the distribution of marks in the test papers because of its under weightage and no weightage marks in some years also. Therefore, this shows that ELT Materials and Practices test papers lack content validity.

Lastly, for **unit eight**, there was 5 marks weightage according to the syllabus but the test papers have been carried different weightage as follows: the weightage of test papers was 11 marks (1 mark for objective and 10 marks for 1 subjective long questions) in 2059. For 2060, there was asked 12 marks questions (2 marks for 2 objective questions and 10 marks for 1 subjective long question) as compulsory. In 2061, only 1 mark for 1 objective question. Similarly, in 2062, the weightage of test papers was 23 marks (1 mark for 1 objective, 10 marks for 1 subjective long question, 6 marks for 1 subjective short question and 6 marks as optional question marks for 1 subjective short question. It has carried 18 marks over than scheduled of the syllabus. The weightage of test papers was 14 marks (2 marks as 2 objective questions and 12 marks as 2 subjective short questions) in 2063. It seems that 9 marks was over than scheduled of the syllabus. Thus, as a whole in unit eight, the researcher found that there was not any norm or tendency in the distribution of marks in the test papers because in one year it was under weightage and in most of the years it was over weightage. This shows that ELT Materials and Practices test papers lack content validity.

As a whole during 5 years (2059 to 2063) in ELT Materials and Practices test papers at B.Ed. second year, the researcher found that in totality "out of 50" full marks (theoretical course) for the ELT Materials and Practices test papers, in 2059, there was asked 53 marks as compulsory questions marks and 12 marks as optional questions marks. In such year optional questions were asked from one unit i.e., unit five. Only 3 marks was over than scheduled syllabus. Anyway, test paper carries 50 marks. In 2060, there was asked 49

marks as compulsory questions marks and only 6 marks as optional questions marks. In such year the optional question was asked from the same unit : five. In 2061, there was asked 53 marks as compulsory questions. Likewise, in 2059, 3 marks was over than scheduled syllabus. In such year there was not asked optional question. Similarly, in the year 2062, 56 marks for compulsory questions marks and 6 marks for optional question marks. In such year optional mark was asked from unit 8. Lastly, in the year 2063, 56 marks was asked as compulsory questions marks. There was 6 marks over than scheduled syllabus in such year.

From the above data and the description the researcher found that there was not any norm or tendency in the distribution of marks in the test papers because in some of the years it was under weightage, in some of the years it was no weightage and in some of the years it was over weightage in the same unit as well. It also seems that there is the negligence in the part of question setters which may lead the lack of content validity and negative washback as well. Although in each unit there is fixed course contents marking (weighting) schedule, it is not found to be followed in setting the question papers. It creates another bad wash back effect because the unit which was given least weightage in test papers in examination, the students do not bother reading such unit because they know this will not be asked in the examination.

Furthermore, we have seen the course contents weightage of language items which is one aspect of content validity. There is no satisfactory finding. Some course language items are found to be over weightaged and some course language items are found to be under weightaged. It seems that there was no specific norm for asking questions. Therefore, it seems that ELT test papers lacked content validity in terms of weightage of the course contents.

# CHAPTER FOUR

## FINDINGS AND RECOMMENDATIONS

This chapter is related with the major findings of the research study. Therefore, after the analysis of the question papers from different perspectives, the researcher has extracted the following findings:

### 4.1. FINDINGS

The major findings of this research work are as follows:

Firstly, according to the coverage /representativeness principle, the ELT Materials and Practices tests have low content validity. It is because out of 149 language items in totality of the course, test items have represented 56 language items i.e., 37.58 percent during 5 years (2059 to 2063).

Secondly, according to the weightage principle, the ELT Materials and Practices tests have also low content validity. It is because the question papers did not follow any norm or tendency in the distribution of marks was given in the weighting schedule of the syllabus. For instance, in 2059 from the unit 5, it carried 21 marks in the test papers whereas its weightage was only 8 marks according to the syllabus.

The unit specification findings of this research work are as follows:

#### **1. In terms of coverage /representativeness between course contents and test contents.**

- a. Regarding in unit one, out of 17 language items of course contents, test items represented only 7 language items during 5 years (2059 to 2063). Ten (10) language items were unasked. It means the coverage of course contents is 41.17(%) percent. Thus, ELT Materials and Practices papers have low content validity in unit one. Because tests contents have been represented less than 50% course contents.

- b. In unit two, out of 12 language items of course contents, test items represented only 7 language items during 5 years (2059 to 2063). 5 language items were unasked. It means the coverage of course contents is 58.33%. That is to say, the content validity of ELT Materials and Practices tests have content validity according to unit two.
- c. In unit three, out of 20 language items of the course contents, test items represented 8 language items during 5 years. Twelve (12) language items were untouched. It means the coverage of course contents is 40(%) percent. That is to say, the content validity of ELT Materials and Practices tests have low content validity.
- d. In unit four (4), there are altogether 18 language items in the course contents but test contents have represented only 4 language items. 14 language items were untouched. It means the coverage of course contents is only 22.23%. Therefore, the content validity of ELT Materials and Practices tests have less content validity according to unit four.
- e. In unit five (5), there are altogether 33 language items in the course contents but test contents have represented only 15 language items. Eighteen (18) language items were unasked. It means the coverage of course contents is 45.46(%) percent. That is to say, the content validity of ELT Materials and Practices tests have not satisfactory according to unit five.
- f. In unit six (6), out of 9 language items of the course contents, test items represented only 3 language items during 5 years (2059 to 2063). Six (6) language items were unasked. It means the coverage

of course contents is only 33.33%. sixteen (16.67) percent contents were not represented in test papers. Thus, it shows that this unit has low content validity.

- g. In unit seven (7), there are altogether 10 language items in the course contents but test contents have represented only 3 language items. Seven (7) language items were untouched. It means the coverage of course contents is only 30%. Seventy (70%) percent contents were not represented in test papers. Thus, it shows that this unit lacks content validity.
- h. In unit eight (8), there are altogether 30 language items in the course contents but test contents have represented only 9 language items. Twenty-one language items were unasked. It means the coverage of course contents is only 30(%) percent. Seventy (70%) percent contents were not represented in test papers. Thus, it shows that this unit has also low content validity.

## **2 In terms of weightage comparison between course contents weightage and the test contents weightage.**

- a. In unit one, out of 8 marks weightage according to syllabus, the test papers have been carried out different weightage as follows: In the year 2059, 2 marks were there as compulsory and as optional question marks. In 2061, 8 marks as compulsory. In 2062, 9 marks for compulsory questions. And in 2063, 8 marks as compulsory.

Thus, as a whole in unit one, it is found that in most of the years the marks weightage in the test papers were proportional to the weightage of its course contents in some years not in all the years. This shows that the ELT materials and practices test papers have content validity.

b. In unit two, out of 8 marks weightage of course contents, the weightage of the test contents were as follows: In 2059, only 6 marks as compulsory question. In 2060, 8 marks as compulsory questions. In the year 2061, 2062 and 2063 the weightage of test papers were 12, 7 and 8 marks respectively as compulsory marks. Thus, wholly in unit two, it is found that, in some of the years the marks weightage in the test papers were not proportional to the weightage of its course contents because of its over and under weightage in the test papers. Thus, the test papers of ELT Materials and Practices lack content validity.

c. In unit three, out of 8 marks weightage of course contents, the weightage of the test contents were as follows:

In 2059, 11 marks as compulsory marks. In 2060 and 2061, 8 and 7 marks as compulsory respectively. In 2062, 7 marks as compulsory and similarly, in 2063, 7 marks as compulsory questions marks. There was not any optional mark in this unit.

Thus, as a whole in unit three, it is found that there was not any norm or tendency in the distribution of marks in the test papers because in some of the years it was under weightage and in one year it was over weightage. This shows that the ELT Materials and Practices test papers lack content validity.

d. In unit four, out of 4 marks weightage of course contents the weightage of the test contents were as follows: In 2059, 4 marks as compulsory. In 2060, 7 marks as compulsory. In 2061, 4 marks as compulsory marks and in 2062, 3 marks as compulsory marks. Similarly, in 2063, there was 11 marks as compulsory marks. There was no marks for optional question.

Thus, wholly in unit four, it is found that in some of the years it has satisfactory weightage and in some of the years it has over



weightage. Thus, the test papers of ELT Materials and Practices lack content validity.

- e. In unit five, out of 8 marks weightage of course contents, the weightage of the test contents were as follows: In the year 2059, 9 marks as compulsory and 12 marks as optional marks. In 2060, 14 marks as compulsory marks and 6 marks as optional marks. In 2061, 2062 and 2063, the weightage of the test papers were: 11, 4 and 7 marks respectively as compulsory.

Thus, as a whole in unit five, it is found that there was not any norm or tendency in the distribution of marks in the test papers because in some of the years it was over weightage and in some of the years it was under weightage. This shows that ELT Materials and Practices test papers lack content validity.

- f. In unit six, out of 5 marks weightage of course contents, the weightage of the test contents were as follows: In the year 2059, 2061 and 2062, the weightage of the test papers were: 3, 3 and 3 marks respectively as compulsory marks where as in 2060 and 2063 there were not any marks as compulsory and optional.

Thus, as a whole in this unit, it is found that there was not also any norm or tendency in the distribution of marks in the test papers because of its under weightage in some years and no weightage in two years. This shows that the test papers of ELT Materials and Practices lack content validity.

- g. Similarly, in unit seven, out of 4 marks weightage of course contents, the weightage of the test contents were as follows:

In the years 2059, 2061, 2062 and 2063 the weightage of the test paper were: 1, 7, 6 and 1 marks respectively as compulsory marks where as in 2060, there was not any marks.

Thus, wholly in unit seven, it is found that there was not any norm or tendency in the distribution of marks in the test papers because of its under weightage, over weightage and no weightage of the test papers. Thus, it shows that the test papers of ELT Materials and Practices lack content validity.

- h. Lastly, in unit eight, out of 5 marks weightage of course contents, the weightage of the test contents were as follows: In the years 2059, 2060 and 2061, the weightage of the test papers were 11, 12 and 1 marks respectively as compulsory only. Similarly, in 2062, 17 marks as compulsory and 6 marks as optional marks and in 2063, 14 marks as compulsory marks.

Thus, as a whole in unit eight, it is found that, there was not also any norm or tendency in the distribution of marks in the test papers because in one year it was under weightage and in most of the years it was over weightage. Thus, this shows that the ELT Materials and Practices test papers lack content validity.

3. It has been found that the ELT Materials and Practices test papers have low content validity in terms of coverage or representation of the course contents and the same case has been seen as low content validity in terms of weightage of the course contents or syllabus.
4. More variation is found in asking questions in all the types of questions (objective, subjective short and subjective long questions) in the different years in same unit as well. For instance, in unit seven, in 2059 and 2063 there were asked only 1 – 1 objective questions and in 2060, there was not asked any

questions. Similarly, from unit 5 in 2059 and 2060, there were asked more marks subjective short questions rather than scheduled the syllabus weightage.

5. According to the weightage of the course contents, it is argued that the ELT Materials and Practices tests have not tested what they have been supposed to test in the testees because weightage system of the course contents in the test contents was not found fully followed. And, according to the representation of the course contents it is also argued that the weightage of the language items seemed to be determined by what is easy to test rather than what is important to test.

## **4.2. RECOMMENDATIONS**

On the basis of the findings of the research work, some of the following recommendations are made:

- (1) It has been found that ELT Materials and Practices test papers have represented only 37.58 percent course contents as a whole which is low content validity in terms of representative principle. Therefore, in order to have high content validity, the related test papers should cover /represent more than 60% language items of the course content.
- (2) It has been found that ELT Materials and Practices test papers have not been followed scheduled weightage of the course contents. They should be followed strictly.
- (3) During the time of construction of questions, all the units should be given emphasis equally as per the weightage given in the curriculum.

- (4) A test would have high content validity, if it follows not only representative principle but also weightage principle. Thus, ELT Materials and Practices tests should follow both principles to get a test high content validity.
- (5) Instead of repeating the same test papers in each and every year examination, it should be touched the different and more language items in the examination.
- (6) To have high content validity, the question setters have to have a minute study on the course objectives, course contents and weightage of the course contents before developing the question papers. Trained and experienced teachers should be assigned to design the question papers.
- (7) There have to be prepared any fixed criteria-for instance 'specification table', for asking different types of question i.e., objective subjective short and long questions in terms of unitwise. Asking different types of questions differently in different years create harmful effects on the validity of the test papers as well as to the testees. Thus, ELT Materials and Practices tests have to follow fixed criteria based on the syllabus of its test papers.
- (8) Finally, the office of the controller of examinations should have positive spirit and have to play a vital role in improving the ELT Materials and Practices tests as a whole. It should care not only content validity of the test papers but also other types of validity as well. For this, the concerned authority should conduct some seminars, workshops and conferences.

## REFERENCES:

- Anastasi, A. (1982). **Psychological Testing Fifth Edition**. New York. Macmillan Publishing Co., Inc.
- Aryal, K.P. (2005). **A Study on Content Validity of Grade Twelve Compulsory English Examination 2061**. An unpublished M.Ed. Thesis, Kathmandu.
- Bachman, L.F. (1998). **Fundamental Considerations in Language Testing**. Oxford University Press.
- Batalia, K.B. (2004). **Validation of the S.L.C. English Examination**. An Unpublished M.Ed. Thesis, Kathmandu: Tribhuvan University.
- Bhattari, Anjana (2006). **'Writing a Research Proposal'** journal of NELTA.6/1:45:51.
- Bhattarai, G.R. (2006). **A Thematic Analysis of Research Reports**. Kathmandu : Ratna Pustak Bhandar.
- Bhattarai, J. (2005). **Content Validity of English Textbook for Grade Eight**. An Unpublished M.Ed. Thesis, Kathmandu: T.U.
- Brown, F.G. (1976). **Principles of Educational and Psychological Testing**. Newyork : Holt, Rinchart and Winston.
- Cronbanch, L.J. (1971). **Test Validation**. In Thorndike (ed.) (1970) : 443.
- Davies, A. (1968). **Language Testing Symposium: A Psycholinguistic Approach**, London : Oxford University Press.

- Davies, A. (1985). **Follow My Leader: Is That What Language Tests Do?** In Lee et al. (eds.) (1985) : 3 - 13.
- Davies, et al. (1999). **Dictionary of Language Testing.** University of Melbourne.
- Hamal, R. (2007). **Content Validity of General English Test Items incase of Reading and Writing** at B.Ed. Level. An Unpublished M.Ed. Thesis, Kathmandu: Tribhuvan University.
- Harrison, A. (1991). **A Language Testing Handbooks.** London: ELTS. Longman.
- Heaton, J.B. (1975). **Writing English Language Tests.** London: Longman.
- Hughes, A. (1989). **Testing for Language Teachers.** Cambridge University Press.
- Khaniya, T.R. (2002). **'Washback: Emerging Validity'**, journal of NELTA. Vol.5, Kathmandu: NELTA.
- Khaniya, T.R. (2005). **Examinations for Enhanced Learners.** Kathmandu: Millennium Publication.
- Lado, R. (1961). **Language Testing: The Construction and Use of Foreign Language Tests.** London. Longman.
- Morrow, K. (1986). **The Evaluation of Tests of Communicative Performance.** Inportal (ed.) (1986) : 1 - 13.
- Neupane, (2005). **Content Validity of English Textbook for Grade Seven.** An Unpublished M.Ed. Thesis, Kathmandu: Tribhuvan University.
- Ojha, D.R. (2005). **Content Validity of ELT Theories and Methods Exam at B.Ed. Level.** An Unpublished M.Ed. Thesis, Kathmandu: Tribhuvan University.
- Richards, et al. (1985). **Longman Dictionary of Applied Linguistics.** London. Longman.