

**CONTENT VALIDITY OF EXAMINATION PAPERS: A
CASE OF READING WRITING AND CRITICAL
THINKING COURSE**

**A Thesis Submitted to the Department of English Education
In Partial Fulfillment for the Master's of Education in English**

**Submitted by
Sita Devi Sharma**

**Faculty of Education
Tribhuvan, University
Kirtipur, Kathmandu, Nepal
2011**

**CONTENT VALIDITY OF EXAMINATION PAPERS: A
CASE OF READING WRITING AND CRITICAL
THINKING COURSE**

**A Thesis Submitted to the Department of English Education
In Partial Fulfillment for the Master's of Education in English**

**Submitted by
Sita Devi Sharma**

**Faculty of Education
Tribhuvan, University
Kirtipur, Kathmandu, Nepal
2011**

T.U. Regd. No: 6-1-53-254-2001

Campus Roll No: 1557

Second Year Exam Roll No: 280721/066

Date of Approval of thesis

Proposal: 2067/09/13

Date of Submission of

Thesis: 2068/03/03

RECOMMENDATION FOR ACCEPTANCE

This is to certify that **Ms. Sita Devi Sharma** has prepared the thesis entitled **Content Validity of Examination Papers: A Case of Reading Writing and Critical Thinking Course** under my guidance and supervision.

I recommend the thesis for approval and acceptance.

Date: 2068/03/03

Dr. Tapasi Bhattacharya (Guide)

Reader

Department of English Language

Education

T.U.Kirtipur, Kathmandu

RECOMMENDATION FOR EVALUATION

This thesis has been recommended for the evaluation by the following 'Research Guidance Committee'.

Dr. Chandreshwar Mishra

Professor and Head

Department of English Education,

T.U., Kirtipur, Kathmandu,

Chairperson

Dr. Tapasi Bhattacharya (Guide)

Reader

Department of English Education,

T.U.Kirtipur, Kathmandu,

Member

Mr.Prem Bahadur Phyak

Lecturer

Department of English Education,

T.U, Kirtipur, Kathmandu,

Member

Date:2068/03/09

EVALUATION AND APPROVAL

This thesis has been evaluated and approved by the following 'Thesis Evaluation and Approval Committee'.

Dr. Chandreshwar Mishra

Professor and Head,
Department of English Education,
T.U.Kirtipur, Kathmandu, Nepal

Member

Dr. Jai Raj Awasthi

Professor
Department of English Education,
T.U. Kirtipur, Kathmandu, Nepal
Chairman
English and Other Foreign Languages
Education Subject Committee

Member

Dr. Tapasi Bhattacharya (Guide)

Reader
Department of English Education,
T.U. Kirtipur, Kathmandu, Nepal
Date: 2068/03/09

Member

DECLARATION

I hereby declare to the best of my knowledge that this thesis is original; no part of it was earlier submitted for the candidature of research degree to any University.

Date: 2068/03/02

.....
Sita Devi Sharma

DEDICATION

**Affectionately Dedicated
To
My parents who devoted their entire life to
make me what I am today.**

ACKNOWLEDGEMENT

I would like to express my sincere, respectful and profound gratitude to my thesis supervisor **Dr. Tapasi Bhattacharya**, Department of English Education, T.U. Kirtipur for making constant supervision, and guiding me with regular inspiration, encouragement, and insightful suggestions throughout the study. Her cooperative nature and valuable guidance from the very beginning helped me to accomplish this task.

I would like to express my warm respect to the prominent Prof. **Dr. Chandreshwar Mishra**, Head of the Department of English Education, T.U. Kirtipur for his encouragement and academic support.

I owe a debt of profound gratitude to **Dr. Jai Raj Awasthi**, Professor and Chairperson of the English and Foreign Languages Education Subject Committee, University Campus T.U. Kirtipur, for his regular inspiration, encouragement, and valued suggestion.

I am equally grateful to **Dr. Anjana Bhattarai, Dr. Govinda Raj Bhattarai, Dr. Balmukunda Bhandari, Dr. Anju Giri, Mr. Bishnu Singh Rai, Mr. Raj Narayan Yadav, Mr. Prem Bahadur Phyak, Ms. Hima Rawal, Ms. Madhu Neuphane, Ms. Saraswati Dawadi**, and other faculty members of the Department for their valuable suggestions and encouragements.

Last but not the least, I would like to express my warm gratitude to my husband **Nabaraj Subedi** for his continuous encouragement, inspiration, and guidance in my study. I want to express special thanks to **Ravi Kumar Subedi** for computer work. I am also Grateful towards my grandfather, grandmother, Dad, Mom, brothers and sisters, for their cooperation in my research study.

Ashad, 2068

Sita Devi Sharma

ABSTRACT

This thesis entitled "**Content Validity of Examination Papers: A Case of Reading, Writing and Critical Thinking Course**" attempts to determine the Validity of Year End Examination of B. Ed First Year Course. The main objective of this research study is to determine the validity of the Year End Examination of B. Ed. First Year's Course (Reading Writing and Critical Thinking, Eng. Ed.317). In order to fulfill this purpose, the researcher has analyzed only subjective question papers of 'Reading Writing and Critical Thinking ' from 2066-2067. Only the secondary sources of data were used for this purpose. The question papers were compared and analyzed with course contents in terms of content coverage and weighting .It was found that the Examination papers have good content validity in terms of content coverage but low content validity in terms of content weighting.

The study is divided into four chapters. The first chapter deals with the introduction which consists of general background (language testing, Achievement tests, language teaching and language testing, qualities of good tests, language skills, introduction of course, reading vs. writing, critical thinking. It also includes review of related literature, objectives and significance of the study. The second chapter is related to the methodology which consists of sources of data, tools for data collection, process of data collection and limitations of the study. The third chapter deals with the analysis and interpretation of the collected data which were analyzed and interpreted in terms of content coverage and weighting. The last chapter deals with the findings, recommendations of the study. The findings have been made on the basis of the analysis and interpretation of data. It also includes appropriate recommendations or pedagogical implications.

TABLE OF CONTENTS

	Page
<i>Declaration</i>	<i>i</i>
<i>Recommendation for Acceptance</i>	<i>ii</i>
<i>Recommendation for Evaluation</i>	<i>iii</i>
<i>Evaluation and Approval</i>	<i>iv</i>
<i>Dedication</i>	<i>v</i>
<i>Acknowledgement</i>	<i>vi</i>
<i>Abstract</i>	<i>vii</i>
<i>Table of Contents</i>	<i>viii</i>
<i>List of Tables</i>	<i>x</i>
<i>List of Figures</i>	<i>xi</i>
<i>List of symbols and Abbreviations</i>	<i>xii</i>
CHAPTER ONE: INTRODUCTION	
1.1 General Background	1
1.1.1 Language Testing	1
1.1.2 Achievement test	4
1.1.3 Language skills	5
1.1.4 ELT in Nepal	7
1.1.5 Strength and Weakness of written tests	8
1.1.6 Language teaching and testing.	10
1.1.7 Qualities of good test	11
1.1.8 Introduction of New Direction: Reading writing and critical thinking	20
1.1.9 Reading Vs Writing	24
1.1.11 Critical thinking	27
1.2 Review of Related Literature	30
1.3 Objectives of the study	35
1.4 Significance of the study	35

CHAPTER TWO :METHODOLOGY

2.1 Sources of Data	36
2.2 Secondary Sources	36
2.3 Tools for Data Collection	36
2.4 Process of Data Collection	36
2.5 Limitations of the Study	37

CHAPTER THREE : ANALYSIS AND INTERPRETATION

3.1 Analysis of Content Coverage	38
3.1.1 Analysis of Content Coverage from Unit One	39
3.1.2 Analysis of Content Coverage from Unit Two	43
3.1.3 Analysis of Content Coverage from Unit Three	47
3.1.4 The Whole Content Validity in Terms of Content Coverage	51
3.1.5 Unit Wise Comparison in Terms of Content Coverage	53
3.2 Content Weighting	53
3.2.1 The Whole Content Validity in Terms of Content Weighting	54
3.3 Comparison of Content Validity in Terms of Content Coverage and Weighting	56

CHAPTER FOUR: FINDINGS AND RECOMMENDATIONS

4.1 Findings	57
4.2 Recommendations	59
References	
Appendix	

LIST OF TABLES

	Page
Table No.1: Analysis of Content Coverage from Unit One	39
Table No. 2: Analysis of Content Coverage from Unit Two	43

Table No. 3: Analysis of Content Coverage from Unit Three	57
Table No. 4: The Whole Content Validity in Terms of Content Coverage	51
Table No. 5: The Whole Content Validity in Terms of Content Weighting	54

LIST OF FIGURES

Figure No.1 Content Coverage from Unit One in 2066 B.S.	41
Figure No.2 Content Coverage from Unit One in 2067 B.S.	42
Figure No.3 Content Coverage from Unit Two in 2066 B.S.	45
Figure No.4 Content Coverage from Unit Two in 2067 B.S.	46
Figure No.5 Content Coverage from Unit Three in 2066 B.S.	49
Figure No.6 Content Coverage from Unit Three in 2067 B.S.	50
Figure No.7 Content Coverage on the Whole During 2 Years	52
Figure No.8 Unit Wise Coverage During 2 Years	53

LIST OF SYMBOLS AND ABBREVIATIONS

alt	=	alternative
w.u.	=	within a unit
Eng	=	English
Ed.	=	Education
Mr.	=	Mister
Ms.	=	Misses
VS	=	Versus
B. Ed.	=	Bachelor in Education
B.S.	=	Bikram Sambat
Dr.	=	Doctor
e.g.	=	For example; for instance
ELT	=	English Language Teaching
et al	=	And other people
etc.	=	etcetera
i.e.	=	That is
No.	=	Numbers
%	=	Percentage
°	=	Degree
Prof.	=	Professor
T.U.	=	Tribhuvan University

CHAPTER - ONE

INTRODUCTION

1.1 General Background

Defining validity is a difficult job. However, as a researcher, I have attempted to find out the validity of the Final Examination of 'Reading Writing and Critical Thinking' Course at B. Ed First Year Level. The main aim of my study is to determine the validity of the Year End Examination of 'Reading Writing and Critical Thinking' (Eng.Ed.317). I hope this study primarily will help and provide the constructive feedback/guideline for question designers to make the test papers more valid. Moreover, this study could also be of great help for the students to prepare for the examination. This chapter includes general background, language testing, achievement test, language teaching and testing, qualities of good tests, 'introduction of new direction: reading writing and critical thinking', syllabus of reading writing and critical thinking, reading vs. writing, critical thinking, literature review, objective of the study, significance of the study.

1.1.1 Language Testing

Generally text and exam are taken synonymously in the testing literature. Brereton (1969) defines test as "a test is designed to provide an accurate standardized measurement of certain abilities or skills without influencing teaching or student and without creating any tension in the student"(as cited in Hamal, 2007, p.6). Similarly, Harrison (1883) says "a test is a natural extension of the classroom work providing teaching, teacher and students that can serve as a basis for improvement"(p.1). Various scholars define test variously so, there is no common definition of a test. However, a test is meant to measure the learner's knowledge. It is a device for providing the learner into showing what he knows or it is set of questions or problems for determining a person's knowledge or ability.

It is generally assumed that the history of English testing is as old as the history of English language teaching itself. Testing has always been an integral part of any English language teaching programme, which probably began in the 15th century with the ordinance promulgated by Henry the 5th that English should be adopted as the language of royal correspondence in the place of French. Giri (2003) writes in his articles:

Testing started much earlier in Britain than in America. In fact the University of Cambridge became involved in overseas testing within a few years of establishment of the public examinations. The university sent out Examination papers overseas for the first time in 1963 to examine ten candidates and continued to develop its overseas examination in the next four decades of the 19th century. By 1898, the University of Cambridge had 36 colonial centers with 1220 candidates. However its formal entry in testing the English of foreigners was not until 1913, when it instituted the examination for the Certificate of Proficiency in English meant to be for "foreigners" who sought proof of their practical knowledge of the language with a view to teaching it in foreign schools (p.49-51).

Later America and Australia adopted different test approaches for testing the language proficiency of learners. Modeling in second or foreign language testing is not a recent concept. Theorists and language teaching methodologists have developed and used language testing models, which matched the language teaching models existing in a particular point of time. How it has been changed differently over time can be discussed as follows.

The pre-discrete point approach to language testing was based on the assumption that no special expertise was required for testing. Language testing was viewed as entirely subjective judgment of the examiner.

The discrete point approach to language testing came into existence when pre - discrete point approach was highly criticized for not being reliable and valid. This approach was based on the assumption that language is a set of habits. This approach put emphasis on reality more than needed, may at the cost of validity. The aspects of language to be tested are short answers questions and multiple choice items. This approach became the most widely used tests around the world in 1960s to 1970s and still popularly practiced in many parts of the world (Giri, 2003, p.55).

Integrative test approach came into existence when discrete- point test could not be far from criticisms. It is based on the assumption that knowledge of language is more than just the sum of a set of discrete parts, and also on the belief that integrative tests attempt to assess a learner's capacity to use many bits all at the time. Close test and dictation were exercised during this period.

When close test and dictation were criticized for not being communicative, the functional approach to testing emerged. "It was argued that the knowledge of language is best captured by detailing the various uses to which the language can be put" (Giri, 2003, p.60). This approach to testing is based on the assumption that knowledge of language should be viewed in terms of language related functions but not in terms of underlying grammatical structures. This approach places importance on performance rather than the linguistic ability of the examinee. However, this approach could not satisfy the thrust for

determining what is to be tested. Consequently communicative approach came into existence.

Communicative approach views that language test should not evaluate in terms of the learners knowledge of the elements and skills, but in terms of their ability to comprehend and produce utterances that are both situational and contextual appropriate. Hymes (1972) developed the first model of communicative test which consisted of both linguistic and sociolinguistic elements. Then several models of communicative language teaching and testing appeared in Europe as well as in North America by different scholars in different time (as cited in Hamal, 2007, p.12).

1.1.2 Achievement Test

Achievement test is also called summative test. Simply speaking an achievement test looks back on what should have been learnt. Such tests are directly based on predetermined course. Most annual examinations take the form of achievement tests. All public tests which are intended to show mastery of a particular syllabus are also called achievement tests. Thus, an achievement test indicates how much of a syllabus has been learned / taught. Richards et al. (1993) defines achievement test as:

A test which measures how much of a language someone has learnt with reference to a particular course of study or programme of instruction...It is based on curriculum. It is administered to find out how much of language course has actually been mastered (as cited in Hamal, 2007, p.7).

Achievement test can further be divided into class progress tests and Standardized (Achievement) tests .The class progress test is a teaching device which enables the teacher to become more familiar with the work of each

student and with the progress of class in general. Its backwash effect on teaching and motivation being important features. Heaton (1988) says:

The class progress test is designed to measure the extent to which the students have mastered the materials taught in the classroom. It is based on the language programme, which the class has been following and is just as important as an assessment of the teacher's own work as the student's own learning... a good progress test should encourage the students to perform well in the target language and to gain additional confidence (p.172)

The main aim of class progress test is to stimulate learning and to reinforce what has been taught. Several achievement tests on the other hand are standardized. A good achievement test should reflect the particular approach to learning and teaching that has previously been adopted.

1.1.3 Language Skills: Introduction

Mainly there are four basic skills of language they are as follows:

- a. Listening Skill
- b. Speaking Skill
- c. Reading Skill
- d. Writing Skill

However, this research study is mainly concerned with reading and writing so I attempt to discuss on these two skills only.

a. Reading Skill

Reading is the receptive skill in the written mode. Reading can be defined as understanding or making sense of given text. The reader should actively involve in reading text to get the meaning. It can help to build vocabulary that helps listening comprehension at the later stages particularly. According to Sharma & Phyak (2007) the following are sub-skills involved in reading:

- i. Recognizing the letters by the alphabet.
- ii. Reading groups of letters as words
- iii. Understanding meaning of punctuation and vocabulary items.
- iv. Distinguishing main ideas from supporting details.
- v. Recognizing indicators or markers in discourse.
- vi. Recognizing the organization of the text and the effect of style etc.(p.33)

b. Writing Skill

Writing is defined as the use of the visual medium to manifest the graphological and grammatical system of the language. It is the most complicated or complex skill. It comes under productive skill which involves manipulating, structuring and communicating.

Heaton (1975) identifies following sub skills involved in writing:

i. Grammatical Skills

It is the ability to write grammatically correct sentences.

ii. Stylistic Skills

It is the ability to manipulate the sentences and language effectively.

iii. Mechanical or Graphological Skills

The ability to use the spellings, punctuation marks, capitals, abbreviations effectively,

iv. Judgmental Skills

The ability to write in an appropriate manner for a particular purpose with a particular audience in mind, together with ability to select, organize and order relevant information (p.135).

1.1.4 ELT in Nepal

Formally, the English language entered in the Nepalese Education in 1854 when the Prime Minister Jung Bahadur Rana opened a high school in Katmandu. However it was not introduced in the higher education system until 1918 when tri Chandra college, the first college in the kingdom was established. Awasthi (2003) writes in his articles:

The introduction of ELT in Nepalese education started only in 1971 with the implementation of National Education System Plan (NESP). Later all round national education commission recommended English as a subject which to be included in the primary level from class three. The three year B. Ed. programme run by T.U and P.U. differs in terms of the weightage given to the courses. T.U. offers a balanced combination of theory, practice and pedagogy whereas the courses offered by P.U. seem to be much more theoretical. In case of one year B. Ed. the courses offered by MSU and KU seem to be heavier than the courses offered by T.U...At the Master's level T.U. is the only university running two years M.Ed. with a focus on ELT and Applied Linguistics. The one year M.Ed. KU mainly concerns with some linguistic input for the people who have an MA in English literature (p. 17-25).

In present time English is taught from class one to bachelor level as a compulsory subject. But, in master's level it is taught as a specialized subject. The course includes four language skills (Reading, speaking, listening, writing), language functions, and language aspects from school to Bachelor level. However, it is not adequate. Most of the private schools have started using English as a medium of instruction but the situation of government aided schools have not been changed. In our context, most of the learners take English as a subject which they can hardly pass or learn. In school level writing has got more priority than reading, listening and speaking. This condition is same in higher level too.

The assessment system of our country is exam oriented. The continuous assessment system (CAS) has not been in practice yet. Lack of trained teachers, supplementary materials, sophisticated libraries, financing and management system create hindrance for professional development in ELT. So the examination system should be changed. It should be practical than theoretical.

1.1.5 Strengths and Weaknesses of Written Tests

Written test demands the response in the form of writing. The Examinees should write their answers using pen and paper in written test. The answer papers in written test are collected and marked later to evaluate. Anything has its two sides, positive and negative aspect .There are both strengths and weaknesses of written tests, which are as follows:

a. Strengths

- i. Written tests are useful to measure both what and how aspect.
- ii. Essay type questions are useful for testing higher mental process.
- iii. Essay type questions can test personal ideas, thoughts, feelings, experiences, and creativity of students.

- iv. These types of tests help to determine specific areas of weaknesses of learners. So, that the teacher is able to plan the most remedial programmes.
- v. Written tests help to evaluate the effectiveness of curriculum, course book, and textbook, teaching materials, methods and evaluation process.
- vi. It motivates students towards their study.
- vii. It helps to find out what is taught and determines what should be taught.
- viii. It is useful for shy students too.
- ix. It is easy in administration.
- x. The learner does not feel frustration and hesitation in written tests etc.

b. Weaknesses

- i. Written test does not evaluate the individual's performance.
- ii. Written test only focuses in developing writing and reading skills but not in speaking and listening one.
- iii. Written test does not help to teach pronunciation or to produce correct form of a language.
- iv. Written test demands writing and therefore handwriting, style and common errors in structure may exert influence to score.
- v. Written test may have less validity than oral one.
- vi. Written test demands experts and takes longer time to score.
- vii. Only written test is not adequate to diagnose strengths, weakness, and difficulties of learners.

viii. Written test does not help to test the Suprasegmental Features (pitch, tone, intonation, stress, rhythm) and paralinguistic features eg. gestures, postures, facial expression, body language) of an examinee which can convey more meanings what we say or hear.

1.1.6 Language Teaching and Testing

Both language teaching and testing come under language education. Language teaching is an art. It is a highly skilled activity which is learned by careful observation and patient practice. "Generally language teaching refers to the teaching and learning activities in the classroom, whereas language testing refers any means of checking what students can do and can not do with language taught"(Sharma, 2009,p.1). Testing is an instrument for determining the effectiveness of different teaching approach. In broad sense "testing is used as a process of scrutinizing how far learners have learned what the teacher wishes them to learn. It has always been an inherent and inseparable part of teaching"(Jigyashu, 2009, p.1). It is said to be an important part of teaching only because it measures someone's language ability, linguistic capacity, and serves as valuable source of information about how effective the teaching learning activities have been. In this modern age of science and technology testing has obtained many scientific methodology and strategic techniques to be more effective tool. It offers bountiful inputs to the teacher to be aware of the effect of his/her teaching. Hence, testing has emerged as an inseparable part of teaching.

Language teaching and testing function like the combination of pick and shovel to dig deep in to the language education. "Language testing is an integral part of language teaching and both testing and teaching are so clearly interrelated that it is virtually impossible to work in either field without being constantly concerned with the others"(Heaton,1988.viii). So, they are the two sides of the same coin. Without testing, teaching is like the horse without rein. There are two way traffic relationships between language teaching and testing

regarding their influences to each other. Testing without teaching is meaningless and teaching without testing does not guarantee the learning on the part of the students. Moreover, evaluation is one of the most important facts of language education and testing is one of the means of evaluation. We can evaluate the learner's performance, the effectiveness of curriculum, syllabus, textbook, teacher's performance, effectiveness of teaching materials and methods through testing. So, a good means of evaluation should not lead to a separation of language education and testing.

1.1.7 Qualities of Good Test

Any test to be a good one, must have the following qualities. They are as follows:

- i. Reliability
- ii. Administrability /Practicality
- iii. Scorability
- iv. Economy
- v. Wash back effect
- vi. Validity

i. Reliability

It is an important quality of good test, which refers to the degree of consistency with which it measures what is intended to measure. It means the term reliability refers to the consistency of scores or administered within a reasonable time. "Any test to be a valid should be reliable but a reliable test is not necessarily to be a valid one" (Sharma, 2009, p.32). Moreover, a test is said to have reliability if there is consistency in the performance of the examinee. In short, it is defined as consistency of measurement. Following are the common methods of determining reliability.

- Test-retest method

- Alternative forms method
- Split half method
- Internal consistency method

ii. Administrability

The term 'Administrability' refers to the degree to which a test is adaptable in varying situations. "The test should be fairly straightforward to administer. It should be practically fit for the situation so that it can easily administer. For example using video for testing at primary level in present context of Nepal is not practical"(Sharma, 2009, p.35). Administrability also indicates the well and advanced organization of how long the test will take, what special arrangements have to be made, what equipment are needed and how the test materials will be reproduced in quantity.

iii. Scorability

Scorability refers to the quality of a test in which a test is shown to have the ability to qualify the learner's performance in terms of certain marking. Scorability also refers to objectively; the test should be objectively scorable. "A test having scorability will have high relevance, practicality and acceptability"(Sharma, 2009, p.35). Subjective tests have low scorability whereas objective tests have high scorability.

iv. Economy

It is another quality of a good test." Economy is the demand of modern time and one of the principle of science and technology any test to be a good one"(Sharma, 2009, p.35). It should be economical in construction, administration, scoring and so on. It is implausible to spend unnecessary time, money and effort.

v. Wash Back Effect

Wash back effect is also called backwash effect. It is the minor quality of a good test. The effect of test on teaching may be good and bad or positive or negative. "Any test is said to be having good backwash effect if it exerts of good influences on teaching and learning that occurs before the administration of the test"(Sharma, 2009, p.36). So, any test to be a good one, it should exerts good backwash effect and should try to minimize the bad wash back effect.

vi. Validity

The term 'valid' is derived from the Latin word 'validus' which mean strong or computing. Generally the term validity refers to the extent to which a concept, conclusion or measurement is well founded and corresponds accurately to the real world. Validity is the degree to which a test measures what it is intended to measure. Lado (1961) defines validity as "a test which measures what it is intended to measure. It is a test of pronunciation which measures pronunciation and nothing else" (p.30). Richards et al. (1999) defines 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" (p.396). Heaton (1988) writes about validity:

The validity of a test is the extent to which it measures what it is supposed to be measured and nothing else...every test, whether it is 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 intended to measure: to the extent that it measures external knowledge and other skills at the same time; it will not be a valid test (p.159).

Similarly, Harrison (1991) also stands at right behind Hughes in defining the term validity. He defines the term validity as “the validity of a test is the extent to which the test measures what it is intended to measure...the three most important characteristics of a good test are validity, reliability and practicality” (p.10).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 purpose for which it is administered. The common types of validity are as follows.

- i Concurrent validity
- ii. Predictive validity
- iii. Construct validity
- iv. Face validity
- v. Content validity

Validity in language test depends on the linguistic content of the test and on the situation or technique used to test this content. It means if a test measures what is actually supposed to be measured, then it becomes a valid test and the situation of such test becomes its validity.

i. Concurrent validity

It refers to 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 two tests can take a form either a well known test or the rating of the teachers. Establishing concurrent validity of a test is very common. Most test developers would be interested to find out the extent to which the test correlates with some other standardized tests. According to Heaton (1988) the validity is obtained as a result of comparing the results of the test with the results of some criterion measure such as anyone or all of the following:

-) An existing test, known or believed to be valid and given at the same time.
or
-) The teacher's ratings or any other such form of independent assessments given at the same time. or
-) The subsequent performance of the testee on a certain task measured by some valid test. or
-) The teacher's rating or any other such form of independent assessment given later .(p.161)

ii. Predictive validity

Predictive validity of a test is concerned with the extent to which the test can predict the future performance of the testees. This type of validity is established by comparing test results with another criterion such as success in a particular job or in higher education. Heaton (1988) says "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" (p.162). One of the difficulties in dealing with predictive validity is to find a satisfactory criterion with which the exam results are to be correlated. To establish the predictive validity of an exam, some standard measure of performance must be pre-specified. It is this pre-specified performance which serves as a criterion.

Thus, 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. If the exam scores are validated against the criterion after a stated time interval, it is predictive validity.

iii. Construct Validity

Construct validity refers to the extent to which a test represents an underlying theory of language learning. The term 'construct' refers to any domain of

knowledge skills or ability . Davies (1978) says "the construct validity provides a baseline to content and predictive validity"(p.62). Likewise, Heaton (1988) writes about construct validity:

If a test has construct validity, it is capable of measuring certain specific characteristics in accordance with a theory of language behaviors and learning ...This type of validity assumes the existence of certain learning theories or constructs underlying the acquisition of abilities and skills... It can be argued that a speed reading test based on a short comprehension passage is an adequate measure of reading ability and thus has low construct validity unless it is believed that the speed reading of short passages relates closely to the ability to read a book quickly and efficiently and is a proven factor in reading ability. If the assumption is held that systematic language habits are best acquired at the elementary level by means of the structural approach, and then a test which emphasizes the communicative aspects of the language will have low construct validity (p.161).

iv. Face validity

A test is said to have face validity if it looks as it measures what is supposed to measure. Heaton (1988) says

If a test items looks right to other testers, teachers, moderators etc, it can be described as having face validity...Students' motivation is maintained if a test has good face validity for most students will try harder if the test looks sound. ...This is a subject judgment rather than based on any

objective analysis of the test and face validity is often considered not to be a true form of validity (p.159).

Face validity is also called pseudo validity. The importance of face validity can not be underestimated. A test which is considered valid by the learners and teachers can be regarded to have face validity.

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. However, today most designers of communicative test regard face validity as the most important of all types of test validity. Bachman (1989) says "the bottom line in any language testing situation, in a very practical sense is whether test takers will take the test seriously enough to by their best and whether test users will accept the test and find it useful"(p.298). The only way to find out face validity is to ask the teachers and students concerned for their opinions through formally or informally. Anastasi (1982) says "If a test does not have face validity though, it may not be acceptable to the students taking it or the teacher and receiving institution who may make use of it"(as cited in Gurung, 2008, p.11).

v. Content validity

Content validity is mainly concerned with what goes into the test. It is the degree to which the test items represent or cover the whole text or course. Content validity is a non statistical type of validity that involves the systematic examination of the test content to determine whether it covers a representative sample of the behavior. Richard et al (1999) defines content validity as "a form of validity which is based on the degree to which a test adequately and sufficiently measures the particular skills or behaviors it sets out to measure"(as cited in Gurung, 2008, p.11). Heaton (1988) says:

Content validity depends on a careful analysis of the language being tested and particular course objectives. The test should be so constructed as to contain a representative sample of the course. The relationship between the test items and the course objectives always being apparent. The test writer should first draw up a table of test specifications describing in every clear and precise terms the particular language skills and areas to be included in the test (p.160).

Content validity is particularly suitable for an achievement tests. They are concerned with previous learning of a syllabus. Davies (1978) writes:

Content validity evidence involves the degree to which the content of the test matches a content domain associated with the construct. For example a test of the ability to add two numbers should include a range of combination of digits. A test with only one or even digit numbers would not have a good coverage of the content domain. Content related evidence typically involves subject matter experts evaluating test items against the test specifications (p.62).

After analyzing the above definition we came to the conclusion that content of a test should be decided by considering the purpose of the assessment to conduct content being specification procedure. The greatest a test's content validity the more likely it is to be an accurate or measures what it is supposed to measure. To achieve this validity the language or materials being tested and the particular course objectives should be carefully analyzed by the expert or researcher. Moreover, the relationship between the test items and the course objectives should be apparent. A test to have good content validity, the test must reflect not only the content of the course, but also demonstrates the

balance of the test items in terms of weightage given to each unit or area. Content coverage and content relevance are the two important components of content validity.

a. Content Coverage

It is really difficult to cover all the contents from the syllabus or course book within three or four hour. So the selection of task to be included in the test is indispensable. Brown (1994) says "...The basic question is therefore whether or not the test items that compose an exam constitute an appropriate representative sample of Behavioral domain under consideration"(p.124). According to Hughes (1989) "Content validity depends on how many of the functions are tested in the component and how representatives they are of the complete set of functions included in the objectives"(as cited in Gurung, 2008, p.12). When a test has content validity, the items on the test represent the entire range of possible items the test should cover. Individual test questions may be drawn from a large pool of items that cover a broad range of topics. That is why, the question designer must as clearly as possible specify what skills the exam is designed to cover. A test can not be fully valid or invalid. So, a test is judged as having content validity when the test items represent the course contents and the course objectives. Content coverage is amount or the way that the subject matter of the course covers an area.

b. Content Relevance

Content relevance means whether or not the tests tasks included in tests are relevant to the language activities that are expected to be exercised under the given course. According to Lado (1961)"a test to be a valid we accept the content and conditions to be relevant and that there will be no irrelevant problems which are more difficult than the problems being tested"(p.119). The content validity of a test is examined in relation to its relevance to the given course of study. The content relevance involves the specification of the ability domain and the test method facet. Moreover, the amount or the way that

subject matters are closely related with the course is content relevance. The more test items, the more content validity. In other words if more of the contents are covered in test paper, certainly there would be more content relevance.

1.1.8 Course Overview: Reading Writing and Critical Thinking

New Direction: Reading Writing and Critical Thinking is a recommended book for B. Ed first year students. It is written by Peter, S. Gardner and published by Cambridge University Press, New York in 2005. This course is revised in 2009 which offers a number of challenging reading and writing activities that encourage the higher order thinking skills of analysis, synthesis, interpretation, evaluation and application necessary for academic success .Through integrated reading, writing ,speaking and listening activities, students learn to generate hypotheses, argue, analyze critically, interpret a writer's meaning, discriminate between opinion and fact, reach conclusions and judgments based on supportable criteria, and propose new ideas. The book also stresses the development of students' academic vocabulary. More information about the book can be given in the following points.

a. Purpose and Audience: New direction is written for students who are studying at college/ university. It may be used by advanced ESL or EFL learners who wants to improve his/her reading writing and critical thinking skills.

b. Writing skills: New Directions takes students through the major stages of the writing process (assessing the writing situation, exploring and planning, drafting, revising, editing and proofreading). The book includes strategies that will help learners compose expository and argumentative essays and use sources effectively in research writing. It provides many opportunities for formal and informal writing, including journal entries, free writing, summaries, reports, and essays.

c. The Readings: New Directions contains 35 readings of varying length and difficulty, though all are appropriate for the college level. The selections are interdisciplinary and include excerpts from college texts and nonfiction books, newspaper and magazine articles, personal essays, letters, short stories, folktales, fables and poems. The reading texts are written by various authors from wide range of cultural backgrounds which helps to promote a stimulating context for developing reading, writing, and critical thinking skills.

d. Structure of the Book: New Direction is divided into five thematically based chapters. An informational section on writing skills, titled "The Essential of Writing" is placed between chapters 2 and 3

i. The Thematic Chapters: The five chapter themes in the book were selected because of their relevance and interest to students. The five chapters and the five writings techniques are as follows.

Chapter 1: Intercultural Communication

Main ideas and supporting details

Chapter 2: Education

Purpose and audience

Chapter 3: Mass Media and Technology

Figures of speech

Chapter 4: Gender Roles

Summarizing and paraphrasing

Chapter5: Work

Tone

ii. The Essentials of Writing: Beside the thematic content, each of the five chapters also focuses on an important writing technique. The writing section appears between chapters 2 and 3. It contains essential information about writing and plentiful examples. The writing section is in three parts:

Part One: The structure of an Essay

The introduction, Body paragraphs, The Conclusion

Part Two: The Writing Process

*Assessing the Writing Situation, Exploring and Planning,
Drafting, Revising, Editing and Proofreading*

Part Three: Writing with Sources

*Types of sources, Locating sources, Evaluating Sources,
Taking Notes from sources, Documenting Sources.*

e. The Organization of a Chapter: Each chapter has an opening activity which raises student's awareness to the chapter theme a series of core readings, with pre and post reading activities for each.

i. Unit Opener: To raise the awareness and interest to the chapter topic, each chapter begins with a brief introduction describing the major theme, followed by a list of question and quotations relating to the readings .Here students discuss the questions and quotations in a small group and share personal opinions and experiences.

ii. Core Readings: Each chapter has three core reading containing a balance of academic and personal writing. The core reading has a full range of pre and post reading activities which are journal writing, previewing the topic, agreeing and disagreeing, taking notes while reading, reading journal, main ideas, reflecting on content, a writer's technique ,vocabulary ,vocabulary in context, discussion, writing follow-up etc.

iii. Making Connection: This section may be used as the basis of a writing or discussion activity. The question help students to synthesize the information presented in the core reading. Students compare and contrast the author's ideas and writing techniques, imagines, how the authors might respond to each other.

iv. Additional Reading: Each chapter contains five addition reading (two proses, a story, a poem, a joke) and a cartoon. These reading have only a few accompanying tasks. The additional material allows students to explore issues

in greater depth, to apply the writing technique taught earlier in the chapter, and to practice their intensive and extensive reading skills.

v. Essay Topics: This section includes five essay topics of each chapter that require students to use various rhetorical strategies, such as compare and contrast, cause and effect, and division and classification, and higher order thinking skills, especially synthesis, interpretation, and application. The assignments include expository, argumentative, and narrative writing for each assignment. Here students support their ideas with references to the chapter readings, library and internet sources, and personal experience.

f. Some Important Activities Accompanying the Core Reading: The core readings have full set of post and pre reading activities which are described below.

i. Reflecting on Content: Here, students think critically about major issues raised in the reading, relating them to their own knowledge and experience. These questions help students develop the skills of analysis, synthesis, evaluation, interpretation, inference, to support writer's ideas.

ii. Main Ideas: This section helps students to understand the main ideas by asking the questions to summarize the main idea of the reading in one or two sentences.

iii. Reading Journal: After reading a selection, students write an entry in their reading journal. These entries include discussions of specific topics relating to the reading, interpretations, points of agreement and disagreement with author, likes and dislikes, and personal experiences.

iv. Taking Notes While Reading: While reading the selection, students are asked to underline, highlight the ideas in the piece, often indicating points of agreement and disagreement and relationships to their own culture and experience.

v. Journal Writing: Students write journal entries about topics relating to the readings, focusing on personal opinions and experiences. They are often asked to respond to a brief quotation, agree or disagree with statement, or make a prediction about an issue raised in the reading. Here students also do practice the pre-writing techniques (brainstorming, free-writing, clustering, the journalist's questions).

vi. A Writer's techniques: In this section Students learn about the major writing technique introduced in the chapter, focusing on the ways in which the author of the reading uses the strategy. Students think critically about such of writing as main ideas, supporting details, purpose, audience, and tone.

vii. Writing follow-up: In this section the students are encouraged to write a summary, reports, letters, and case studies.

1.1.9 Reading Vs Writing

Reading is the receptive skill in the written mode. It can be defined as understanding or making sense of a given text. In other words, reading refers to information and increasing one's professional knowledge from the graphic symbols. The reader should actively involve in reading text to get the meaning. It can help to build vocabulary that helps listening comprehension at the later stages. Reading is a process that involves recognizing graphic symbols, vocalizing them and getting the message that the writer has expressed by means of these symbols. Reading without understanding is just barking at prints. So, reading involves both understanding and comprehension. It involves various aspects and sub-skills in various stages:

i. Mimicry: The term 'mimicry' refers to the action of imitation. It is the first stage in which students are given practice in recognizing and vocalizing words by imitating their teachers.

ii. Controlled Reading: At this stage of reading, students are helped with reading for meaning. They are expected to read the letters aloud to get the meaning.

iii. Guided Reading: At this stage the teacher exercises less control and students are given more freedom to interpret the text they read. Texts used for this purpose are simple narratives and conversational materials which can develop uncomplicated and entertaining themes.

iv. Intensive Reading: Intensive reading plays a vital role for further progress in language learning under the teacher's guidance. It is done not only for the detailed comprehension of the ideas, feelings and language context, but also for mastering the structure and vocabulary.

v. Extensive Reading: Extensive reading helps to equip students with different sorts of information. The main aim of extensive reading is to create interest in students for reading. It is also known as reading for fluency.

vi. Reading for Pleasure: At this stage, students can read variety of reading materials to meet their requirements especially; literary texts (poems, essays, novels, articles, letters etc) are the best materials to have reading for pleasure.

vii. Skimming: At this stage, we go through the reading materials quickly in order to get the gist of it, or to know how it is organized and to get an idea of the intentions/ attitudes of the writer.

viii. Scanning: At this stage we only try to locate specific information. It may be a name, a date or some other specific piece of information.

Writing on the other hand, is one of the most important skills in learning a language. It comes under productive skills. Writing conveys meaning through the use of graphic symbols that represents a language. The job of writing is not an easy. While we do the activity of writing, we write letters, words, phrases, sentences, paragraphs, and varieties of compositions such as essay, story,

poem, dialogues, letters, reports, notice etc. Through writing we can communicate with one another, and transmit our accumulated culture from one generation to another. Writing is a very complex process requiring many composite skills viz. mental, psychological, rhetorical and critical. According to Bhandari (2007, p.1) while writing a piece of language for purpose the following process and skills are essential:

- i. Collecting the ideas.
- ii. Putting the ideas in order.
- iii. Analyzing and synthesizing the ideas.
- iv. Modifying the ideas with suitable modifiers.
- v. Selection of right words.
- vi. Consideration of grammar.
- vii. Consideration of spelling and punctuation.
- viii. Drafting.

According to Gardner (2009) an experienced writer most passes the following stages involved in writing:

i. Assessing the Writing Situation: Reflecting on the subject, ones' attitude towards the subject, purpose, audience, source of available information, and writing assignments.

ii. Exploring and Planning: Discovering, refining, finding support for and organizing ideas.

iii. Drafting: Expressing and developing, ideas and supporting details in rough form.

iv. Revising: Rethinking and rewriting drafts to improve the content, focus, and structure.

v. Editing and Proof Reading: Checking for effective word choice and sentence structure and for correct grammar, spelling, punctuation, and mechanics (p.111).

The common stages involving in writing are as follows:

a. Copying: Copying is essential and useful for the recognition and production of shapes of letters for improving handwriting and spelling .In this stage, the students are required to copy the model given by the teachers or textbooks.

b. Reproduction: Reproduction helps students to improve their memory. At this stage, students are required to write sentences what they have mastered.

c. Recombination: At this stage, students are required to write sentences which they have learnt previously with slight changes. Completing sentences, using alternative words given in the test, transforming sentences on the basis of given clues etc are some examples of recombination. These exercises are primarily for the manipulation of grammatical patterns rather than for the communication of new ideas.

d. Guided Composition: In this stage students are required to write sentences in combination to produce a text, keeping to the subject matter and guidelines given by the teachers, such as transforming information from a chart, table, graph etc into orthographic texts.

e. Free Composition: This is the final stage of the writing. At this stage students should be required to express their ideas, thoughts freely with suitable reasons. This stage comes after long practice. It requires a careful planning and systematic writing process.

1.1.11 Critical Thinking

Critical thinking is a broad and abstract concept. It is an art of analyzing and evaluating thinking with a view to improving it. It is self-guided, self-disciplined thinking which attempts to reason at the highest level of quality in a

fair-minded way. The Critical Thinking Program in Nepal is being implemented by Alliance for Social Dialogue /Social Science Baha with the support of Education Support Program of Open Society Institute New York. The Critical Thinking program started in Nepal in December 2008 with the visit of ESP program officer Svetlana Batrak along with International critical thinking trainers Nicu and Veronica. During their visit they conducted a demonstration workshop on 'Critical Thinking Methodology' and consulted different stakeholders working in the education sector to assess the feasibility of implementing the critical thinking program in Nepal.

By presenting the paper at the 8th Annual International Conference on Critical Thinking and Education in New York, Scriven & Paul (1987) say:

Critical thinking is the intellectually disciplined process of skillfully conceptualizing, applying, analyzing, synthesizing, evaluating information gathered/generated by observation, experience, reflection, reasoning, communication, as a guide to belief and action... It is based on universal intellectual values that transcend subject matter divisions: clarity, accuracy, precision, consistency, relevance, sound evidence, good reasons, depth, breadth, and fairness...It is believed the people who think critically consistently attempt to live rationally, reasonably and empathically. They are keenly aware of the inherently flawed nature of human thinking when left unchecked. They strive to diminish the power of their egocentric and sociocentric tendencies. They use the intellectual tools that critical thinking offers concepts and principles that enable them to analyze, assess, and improve thinking. They work diligently to develop the intellectual virtues of intellectual integrity,

intellectual humility, intellectual civility, intellectual empathy,
intellectual sense of justice and confidence in reason
([http://www.critical thinking.org](http://www.criticalthinking.org)).

Critical thinking has great scope in educational system, mainly in teaching learning field. It is not new topics for teachers "... All the teachers who are trained or hold degrees in education, they have knowledge of problem solving methods, deductive and in-deductive methods, and cooperative learning methods and these all methods come under critical thinking"(Allister,2009,p.4). If critical thinking strategies are adopted in the classroom teaching, it obviously promotes active inquiry, student-initiated learning, problem solving, cooperative learning, critical thinking, writing and reading processes and alternative assessment.

However, the existing pedagogical practices in classroom teaching have been limited only to providing more information to students about the content. Some teachers think that it is difficult to implement such strategies in classroom because of crowded classrooms, lack of necessary materials, and the infrastructures, limitation of allocated time and evaluation system etc.

In case of B.Ed. level, the curriculum includes reading writing and critical thinking as a specialized subject. However, the course is same to the other courses in terms of nature and allotment of time frame. The evaluation system is also the similar, which does not help to develop critical thinking of students. It is clear that the existing curriculum, text books, supplementary materials, contemporary political situations, lack of teachers training, lack of access in science and technology, examination system, previous back ground of students etc. are the influencing factors for critical thinking in the classroom teaching.

This is the era of science and technology i.e. Computers, E-mail, and Internet. Due to the globalization of nations and markets, emergence of entirely new fields of scholarly investigation make life more complicated. So we should be

open-minded from a variety of perspectives, often global and cross-cultural. The essence of any critical thinking strategies is to active engagement of students in learning. There is a need to revise pedagogical practices in classroom teaching to prepare competitive students .Providing critical thinking training for all the teachers is not sufficient, however, the positive attitude of teacher to implement critical thinking strategies is much more important than other aspects. So, the vital aspect is to change the attitude of teachers towards critical thinking strategies to implement in classroom situation. Therefore, the most successful classrooms are those that encourage students to develop critical thinking skills and power to solve the problems of everyday life.

1.2 Review of Related Literature

Many researches have been carried out in the field of language testing and content validity which were related for this study are reviewed here.

Giri (2003) summarize in his article **Language Testing: Then and Now** that the trends in the second or foreign language testing may be described as analytical, global, communicative. Almost two decade ago, discrete-point and integrative test approach was viewed necessary. But today communicative approach to language testing is becoming increasingly popular. It is assumed that no single language testing model or test can accommodate the wide variety of test scenario .The present research ,so is an attempt to justify that an adaptation or combination of models or tests in different language areas and target situations should be made available in order to meet the local language testing requirements.

Similarly, Batala (2004) conducted research entitled **A Study on the Validity of the SLC Examination English Question Papers** .The objectives of his research was to examined the predictive and content validity of the SLC examination English Question paper. He administered three different sets of questionnaires, one to the secondary level English teachers, the next the students and the last to the question setters. He found out that SLC examination

English question paper had very low predict validity. The coefficient of correlation between two sets of scores on the SLC and grade II English examinations was +0.1. He also found that the content representation of the test is high (72.86 percent). In terms weighting, the marks weighting to each skill in the test paper is now proportional to the weighting of its course contents.

On the other hand, Neupane (2004) conducted research entitled **Wash back Effect of Examination : A Case of Communicative English**. The objectives of the study were to find out the wash back effect of the examination of Communicative English and to analyze the test paper of communicative English in term of their validity and practicality. She has used three set of questionnaire, one for subject teachers and two for students (one set for first year students and another for second year students) to collect the data. She concluded that the test of communicative English had negative wash back effect as the students could pass the examination without meeting the objectives of the course. The test was not relevant to the course objectives and the test item did not represent the course objectives too. The tasks to be performed in answering the question did not resemble the tasks which the students face in the real situation.

Similarly, Aryal (2005) studied on **Quality of English Exam: A Case of Content Validity of Grade Twelve Compulsory Exam 2061**. He has analyzed the question papers from the view points of rubrics, length difficulty level and content coverage. He found out that the rubrics of the questions were simple, scientific except few items; the length of the question matches to the allotment of time for the examinees; the questions of English are of moderate difficulty level; the question relation to Heritage of Words lack content validity and are unscientific. They do not cover all the genres equally. The content validity of Meaning in to Words is nearer to the coverage of units and teaching items in comparison with Heritage of Words.

Similarly, Ojha (2005) conducted research entitled **Content Validity of ELT Theories and Methods Exam at B. Ed Level**. The objective of his research was to examine the content validity of ELT Theories and Methods tests at B. Ed. second year in terms of content coverage and content weighting. To fulfill the objectives, he analyzed the test paper administered in the annual examination from 2057 to 2061. The finding of his research work is that in terms of content coverage, the test paper has good content validity (57.61 percent). But in terms of content weighting the test papers have low content validity because the question did not show the expected tendency to follow the distribution of marks specified by the evaluation scheme of the syllabus.

Similarly, Nepal (2006) carried out a research work on **Content Validity of Examinations: A Case of Fundamentals of Language and Linguistics at B. Ed Level**. The objectives of the study were to examine the content validity of fundamentals of language and linguistics at B. Ed first year level. He has used only secondary sources of data. It means the collection of old questions of above course administered during 2057 to 2062 were the main tools of his study. He found that the content validity was high (80.28 percent) in terms of content coverage but was low content validity in terms of content weighting.

Similarly, Subedi (2006) carried out a research work on title **Content Validity of B. Ed Questions: A Case of English Sound and Structures (302)**. The main objective of the study was to find out the content validity of B. Ed questions of English Sounds and structures (302). He further aimed to compare the questions of different years in terms of their content validity. He has used a set of questionnaire for 20 subject teachers around Katmandu, Lalitpur, Bhaktapur and Kabre districts. He followed purposive sampling procedure to specify the sample population of study. He found that in term of content Coverage the question paper of eight years (2054 to 2061 B.S) has 66.75 percent valid in average whereas 93.12 percent was valid on the basis of the distribution of weightage.

Similarly, Hamal (2007) conducted a research entitled **Content Validity of Compulsory English Test Items at B. Ed. Level: A Case of Reading and Writing**. The main objectives of the study was to examine the content validity of Reading and Writing test items at B. Ed. level first year in terms of content coverage and content weighting. The test items designed to test reading and writing skills from the year 2059 to 2063 were the main sources of secondary sources of data. To accomplish this objective, the researcher has compared the test contents in relation to course contents. For this purpose, the researcher has used only descriptive method and tabulation method to analyze the data. He concluded that according to the coverage/representative principle 'Reading and Writing' tests have low content validity. Because there are altogether 31 sub-skills as prescribed in the syllabus to develop reading and writing skills. Among which only 7 sub-skills have been represented in the question papers of 5 academic years. But in terms of content weighting that reading and writing tests have the required content validity as it has followed the norms and tendency in the distribution of marks in the test papers what it was given in the weighting schedule of the syllabus.

Similarly, Koirala (2008) conducted a research entitled **The Content Validity of English Textbook for Grade VI**. The main aim of his study was to find out the content validity of English Textbook for grade six in terms of general aspects of the content, language skills, language functions, and Language structures. The primary sources for collecting the data were the responses made by the lower secondary level English teachers teaching in Katmandu valley, the experts, related to the field of ELT and analysis of the textbook made by the researcher himself. To achieve above objectives, the researcher randomly selected thirty lower secondary English teachers with at least one year teaching experience at lower secondary level, and five experts who are concerned in the field of ELT. He prepared both close ended and open-ended questions for the target population. He asked forty-two close-ended and three open-ended questions for the teachers and twenty-one close-ended and two open-ended

questions for the experts. He concluded that the book has content validity in terms of speaking exercise, language structure, selection and gradation, but in terms of language functions, listening exercises, Interest and level of the students it has less valid .

Gardner (2009) summarizes that his textbook **New Directions: Reading Writing and Critical Thinking** includes five chapters which offers a number of challenging reading and writing activities that encourage the higher-order thinking skills of analysis, synthesis, interpretation, evaluation and application necessary for academic success. .Through integrated reading ,Writing ,speaking and listening activities, students learn to generate hypotheses, argue, analyze critically, interpret a writer's meaning inferentially as well as literally , discriminate between opinion and fact ,detect fallacies in reasoning, reach conclusions and judgments based on supportable criteria ,and propose new ideas.

The **Syllabus of Reading Writing and Critical Thinking (2066)** includes three units. Unit one deals reading which takes 60 periods and carries 30 marks. The main aim of this unit to read wide variety of reading materials and comprehend/ analyze them. Unit two deals writing which takes 60 periods and carries 50 marks. The main aim of this unit is to develop writing skills of students by involving different types of activities. Unit three deals critical thinking which takes 30 period and 20 marks. The main aim of this unit is to enhance the student's abilities to argue with reasons, analyze and write critically. It is revised and specialized course for B.Ed first year students. It has theoretical nature and carries 100 marks with subjective and objective questions.

However, none of the researches has been carried out on Content Validity of Examination Papers: A case of Reading Writing and Critical Thinking Course. So, this study has been the first work to determine the content validity of the Year End Examination of B. Ed. First Year Course in the Department.

1.3 Objectives of the Study

The objectives of the study are as follows:

- i. To determine the content validity of the Year End Examination of B. Ed. first year Course (Reading Writing and Critical Thinking, Eng. Ed. 317) .
- ii. To suggest some pedagogical implications.

1.4 Significance of the Study

The research findings will be beneficial to those who are involved in teaching and testing. Moreover, it will be equally useful to the curriculum designer, text book writers, subject experts, language teachers, learners, question setters, policy makers, and those persons who are directly or indirectly involved in the field of teaching learning and testing.

CHAPTER -TWO

METHODOLOGY

The researcher followed the following methodology during her study in order to achieve the specific objectives. This section includes sources of data, tools for data collection, process of data collection and limitations of the study.

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 used only the subjective question papers of the 'Reading, Writing and Critical Thinking Course of B. Ed. First Year from 2066 to 2067 B.S. Then, she also studied the various books, journals, research reports and articles written by various writers, such as Lado (1961), Davies(1978), Scriven & Paul(1987) , Heaton (1988), Harrison(1991), Brown (1994), Bachman (1998), Giri(2003), Catford (2005), Gardner (2009) related to the topic.

2.3 Tools for Data Collection

The tools for data collection were the question papers of 'Reading Writing and Critical Thinking Course '(Eng, Ed.317) B.Ed. First Year Level, during two years (2066 to 2067).

2.4 Process of Data Collection

The researcher collected only subjective question Papers of Reading, Writing and Critical Thinking (Eng.Ed.317) of B.Ed. first year from 2066 to 2067 B.S. Then she has analyzed the question papers as to whether they are valid or not in terms of content coverage and content weighting based on the course

objectives, course contents, test contents, evaluation scheme and time allotment etc.

2.5 Limitations of the Study

The study had the following limitations:

- i. The study was limited to determine the content validity of Examination papers of Reading, Writing and Critical Thinking (Eng.Ed.317).
- ii. The study was limited to the subjective question papers of 'Reading Writing and Critical Thinking of B. Ed. First Year Course' asked in annual examination during two years (2066 to 2067).
- iii. The study was limited only to in secondary sources.

CHAPTER -THREE

ANALYSIS AND INTERPRETATION

This chapter deals with the analysis and interpretation of the raw data that have been collected. The main objective of this study is to determine the validity of the Year End Examination of B. Ed Course of 'Reading, Writing and Critical Thinking' in terms of content coverage and weighting. The researcher has analyzed content coverage and weighting of the questions with her own judgment. For this purpose this chapter has been divided into two parts. The first part deals with the analysis and interpretation of the content validity in terms of content coverage and the second part deals with the content validity in terms of content weighting.

3.1 Analysis of Content Coverage

For the purpose of analyzing content coverage of the question papers the researcher has compared the test content with course content. So, the researcher examined the two years' subjective question papers of 'Reading, Writing and Critical thinking' to find out whether or not the test items are the representative sample of the course objectives/contents. The course objectives and contents are given in appendix I. To determine the content validity, the test items and particular course objectives/contents should be carefully analyzed by the expert or researcher. A test to have good contents coverage, it should represent all the content of the course.

3.1.1 Analysis of Content Coverage from Unit One

Table No. 1

Analysis of Content Coverage from Unit One

Unit	Course Contents	Content of the Examination Papers (Subjective only)	
1	Reading	2066 B.S.	2067 B.S.
1.1	Reading with a purpose	no.5(alt)	
1.2	Reading for specific information	no.5(alt)	no. 6
1.3	Reading for general information	×	×
1.4	Reading for main idea	no.5(alt)	
1.5	Reading critically and analytically	no.5 (alt)	
1.6	Reading for pleasure	×	×
1.7	Reading and taking notes	×	×
	Total number of question asked from unit one	1 short question (alt. w.u.)	1 short question
	Total Marks:	7 marks	7 marks

Note: The questions asked in examination are given in appendix I

Where,

alt = alternative

no = number

S.L. = subjective long question

× = the sub- units were not represented

= the sub- units were represented

w.u. = within a unit

A question without bracket is compulsory.

There are altogether 3 units in 'Reading Writing and Critical Thinking Course'.

Unit one includes Reading. Under reading there are 7 language items; reading with purpose, reading for specific information , reading for general information, reading for main idea, reading critically and analytically, reading for pleasure ,reading and taking notes. The specific objectives of this unit are as follows:

- To read and comprehend the purpose to text
- To read and take notes of the important points.
- To comprehend details of the texts.
- To read and analyze the texts.
- To find the main points in the text.
- To identify the general ideas of the texts. etc.

The table given above indicates the representation of units and sub units asked in annual examination (2066-2067). The researcher has analyzed subjective questions only. The question papers of those academic years are given in appendix I .Unit one includes seven different sub-units which are spread over 1 to 1.7. Within two years only 2 questions were asked. In 2066 one short question which carries 7 marks each with alternative was asked from 1.1(reading with a purpose), 1.2 (reading for specific information). 1.4 (reading for main idea), 1.5 (reading critically and analytically).After analyzing the above data the researcher found that the question paper represented 4 sub-units out of 7 in 2066 B.S. by asking 1 alternative question from this unit.

Here,

$$\text{Total area of content} = 7$$

$$\text{Covered content} = 4$$

$$\text{Covered content in percentage} = \frac{4}{7} \times 100$$

$$= 57.15\%$$

$$\text{Uncovered content in percent} = 42.85\%$$

This can be presented in the form of pie chart as given below.

Here,

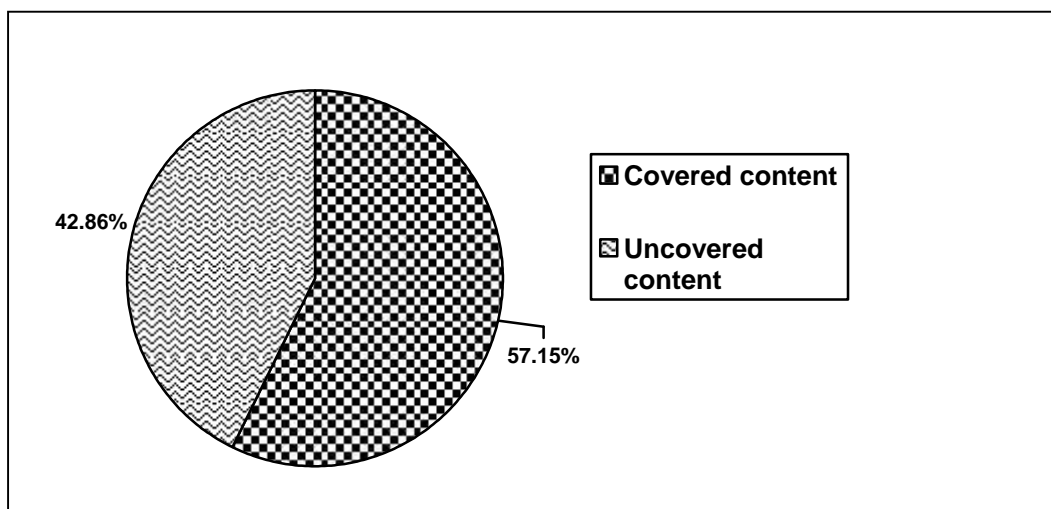
$$\text{Covered content in degree} = \frac{57.15}{100} \times 360$$

$$= 205.72^\circ$$

$$\text{Uncovered content in degree} = 154.28^\circ$$

Figure No. 1

Content Coverage from Unit One in 2066 B.S.



In 2067, one short question each was asked from 1.1 (reading with a purpose), 1.2 (reading for specific information), and 1.5 (reading critically and analytically), whereas 1.3 (reading for general information), 1.6 (reading for pleasure), 1.7 (reading and taking notes) were not represented in both of the

years. After analyzing the above data the researcher found that the question papers had represented 3 sub-units out of 7 from this unit in 2067 B.S.

Here,

$$\text{Total area of content} = 7$$

$$\text{Covered content} = 3$$

$$\text{Covered content in percentage} = \frac{3}{7} \times 100$$

$$= 42.85\%$$

$$\text{Uncovered content in percent} = 57.14\%$$

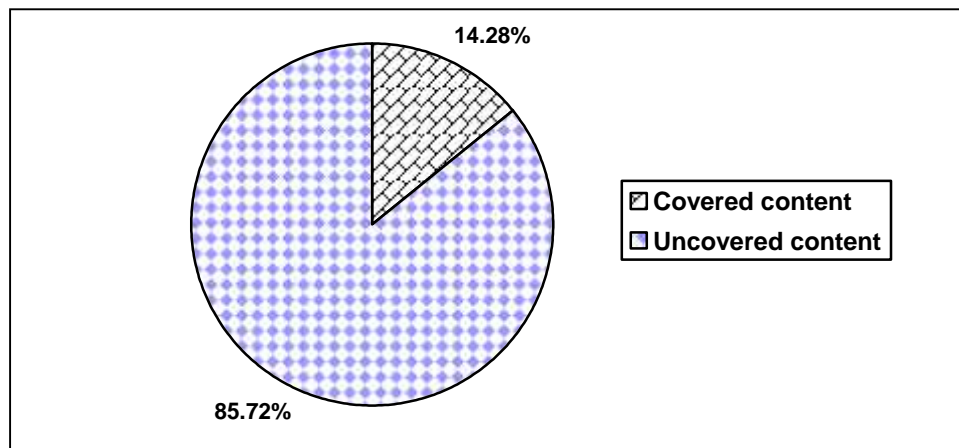
This can be presented in the form of pie chart as given below.

$$\begin{aligned} \text{Covered content in degree} &= \frac{42.85}{100} \times 360 \\ &= 154.28^\circ \end{aligned}$$

$$\text{Uncovered content in degree} = 205.71^\circ$$

Figure No. 2

Content Coverage from Unit One in 2067 B.S.



From the above data and description, it is clear that there are altogether 7 sub-units in unit one to reflect the above course objectives under reading. Among them only 4 sub -units were represented in the examination during two years. However, the language items 1.3 (reading for general information), 1.6 (reading for pleasure), 1.7 (reading and taking notes) were not represented at

all. The data shows that the content coverage of this unit is only 57.15 percent, which shows the lack of content validity in terms of content coverage.

However, we cannot say that it lacks content validity because our examination system is written oriented, which emphasizes the theoretical aspect rather than practical one. The course 'Reading Writing and Critical Thinking' has adopted 100 percent written examination. So, it lacks practical. That is why it can test only writing and critical thinking or analytical skills than reading.

3.1.2 Analysis of Content Coverage from Unit Two

Table No. 2

Analysis of Content Coverage from Unit Two

Units	Course Contents	Contents of the Examination Papers (Subjective only)	
		2066 B.S.	2067 B.S.
2	Writing	2066 B.S.	2067 B.S.
2.1	Purpose and audience		
2.2	Main idea and supporting details	no.7(alt)	no.7, no.8(alt)
2.3	The essential of writing		
2.4	Writing an essay	×	S.L.no.9a(alt)
2.4.1	The structure of an essay	×	
2.4.2	The writing process	×	×
2.4.3	Writing with sources	no. 4	×
2.5	Creative writing		
2.5.1	Different genres of creative writings	no.7(alt), S.L.no.9(alt)	no. 3, no.8(alt)
2.5.2	Writing anecdotes	no.6(alt)	×
2.5.3	Writing stories	×	S.L.no.9b(alt)
2.5.4	Writings poems	no.6(alt)	no.4 (alt)
	Total number of questions asked from unit two	1 long (alt), 3 short (2 alt.w.u.)	1 long (alt. w.u.), 4 short
	Total Marks:	7+(26)	14+(26)

Note: The questions asked in examination are given in appendix I

Where,

alt = alternative

no = number

S.L. = subjective long question

× = the sub- units were not represented

= the sub- units were represented

w.u. = within a unit,

A question without bracket is compulsory

Unit two includes writing. The specific objectives of the course under writing are as follows:

- To write creatively
- To write reports and letters
- To write different types of essays.
- To write narratives and anecdotes.
- To organize main ideas with supporting details.
- To write keeping in mind the purpose and audience etc.

The table given above indicates that in unit two there are altogether 12 language items which are spread over 2 to 2.5.4. The diachronical analysis shows that most representative sub- units were 2.2 (main ideas and supporting details), 2.5.1 (different genres of creative writing), 2.5.4 (writing poems) which were represented in both years. Similarly 2.4.3 (writing with sources), 2.4 (writing an essay) 2.5.2 (writing anecdotes) and 2.5.3 (Writing stories), 2.1 (purpose and audience) were represented in only one year.

If we see synchronically, one long question (alternative) and 3 short questions (2 alternatives) were asked in 2066 B.S. from unit two. One long question which carries 12 marks was alternative within the unit was asked from 2.5.1 (different genres of creative writing). Among the three short questions the first (no.4) was from 2.4.3 (writing with sources), the second (no.7) was alternative within the unit was from 2.2 (main ideas and supporting details) and 2.5.1 (different genres of creative writing) respectively. The last (no. 6) was also optional within the unit was from 2.5.2 (Writing anecdotes) and 2.5.4 (writing

poems) respectively. After analyzing the above data the researcher found that the question paper had represented 8 sub-units out of 12 from unit two in 2066 B.S.

Here,

$$\text{Total area of content} = 12$$

$$\text{Covered content} = 8$$

$$\begin{aligned} \text{Covered content in percentage} &= \frac{8}{12} \times 100 = \\ &= 66.66\% \end{aligned}$$

$$\text{Uncovered content in percent} = 33.33\%$$

This can be presented in the form of pie chart as given below.

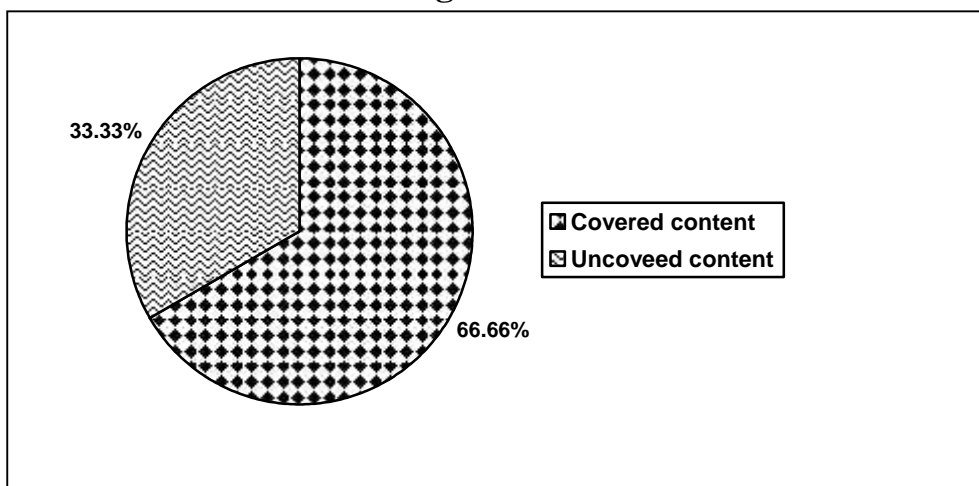
Here,

$$\begin{aligned} \text{Covered content in degree} &= \frac{66.66}{100} \times 360 \\ &= 240^\circ \end{aligned}$$

$$\text{Uncovered content in degree} = 120^\circ$$

Figure No. 3

Content Coverage from Unit Two in 2066 B.S.



In 2067, one long, four short questions were asked from unit two. A long question (no.9) was alternative within the unit asked from 2.4 (writing an essay) and 2.5.3 (writing stories) respectively. In case of four short questions, the first (no.7) was from 2.2 (main ideas and supporting details). The second one (no.3) was from 2.5.1 (different genres of creative writing), the third one (no.4) was alternative asked from 2.5.4 (writing poem), and the last one was

(no.8) was alternative asked from 2.2(main ideas and supporting details) and 2.5.1 (different genres of creative writing). . After analyzing the above data the researcher found that the question paper had represented 9 language sub- units out of 12 from this unit two in 2067 B.S.

Here,

$$\text{Total area of content} = 12$$

$$\text{Covered content} = 9$$

$$\text{Covered content in percentage} = \frac{9}{12} \times 100$$

$$= 75\%$$

$$\text{Uncovered content in percentage} = 25\%$$

This can be presented in the form of pie chart as given below.

Here,

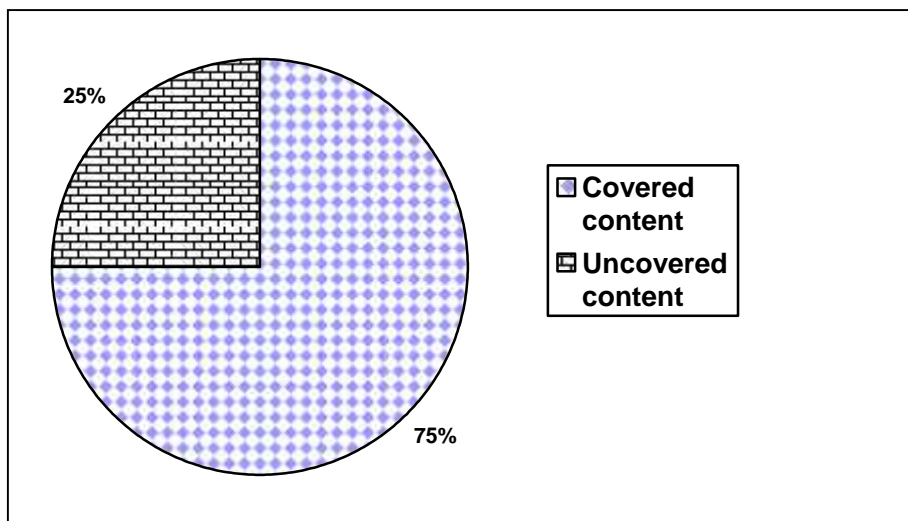
$$\text{Covered content in degree} = \frac{75 \times 360}{100}$$

$$= 270^\circ$$

$$\text{Uncovered content in degree} = 90^\circ$$

Figure No. 4

Content Coverage from Unit Two in 2067 B.S.



3.1.3 Analysis of Content Coverage from Unit Three

Table No. 3

Analysis of Content Coverage from Unit Three

Units	Course Contents	Contents of the Examination Papers (Subjective only)	
3	Critical Thinking	2066 B.S.	2067 B.S.
3.1	Developing a Critical Mind	S.L.no.10	S.L.no.10b
3.2	Analyzing	S.L.no.10	S.L.no.10c
3.3	Arguing	S.L.no.10, no.3, S.L.no.9a (alt),	no.2(alt)
3.4	Reflecting	S.L.no.9a (alt) no.8, no.1, no.2	S.L.no.10a, no.1, no.4(alt) no.8(alt), no.2(alt)
3.5	Reasoning	S.L.no.10, no.3, no.8 S.L.no.9a (alt)	×
	Total number of questions asked from unit 3	2 long,4 short	1 long ,four short(3 alt)
	Total Marks:	52 marks (12 marks as alt)	40 marks (7 as alt)

Note: The questions asked in examination are given in appendix I

Where,

alt = alternative

no = number

S.L. = subjective long question

× = the sub- units were not represented

A question without bracket is compulsory.

Unit three deals critical thinking. The specific objectives of the course under critical thinking are as follows:

- To read and analyze the text critically
- To argue with reasons
- To think independently.
- To debate confidently. etc.

The table given above indicates that there are altogether 5 language sub- units under critical thinking which are spread over 3 to 3.5. Within 2 years (2066 - 2067) almost all the sub-units were represented in the examination. The diachronical analysis shows that the most represented language items were 3.4 (reflecting), 3.3 (Arguing), 3.1(developing a critical mind), 3.2(analyzing) respectively, which were represented in both years. However, 3.5(reasoning) was asked only in one year.

If we see synchronically, 2 long and four short questions were represented in 2066 B.S. from this unit. In case of long questions, the first (no.9) was alternative asked from 3.3 (arguing), 3.4 (reflecting) and 3.5 (reasoning). The second (no.10) was alternative asked from 3.1(developing a critical mind), 3.2 (analyzing), 3.3 (arguing), 3.4 (reflecting) and 3.5(reasoning). Among the four short questions, the first and second one (no.1 and 2) were from the same sub-unit i.e. 3.4 (reflecting), the third one (no.3) was partial from 3.3 (arguing) and 3.5(reasoning).The fourth or last one (no.8) was also alternative asked from 3.4 (reflecting) and 3.5 (reasoning). After analyzing the above data the researcher found that the question papers had represented 5 language sub-units out of 5 from this unit in 2066 B.S. which shows it was very good content validity of question papers in terms of content coverage.

Here,

Total area of content = 5

Covered content = 5

Covered content in percentage = $\frac{5 \times 100}{5}$

= 100%

Uncovered content in percent = 0%

This can be presented in the form of pie chart as given below

Here,

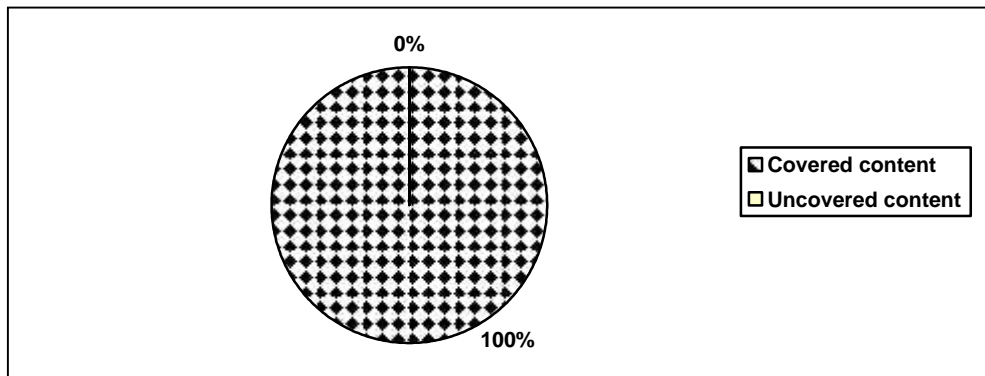
Covered content in degree = $\frac{100 \times 360}{100}$

= 360°

Uncovered content in degree = 0°

Figure No. 5

Content Coverage from Unit Three in 2066 B.S.



In 2067, one long and four short questions were asked from unit three. The long question (no.10) was alternative asked from 3.1 (developing a critical mind) 3.2(analyzing), 3.4 (Reflecting). Among four short questions the first (no.1) was from 3.4(reflecting), the second one (no.2) was alternative within the unit asked from 3.3(arguing) and 3.4(reflecting) respectively. The third one (no.4) was also alternative from 3.4 (reflecting), and the last one (no.8) was

also alternative asked from 3.4 (reflecting). After analyzing the above data the researcher found that the question papers had represented 4 language sub-units out of 5 from unit this unit in 2066 B.S. which shows it was very good content validity of question papers in terms of content coverage.

Here,

Total area of content = 5

Covered content = 4

$$\begin{aligned}\text{Covered content in percentage} &= \frac{4 \times 100}{5} \\ &= 80\%\end{aligned}$$

Uncovered content in percent = 20%

This can be presented in the form of pie chart as given below.

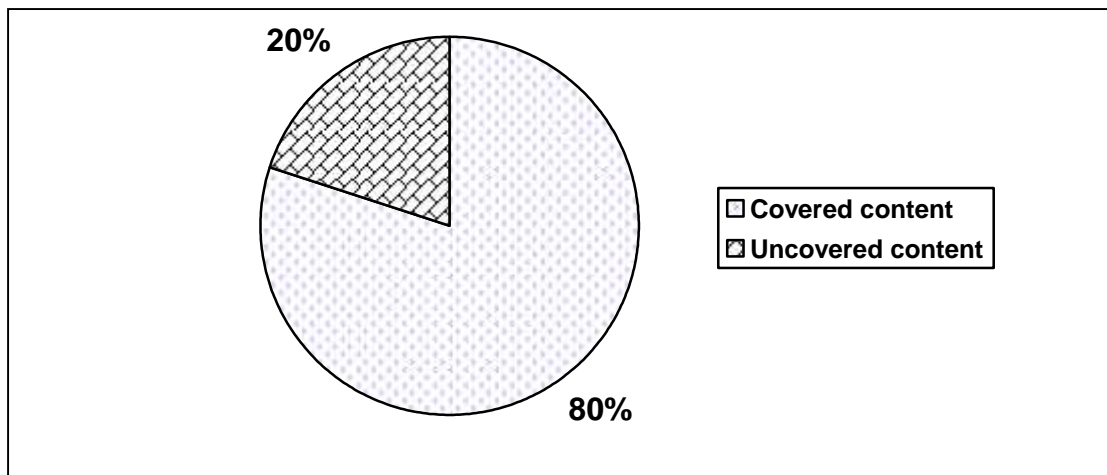
Here,

$$\begin{aligned}\text{Covered content in degree} &= \frac{80 \times 360}{100} \\ &= 288^\circ\end{aligned}$$

Uncovered content in degree = 72°

Figure No. 6

Content Coverage from Unit Three in 2067 B.S.



3.1.4 The Whole Content Validity in Terms of Content Coverage

Table No. 4

The Whole Content Validity in Terms of Content Coverage

S.N.	Units	Sub-Units	Covered Sub-Units		Covered Sub-Units in Percentage		Average Percentage 2066-2067
			2066 B.S.	2067 B.S.	2066 B.S.	2067 B.S.	
1.	Reading	7	4	3	42.86	57.14	50
2.	Writing	12	8	9	66.66	75	70.83
3.	Critical Thinking	5	5	4	100	80	90
Total		24	17	16	69.84	70.71	

The average content coverage during two years = 42.86% +57.14% +66.66% +75% +100% +80% (from unit 1, 2, and 3 in 2067-2067 respectively)

We have,

$$M = \frac{\sum X}{N}$$

Where,

$\sum X$ = sum of items in percentage

N =number of years,

M= mean,

$$= \frac{42.86\% +57.14\% +66.66\% +75\% +100\% +80\%}{6}$$

$$= 70.27\%$$

So, covered content in percentage = 70.27%

Uncovered content in percentage = 29.73%

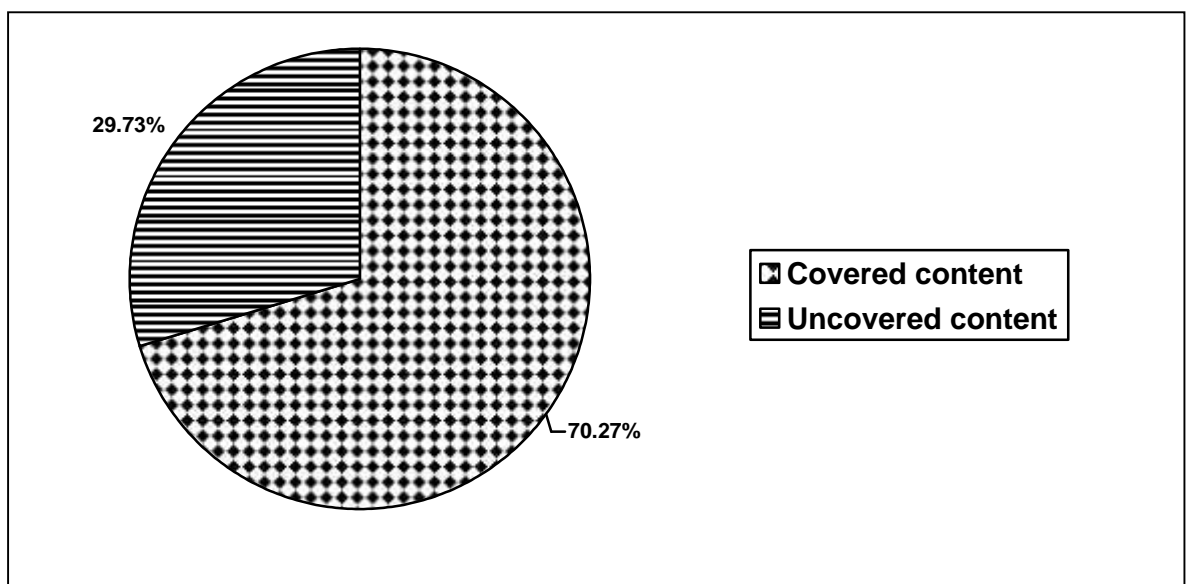
This can be presented in the form of pie chart as given below.

$$\begin{aligned}\text{Covered content in degree} &= \frac{70.27 \times 360}{100} \\ &= 252.97^\circ\end{aligned}$$

$$\text{Uncovered content in degree} = 107.02^\circ$$

Figure No. 7

The Whole Content Validity in Terms of Content Coverage



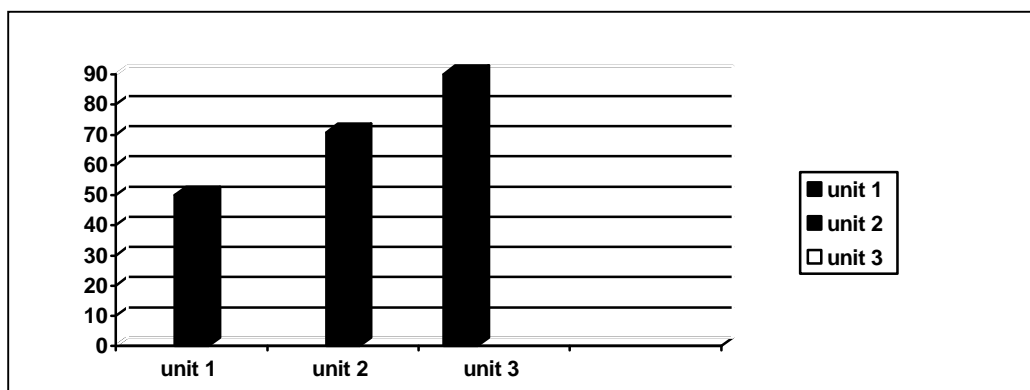
The table given above indicates that there are altogether 24 sub- units in the reading writing and critical thinking course. Among them 16.5 were represented during two years (2066 to 2067). It means the covered content is 70.27 percentages, but 29.73 percentages contents were not represented in the question papers. After analyzing the above given table and pie chart the researcher concludes that the content validity of the question papers on the whole in terms of content coverage was good (70.27%) which is considered as good. So, it is right to claim that the designed question papers were the representative sample of the contents/objectives.

3.1.5 Unit Wise Comparison in Terms of Content Coverage

The unit wise comparison of content representatives shows that the most representative unit is three, which represented 90% of the course contents in (2066 to 2067). In the hierarchy of comparison from most to least representation, unit two comes in the second place which represented 70.83% of the course contents and finally unit one lies in the last place of the hierarchy which covers only 50% of course contents. This can be shown in the following diagram as follows.

Figure No. 8

Unit wise Content Coverage 2066 to 2067



3.2 Content Weighting

This is the second part of the chapter which deals with content weighting.

Content weighting is the comparison between weighting of the course contents and contents of the question papers. For the purpose of examining content validity of the question papers in terms of content weighting, the researcher has compared whether the marks allocated by the syllabus is proportional or not in comparison to the weighting of question papers of 2066 to 2067.

3.2.1 The Whole Content Validity in Terms of Content Weighting

Table No. 5

The Whole Content Validity in Terms of Content Weighting

Weighting of the Course Content			Weighting of the Examination Papers		
S.N	Units	Full Weighting (Subjective only)	2066 B.S	2067 B.S.	Average Weighting of 2 years
1	Reading	24	7(alt)	7	7
2	Writing	40	7 (26)	14(26)	36.5
3	Critical Thinking	16	38(14)	19(21)	46
	Total Marks	80	45(47)	40 (47)	89.5
	Total	80	92	87	

Note: The allocated marks according to syllabus are given in appendix II.

alt = alternative marks

The number without bracket is compulsory marks

The number with bracket is alternative marks.

Here, the researcher has analyzed only subjective questions for a research purpose.

The table given above shows that the full marks of the 'Reading Writing and Critical Thinking Course' is 80 (subjective only). According to unit wise allocation of marks given in the syllabus, unit one deals 'Reading' which carries 24 marks (subjective only). It means from unit one 24 marks of questions (no more no less) can be asked in the examination. However, in 2066 B.S., the question paper carried 7 marks each with one alternative. It means 17 less marks

was asked in that year. Finally, in 2067 B.S., the test items carried 7 marks which was compulsory. It means 17 less marks was asked in that year from this unit. After analyzing the above data and description, we can come to conclusion that the average weighting of unit one during two years was only 7 marks. It means 70.83 percent less marks was asked specified in the syllabus. So, unit one was less valid (29.16%) in terms of content weighting.

Unit 2 includes 'Writing', which carries 40 marks (subjective only) according to unit wise allocation of marks specified in the syllabus. The above table shows that in 2066 B.S., the test items carried 33 marks (7 marks as compulsory and 26 marks as alternative) were asked. It means the weighting of the question papers was underrepresented by 7 marks. Finally, in 2067 B.S., the test items carried 40 marks from this unit, which was exactly. It means the question designers had strictly followed the weighting percentage given in the syllabus. The average weighting of that unit seems to be 36.5. It means the weighting of the question paper was underrepresented by 3.5 marks (8.75%) during two years from this unit. Therefore, the researcher concludes that unit two was less valid in terms of content weighting.

Unit three deals 'Critical Thinking', which carries 16 marks (subjective only) according to the unit wise allocation of marks specified in the syllabus. The table given above shows in 2066, amazingly that test items carried altogether 52 marks (38 as compulsory and 14 as alternative) were asked. It means the weighting of the question paper was more than triple specified in the syllabus. Finally in 2067, the test items carried 40 marks (19 as compulsory and 21 as alternative) were asked. It means the weighting of the question paper was more than double specified in the syllabus. The average weighting of this unit during two years was 46 marks, which shows that 30 marks was overrepresented. So, the researcher concludes that unit three was less valid in terms of content weighting.

It is believed that a good test should not provide options to the testee because freedom in the selection of the question reduces validity. In other words it is a good idea to give options from the same unit/sub-units if alternative to be given. In case of 'Reading Writing and Critical Thinking', the data shows that four alternative questions were given out of 10 (40%) during two years. The question paper of 2066 was better than 2067 to provide the choose options, which were represented from the same unit/sub-units.

After analyzing the above presented data and description the researcher concludes that the question papers of 'Reading writing and Critical thinking' from 2066 to 2067 was less valid in terms of content weighting. It means the question designers were not serious towards the weighting percentage given in the syllabus. As a result, the question papers were less valid in terms of content weighting and such questions also exert negative wash back effect.

3.3 Comparison of Content Validity in Terms of Content Coverage and Weighting

After analyzing the various data presented above, the researcher found that the question papers or 'Reading Writing and Critical Thinking' was good content validity in terms of content coverage. This is proved by the fact that 16 and 17 sub-units were represented out of 24 in the examination in the year 2066 and 2067 respectively, which means 70.27 % and considered as good coverage. However, same question papers were less valid in terms of content weighting. It is proved by the fact that the average weighting of the question papers during 2 years was 7 (unit one), 36.5 (unit two) and 46 (unit three) respectively during two years. But the accurate weighting specified in the syllabus of those units is 24, 40 and 16 respectively. This shows that the weighting of the examination papers is underrepresented from unit one and two and overrepresented from unit three. Therefore, the researcher concludes that the question papers of 'Reading Writing and Critical Thinking' was good content validity (70.27%) in terms of content coverage but was less valid in terms of content weighting.

CHAPTER - FOUR

FINDING AND RECOMMENDATIONS

4.1 Finding

This is last but not least chapter of this research report which deals with the findings and recommendations made by researcher after analyzing the data. This chapter is divided in to two subparts. The first chapter deals with the finding made by the researcher and the second chapter deals with the recommendations made on the basis of the collected and analyzed data and findings. After analyzing the questions papers the researcher found the following findings.

1. The researcher found that the question papers of 'Reading Writing and Critical Thinking Course'(2066 B.S to 2067 B.S) was good (70.27%) content validity in terms of content coverage.
2. It was found that same question papers were less valid in terms of content weighting.
3. From unit one (reading) 3.5 sub-units were covered /represented out of 7 in the examination during two years. It means the average coverage of course content is 50 percent. So, this unit was less valid in terms of content coverage.
4. From unit two (writing) 8.9 sub- units were covered /represented out of 12 in the examination during two years. It means the average coverage of course content is 70.83 percent. Therefore, this unit was good content validity in terms of content coverage.
5. From unit three (critical thinking) 4.5 course sub-units were represented out of 5 in the examination during two years. It means the coverage of

course content is 90 percent. Therefore, unit three was very high content validity in terms of content coverage.

6. According to unit wise allocation of marks specified in the syllabus , unit one (Reading) carries 24 marks (subjective only) ,but it is not found to have followed in designing the question papers, which is proved by the fact that only 7 marks in average was asked out of 24 .It means 17 marks (70.83%) were underrepresented during two years. So, this unit lacks content validity in terms of content weighting.
7. According to the unit wise allocation of marks specified in the syllabus, unit two (Writing) carries 40 marks (subjective only) .In 2067, weighing of the question papers was exactly 40. However, in 2066 it was underrepresented. The average weighting of question papers during two years was 36.5 marks. It means 3.5 marks (8.75%) was underrepresented. Therefore, this unit also lacks content validity in terms of content weighting.
8. According to unit wise allocation of marks specified in the syllabus, unit three (Critical Thinking) carries 16 marks (subjective only). However it is not been found to have followed by the question designer. The weighting of the question paper was overrepresented by carrying 46 marks in average during two years. It means 187.5 percent marks were overrepresented. Therefore, unit three also lacks content validity in terms of content weighting.
9. It was found that 4 optional questions (3 short and 1 long) were asked out of 10(40%) during two years.
10. The question papers of 2066 was better than 2067 on the basis of the alternatives.
11. It was found that some language items were neglected totally and some others were given more emphasizes in designing the questions.

12. It was found that most of the questions were designed on the basis of its importance.

4.2 Recommendations

The researcher has made some recommendations on the basis of the analyzed data, which are as follows.

1. The researcher found that the content coverage of question papers of 'Reading writing and Critical thinking' during two years is 70.27 percent. It is considered as good but not maximum level of representation. Therefore, the question designer should design the questions by ensuring more content coverage.
2. The researcher found that same question paper has very less content validity in terms of content weighting. So, the question designer should design the questions by strictly following the weighting specified in the syllabus.
3. To enhance high content validity, the question designer should provide equal emphasize to all units and language items.
4. Due to the lack of continuous assessment system (CAS) and practical examination, unit one (Reading) was neglected final examination. Therefore, CAS or practical examination system should focus to obtain good content validity.
5. The question designer should be highly educated, experienced, and trained.
6. The question designer should prepare specification chart compulsory and strictly to be followed it in designing the question papers.

7. To enhance high content validity, the optional questions should reduce and the given choice should be asked from the same unit / language items.
8. The office of the controller of examination, T .U has the sole authority in conducting examinations. So, all the credits and criticisms go to it. Therefore, it should play vital role in improving tests by organizing workshops and training program.
9. Finally, in the name of maintaining content validity other types of validity should not be neglected.
10. To have high content validity, the question setters have to have minute study on course objectives, course contents and weighting of the course contents before developing the question papers. The question papers have to be piloted if possible.

References

- Allister, S.M. (2009). *Critical thinking development report: A review of current best practice, research in the field and selected programs*.
<<http://www.criticalthinking.org>.
- Aryal, K. P. (2005). *A Study on content validity of grade twelve compulsory English examination 2061*. An unpublished M. Ed. thesis, Kathmandu. Tribhuvan University.
- Awasthi, J.R. (2003). Teacher education with special references to English language teaching in Nepal. In *Journal of NELTA, vol.8,no.1.p.22-23*.
- Bachman, L. F. (1998). *Fundamental considerations in language testing*. London: OUP.
- Batala, K.B. (2004). *Validation of the SLC English examination*. An unpublished M.Ed. thesis, Kathmandu. Tribhuvan University.
- Bhandari, B.M. (2007). *Writing skills in English*. Kathmandu: Vidyarthi Prakshan.
- Brown, D.H (1994). *Principles of language learning and teaching*. London: Prentice Hall.
- Byrne, D. (1972). *Teaching writing extracts*. London: Longman.
- Crawford, A. et al. (2005). *Teaching and learning strategies for the thinking classroom*. Kathmandu: Social Science Baha.
- Davies, A. & J.P.B, Allen (1978). *The Edinburgh course in applied linguistics*. vol.4, London: OUP
- Faculty of Education (2008). *Revised course of three year B.Ed. for first year*. Kathmandu: TUP

- Gautam, R. B. (2065). *Research methodology in language education*. Kathmandu: Sunlight Publication.
- Gardner, P. S. (2009). *New directions: reading, writing and critical thinking*. New York: Cambridge University Press.
- Giri, R. A. (2003). Language testing: then and now. In *Journal of NELTA*, vol.8, no.1, p. 49-65.
- Gurung, R.C. (2008). *The content validity of semantics and pragmatics*. An unpublished M.Ed. thesis, Kathmandu: Tribhuvan University.
- Harrison, A. (1991). *Language testing handbook*. London: ELTS. Longman.
- Harmer, J.(2007). *How to teach English*. London: Pearson Longman.
- Heaton, J .B. (1988). *Writing English language test*. London: Longman.
- Jigyashu (2010). *A solution of language testing*. Kathmandu: Sunlight Publication.
- Koirala, T.N. (2008).*The content validity of English textbook for grade six*. An unpublished M. Ed. thesis, Kathmandu. Tribhuvan University.
- Kumar, R. (1996). *Research methodology*. New Delhi: SAGE Publication.
- Lado, R. (1961). *Language testing: The construction and use of foreign language test*. London: Longman.
- Nepal,P.(2006). *Content validity of examination: A case of fundamentals language and linguistics at B. Ed level*. An unpublished M.Ed. thesis, Kathmandu. Tribhuvan University.
- Neupane, M. (2004) .*Wash back effect of examination: A case of communicative English*. 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.

Phyak, P.B. & Sharma & B.K. (2007).*Teaching English language*. Kathmandu: Sunlight Publication.

Scriven & Paul (1987).*Critical thinking and education*. paper presented at the 8th annual international conference: New York, America.

Subedi, R. (2006). *Content validity of B. Ed. questions: A case of English sound and structure (302)*. An unpublished M.Ed. thesis, Kathmandu. Tribhuvan University.

Sharma, U. N. (2009). *Language testing*. Kathmandu: Sunlight Publication.

Appendix

Appendix I

Tirbhuvan University

2066

Bachelor Level/ Education/I Year

Full Marks: 100

Reading, Writing and Critical Thinking (Eng.Ed.-317)

Time: 3 hrs

Candidates are required to give their answers in their own words as far as practicable.

The figures in the margin indicate full marks.

Attempt All the questions.

GROUP "B"

8×7=56

1. What do you understand by non-verbal communication? Explain with examples from *where Do we Stand?*
2. Why does Rothenberg think that much of the fault for the declining quality of his students' research paper is his own? (*How the web destroys the Quality of students Research papers*)
3. " A person who hasn't attended school for many years cannot be intelligent". Do you agree or disagree with the statement? Justify your answer with reference to the reading Multiple Intelligences and Emotional Intelligence.
4. What are the propaganda technique discussed by Cc Clintock ? With reference to either school, government, political candidates, and advertisers, discuss their use in your own culture. (Propaganda Techniques in Today's Advertising)
5. Discuss whether you agree or disagree with Anderson that "television is almost surely having a major social impact on the kids, as oppose to a cognitive impact." (Don't Touch That Dial)

OR

Read the following passage and answer the question.

She is the prototype of today's young women - confident, outgoing, knowledgeable, and involved. She is active in her school, church, or community. She may have a wide circle of friends or simply a few close ones, but she is committed to them and to their friends. She is sophisticated about the central issue facing young people today - planning for the future, intimacy, sex, drugs, and alcohol - and discusses them seriously, thoughtfully, and forthrightly. She wants to take control of her life and is trying to figure out how to get from where she is to where she wants to go. Above all, she is convinced that if she plans carefully, works hard, and makes the right decision she will be a success in her chosen field. Have the material goods she desires. In time, marry if she wishes; and in all probability, have children. She plans, as the expression goes, to "have it all".

Questions:

- a) What are the specific issues that the young women are fond of?
 - b) What is the main idea of the text?
 - c) Write the summary of the paragraph.
6. Write an anecdote describing your interesting or unforgettable event during your school life.

OR

Write a poem about the present political situation of Nepal.

7. What is the main point, or message, that Frank wishes to convey in the "Conceptual Fruit"? Explain with examples from the story.

OR

8. Who do you blame for the death of Prema? Justify your answer. (A Coward)

GROUP "C"

2×12=24

9. What are the major criticisms Holt makes against the formal education system in the essay "School Is Bad For children "? Do you agree or disagree? Justify your ideas with reference to the school education of Nepal.

OR

Imagine that you are applying for the one of the positions advertised in a local newspaper. You already have a resume. Write a cover letter explaining why you are interested in the job and what your qualifications are Remember that you will probably be one of many applicants for the position. What makes you a special candidate?

10. Answer any one of the following questions. Support your arguments with the detail information from the relevant readings from the text book, other sources, your own experience and observations of your own society and culture.
- a) Which Factor - biology or socio-cultural expectation - are more important to determine our gender roles? Justify your answer by discussing the traditional roles assigned to males and females in your own society or culture?
- b) What professions or job are valued in your country and why? Are people discriminated in terms of their jobs in your country? Justify your answers.
- c) Some people argue that technology promotes creativity of students whereas others say that it kills their creativity. What do you think about this? Explain.

Tirbhuvan University

2067

Bachelor Level/ Education/I Year

Full Marks: 100

Reading, Writing and Critical Thinking (Eng.Ed.-317)

Time: 3 hrs

Candidates are required to give their answers in their own words as far as practicable.

The figures in the margin indicate full marks.

Attempt All the questions.

GROUP "B"

8×7=56

1. How do the Gardner's and Goleman's concept of intelligence differ from the tradition ones? Elaborate two important differences (Multiple Intelligence and Emotional Intelligence).
2. How do you think Davis would recommend avoiding intercultural misunderstanding that result from difference in body language? (Where Do you Stand?)

OR

Do you agree with the statement most students formal education has little connection in real life. Elaborate your answer with reference to Nepal.

3. Write a letter to your friends in the USA suggesting him/her what she/he should do in order to avoid intercultural misunderstanding in Nepal.
4. What are the Rothenberg's major criticisms of student's use of the World Wide Web in writing research paper? (How the web destroys the Quality of Students Research Paper)

OR

Write a poem on "student life"

5. What is the main point McClintock makes the media "Propaganda Techniques in Today's Advertising?" summaries her main idea in the propaganda.
6. Discuss a point that Venture makes with which you agree or disagree? (Someone is Stealing Your Life).
7. Explain the main idea of "A Coward"
8. How realistic does Sidel believe the New American Dream is? What are the chances the youngest women and men will achieve it?

OR

Write a letter to the editor of a newspaper describing how people can make their jobs joyful or less stressful.

GROUP "C"

2×12=24

9. Write an essay on "Gender Discrimination in Nepal" in minimum 300 words. In your essay incorporate the ideas given in the reading "Sex Roles" and "Boys Will Be Boys"

OR

Write a story on "Miseries or Problems of People Living in Villages".

You can select one or more person as a character of the story. Limit your story to 300 words.

10. Answer any one of the following questions. Support your argument with detail information from the relevant reading from the prescribed textbook, other sources e.g. radio, TV newspaper, journals, book, your own experience, and observation of your own society and culture.
 - i. What are the major values and assumptions discussed in "American Values and Assumptions?" How are they different from and similar to Nepalese contexts? Explain with examples.
 - ii. Examine how well the educational system in your country prepares students for the challenges they will face in the real world. What are the

strengths and weaknesses of the education that the students receive? Do you have any recommendations for its improvements?

iii. Discuss the advantage and disadvantage of information technology of the students.

Appendix II

A Syllabus of Reading Writing and Critical Thinking

Course No .Eng .Ed. 317 (specialization II)	Full marks: 100
Nature of the Course: Theory	Pass marks: 35
Periods per week: 6	Total periods: 150
Time per period: 50 minutes	

1. Course Description

This course exposes the students to the varieties of reading materials and writing strategies in order to enhance their academic skills. It adopts a content - based approach to the development of reading, writing and critical thinking abilities. Furthermore, it focuses on stages of the writing process and the structure of academic writing.

2. General Objectives

The general objectives of the course are as follows:

-) To expose the students to a wide variety of contemporary reading materials.
-) To facilitate them to read critically, and write logically.
-) To expose them to a variety of writing activities.
-) To encourage students to think independently.
-) To enhance their abilities to argue with reason and confidence.

3. Specific Objectives

-) Read and comprehend the purposes of texts.
-) Identify the general idea of the texts.
-) Find the main points in the texts.
-) Read and take notes of the important points.
-) Comprehend details of the texts.
-) Read and analyze the text.
-) Write keeping in mind the purpose and audience.
-) Organize main ideas with supporting details.
-) Write narratives and anecdotes.

-) Write reports and letters.
-) Write different types of essays.
-) Write creatively.
-) Read and analyze the text critically.
-) Argue with reason.
-) Think independently.
-) Debate confidently.

4. Contents

Unit One: Reading

(60 periods)

- 1.1 Reading with a purpose
- 1.2 Reading for specific information
- 1.3 Reading for general information
- 1.4 Reading for main idea.
- 1.5 Reading critically and analytically
- 1.6 Reading for pleasure
- 1.7 Reading and taking notes.

Unit Two: Writing

(60 periods)

- 2.1 purpose and audience
- 2.2 Main ideas and supporting details
- 2.3 The essential of writing
- 2.4 Writing an essay
 - 2.4.1 The structure of an essay
 - 2.4.2 The writing process
 - 2.4.3 Writing with sources
- 2.5 Creative writing
 - 2.5.1 Different genres of creative writing
 - 2.5.2 Writing anecdotes
 - 2.5.3 Writing stories
 - 2.5.4 Writing poems

Unit Three: Critical Thinking

(30 periods)

- 3.1 Developing a critical mind

3.2 Analyzing

3.3 Arguing

3.4 Reflecting

3.5 Reasoning

5 Methods\ Instructional Techniques

) *Lecture and discussion*

) *Explanation and illustration*

) *Project work*

) *self study*

) *Demonstration*

) *Group and individual work*

) *presentation*

6 Evaluation Scheme and Time Allotment

The course is for one academic year and it carries 100 marks. The distribution of marks and time allotment is as follows:

Unit	Marks	Time
1 Reading	30	30%
2 Writing	50	50%
3 Critical Thinking	20	20%
Writing Examination	100%	

The type of questions and number of items to be asked in the final exam is as follows:

Nature of Questions	Total questions to be asked	Number of question to the answered	Weight age
Multiple choice items	20	20 marks	20marks
Short answer questions	8 with 3 or questions	8×7 marks	56marks
Long answer questions	2 with 1 or question	2×12marks	24 marks
Total			100 marks

Appendix III
Analysis of Content Coverage from Unit One
Table No. 1

Units	Course Contents	Content of the Examination Papers (Subjective only)	
1	Reading	2066 B.S.	2067 B.S.
1.1	Reading with a purpose	no.5(alt)	
1.2	Reading for specific information	no.5(alt)	no.6
1.3	Reading for general information	×	×
1.4	Reading for main idea	no.5(alt)	
1.5	Reading critically and analytically	no.5 (alt)	
1.6	Reading for pleasure	×	×
1.7	Reading and taking notes	×	×
	Total number of question asked from unit 1	1 short question (alt. w.u.)	1 short question
	Total Marks:	7 marks	7 marks

Note: The questions asked in examination are given in appendix I

Where,

alt = alternative

no = number

S.L. = subjective long question

×

= the sub- units were represented

w.u. = within a unit

A question without bracket is compulsory.

Analysis of Content Coverage from Unit Two

Table No 2

S.N.	Course contents	Content of the Examination Papers (Subjective only)	
2	Writing	2066 B.S.	2067 B.S.
2.1	Purpose and audience		
2.2	Main idea and supporting details	no.7(alt)	no7, no.8(alt)
2.3	The essential of writing		
2.4	Writing an essay	×	S.L.no.9a(alt)
2.4.1	The structure of an essay	×	
2.4.2	The writing process	×	×
2.4.3	Writing with sources	no. 4	×
2.5	Creative writing		
2.5.1	Different genres of creative writings	no.7(alt), S.L.no.9(alt)	no.3, no.8(alt)
2.5.2	Writing anecdotes	no.6(alt)	×
2.5.3	Writing stories	×	S.L.no.9b(alt)
2.5.4	Writings poems	no.6(alt)	no.4 (alt)
	Total number of questions asked from unit two	1 long(alt) 3 short(2 alt)	1 long(alt) 4 short(2 alt)
	Total Marks:	7+(26)	14+(26)

Note: The questions asked in examination are given in appendix I

Where,

alt = alternative

no = number

S.L. = subjective long question

× = the sub- units were not represented

= the sub- units were represented

w.u. = within a unit,

A question without bracket is compulsory

Analysis of Content Coverage from Unit Three

Table No 3

Units	Course Contents	Contents of the Examination Papers (Subjective only)	
3	Critical thinking	2066 B.S.	2067 B.S.
3.1	Developing a critical mind	S.L.no.10	S.L.no.10b
3.2	Analyzing	S.L.no.1o	S.L.no.10c
3.3	Arguing	S.L.no.10 S.L.no.9a (alt),no.3	no.2(alt)
3.4	Reflecting	S.L.no.9a(alt),no.8,no.1, no.2	S.L.no.10a,no.1 no.4(alt),no.8(alt) no.2(alt)
3.5	Reasoning	S.L.no.10 S.L.no.9a(alt),no.3,no.8	×
	Total Number of questions asked from Unit 3	2 long, 4 short(2 alt)	1 long , four short(3 alt)
	Total Marks:	52 marks (12 marks alt)	40 marks (7 alt)

Note: The question numbers asked in examinations are given in appendix I

Where,

alt = alternative

no = number

S.L. = subjective long question

× = the sub- units were not represented

A question without bracket is compulsory.

The Whole Content Validity in Terms of Content Coverage

Table No. 4

Units	Contents	Sub-Units	Covered Sub-Units		Covered Sub-Units in Percent		Average Percent 2066- 2067	Remarks
			2066 B.S.	2067 B.S.	2066 B.S.	2067 B.S.		
1.	Reading	7	4	3	42.86	57.14	50	Less Valid
2.	Writing	12	8	9	66.66	75	70.83	Good Validity
3.	Critical Thinking	5	5	4	100	80	90	High Validity
Total		24	17	16	69.84	70.71	70.27	

The Whole Content Validity in Terms of Content Weighting

Table No 5

Weighting of the Course Contents			Weighting of the Examination Papers				
S.N	Units	Full Weighting (Subjective only)	2066 B.S	2067 B.S.	Average Weighting Marks	Average percent 2066-2067	Remarks
1	Reading	24	7(alt)	7	7	29.16	Underrepresented by 70.83%
2	Writing	40	7 (26)	14(26)	36.5	91.25	Underrepresented by 8.75%
3	Critical Thinking	16	38(14)	19(21)	46	287.5	Overrepresented by 187.5%

Note: The subjective marks 24, 40, and 16 (unit 1, 2 and 3 respectively) are carried out by converting 80 marks in to 100 percent.

e.g. 100% = 80 marks.

Where,

alt= alternative

The number without bracket is compulsory marks

The number with bracket is alternative marks.

Here, the researcher has analyzed only subjective questions for a research purpose.