

**ITEM ANALYSIS OF MULTIPLE CHOICE ITEMS:
A CASE OF TEACHING ENGLISH LITERATURE
AT B. ED. LEVEL**

**A Thesis Submitted to the Department of English Education,
Tribhuvan University, Kirtipur, Kathmandu
In Partial Fulfillment for Master's Degree in Education
(Specialization in English Education)**

**By
Hari Prasad Subedi
M.Ed. Second Year**

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

2007

**ITEM ANALYSIS OF MULTIPLE CHOICE ITEMS:
A CASE OF TEACHING ENGLISH LITERATURE
AT B. ED. LEVEL**

**A Thesis Submitted to the Department of English Education,
Tribhuvan University, Kirtipur, Kathmandu
In Partial Fulfillment for Master's Degree in Education
(Specialization in English Education)**

By

**Hari Prasad Subedi
M.Ed. Second Year**

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

2007

T.U. Regd. No.: 9-2-48-1471-98

Campus Roll No. : 696/2059

2nd Year Exam Roll No.: 280205/061

Date of Approval of the Thesis

Proposal: 2063-4-11

Date of Submission of the

Thesis: 2064-1-9

RECOMMENDATION FOR ACCEPTANCE

This is to certify that **Mr. Hari Prasad Subedi** has prepared this dissertation entitled “**Item Analysis of Multiple Choice Items: A Case of Teaching English Literature at B. Ed. Level**” under my guidance and supervision.

I recommend this dissertation for acceptance.

Date:

.....

Mr. Umanath Sharma

(Guide)

Lecturer

Faculty of English Education

Tribhuvan University

Kirtipur, Kathmandu, Nepal

RECOMMENDATION FOR EVALUATION

This dissertation has been recommended for evaluation by the following Research Guidance Committee:

Signature

Prof. Dr. Tirth Raj Khaniya

Professor and Head,

Department of English Education

Chairperson

Mr. Umanath Sharma

Lecturer

Department of English Education

Member

Mrs. Madhu Neupane

Lecturer

Department of English Education

T.U., Kirtipur

Member

Date:

EVALUATION AND APPROVAL

This dissertation has been evaluated and approved by the following
Thesis Evaluation Committee:

Signature

Dr. Chandreshwar Mishra

Reader and Head

Department of English Education

T.U., Kirtipur

Chairperson

Dr. Jai Raj Awasthi

Professor of English

Department of English Education

T.U., Kirtipur

Member

Mr. Umanath Sharma (Guide)

Lecturer

Department of English Education

T.U., Kirtipur

Member

Date:

DEDICATION

*To my parents
who spent their entire life
to make me what I am today*

ACKNOWLEDGEMENT

This dissertation has been prepared for the partial fulfillment of Master Degree in English Language Education. First of all, my heartfelt gratitude goes to **Prof. Dr. Shishir Kumar Sthapit**, the senior most professor of the Department and **Dr. Chandreshwar Mishra**, Head of the Department of English Language Education for providing a vital support to this work. I would like to express my sincere gratitude to my thesis supervisor **Mr. Umanath Sharma**, Lecturer at the Department of English Language Education T.U. who provided me with continuous guidance, enlightening ideas, invaluable suggestions and encouraged me constantly.

Similarly, I am grateful to **Prof. Dr. Shanti Basnyat**, Professor of the Department of English Language Education and Chairperson of English and other Foreign Languages Education subject committee, who provided me enlightening ideas and encouraged me to develop this work.

I would like to express my sincere gratitude to **Prof. Dr. Jai Raj Awasthi** for providing me with a lot of encouragement. I would also like to express my sincere gratitude to Prof. **Dr. Tirtha Raj Khaniya**, **Prof. Dr. Govinda Raj Bhattarai**, Assist. Dean, FOE, **Dr. Anjana Bhattarai**, Lecturer of the Department and other Gurus for their kind support, encouragement and continuous inspiration during my study period in the department.

My thanks also go to all the teachers and students for their help in my research study and enthusiastically participated in this research.

My deepest appreciation also goes to my wife Shanta Subedi and my daughter Samikshya and son Saphal for their patience, I would like to thank my close friend Mr. Prem Prasad Poudel, Chhabi Lal Subedi and Bishnu Prasad Subedi for their kind help.

At last, I would like to express my special thanks to my relatives Pushpa Raj Adhikari, Krishna Tiwari and other well-wishers for their serious encouragement and kind support for this research.

ABSTRACT

This research attempts to make item analysis of the multiple choice items of Teaching English Literature asked in the final exam of the years 2056 to 2060 in Tribhuvan University. So the study was focused to analyze the facility value and discrimination index of the test items. In order to carry out this analysis, the researcher collected the original test items asked in the year 2056, 2057, 2058, 2059 and 2060. The researcher prepared the questionnaire on the basis of these questions. For the collection of data, the researcher requested the students who had just finished their bachelor level specializing in the English Language Education – from different colleges (both TU Constituent and TU TU affiliated campuses) of Kathmandu, Pokhara and Baglung. The researcher collected data from the students selecting them purposefully. Altogether there were hundred students, fifty from private and fifty from TU Constituent campuses. They were requested to sit for the test for one and half an hours. Their scores were tabulated and analyzed statistically calculating averages and percentages. On the basis of this analysis and interpretation of the data, this study found out that these test items collectively from the years 2056 – 2060 were not functioning very efficiently since they have the moderate level of discrimination index and average level facility value. But the test items in some of the years had very poor discrimination power since the year 2056, the discrimination index is found to be 0.27.

This study consists of four chapters. Chapter one consists of general background, review of related literature, objectives of study and significance of the study.

Chapter two consists of methodology of study. It encompasses the sources of data, population of the study, sample population, sampling procedure, tools for data collection, process of data collection and limitation of the study.

Chapter three consists of analysis and interpretation of the collected data. The data were analyzed and interpreted on the basis of the variables that were specified in the objectives, i.e. analysis on the basis of private or TU Constituent colleges, males Vs females in terms of facility value and discrimination index.

Chapter four consists of the major findings of the study and recommendations of the study.

ABBREVIATIONS

List of the abbreviations used in the study

B.Ed.	-	Bachelor of Education
FV	-	Facility Value
i.e.	-	that is
etc.	-	Etcetera
DI	-	Discrimination Index
M.Ed.	-	Master of Education
MCQ	-	Multiple Choice Question
S.N.	-	Serial Number
F.M.	-	Full Mark
T.U.	-	Tribhuvan University
P.N.C.	-	Prithivi Narayan Campus
M.	-	Male
F.	-	Female
MRC	-	Mahendra Ratna Campus
COE	-	College of Education
GAC	-	Gramin Adarsha Campus
MMC	-	Mahendra Multiple Campus
KSC	-	Kathmandu Sikshya Campus
QN	-	Question Number
No.	-	Number
MC	-	Machhapuchhre Campus
U	-	Upper
L	-	Lower

TABLE OF CONTENTS

	<i>Page No.</i>
RECOMMENDATION FOR ACCEPTANCE	I
RECOMMENDATION FOR EVALUATION	II
EVALUATION AND APPROVAL	III
DEDICATION	IV
ACKNOWLEDGEMENTS	V
ABSTRACT	VI
ABBREVIATIONS	VIII
CONTENTS	IX
LIST OF TABLES	XII
CHAPTER ONE: INTRODUCTION	1-12
1.1 General Background	1
1.2 Qualities of a Good Test	3
1.3 Types of Test: Subjective and Objective	5
1.4 Multiple Choice Items	5
1.5 Item Analysis	7
1.6 Review of Literature	10
1.7 Objectives of the Study	12
1.8 Significance of the Study	12
CHAPTER TWO: METHODOLOGY	13-15
2.1 Sources of Data	13
2.1.1 Primary Sources of Data	13
2.1.2 Secondary Sources of Data	13
2.2 Population of the Study	13
2.2.1 Sample Population	13
2.2.2 Sampling Procedure	14
2.3 Tools for Data Collection	14
2.4 Process of Data Collection	14
2.5 Limitations of the Study	15
CHAPTER THREE : ANALYSIS AND INTERPRETATION OF DATA	16-41
3.1 The multiple choice test items asked in the year 2056	17
3.2 The Multiple Choice test items asked in the year 2057	19

3.3	The multiple choice test items asked in the year 2058	22
3.4	The Multiple Choice Test Items asked in the year 2059	24
3.5	The Multiple Choice Items asked in the year 2060	27
3.6	Sex-wise performance of the Students	29
3.7	The Performance of Males and Females as a Whole	37
3.8	The Scores of the students of TU affiliated campuses	38
3.9	The scores of the students of TU Constituent colleges	39
3.10	The scores of the TU affiliated and the TU Constituent colleges	40
3.11	Comparative Analysis of Facility Value and Discrimination Index as a Whole	40
CHAPTER FOUR: FINDINGS AND RECOMMENDATIONS		42-45
4.1	Findings	42
4.2	Recommendations	44
REFERENCES		46-47
APPENDICES		48-75
APPENDIX-I:		
	Model of questionnaire of multiple choice questions asked in the year 2056 to 2060	48-53
APPENDIX-II:		
	Response of the students of multiple choice questions asked in the year 2056 to 2060	54-59
APPENDIX-III:		
	Response of the students of multiple choice questions asked in the year 2056 to 2060	60-65
APPENDIX-IV:		
	Raw Score of the Students as a Whole	66-68
APPENDIX-V:		
	Raw Score of the Students of Prithivi Narayan Campus	69
APPENDIX-VI:		
	Raw Score of the Students of Mahendra Ratna Campus	70

APPENDIX-VII:	
Raw Score of the Students of College of Education	71
APPENDIX-VIII:	
Raw Score of the Students of Gramin Adarsha Campus	72
APPENDIX-IX:	
Raw Score of the Students of Mahendra Multiple Campus	73
APPENDIX-X:	
Raw Score of the Students of Kathmandu Sikshya Campus	74
APPENDIX-XI:	
Raw Score of the Students of Machhapuchhre Campus	75

LIST OF TABLES

	<i>Page No.</i>
Table 1: Analysis of the Multiple Choice Test Items Asked in the Year 2056	17
Table 2: Analysis of the Multiple Choice test items asked in the year 2057	19
Table 3: Analysis of the Multiple Choice test items asked in the year 2058	22
Table 4: Analysis of the Multiple Choice Test Items asked in the year 2059	24
Table 5: Analysis of the Multiple Choice Items asked in the year 2060	27
Table 6: Analysis of Sex-wise Performance of College of Education	29
Table 7: Analysis of Sex-wise Performance of Prithivi Narayan Campus	30
Table 8: Analysis of Sex-wise Performance of Mahendra Ratna Campus	32
Table 9: Analysis of Scores of the Students of Machhapuchhre Campus	33
Table 10: Analysis of the Scores of the Students of Kathmandu Sikshya Campus	34
Table 11: Analysis of the Scores of the Students of Gramin Adarsh Campus	35
Table 12: Analysis of the Scores of the Students of Mahendra Multiple Campus	36
Table 13: Comparative Analysis of the Performance of males and females as a whole	37
Table 14: Analysis of Scores of the Students of TU affiliated campuses	38
Table 15: Analysis of Scores of the Students of TU Constituent Colleges	39
Table 16: Comparative Analysis of the Scores of the Public/Private and the TU Constituent Colleges	40
Table 17: Facility Value and Discrimination Index as a Whole	40

CHAPTER – ONE

INTRODUCTION

1.1 General Background

Language is purely a human phenomenon. It is a non-instinctive possession of humans. It is the feature that makes humans different from other creatures. Cambridge International Dictionary of English defines language as "A system of communication consisting of a set of small parts and a set of rules which decide the ways in which these parts can be combined to produce messages that have meaning". It is language through which people could exchange their ideas, emotions, experiences, and thoughts. Language plays one of the instrumental roles in the development of the whole present day world. The speed of development in the multiple aspects can be accelerated by the knowledge and experiences they share which is possible only through the use of a language. It has played a dominant role in all areas of life like trade, science, foreign affairs, information technology, education etc.

Among many languages in the world, English has been one of the world languages. It is the gateway to the world body of knowledge. In view of these facts many countries in the world are accepting English as their language of education media and other areas of governance. The value and the use of English language has been increasing tremendously every next moment. English language has become a passport to the world body of knowledge and opportunities. It is an international lingua Franca. According to UNESCO, more than sixty countries in the world use English as an official language, about 160 million people listen to English radio programmes, over 60 million children study English at the primary level. A vast store of knowledge and technology is explored through English. English is taught as one of the foreign/second language in many

non-native countries. In Nepal too, it is taught as the chief foreign language.

The beginning of English language teaching can be traced back to the period of Rana Prime Minister Junga Bahadur Rana in 1910 B.S. At first it was confined to the Rana families and slowly and gradually it came to the public access. These days, English language teaching is growing as one of the professions. Many English medium schools have been established. It is taught as a compulsory subject from grade one in the TU Constituent and the private schools to the diploma level. The advanced studies are carried out in English language.

Language teaching is an art. It includes many skills and sub skills. Language teaching started along with the felt need of many languages by human beings. Since language is the means of communication, it is to be taught and learnt to have good social relationship. Whenever and wherever there is language teaching there emerges the need of language testing. This means language teaching and testing are related to each other. Khaniya (2005:3) says that testing has been an inherent part of teaching from the time when teaching began the teacher has always been keen to know the extent to which his teaching has been effective in making the learner understand what has been taught. Testing is quite necessary to decide how much has been learnt by the learners and sometimes to determine the remedial measures. Testing can be carried out in formal as well as informal setting. This can be done by using various types of tools i.e. test items.

Test may be primarily constructed as a device to reinforce learning and to motivate the students or as a means of assessing student's performance in language. It is invariably associated with making of decisions, a way of arriving at meaningful conclusion. The development

of language testing as a discipline has undergone various ups and downs. Functional and communicative language testing have been the major approaches and issues at present time.

Language testing is a very complex task. Construction of a genuine test is a far more complex matter for instance, a discrete point test can be good for ascertaining grammatical competence but not for communicative competence. The quality of a test can be examined in light of its usefulness. An efficient test should have certain qualities and these are to be considered while designing a test. Three basic qualities of a test are validity, reliability, and practicality without any one of which a test would be a mere investment of time and money. Harris (1994) points out that whether the teacher is constructing his own test or selecting a standard instrument for use in his class or school he should bear in mind the above mentioned test qualities.

1.2 Qualities of a Good Test

The common agreement is that a test should be of good quality. The quality of the test is the prerequisite to fulfill the purposed goal of the test. We can't easily say which test is better than the other on the basis of an overt observation. A test should be judged practically, we should think for the efficiency of a test. Therefore, the qualities of a test are pointed out as follows:

- a. **Validity:** If a test tests what is intended to be measured, then this test becomes a valid test. That means, a test should fulfill/ satisfy the desired goals of any academic programme. Validity of a test is determined by the type of test. That means a valid test for one exam may not be valid for another.
- b. **Reliability:** In testing, a measure of the degree to which a test gives consistent results. A test is said to be reliable if it gives the

same results when it is given on different occasions or when it is used by different people.

- c. **Practicality:** Practicality is the most important quality of a test. It can be helpful whether a test is too easy or too difficulty. Heaton (1975:158) has explained the practicality as or the exam must be fairly straightforward to administer. Generally, practicality involves the cost, ease of administration and scoring.
- d. **Wash back effect:** Any effect of test either positive or negative in teaching and learning is called wash back effect. Any test is said to be having good backwash effect it exerts a good influence on the teaching and learning that take place before the administration of the test. Each elements and skill has been treated in the book in relation to its potential influence teaching. Testing has been one of the greatest single beneficial forces in changing the direction of language teaching in many areas and in encouraging the more responsive teachers to examine not only their own teaching methods but also the language they are teaching.

The fundamental qualities of a good test (Validity, reliability, practicality and wash back effect) are possessed by any type of test. But the major weakness subjective tests is lack of validity and difficult to achieve reliability because there are so many degrees of acceptability and ways of scoring all the possible responses. On the contrary, objective tests have high confound validity and reliability is not difficult to achieve. Objective tests, usually have only one correct answer, they can be scored mechanically. So, the researcher analyzed the quality of the multiple choice objective test on the basis of item of difficulty and item of discrimination.

1.3 Types of Test: Subjective and Objective

Khaniya (2005: 92) points out that subjective and objective tests are two contrasting procedures of testing. Basically they are different in terms of how the marking of the performance of the students is done, a test with subjective marking is a subjective test and a test with objective marking is an objective test. Heaton (1975: 25) has the similar claim, he says subjective and objective are terms used to refer to the scoring of tests. All test items, no matter how they are devised, require candidates to exercise a subjective judgment. In any essay test, for example, candidates must think of what to say and then express their ideas as well as possible; in a multiple-choice test they have to weigh-up carefully all the alternatives and select the best one. The objective test can be answered by selecting a correct one out of some choices given, whereas in subjective test it is the test in which the testee has to supply the answer.

These days, preference is on objective test items may be because the objective test item can have a wide coverage of the course of study. The birth of objective test took place in response to the reliability problem with the subjective test during the discrete point or scientific era of testing. So that objective test is useful for both coverage and reliability. Davies et al. (1999) defines objective item as the item that follows the format-multiple choice, true-false, matching, etc. The objective tests have mechanical scoring. On the whole objective test requires far more careful preparation than subjective test.

1.4 Multiple Choice Items

Multiple choice items are undoubtedly one of the most widely used types of items in objective tests and other being true false, matching and fill in the blanks. Though they are frequently criticized as not leading to the testing of language as communication. Richard et al. (1999:239)

define multiple choice items as “A test item in which the examinee is presented with a question along with four or five possible answers from which one must be selected. Usually the first part of a multiple choice item will be a question or incomplete statement. This is known as stem. The different possible answers are known as alternatives. The alternatives contain (usually) one correct answer and several wrong answers or distractors”.

Heaton (1988:28) points out the following general principles to be considered while constructing multiple choice Items:

1. Each multiple choice item should have only one answer. This answer must be absolutely correct unless the instruction specifies choosing the best option. Although this may seem an easy matter, it is sometimes extremely difficult to construct an item having only one correct answer.
2. Only one feature at a time should be tested: it is usually less confusing for the testees and it helps to reinforce a particular teaching point. Obviously, few would wish to test both grammar and vocabulary at the same times, but sometimes word order and sequence of tenses are tested simultaneously.
3. Each option should be grammatically correct when placed in the stem except of course in the case of specific grammar test items. For examples terms ending with the determiner, followed by options in the form of nouns or noun phrases, sometimes trap the unwary test constructor.
4. All multiple choice items should be at a level appropriate to the proficiency level of the testees. The context, itself, should be at a lower level than the actual problem which the item is testing: a grammar test item should not contain other grammatical features as

difficult as the area being tested, and a vocabulary item should not contain more difficult semantic features in the stem than the area being tested.

5. Multiple choice items should be as brief and as clear as possible (though it is desirable to provide short contexts for grammar items).
6. In many tests, items are arranged in rough order of increasing difficulty. It is generally considered to have one or two simple items to "Lead in" the testees, especially if they are not too familiar with the kind of test being administered. Nevertheless, areas of language which are trivial and not worth testing should be excluded from the test.

Multiple choice item is one of the most widely used types of items in objective tests. In the course Teaching English Literature at B.Ed. level, multiple choice items have been used as objective test. Other items of objective tests like true false, matching, and fill in the blanks have not been used. So, the researcher analyzes the multiple choice items asked in the examination of the course teaching English Literature at B.Ed. level.

1.5 Item Analysis

The performance of the test items, themselves, is of obvious importance in compiling future tests. That's why item analysis allows us to observe the characteristics of a particular question (item) and can be used to ensure the questions are of an appropriate standard and select items for test inclusion. With the help of statistical analysis, it can be measured how the effectiveness of an individual test item is Richard et al. (1999:192) define item analysis as “the analysis of the responses to the

items in a test in order to find out how effective the test items are and to find out if they indicate difference between good and weak students”.

Baker (1989:5) says “One of the immediate benefits of item analysis is to draw the attention of the test designers to the items which contain problems or ambiguities in their construction which escaped notice during the construction of the test”. Hughes (1986:169) says “Even individual tests make their own contribution to the total tests, some contribute more than others, and it is the purpose of item analysis to identify those that need to be changed.” According to Heaton (1988:178) all items should be examined from the point of view of (i) their difficulty level and (2) their level of discrimination.

There are mainly three statistics (The facility value, discrimination index and distracter analysis) to analyze the objective test items. That is to say it is useful to identify those items which are answered correctly by the more able students taking the test and badly by the less able students. The identification of certain difficult items in the test, together with the knowledge of the performance of the individual distractors in multiple choice items can be proved just as valuable in its implications for teaching as for testing.

For this study the researcher analyzed the objective test items. The researcher found the difficult item, discrimination between the testees of the item of the test in the subject Teaching English Literature which is included in the course of B.Ed. 3rd year.

A. Distractor Analysis

The incorrect options of the multiple choice items are called distractors whereas correct one is correct option or key and all together is called alternatives of the multiple choice item. Each distractor should be reasonably attractive and plausible. It should appear right to any testee

who is unsure of the correct option. For many purposes, each distractor should be grammatically correct when it stands by itself: otherwise testees will be exposed to incorrect forms. Unless a distractor is attractive to the student who is not sure of the correct answer, its inclusion in a test item is superfluous. Plausible distractors are based on:

- a. Mistakes in the students own written work.
- b. Their answers in previous tests.
- c. The teacher's experience
- d. A contrastive analysis between native and the target languages.

Distractors should not be too difficult nor demand of their proficiency in the language than the correct option. If they are too difficult, they will succeed only in distracting the good student who will be led into considering the correct option too easy.

B. Item Difficulty (FV)

The index of difficulty or facility value (FV) of an item simply shows how easy or difficult the particular item is to be proved in the test. The index of difficulty (FV) is generally expressed as the fraction (or percentage) of the students who answered the item correctly. It is calculated by using the formula:

$$FV = \frac{R}{N}$$

In this formula, R represents the number of correct answers and N shows the number of students involved in the test.

C. Item Discrimination

The discrimination index of an item indicates the extent to which the item discriminates between the testees, separating the more able

testees from the less able. It measures how performance on one test item correlates to the performance in the test as whole. Heaton (1988:180) points out the following statistical method for calculating item discrimination.

1. Arrange the scripts in rank order of total score and divide into two groups of equal size (i.e. the top half and the bottom half). If there is an odd number of scripts, dispense with one script chosen at random.
2. Count the number of those candidates in the upper group answering the first item correctly; then count the number of lower group candidates answering the item correctly.
3. Subtract the number of correct answers in the lower group from the number of correct answers in the upper group i.e. find the difference in the proportion passing in the upper group and the proportion passing in the lower group.
4. Divide this difference by the total number of candidates in one group:

$$D = \frac{\text{CorrectU} - \text{CorrectL}}{n}$$

1.6 Review of Literature

Many researches have been carried out in the field of language testing. There are various aspects of language testing. Some of the aspects are studied by some of the researchers such as comparison of objective and close test, testing of reading skill, wash back effects, analysis of subjective marking, comparison of discrete point test and integrative test and there is only one research work about multiple choice objective test. These studies are briefly introduced below:

Khanal (1997) Studied on the effectiveness of Close Test over conventional objective test in testing reading comprehension in English. The researcher found out that the private schools' students' performance was better than public schools students' performance in both objective and close test. The researcher also concluded that the close test is better than the objective test in testing reading comprehension.

Dhakal (2000) carried out a work on "Effectiveness of Discrete Point Test and Integrative Test as Measures of English Language Proficiency". This study compares the two types of test- Discrete point test and Integrative test. The researcher concluded that the score in the seen text was naturally higher than unseen text and girl's performance was better than boys' performance.

Subedi (2000) carried out his M.Ed. thesis on reading comprehension of Grade IX students of Jhapa and Kathmandu districts. The researcher concluded that the students of urban schools had better reading skills than those of rural schools and students could understand magazines better than newspaper.

Gautam (2001) analyzed subjective marking. The researcher observed the various techniques of marking in subjective test followed by the teachers and he found that some clues to mark subjective test to maximize the reliability.

Neupane (2004) studied on "Wash Back Effect of Examination: A Case of Communicative English". The researcher found that the teacher centered method has been used and examination hasn't helped to promote the communicative competence of the students.

Baral (2004) carried out the research on "Item Analysis of Multiple Choice Objective Test: A Practical Study". He found that items with good facility value need not necessarily have good discrimination index and the

students at upper level guess more correctly than the students at lower level. He also concluded that only 62% items average in the years from 2054 to 2058 of fundamentals of language and linguistics are statistically satisfactory.

This research differs from the above mentioned studies that have been done so far under the department of English language education. This study carries out a purely practical study of multiple choice objective test items of the course ‘Teaching English Literature of B.Ed. 3rd year (Major English) from 2057 to 2061.

1.7 Objectives of the Study

This study has the following objectives:

- To determine the quality of multiple choice test items asked in the years from 2056-2060 of the course ‘Teaching English Literature’ in terms of
 - the index of difficulty (facility value)
 - the index of discrimination
- To suggest some measures for improving multiple choice test items.
- To suggest some pedagogical implications.

1.8 Significance of the Study

This study attempted to find out the cause of common difficulty with the help of the statistical efficiency of multiple choice objective test that the students point out in the above mentioned subjects. This study will be fruitful in the field of language testing. This study will be very significant for the test designers, the curriculum designers, the teachers and the students involved directly in the teaching and learning process. It will also be very useful for the researcher and others involved directly indirectly in this field.

CHAPTER TWO

METHODOLOGY

To carry out this research work on the basis of the aforementioned objectives, the researcher used the following methodology.

2.1 Sources of Data

This study was entirely field based and practical. Mainly two types of sources of data were utilized.

2.1.1 Primary Sources of Data

The students who had just completed B. Ed. 3rd year were the primary sources of data. There were 100 students from different campuses including TU affiliated campuses and TU Constituent of Kathmandu, Pokhara and Baglung.

2.1.2 Secondary Sources of Data

The researcher consulted various books on language testing, articles, journal and the research works as secondary sources of data. They were used to facilitate this study.

2.2 Population of the Study

The students who had just finished B. Ed. 3rd year from different campuses of Kathmandu, Pokhara and Baglung were the population of the study. There were altogether 100 students from both the TU affiliated campuses and TU Constituent campuses

2.2.1 Sample Population

The population of this study were 100 students from different campuses (both public/private and TU Constituent). The researcher had taken 20 students from Prithvi Narayan Campus (10 girls and 10 boys), 20 students from college of Education (10 girls and 10 boys), 20 students

from Mahendra Ratna Campus Tahachal (10 girls and 10 boys), 10 students from Gramin Adarsha (5 girls and 5 boys), 10 students from Mahendra Multiple Campus Baglung (5 girls and 5 boys) and 10 students from Machhapuchhre Campus, Pokhara (5 girls and 5 boys) and 10 students from Kathmandu Sikshya Campus, Kathmandu. All together, 50 students were from TU Constituent campuses and 50 students were from the private campuses. There were equal number of the students from both sex (male and female).

2.2.2 Sampling Procedure

Seven different campuses including TU affiliated and TU constituent were selected from the different areas Kathmandu, Pokhara and Baglung. The researcher used non- random judgmental sampling procedure to select the population. These population were selected 50 students from TU Constituent and 50 students from private campuses.

2.3 Tools for Data Collection

As this was a practical study, the researcher prepared a question set consisting of 100 multiple choice items asked in the years 2056-2060 as the tool for the data collection.

2.4 Process of Data Collection

In this practical study, the researcher collected multiple choice items of the subject 'Teaching English Literature' at B.Ed. third year from the years 2056 to 2060. After preparing the question set, the researcher himself visited the concerned campuses to meet the students who had just finished their 3rd year of B.Ed. Sometimes the researcher had to consult the students individually since some of the students were out of the campuses after they finished 3rd year. The researcher then requested the students to help in this study. Then after the researcher fixed the time of test and administered the test for 1:40 hours. The

researcher administered the test to the required population as sampled above. He provided the general instruction about the test and collected the papers after the completion of the test. He checked and tabulated the marks of students and analyzed them systematically by using some statistical tools.

2.5 Limitations of the Study

The study has the following limitations:

- The study covered only the performance of the sampled students of the selected campuses.
- This study was only based on multiple choice items of 'Teaching English Literature' at B.Ed. third year.
- This study was only based on the questions asked in the year from 2056 to 2060.
- This study was limited to Kathmandu, Pokhara and Baglung.

CHAPTER – THREE

ANALYSIS AND INTERPRETATION OF DATA

This chapter deals with the analysis and interpretation of the data. While analyzing the data the responses were assigned marks. For one correct response one mark was awarded. The researcher determined the student's responses correct or incorrect on the basis of the answer sheet he had already prepared. The marks were tabulated under various headings. The objectives were considered constantly to tabulate and carry out statistical analysis. Various statistical tools and diagrams or charts were used to analyze the data illustratively.

This part of the study is divided into two sections. The first one deals with the performance of the students in the MCQ asked in the examination of 2056-2060. The analysis of multiple choice test items was based on the following statistical tools:

- Index of the difficulty.
- Index of discrimination.
- Power of the distractors.

To determine the index of difficulty (facility value) the items were categorized into three categories.

- Easy items $\rightarrow < 0.7 F_v$
- Average items $\rightarrow 0.3 - 0.7 F_v$
- Difficult items $\rightarrow > 0.3$

Similarly the discrimination index of the items was calculated by dividing them into five categories.

- Perfect discrimination $\rightarrow DI \text{ of } + 1.$
- Good discrimination $\rightarrow DI \text{ above } 0.7.$
- Moderator discrimination $\rightarrow DI \text{ between } 0.3 - 0.7.$
- Poor discrimination $\rightarrow DI \text{ between } 0 - 0.3.$
- Negative discrimination $\rightarrow DI \text{ in minus between } 0 - 1.$

Similarly, the distractor analysis is carried out on the basis of the responses of the students. If very few students tick that distracter, then this one is less stronger that which many students tick. In each question there are four options and one of them is correct answer and other three are the distracters.

3.1 The multiple choice test items asked in the year 2056

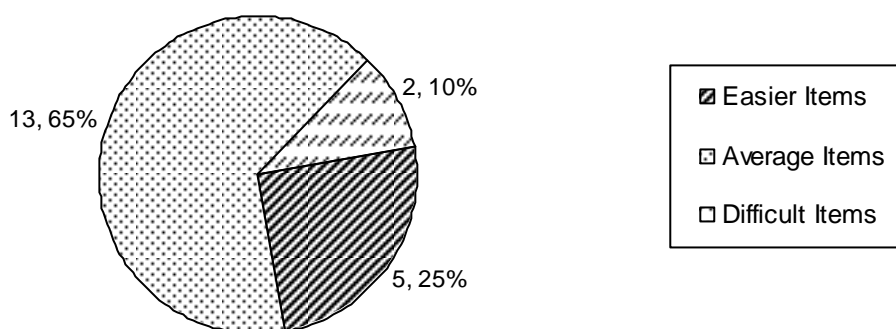
Table 1: Analysis of the Multiple Choice Test Items Asked in the Year 2056

S.N.	No. of Students answering in each options				FV	DI	No. of students who didn't answer
	A	B	C	D			
1	9	15	4	72	0.72	0.16	
2	22	65	8	0	0.68	-0.10	5
3	30	30	52	14	0.52	0.56	
4	18	46	9	27	0.46	0.2	
5	22	31	8	39	0.391	0.62	
6	5	14	73	8	0.73	0.02	
7	80	8	4	8	0.8	0	
8	30	11	0	59	0.59	0.34	
9	48	7	3	42	0.48	-0.16	
10	20	5	11	64	0.64	0.52	
11	82	6	3	9	0.82	0.04	
12	7	19	72	2	0.72	0.16	
13	26	0	7	67	0.67	0.1	
14	31	42	16	11	0.42	0.56	
15	32	45	16	7	0.16	0.32	
16	37	0	43	20	0.37	0.14	
17	57	4	17	12	0.63	0.6	10
18	35	23	17	25	0.17	0.26	
19	47	43	0	10	0.43	0.54	
20	50	12	13	25	0.5	1	
Total Average					0.54	0.27	

In the analysis of the test items of the year 2056 it has been found that most of the items belong to the average difficulty level since the test items were categorized into three categories as easy items, average items and the difficult items.

Out of twenty test items (Multiple choice questions) only five questions were found to be too easy since the facility value of these questions was above 0.7. In the analysis it was seen that two questions were very much difficult since the facility value is 0.16 and 0.17 of question no. 15 and 18 respectively. That reveals that only 10% questions are too difficult and 25% were too easy. Rest of the items, i.e. 13 items, are of average difficulty level.

Easier Items – 5, → 25%
 Average Items – 13 → 65%
 Difficult Items – 2 → 10%

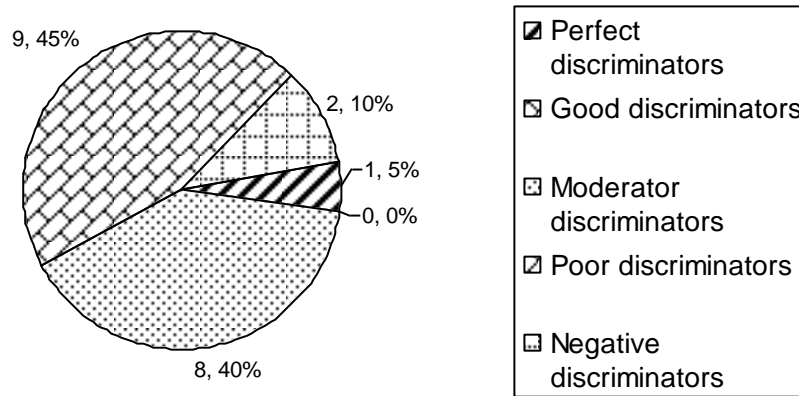


Pie Chart No. 1

The analysis of the discrimination index of the items revealed that only one item, i.e. item no. 20 was the perfect discriminator. None of the items were found to be good discriminators and eight out of twenty items were moderate discriminators. Nine items out of twenty were poor discriminators and two items out of twenty were negative discriminators.

Therefore,

Perfect Discriminators : 1, → 5%
 Good Discriminators : 0 → 0%
 Moderate Discriminators : 8 → 40%
 Poor Discriminators : 9 → 45%
 Negative Discriminators : 2 → 10%



Pie Chart No. 2

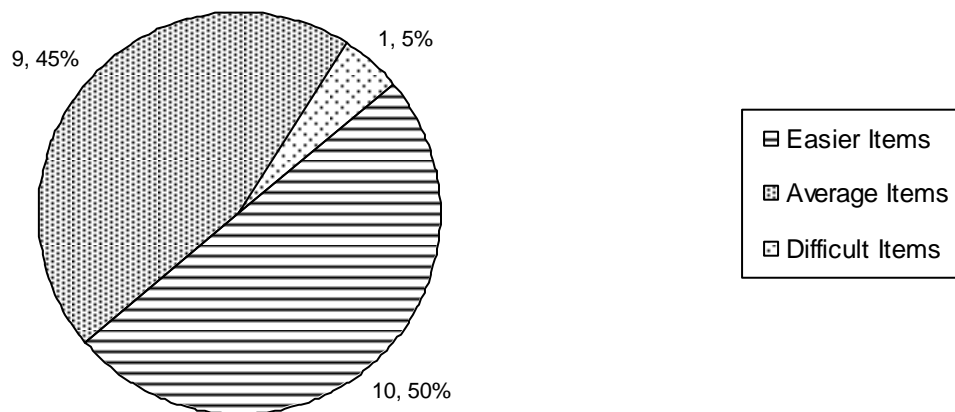
3.2 The Multiple Choice test items asked in the year 2057

Table 2: Analysis of the Multiple Choice test items asked in the year 2057

S.N.	No. of Students answering in each options				FV	DF	No. of students who didn't answer
	A	B	C	D			
1	29	13	31	27	0.29	0.22	
2	0	17	3	80	0.80	0.40	
3	75	14	5	6	0.95	0.42	
4	47	7	46	0	0.47	0.66	
5	0	69	8	20	0.71	0.43	3
6	27	17	4	37	0.42	0.52	13
7	7	13	10	70	0.70	0.52	
8	27	65	0	8	0.65	0.7	
9	19	7	25	49	0.49	0.98	
10	20	61	3	16	0.61	0.58	
11	0	3	81	11	0.85	0.27	5
12	7	27	62	4	0.62	0.16	
13	20	23	27	30	0.30	0.52	
14	6	79	8	7	0.79	0.34	
15	18	0	0	80	0.81	0.38	2
16	5	0	62	33	0.62	0.76	
17	1	2	11	86	0.86	0.28	
18	3	42	6	49	0.42	0.76	
19	78	13	9	0	0.78	0.36	
20	0	2	6	98	0.98	0.16	
Total Average					0.65	0.47	

Table 2 represents the analysis of the test items asked in the year of 2057. It revealed that only one item out of twenty was difficult since the percentage of these difficult items is 5. Item no. 1 was the difficult item since their facility values are 0.29. Similarly, item no. 5, 6, 11 and 15 were not totally answered by all the students. Ten items out of twenty were found to be above the 0.7 level of facility value that revealed that almost 50% of the items were very easier. Nine items out of twenty were in the average level facility value since their facility value was within 0.3 – 0.7. Only one item out of 20 was a difficult item since its facility value was 0.29.

Easier Items – 10	→ 50%
Average Items – 9	→ 45%
Difficult Items – 1	→ 5%



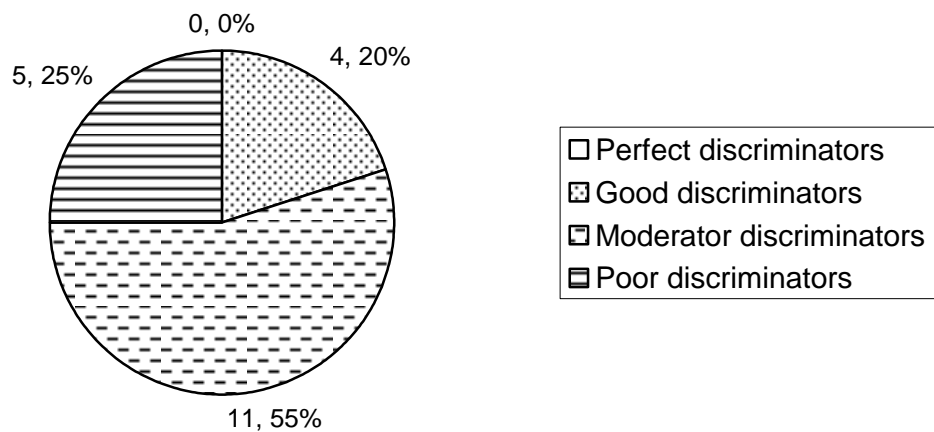
Pie Chart No. 3

The analysis of the discrimination index of the items, revealed that there was no perfect and negative discriminator. There were four good discriminators out of 20. These items were item No. 8, 9, 16 and 18 since their discrimination index were 0.7, 0.98, 0.76 and 0.76 respectively. Eleven items were moderate discriminators. They were item No. 2, 3, 4,

5, 6, 7, 10, 13, 14, 15 and 19 since their discrimination index were 0.40, 0.42, 0.66, 0.43, 0.52, 0.58, 0.52, 0.34, 0.38 and 0.36 respectively. There were five poor discriminator items i.e. the items number 1, 11, 12, 17 and 20 whose DI was 0.22, 0.27, 0.16, 0.28 and 0.16 respectively.

Similarly the condition of discriminators is as follows:

Perfect discriminators – 0	→ 0%
Good discriminators – 4	→ 20%
Moderate discriminators – 11	→ 55%
Poor discriminators – 5	→ 25%



Pie-Chart No. 4

3.3 The multiple choice test items asked in the year 2058

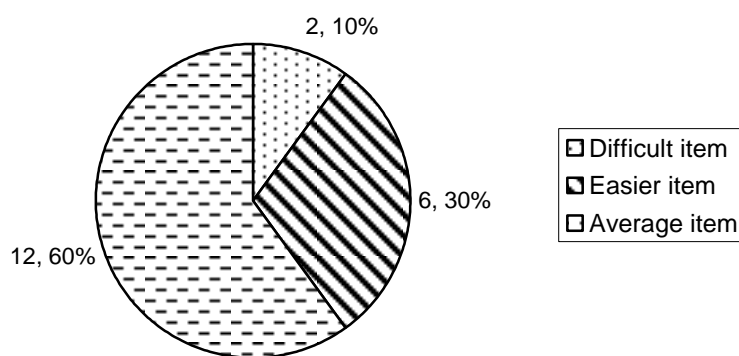
Table 3: Analysis of the Multiple Choice test items asked in the year 2058

S.N.	No. of Students answering in each options				FV	DI	No. of students who didn't answer
	A	B	C	D			
1	55	×	21	19	0.58	0.55	5
2	17	27	37	19	0.19	0.34	
3	85	×	×	15	0.85	0.22	
4	91	0	0	9	0.91	0.34	
5	0	0	97	3	0.97	0.06	
6	35	15	7	43	0.43	0.34	
7	20	41	18	1	0.45	0.58	10
8	32	4	47	17	0.47	0.66	
9	15	29	5	51	0.29	0.42	
10	11	8	73	8	0.73	0.46	
11	65	4	23	6	0.66	0.52	2
12	12	21	42	25	0.42	0.8	
13	10	26	54	10	0.54	0.72	
14	18	13	21	48	0.48	0.64	
15	7	4	83	6	0.83	0.14	
16	21	47	23	9	0.47	0.66	
17	0	85	7	8	0.85	0.22	
18	52	24	14	6	0.54	0.82	4
19	28	19	46	7	0.46	0.6	
20	5	20	60	15	0.60	0.76	
Total Average					0.58	0.49	

Table 3 revealed the analysis of the test items asked in the year 2058. In this test, two of the items were found to be the difficult items. The most difficult one was item no. 2 because the facility value of this item was 0.19. The another difficult item was item no. 9 since its facility value was 0.29. Among twenty items, six items were found to be the easier. Item no. 3, 4, 5, 10, 15 and 17 were easier since their facility values were 0.85, 0.91, 0.97, 0.78, 0.83 and 0.85 respectively. The most

item was item no. 5 since 97 out of 100 students answered it correctly. Other twelve items belong to the average level of difficulty. Four different items, item no. 1, 7, 11 and 18 were not answered by 5, 10, 2 and 4 students respectively.

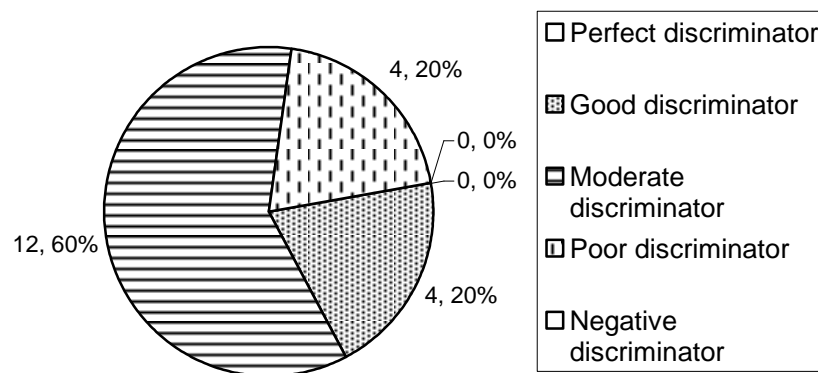
Difficult item	– 2	→ 10%
Easier item	– 6	→ 30%
Average item	– 12	→ 60%



Pie-Charts No. 5

The analysis of the discrimination index of the items revealed that item no. 12, 13, 18 and 20 were good discriminators whose DI was 0.8, 0.72, 0.82 and 0.76 respectively. There were 12 moderate discriminator items out of 20. These items were 1, 2, 4, 6, 7, 8, 9, 10, 11, 14, 16 and 19 whose DI values were 0.55, 0.34, 0.34, 0.34, 0.58, 0.66, 0.42, 0.46, 0.52, 0.64, 0.66 and 0.6 respectively. Similarly, four items were found in poor discriminators. These items were item no. 3, 5, 15 and 17 whose DI values were 0.22, 0.06, 0.14 and 0.22 respectively. None of the items were perfect and negative discriminators.

Perfect discriminator	– 0,	→ 0%
Good discriminator	– 4	→ 20%
Moderate discriminator	– 12	→ 60%
Poor discriminator	– 4	→ 20%
Negative discriminator	– 0	→ 0%



Pie-Charts No. 6

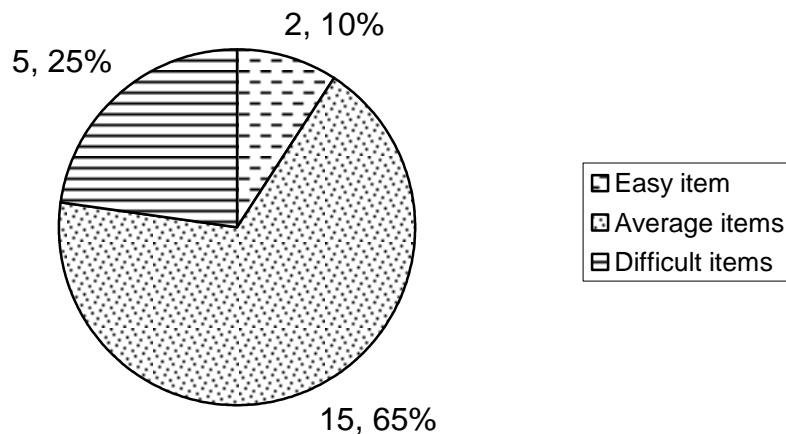
3.4 The Multiple Choice Test Items asked in the year 2059

Table 4: Analysis of the Multiple Choice Test Items asked in the year 2059

S.N.	No. of Students answering in each options				FV	DI	No. of students who didn't answer
	A	B	C	D			
1	22	21	44	13	0.44	0.52	
2	56	8	5	31	0.56	0.68	
3	36	32	29	3	0.29	0.58	
4	0	15	36	49	0.49	0.62	
5	12	21	45	22	0.45	0.90	
6	25	40	15	20	0.40	0	
7	24	60	6	9	0.61	0.73	1
8	31	4	46	19	0.46	0.68	
9	25	32	31	12	0.25	0.50	
10	91	0	0	9	0.91	0.10	
11	0	18	22	60	0.60	0.60	
12	5	31	15	49	0.49	0.62	
13	45	5	8	42	0.45	0.70	
14	65	10	15	10	0.65	0.70	
15	7	67	5	21	0.67	0.58	
16	12	2	58	28	0.12	0.24	
17	0	0	29	71	0.71	0.46	
18	20	30	0	50	0.5	1	
19	64	4	22	10	0.64	0.64	
20	36	6	48	10	0.36	0.48	
Total Average					0.501	0.56	

Table 4 revealed the analysis of the test item asked in the year 2059. In this test, three of the test items were found to be difficult among the twenty items. Test item no. 3, 9 and 16 were the difficult items since their facility values were 0.29, 0.25 and 0.12. Among these, item no. 16 was the most difficult item. Two of the items were found to be more easy. Item no. 10 and item no. 17 were easy items. Since the difficulty level of these two was 0.91 and 0.71. The most easy item was item no. 10 since 91 out of 100 students answered it correctly. All others were the average level items for difficulty. This test item has the following percentage in terms of difficulty.

Easy item	– 2 out of 20,	Percentage : 10%
Average items	– 15 out of 20,	Percentage : 65%
Difficult items	– 5 out of 20,	Percentage : 25%

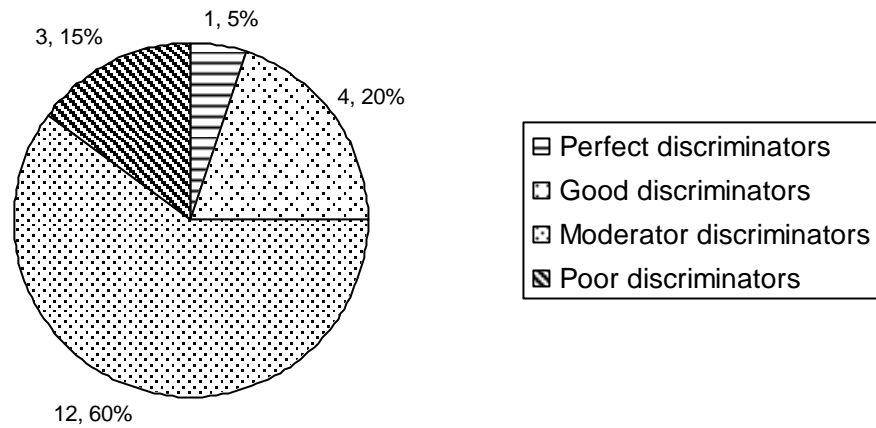


Pie-Chart No. 7

The analysis of the discrimination index of the items revealed that item no. 18 was the perfect discriminator since its DI was 1. Four different items were good discriminators. They were – item no. 5, 7, 13 and 15 whose DI was 0.90, 0.73, 0.70 and 0.70 respectively. Similarly, twelve items were found to be the moderate discriminators. They were item no. 1, 2, 3, 4, 8, 9, 11, 12, 15, 17, 19 and 20 whose DI values were

0.52, 0.68, 0.58, 0.62, 0.68, 0.50, 0.60, 0.62, 0.58, 0.46, 0.64 and 0.48 respectively. Three items were poor discriminators, they were item no. 6, 10 and 16 whose DI was 0, 10 and 0.24 respectively. None of the items were found to be negative discriminators.

DI type	No.	Percentage
Perfect discriminator	1	5%
Good discriminator	4	20%
Moderate discriminator	12	60%
Poor discriminator	3	15%



Pie-Chart No. 8

3.5 The Multiple Choice Items asked in the year 2060

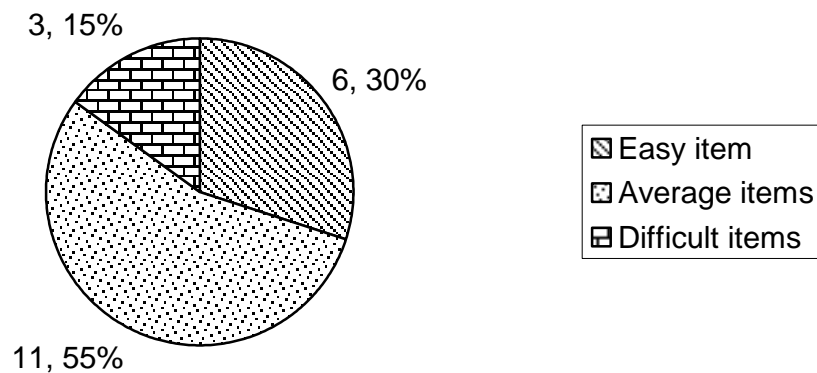
Table 5: Analysis of the Multiple Choice Items asked in the year 2060

S.N.	No. of Students answering in each options				FV	DI	No. of students who didn't answer
	A	B	C	D			
1	40	23	13	24	0.4	0.8	
2	24	4	8	61	0.62	0.81	3
3	0	7	82	11	0.82	0.36	
4	3	12	31	54	0.54	0.92	
5	42	34	3	16	0.44	0.84	5
6	12	75	7	6	0.75	0.50	
7	81	4	8	7	0.81	-0.02	
8	25	45	22	8	0.45	0.70	
9	32	48	5	15	0.48	0.84	
10	28	25	12	31	0.29	0.47	4
11	9	5	0	86	0.86	0.28	
12	2	92	5	1	0.92	0.16	
13	14	23	36	25	0.37	0.69	2
14	46	44	10	0	0.44	0.88	
15	85	2	6	7	0.85	0.3	
16	25	36	29	10	0.25	0.5	
17	62	18	10	5	0.65	0.79	5
18	37	42	15	6	0.37	0.74	
19	15	25	32	28	0.15	0.18	
20	10	56	16	18	0.56	0.68	
Total Average					0.55	0.57	

Table 5 revealed the analysis of the test item asked in the year 2060. Out of the twenty test items, three were found to be difficult items, i.e. item no. 10, 16 and 19 since their facility value was 0.29, 0.25 and 0.15 respectively. The most difficult item was item no. 19. Similarly, six other items belong to the category of the easier items. They were item no. 3, 6, 7, 11, 12 and 15 whose facility value was 0.82, 0.75, 0.81, 0.86, 0.92

and 0.85 respectively. Other items belonged to the average level of difficulty. The most easy item was item no. 12 since 92 students out of hundred answered it correctly. It was found that 19 students had not answered some questions. That means, item no. 5 by 5 students, item no. 10 by 4 students, item no. 13 by 2 students and item no. 17 by 5 students.

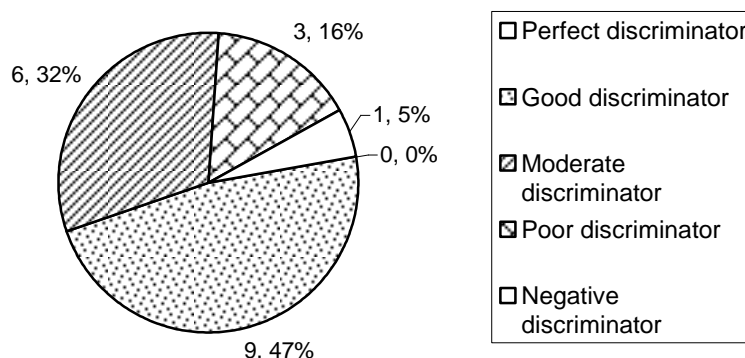
Easier items – 6, Percentage : 30%
 Average items – 11, Percentage : 55%
 Difficult items – 3, Percentage : 15%



Pie-Chart No. 9

The analysis of the discrimination index of the items revealed that none of the items were found to be perfect discriminator out of twenty items, nine items were found to be good discriminators, they were – item no. 1, 2, 4, 5, 9, 10, 14, 17 and 18. Since their discrimination index was 0.8, 0.81, 0.92, 0.84, 0.84, 0.70, 0.84, 0.88, 0.79 and 0.74 respectively. Six other items were moderate discriminators – they were, item no. 3, 6, 10, 15, 16 and 20 whose discrimination index was 0.36, 0.50, 0.47, 0.30, 0.50 and 0.68 respectively. Similarly, item no. 11, 12 and 19 were poor discriminators. Finally, item no. 7 was a negative discriminator since the discrimination index was 0.02.

Perfect discriminator – 0,	Percentage: 0%
Good discriminator – 9	Percentage: 47%
Moderate discriminator – 6	Percentage: 32%
Poor discriminator – 3	Percentage: 16%
Negative discriminator – 1	Percentage: 5%



Pie-Charts No. 10

3.6 Sex-wise performance of the Students

Table 6: Analysis of Sex-wise Performance of College of Education

Female - 10

S.N.	2056	2057	2058	2059	2060	Average
1	13	14	12	11	13	11.8
2	13	13	12	11	13	12.4
3	11	12	13	14	15	13.0
4	15	12	13	14	11	13.0
5	12	12	11	13	14	12.4
6	14	12	13	11	14	12.8
7	15	15	14	13	14	14.2
8	13	12	12	13	14	12.8
9	12	13	13	14	14	13.2
10	13	14	15	14	15	14.2
Average Total	13.1	12.9	12.8	12.8	13.7	13.08

Male - 10

S.N.	2056	2057	2058	2059	2060	Average
1	12	13	11	12	14	12.40
2	14	15	12	11	13	13.00
3	13	14	15	12	11	13.00
4	13	10	11	12	13	11.80
5	12	12	12	13	14	12.60
6	11	11	12	13	13	12.00
7	14	10	13	14	15	13.20
8	14	14	13	14	12	13.40
9	14	15	15	13	12	13.80
10	14	13	14	15	14	14.00
Average Total	13.1	12.7	12.8	12.9	13.1	12.95

To see the level of performance of the students of college of education the females were found to have done better than the males. The average performance score of females was 13.08 and the average performance score of males was 12.95. However, it also revealed that some females also had done better in some of the years (As seen in the table, females had done better in 2059). The level of performance between males and females were equal in some years i.e. 2056 and 2058.

Table 7: Analysis of Sex-wise Performance of Prithivi Narayan Campus

Males

S.N.	2056	2057	2058	2059	2060	Average
1	12	10	14	9	16	12.2
2	10	12	13	15	18	13.6
3	15	14	11	10	15	13.0
4	9	13	12	11	16	12.2
5	14	11	13	11	10	11.8
6	15	14	12	11	11	12.6
7	13	12	11	10	14	12.0
8	16	15	12	13	12	13.6
9	12	13	14	10	12	12.2
10	11	12	13	10	14	12.0
Average Total	12.7	12.6	12.5	11.0	13.8	12.52

Females

S.N.	2056	2057	2058	2059	2060	Average
1	15	10	12	11	14	12.4
2	16	17	10	12	13	13.6
3	10	11	13	12	15	12.2
4	15	10	12	14	10	12.2
5	12	13	14	11	10	12.0
6	16	12	12	13	13	13.2
7	16	14	12	11	13	13.2
8	12	13	14	15	16	14.0
9	16	14	13	12	11	13.2
10	10	12	14	13	10	11.8
Average Total	13.8	11.6	12.6	12.4	12.5	12.58

Table – 7 revealed the analysis of the scores of the test items of the students of PNC, Pokhara. To see wholistically, the average performance of the females was slightly better in comparison to males. Since the former had the average score of 12.52 and the latter had the average score of 12.58. It also revealed the year-wise difference of the scores according to the sex variable. In the year 2057, the males had done better than the females since the former had got 12.6 and latter one had got 11.6. In the year 2058, the females were slightly better than the males since their scores were 12.6 and 12.5 respectively. In the year 2059 also, the females were better than males since the former had 12.4 and latter had 11.4. Similarly, in the year 2060, the males had done better than the females since the former had got 13.8 and the latter had got 12.5.

**Table 8: Analysis of Sex-wise Performance of Mahendra Ratna
Campus**

Males

S.N.	2056	2057	2058	2059	2060	Average
1	12	12	13	14	11	12.4
2	14	13	12	10	11	12.0
3	16	15	10	12	13	13.2
4	15	10	12	13	12	12.4
5	14	14	13	12	12	13.0
6	14	13	12	10	14	12.6
7	14	12	12	11	13	12.4
8	10	12	12	13	12	11.8
9	13	14	13	12	11	12.6
10	14	13	15	12	10	12.8
Average Total	13.6	12.8	12.8	10.9	11.9	12.4

Females

S.N.	2056	2057	2058	2059	2060	Average
1	13	12	10	11	15	12.2
2	15	12	14	13	12	13.2
3	14	13	12	10	11	12.0
4	14	12	13	12	10	12.2
5	12	13	12	14	15	13.2
6	15	10	12	13	14	12.8
7	15	10	12	13	12	12.4
8	13	12	11	14	14	12.8
9	15	12	14	13	12	13.2
10	12	13	14	12	13	12.8
Average Total	13.8	11.9	12.4	12.5	12.8	12.68

Table – 8 represents the analysis of the test scores of the students of MRC. According to the sex variable. Wholistically, the females had performed slightly better than the males. Since the former had the average score of 12.68 and the latter had 12.4 respectively. But it also revealed that there was difference in the performance of the males and females in year-wise way. In the year 2056, the females had done slightly better than males. In the year 2057, the males were far better than females since the former had 12.8 but latter had average score of 11.9. Similarly males had done better in 2058 but in the year 2059 and 2060, the females had performed better.

Table 9: Analysis of Scores of the Students of Machhapuchhre Campus

Males

S.N.	2056	2057	2058	2059	2060	Average
1	13	12	14	13	15	13.4
2	12	14	12	13	11	12.4
3	14	15	14	16	11	14.0
4	14	14	13	15	13	13.8
5	14	14	15	14	12	13.8
Average Total	13.4	13.8	13.6	14.2	12.4	13.48

Females

S.N.	2056	2057	2058	2059	2060	Average
1	14	14	13	13	12	13.2
2	13	11	12	12	14	12.4
3	15	12	13	13	14	13.4
4	15	14	13	13	13	13.6
5	12	12	14	14	12	12.8
Average Total	13.8	12.6	13.0	12.8	13	13.04

Table-9 represents the analysis of the scores of the test items according to the sex variable. It revealed that males were slightly better in comparison to the females since the total average achievement score of the former was 13.48 and of the latter was 13.04. But in some years the females were better than the males since in the year 2056 and 2060, the females had performed better and in the year 2057, 2058 and 2059, the males had done better.

**Table 10: Analysis of the Scores of the Students of Kathmandu
Sikshya Campus**

Males

S.N.	2056	2057	2058	2059	2060	Average
1	12	14	14	13	13	13.2
2	15	11	15	12	13	13.2
3	15	10	16	13	12	13.2
4	12	13	11	10	15	12.2
5	9	13	14	13	12	12.2
Average Total	12.6	12.2	14.0	12.2	13.0	12.8

Females

S.N.	2056	2057	2058	2059	2060	Average
1	13	12	13	14	15	13.4
2	13	12	14	14	15	13.6
3	14	11	12	13	13	12.6
4	11	10	14	15	12	12.4
5	10	12	11	13	14	12.0
Average Total	12.2	11.4	12.8	13.8	13.8	12.8

Table -10 revealed the analysis of the scores of the students of KSC in the multiple choice test items. It was found out that both the males and females had got equal score of 12:8 in total average. However, we found some differences in the average scores of the students in year-wise way. The males were better in the test items of 2056, 2057 and 2058 and the males were better in the year 2059 and 2060.

Table 11: Analysis of the Scores of the Students of Gramin Adarsh Campus

Males

S.N.	2056	2057	2058	2059	2060	Average
1	12	13	14	10	12	12.20
2	13	13	14	13	15	13.60
3	13	12	11	15	16	13.40
4	13	13	14	14	12	13.20
5	12	14	16	17	14	14.60
Average Total	12.6	13.0	13.8	14.4	13.8	13.52

Females

S.N.	2056	2057	2058	2059	2060	Average
1	14	12	13	14	15	13.60
2	14	14	13	12	11	12.80
3	15	15	16	14	14	14.80
4	15	14	13	13	12	13.40
5	14	12	11	13	16	13.20
Average Total	14.4	13.4	13.2	13.2	13.6	13.56

Table No. 11 represents the analysis of the scores of the students of Gramin Adarsha College in the test item asked in the year from 2056 – 2060 B.S. The table revealed that the females were slightly better than the

males in the total average since the former had got 13.56 and latter had got 13.52. But there was difference in the score of the males and females in year-wise analysis. The males had done better in the year 2058, 2059 and 2060. But the females had done better in the year 2056 and 2057 in comparison with the year 2056 and 2057 in comparison with the males.

Table 12: Analysis of the Scores of the Students of Mahendra Multiple Campus

Males

S.N.	2056	2057	2058	2059	2060	Average
1	14	12	14	13	14	13.40
2	13	12	13	14	15	13.40
3	12	13	12	11	11	11.80
4	13	12	11	15	14	13.00
5	13	10	12	14	13	12.40
Average Total	13.0	13.0	13.4	13.4	13.4	13.24

Females

S.N.	2056	2057	2058	2059	2060	Average
1	15	15	14	13	13	14.00
2	14	14	13	14	14	13.80
3	11	14	12	13	14	12.80
4	11	10	14	13	12	12.00
5	14	13	11	12	15	13.00
Average Total	13.0	13.2	12.8	13.0	13.6	13.12

Table No. 12 includes the scores of the students of the Mahendra Multiple Campus in the Multiple choice test items. It revealed that wholistically the males were better than the females since the former had got 13.24 and the latter had got 13.12 respectively. But in some of the

years, the males were weaker than the females and vice versa. In the year 2057 and 2060, the females were better than the males. In the year 2058 and 2059, the males were better than the females. Similarly, in the year 2056, both males and females were equal in terms of the achievement.

3.7 The Performance of Males and Females as a Whole

Table 13: Comparative Analysis of the Performance of males and females as a whole

Sex	Name of Campuses							
	COE	PNC	MRC	MC	KSC	GAC	MMC	Average
Males	12.95	12.52	12.40	13.48	12.8	13.52	13.24	12.98
Females	13.08	12.58	12.68	13.04	12.8	13.56	13.12	12.98
Total	13.01	12.55	12.54	13.26	12.8	13.54	13.18	12.98

Table No. 13 revealed the performance of the students of different campuses in terms of sex. The total performance of the males and females of different campuses were slightly different. The total performance of COE, PNC, MRC, MC, KSC, GAC and MMC was 13.01, 12.55, 12.54, 13.26, 12.8, 13.54 and 13.18 respectively. GAC had got the highest performance (i.e. 13.54) and MRC had got the lowest performance (i.e. 12.54).

The performance between males and females were the same (i.e. 12.98) as a whole.

3.8 The Scores of the students of TU affiliated campuses

Table 14: Analysis of Scores of the Students of TU affiliated campuses

S.N.	College	No. of Students	2056	2057	2058	2059	2060	Total average
1	COE		13.1	12.8	12.8	12.85	13.4	12.99
2	GAC		13.5	13.2	13.5	13.8	13.7	13.54
3	MC		13.6	13.2	13.2	13.5	12.7	13.24
4	KSC		12.4	11.8	13.4	13	13.4	12.8
Total			13.15	12.75	13.22	13.28	13.3	13.14

Table – 14 includes the average scores of the students of the TU affiliated colleges. There were only four difference TU affiliated colleges. The scores in this table can be analyzed from two aspects. The first one was average score of the students of colleges collectively of five years and the other in the average score of all the private colleges in each year.

The table revealed that the students of GAC had got better score in comparison to other colleges since it had the average of 13.54 in comparison to 13.24, 12.8, 12.99 of MC, KSC and COE respectively. Similarly, the students of TU affiliated campuses colleges as a whole had got better scores in the test item of the year 2060. They had secured the lowest in the test year 2057 (i.e. they have scored 12.75 in average). From this we came to the conclusion that the test items of the year 2060 were more facilitative in comparison to the test items of the year 2057 and others. This also revealed that the average performance of the students of

TU affiliated colleges was 13.14. The test items were not so much facilitative and were also not so much non-facilitative.

3.9 The scores of the students of TU Constituent colleges

Table 15: Analysis of Scores of the Students of TU Constituent Colleges

S.N.	College	2056	2057	2058	2059	2060	Total average
1	PNC	13.25	12.1	12.55	11.7	13.15	12.55
2	MRC	13.7	12.35	12.6	11.7	12.35	12.54
3	MMC	13	13.1	13.1	13.2	13.5	13.18
Total average		13.3	12.51	12.75	12.2	13	12.75

Table – 15 represents the scores of the students of the TU Constituent colleges in the years from 2056 to 2060. There were three TU Constituent colleges and the total average scores of the students had been found to be 12.75. This shows that the multiple choice test items were less facilitative to these students in comparison to the former one. The average scores of the students of PNC from the year 2056 to 2060 was 12.55, the scores of the students of MRC was 12.54 and that of MMC was 13.18. From this we derived the conclusion that among these TU Constituent campuses, Mahendra Multiple Campus was facilitated more by these test items. But it also showed that the items were almost equally facilitative to the students of PNC and MRC since their average scores were similar. Similarly, in the year 2056, the TU Constituent campus students had done the best that means the test items of 2056 had best facilitated the students since the average total of the scores is 13.3.

3.10 The scores of the TU affiliated and the TU Constituent colleges

Table 16: Comparative Analysis of the Scores of the Public/Private and the TU Constituent Colleges

S.N.	College	2056	2057	2058	2059	2060	Total average
1	TU affiliated campuses	13.15	12.75	13.22	13.28	13.30	13.14
2	TU Constituent	13.30	12.51	12.75	12.20	1300	12.75

Table – 16 revealed the comparison between the TU affiliated colleges and TU Constituent colleges in terms of the test scores of the students. It was found that the students of TU affiliated colleges had done better in the test items as a whole since their average score was 13.14 and the average score of the students of TU Constituent colleges was 12.75. That means, the test items were facilitative to the TU affiliated campuses college students in comparison to the TU Constituent college students. But the difference was not so much vast.

3.11 Comparative Analysis of Facility Value and Discrimination Index as a Whole

Table 17: Facility Value and Discrimination Index as a Whole

	2056	2057	2058	2059	2060	Total
FV	0.54	0.65	0.58	0.50	0.55	0.56
DI	0.27	0.47	0.49	0.56	0.57	0.46

Table No. 17 revealed the total facility of value and discrimination index of the test items of each year. It was found that the test items of the

year 2056, 2057, 2058, 2059 and 2060 had the facility value of 0.54, 0.65, 0.58, 0.50 and 0.55 respectively. This showed that all the test items of these years had average level difficulty since the total average of the facility value was concluded to be 0.56.

Similarly, the items of 2056, 2057, 2058, 2059 and 2060 had the discrimination index 0.27, 0.47, 0.49, 0.56 and 0.57 respectively. In the year 2056, the test items had poor discrimination power since the discrimination index in 0.27. The test items of 2060 had highest discrimination index i.e. 0.57.

CHAPTER FOUR

FINDINGS AND RECOMMENDATION

4.1 Findings

This was a practical study. The researcher had drawn the following findings of the study.

- In the year 2056, out of twenty items, only 13 items (i.e. 65%) were of average difficulty level, 5 items (i.e. 25%) were easy items and 2 items (i.e. 10%) were difficult items.
- In the year 2056, only one item was found to be perfect discriminator, eight items were moderate discriminators, nine items were poor discriminators, two items were negative discriminators and no items were found to be good discriminators.
- In the year 2057, only eleven items were easy (50%), nine items were of average difficulty level (i.e. 45%) and only one item (i.e. 5%) was difficult item.
- In the year 2057, none of the items were found to be perfect discriminators (i.e. 0%), four items were good discriminators (i.e. 20%), eleven items were moderate discriminators (i.e. 55%) and only five items (i.e. 25%) were found to be poor discriminators.
- In the year 2058, out of twenty items, six items (i.e. 30%) were easier items, twelve items (i.e. 60%) were average items and only two items (i.e. 10%) were found to be the difficult items.
- In the year 2058, none of the items were found to be the perfect discriminators. Only four items (i.e. 20%) were found to be good discriminators, twelve items (i.e. 60%) were moderators, four items

(20%) were poor discriminators and non of the items were found to be the negative discriminator.

- In the year 2059, only two items (i.e. 10%) were found to be easier items. Out of twenty, fifteen items (i.e. 75%) were average items and only three items (i.e. 15%) were found to be the difficult items.
- In the year 2059, out of twenty items, only one item (i.e. 5%) was found to be the perfect discriminator. Only four items (i.e. 20%) were good discriminators, twelve items (i.e. 60%) were moderators, three items (i.e. 15%) were poor discriminators and none of the items were found to be the negative discriminators.
- In the year 2060, out of twenty items, only six items (i.e. 30%) were found to be easier items, eleven items (i.e. 55%) were average items and only three items (i.e. 15%) were found to be the difficult items.
- In the year 2060, out of twenty items, none of the items were perfect discriminators, nine items (i.e. 45%) were found to be good discriminators, seven items (i.e. 35%) were moderators, only three items (i.e. 15%) were poor discriminators and only one item was found to be negative discriminators.
- The average discrimination index of the test items as a whole was 0.46 and the average facility value of test items was 0.56.
- The test items as a whole were found to be functioning no so efficiently since they have moderate level of discrimination index and they were neither too difficult not too easier since they have the average level difficulty.

- Perfect discrimination was impossible.
- It was found that the test item had equally facilitated the males and females since the males and females performed equally in the test.
- The performance of the students of TU affiliated campuses was better than the performance of the students of TU constituent campuses since the former had 13.14 and latter had 12.75.

4.2 Recommendations

On the basis of the findings derived from the study, the researcher had made the following recommendations:

1. Only the statistically satisfactory items should be included in the final version of the test. Items are said to be statistically satisfactory if they are suitable level of difficulty (neither too hard nor too easy) and if they discriminate between those examinees.
2. More powerful protected distractors should replace malfunctioning and weak distractors.
3. Multiple choice items should be brief, clear, unambiguous and grammatically correct.
4. Multiple choice test should be arranged in the order of increasing difficulty. The easy items are desirable to encourage and motivate the poor students while the difficult items are necessary to motivate and encourage the good students.
5. It is desirable to include statistically efficient 25% easy items in the beginning of the test as a 'lead-in' device, 50% items of average FV in the middle of the test and 25% difficult items at the concluding part of the test to motivate the poor, middle and good students.

Though the researcher in this research found out that the items have average difficulty level and moderate level of discrimination, they were not functioning very efficiently. The test items should be developed only after the trial test. The concerned authority should realize the fact that the test items should be efficient and should have good discrimination power. The researcher thought that other different researches should be carried out in the related topic for the betterment of the total testing system.

REFERENCES

- Baral (2004). Item analysis of multiple choice objective text: A Practical study. M.Ed. Thesis Department of English Language Education, T.U. Kirtipur
- Bhattra, G.R. (2001). A Thematic Analysis of Research Report. Kathmandu: Ratna Pustak Bhandar.
- Davies et al. (1999). Dictionary of Language Testing. University of Melbourne.
- Dhakal (2000). Effectiveness of Discrete point test and Integrative test as measures of English Language Proficiency.
- Gautam (2001). An Analysis of Subjective Marking. An unpublished M.Ed. Thesis. T.U., Kathmandu.
- Harris, D.P. (1974). Testing Writing as a Second Language. TMH Edition MC Graw- Hill Book Company, New York.
- Heaton, J.B. (1985). Writing English Language Tests. FLBS: A practical Guide for Teachers of English as a Second or Foreign Language.
- Hughes, A. (1983). Current Development in Language Testing. Academic Press.
- Khaniya (2005). Examination For Enhanced Learning. Kathmandu.
- Neupane (2004). Washback effect of examination: A case of communicative English. An unpublished M.Ed. Thesis. T.U., Kathmandu.
- Richard, J.C., John Platt and Heidi Platt, (1999). Longman Dictionary of Applied Linguistics. Essex: Longman.

Subedi (2000). A Study on Reading Comprehension of Grade IX Students of Jhapa and Kathmandu Districts. An unpublished M.Ed. Thesis. T.U., Kathmandu.

Weir, Cyril, (1998). Communicative Language Testing. U.K.: Prentice Hall.

Cambridge International Dictionary of English (1995). Cambridge University Press. Cambridge.

Khanal (1997). The Effectiveness of Close Test Over Conventional Objective Test in Testing Reading Comprehension in English. An unpublished M.Ed. Thesis. T.U., Kathmandu.

APPENDIX – IV

Raw Score of the Students as a Whole

S.N.	Name of Campus	Score					Total
		2056	2057	2058	2059	2060	
1.	PNC	12	10	14	9	16	61
2.	PNC	10	12	13	15	18	68
3.	PNC	15	10	12	11	14	62
4.	PNC	16	17	10	12	13	68
5.	PNC	15	14	11	10	15	65
6.	PNC	9	13	12	11	16	61
7.	PNC	10	11	13	12	15	61
8.	PNC	15	10	12	14	10	61
9.	PNC	14	11	13	11	10	59
10.	PNC	12	13	14	11	10	60
11.	PNC	15	14	12	11	11	63
12.	PNC	16	12	12	13	13	66
13.	PNC	13	12	11	10	14	60
14.	PNC	16	14	12	11	13	66
15.	PNC	16	15	12	13	12	68
16.	PNC	12	13	14	10	12	61
17.	PNC	11	12	13	10	14	60
18.	PNC	12	13	14	15	16	70
19.	PNC	16	14	13	12	11	66
20.	PNC	10	12	14	13	10	59
21.	MRC	12	12	13	14	11	62
22.	MRC	13	12	10	11	15	61
23.	MRC	14	13	12	10	11	60
24.	MRC	15	12	14	13	12	66
25.	MRC	16	15	10	12	13	66
26.	MRC	14	13	12	10	11	60
27.	MRC	15	10	12	13	12	62
28.	MRC	14	12	13	12	10	61
29.	MRC	14	14	13	12	12	65
30.	MRC	12	13	12	14	15	66
31.	MRC	15	10	12	13	14	64
32.	MRC	14	13	12	10	14	63
33.	MRC	15	10	12	13	12	62
34.	MRC	13	12	11	14	14	64
35.	MRC	14	12	12	11	13	62
36.	MRC	15	12	14	13	12	66

37.	MRC	10	12	12	13	12	59
38.	MRC	12	13	14	12	13	64
39.	MRC	13	14	13	12	11	63
40.	MRC	14	13	15	12	10	64
41.	COE	12	13	11	12	14	62
42.	COE	13	14	12	11	13	63
43.	COE	14	15	12	11	13	65
44.	COE	13	13	12	11	13	62
45.	COE	13	14	15	12	11	65
46.	COE	11	12	13	14	15	65
47.	COE	15	12	13	14	11	65
48.	COE	13	10	11	12	13	59
49.	COE	12	12	12	13	14	63
50.	COE	11	11	12	13	13	60
51.	COE	12	12	11	13	14	62
52.	COE	14	12	13	11	14	64
53.	COE	14	10	13	14	15	66
54.	COE	15	15	14	13	14	71
55.	COE	14	14	13	14	12	67
56.	COE	13	12	12	13	14	64
57.	COE	14	15	15	13	12	69
58.	COE	12	13	13	14	14	66
59.	COE	14	13	14	15	14	70
60.	COE	13	14	15	14	15	71
61.	GAC	12	13	14	13	12	64
62.	GAC	13	13	14	13	15	68
63.	GAC	14	12	13	14	15	68
64.	GAC	13	12	11	15	16	67
65.	GAC	14	14	13	12	11	64
66.	GAC	15	15	16	14	14	74
67.	GAC	15	14	13	13	12	67
68.	GAC	13	13	14	14	12	66
69.	GAC	12	14	16	17	14	73
70.	GAC	14	12	11	13	16	66
71.	MMC	14	12	14	13	14	67
72.	MMC	13	12	13	14	15	67
73.	MMC	15	15	14	13	13	70
74.	MMC	14	14	13	14	14	69
75.	MMC	12	13	12	11	11	59
76.	MMC	11	14	12	13	14	64
77.	MMC	11	10	14	13	12	60
78.	MMC	13	12	11	15	14	65

79.	MMC	14	13	11	12	15	65
80.	MMC	13	16	17	14	13	73
81.	KSC	12	14	14	13	13	66
82.	KSC	13	12	13	14	14	66
83.	KSC	15	11	15	12	12	65
84.	KSC	13	12	14	14	14	67
85.	KSC	15	10	16	13	13	67
86.	KSC	14	11	12	13	13	63
87.	KSC	12	13	11	10	10	56
88.	KSC	11	10	14	15	15	65
89.	KSC	9	13	14	13	13	62
90.	KSC	10	12	11	13	13	59
91.	MC	13	12	14	13	15	67
92.	MC	14	14	13	13	12	66
93.	MC	12	14	12	13	11	62
94.	MC	13	11	12	10	14	60
95.	MC	14	15	14	16	11	70
96.	MC	15	12	13	14	14	68
97.	MC	14	14	13	15	13	69
98.	MC	13	14	13	14	13	67
99.	MC	14	14	15	14	12	69
100.	MC	12	12	14	13	12	63
Total		1324	1268	1288	1270	1307	6457

APPENDIX – V

Raw Score of the Students of Prithivi Narayan Campus

S.N.	Sex	Score					Total
		2056	2057	2058	2059	2060	
1.	M	12	10	14	9	16	61
2.	M	10	12	13	15	18	68
3.	F	15	10	12	11	14	62
4.	F	16	17	10	12	13	68
5.	M	15	14	11	10	15	65
6.	M	9	13	12	11	16	61
7.	F	10	11	13	12	15	61
8.	F	15	10	12	14	10	61
9.	M	14	11	13	11	10	59
10.	F	12	13	14	11	10	60
11.	M	15	14	12	11	11	63
12.	F	16	12	12	13	13	66
13.	M	13	12	11	10	14	60
14.	F	16	14	12	11	13	66
15.	M	16	15	12	13	12	68
16.	M	12	13	14	10	12	61
17.	M	11	12	13	10	14	60
18.	F	12	13	14	15	16	70
19.	F	16	14	13	12	11	66
20.	F	10	12	14	13	10	59

APPENDIX – VI

Raw Score of the Students of Mahendra Ratna Campus

S.N.	Sex	Score					Total
		2056	2057	2058	2059	2060	
1.	M	12	12	13	14	11	62
2.	F	13	12	10	11	15	61
3.	M	14	13	12	10	11	60
4.	F	15	12	14	13	12	66
5.	M	16	15	10	12	13	66
6.	F	14	13	12	10	11	60
7.	F	15	10	12	13	12	62
8.	M	14	12	13	12	10	61
9.	F	14	14	13	12	12	65
10.	M	12	13	12	14	15	66
11.	F	15	10	12	13	14	64
12.	F	14	13	12	10	14	63
13.	M	15	10	12	13	12	62
14.	F	13	12	11	14	14	64
15.	F	14	12	12	11	13	62
16.	F	15	12	14	13	12	66
17.	M	10	12	12	13	12	59
18.	F	12	13	14	12	13	64
19.	M	13	14	13	12	11	63
20.	M	14	13	15	12	10	64

APPENDIX – VII

Raw Score of the Students of College of Education

S.N.	Sex	Score					Total
		2056	2057	2058	2059	2060	
1.	M	12	13	11	12	14	62
2.	F	13	14	12	11	13	63
3.	M	14	15	12	11	13	65
4.	F	13	13	12	11	13	62
5.	M	13	14	15	12	11	65
6.	F	11	12	13	14	15	65
7.	F	15	12	13	14	11	65
8.	M	13	10	11	12	13	59
9.	M	12	12	12	13	14	63
10.	M	11	11	12	13	13	60
11.	F	12	12	11	13	14	62
12.	F	14	12	13	11	14	64
13.	M	14	10	13	14	15	66
14.	F	15	15	14	13	14	71
15.	M	14	14	13	14	12	67
16.	F	13	12	12	13	14	64
17.	M	14	15	15	13	12	69
18.	F	12	13	13	14	14	66
19.	M	14	13	14	15	14	70
20.	F	13	14	15	14	15	71

APPENDIX – VIII

Raw Score of the Students of Gramin Adarsha Campus

S.N.	Sex	Score					Total
		2056	2057	2058	2059	2060	
1.	M	12	13	14	13	12	64
2.	M	13	13	14	13	15	68
3.	F	14	12	13	14	15	68
4.	M	13	12	11	15	16	67
5.	F	14	14	13	12	11	64
6.	F	15	15	16	14	14	74
7.	F	15	14	13	13	12	67
8.	M	13	13	14	14	12	66
9.	M	12	14	16	17	14	73
10.	F	14	12	11	13	16	66

APPENDIX – IX

Raw Score of the Students of Mahendra Multiple Campus

S.N.	Sex	Score					Total
		2056	2057	2058	2059	2060	
1.	M	14	12	14	13	14	67
2.	M	13	12	13	14	15	67
3.	F	15	15	14	13	13	70
4.	F	14	14	13	14	14	69
5.	M	12	13	12	11	11	59
6.	F	11	14	12	13	14	64
7.	F	11	10	14	13	12	60
8.	M	13	12	11	15	14	65
9.	F	14	13	11	12	15	65
10.	M	13	16	17	14	13	73

APPENDIX – X

Raw Score of the Students of Kathmandu Sikshya Campus

S.N.	Sex	Score					Total
		2056	2057	2058	2059	2060	
1.	M	12	14	14	13	13	66
2.	F	13	12	13	14	14	66
3.	M	15	11	15	12	12	65
4.	F	13	12	14	14	14	67
5.	M	15	10	16	13	13	67
6.	F	14	11	12	13	13	63
7.	M	12	13	11	10	10	56
8.	F	11	10	14	15	15	65
9.	M	9	13	14	13	13	62
10.	F	10	12	11	13	13	59

APPENDIX – XI

Raw Score of the Students of Machhapuchhre Campus

S.N.	Sex	Score					Total
		2056	2057	2058	2059	2060	
1.	M	13	12	14	13	15	67
2.	F	14	14	13	13	12	66
3.	M	12	14	12	13	11	62
4.	F	13	11	12	10	14	60
5.	M	14	15	14	16	11	70
6.	F	15	12	13	14	14	68
7.	M	14	14	13	15	13	69
8.	F	13	14	13	14	13	67
9.	M	14	14	15	14	12	69
10.	F	12	12	14	13	12	63