

PREDICTIVE VALIDITY OF INTERNAL ASSESSMENT

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

**Submitted by
Kuber Prasad Bhattarai**

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

2008

PREDICTIVE VALIDITY OF INTERNAL ASSESSMENT

**A Thesis Submitted to the Department of English Education
in Partial Fulfilment for the Master's Degree in Education
(Specialization in English Education)**

**Submitted by
Kuber Prasad Bhattarai
Faculty of Education
Tribhuvan University, Kirtipur
Kathmandu, Nepal
2008**

**TU Reg. No: 9-2-1-888-2002
Second Year Examination
No: 280328/ 2064**

**Date of Approval of the
Thesis Proposal: 2065-03-09 Roll
Date of Submission:2065-08-25**

RECOMMENDATION FOR ACCEPTANCE

This is to certify that Kuber Prasad Bhattarai has prepared this thesis entitled Predictive Validity of Internal Assessment under my guidance and supervision.

I recommend the thesis for acceptance.

Date: 2065-08-25

.....

Dr. Bal Mukunda Bhandari

Reader

Department of English Education

Faculty of Education

TU, Kirtipur, Kathmandu

RECOMMENDATION FOR EVALUATION

**This thesis has been recommended for evaluation from the following
Research Guidance Committee.**

Signature

Dr. Chandreshwar Mishra

.....

Reader and Head

Chairperson

Department of English Education

TU, Kirtipur

Dr. Bal Mukunda Bhandari (Guide)

.....

Reader

Member

Department of English Education

TU, Kirtipur

Mrs. Madhu Neupane

.....

Teaching Assistant

Member

Department of English Education

TU, Kirtipur

Date: 2065-09-03

EVALUATION AND APPROVAL

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

Signature

Dr. Chandreshwar Mishra

.....

Reader and Head

Chairperson

Department of English Education

TU, Kirtipur

Dr. Jai Raj Awasthi

.....

Professor

Member

Department of English Education

Chairperson

English and Other Foreign Languages Education

Subject Committee

TU, Kirtipur

Dr. Bal Mukunda Bhandari (Guide)

.....

Reader

Member

Department of English Education

TU, Kirtipur

Date: 2065-09-06

DEDICATION

Dedicated

to

**My parents and teachers who devotedly tried to make me what I am
today.**

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 the research to any university.

Date: 2065-08-24

Kuber Prasad Bhattarai

ACKNOWLEDGEMENTS

Such a research work undoubtedly owes intellectual debts from many people. First of all, I express my hearty acknowledgement to Dr. Bal Mukunda Bhandari, Reader of English, Department of English Education TU, for his generous encouragement, helpful suggestions and precious guidance throughout the research work as my thesis supervisor.

My sincere gratitude also goes to Dr. Chandreshwar Mishra, Reader and Head of the Department of English Education, and Mrs. Madhu Neupane, Lecturer of the Department of English Education for their enlightening ideas and invaluable suggestions as the members of Research Guidance Committee. I am deeply indebted to Prof. Dr. Jai Raj Awasthi of the Department of English Education for his generous help from the very beginning of this research work. I am equally grateful to the intellectual personalities of the Department of English Education- Prof. Dr. Govinda Raj Bhattarai, Prof. Dr. Tirth Raj Khaniya, Dr. Anjana Bhattarai, Prof. Dr. Shanti Basnyat and all the Gurus and Gurumas for their direct or indirect encouragement and academic supports.

Further, I would like to thank Mr. Ambika Khanal, Arjun Parajuli, Bed Ghimire, Ganga Parajuli, Prem Ghimire, Visan Niroula, Yam Gartoula and Yog Raj Gautam for their cooperation in carrying out this research. Mr. Drona Parajuli and Subash Rai also deserve a lot of thanks for their support to type this thesis.

Finally, I express my deep sense of gratitude to my father Hom Nath Bhattarai, mother Dika Bhattarai and other members of the family who, despite many practical difficulties, encouraged and supported me in my studies since the very onset of my formal academic life.

Kuber Prasad Bhattarai

December, 2008

ABSTRACT

The only one major function of a good examination is and unbiased an objective evaluation of students' performance for the purpose of selection and comparison. In order to carry on this function, any means of assessment must always have validity especially predictive validity. Thus, this research work entitled 'PREDICTIVE VALIDITY OF INTERNAL ASSESSMENT' attempts to find out the predictive validity of internal assessment which is conducted in all the departments at Master's Level. For this, I have selected M. Ed. first year students as the population of the study. Altogether eighty students were selected proportionately from seven departments of the Central Department of Education and their internal assessment scores and final examination scores were analyzed thoroughly. The data were taken only from the students of the academic year 2062/63. From the study I have concluded that there is no uniformity in the degree of predicative validity of internal assessment in different departments; it ranges from very high to low- negative predicative validity.

The study report consists of four chapters. Chapter one introduces the study in terms of general background, review of the related literature, objectives and significance of the study along with the definition of some specific terms. Chapter two deals with the methodological strategies used to conduct the study which includes the sources of data, population of the study, sampling procedure, process of data collection and limitations of the study. Chapter three includes analysis and interpretation of the collected data. The data were analysed and interpreted by using simple statistical tools and presented with the help of tables, charts and diagrams. Chapter four, the final chapter of the study report deals with the findings, recommendations and pedagogical implications of the study.

TABLE OF CONTENTS

Declaration	i
Recommendation for Acceptance	ii
Recommendation for Evaluation	iii
Evaluation and Approval	iv
Dedication	v
Acknowledgements	vi
Abstract	vii
Table of Contents	viii
List of Tables and Diagrams	x
List of Abbreviations and Symbols	xii
CHAPTER-I: INTRODUCTION	1-17
1.1 General Background	1
1.1.1 Internal Assessment in Tribhuvan University	3
1.1.2 Issues Related to Assessment	5
1.1.3 Uses of Assessment	7
1.1.4 Validity of Assessment	9
1.1.5 Correlation Co-efficient as a Statistical Tool	12
1.2 Literature Review	13
1.3 Objectives of the Study	15
1.4 Significance of the Study	16
1.5 Definition of the Specific Terms	16
CHAPTER-II: METHODOLOGY	18-20
2.1 Sources of Data	18
2.1.1 Secondary Sources of Data	18
2.2 Population of the Study	18

2.3 Sampling Procedure	18
2.4 Process of Data Collection	19
2.5 Limitations of the Study	19
CHAPTER-III: ANALYSIS AND INTERPRETATION	21-46
3.1 Position of Education in Examination Result	23
3.2 Degree of Predictive Validity of Different	25
3.3 Department- wise Analysis of Correlation Coefficient	27
3.3.1 Department of English Language Education	27
3.3.2 Department of Mathematics Education	32
3.3.3 Department of Curriculum and Evaluation	34
3.3.4 Department of Nepali Language Education	37
3.3.5 Department of Health and Physical Education	40
3.3.6 Department of Educational Planning and Management	42
3.3.7 Department of Economics Education	44
CHAPTER-IV: FINDINGS AND RECOMMENDATIONS	47-49
4.1 Findings	47
4.2 Recommendations	49
References	
Appendices	

LIST OF TABLES AND DIAGRAMS

Table No.	Page No.
1. No. of Students Selected for the Study from Different Departments	19
2. Interpretation of Correlation Coefficient	22
3. Degree of Predictive Validity of Different Departments	25
4. Analysis Scores of Students from DELE	28
5. Analysis Scores of Students from DME	32
6. Analysis Scores of Students from DCE	35
7. Analysis Scores of Students from DNLE	38
8. Analysis Scores of Students from DHPE	40
9. Analysis Scores of Students from DEPM	42
10. Analysis Scores of Students from DEE	44

Diagram No.

1. Bar-diagram Showing the Degree of Predictive Validity of Different Departments	26
2. Bar-diagram Showing Two Sets of Scores of Students from DELE	30
3. Correlation Coefficient of the Two Sets of Scores of the Students from DELE	31
4. Bar-diagram Showing Two Sets of Scores of Students from DME	33
5. Correlation Coefficient of the Two Sets of Scores of the Students from DME	34
6. Bar-diagram Showing Two Sets of Scores of Students from DCE	36
7. Correlation Coefficient of the Two Sets of Scores of the Students from DCE	37
8. Bar-diagram Showing Two Sets of Scores of Students from DNLE	38
9. Correlation Coefficient of the Two Sets of Scores of the Students	

from DNLE	39
10. Bar-diagram Showing Two Sets of Scores of Students from DHPE	41
11. Correlation Coefficient of the Two Sets of Scores of the Students from DHPE	41
12. Bar-diagram Showing Two Sets of Scores of Students from DEPM	43
13. Correlation Coefficient of the Two Sets of Scores of the Students from DEPM	43
14. Bar-diagram Showing Two Sets of Scores of Students from DEE	45
15. Correlation Coefficient of the Two Sets of Scores of the Students from DEE	46

LIST OF ABBREVIATIONS AND SYMBOLS

Assit.	Assessment
B.Ed.	Bachelor of Education
CDC	Curriculum Development Centre
CUP	Cambridge University Press
DCE	Department of Curriculum and Evaluation
DEE	Department of Economics Education
DELE	Department of English Language Education
DEPM	Department of Educational Planning and Management
DHPE	Department of Health and Physical Education
DME	Department of Mathematics Education
DNLE	Department of Nepali Language Education
Dr.	Doctor
Eds.	Editors
eg.	Exempli gratia (for example)
ELT	English Language Teaching
et al.	et alli/alia (and other people or things)
etc.	et cetera (and other similar things)
F.M.	Full Marks
i.e.	id est (that is to say)
LLM	Master in Law
MA	Mater in Arts
MBS	Master in Business Studies
Mr.	Mister
MO.	Marks obtained
No.	Number
OUP	Oxford University Press
P	Page

PCL	Proficiency Certificate Level
Prof.	Professor
Reg. No.	Registration Number
SLC	School Leaving Certificate
TU	Tribhuvan University
Viz.	Videlicet (namely)
Vol.	Volume
1 st and 2 nd	First and Second
%	Percentage
+	Plus
-	Minus
()	Braces
=	Equals to
∴	Therefore
r	Correlation and Coefficient
$\sqrt{\quad}$	Square root

CHAPTER - ONE

INTRODUCTION

This chapter attempts to introduce the research in terms of its general background, review of related literature, objectives and significance of the study along with the definition of some specific terms related to the study.

1.1 General Background

Educational assessment is as old as education itself. The teachers always want to know how much their students have learned whatever has been taught. They use different techniques they know in order to assess their students' achievement. Assessment of student learning requires the use of a number of techniques. But assessment is more than a collection of techniques and more than the paper-pencil test that takes place at the end of every academic year. It is a systematic and continuous process and integral part of education system which helps to find out the effectiveness of teaching-learning of both teacher-students and education institutions and also, it provides a strong basis for performance comparison among students and educational institutions. It begins with the identification of learning goals and ends with a judgement concerning how well those goals have been attained.

According to Linn and Gronlund (2003,p.31) "Assessment includes the full range of procedures used to gain information about student learning (observations, ratings of performances or projects, paper and pencil tests) and formation of value judgements concerning learning progress". Similarly, in the words of Arends (2001, p.194) "Assessment is the process of collecting a full range of information about students and classrooms for the purpose of making instructional decisions". So, assessment can not be limited to a test or an

examination but it is a full range of procedures used to gain information about student learning. These procedures may include observation, ratings of performances and assignment of homework, classroom oral/written test and project work which are used to make instructional decisions. Information can be gathered on students in informal ways such as through observation and verbal exchange. It can also be gathered through formal means such as homework, tests and written reports. Information about classrooms and the teacher's instruction can also be part of assessment.

Assessment is generally taken as synonymous to evaluation. But technically, assessment and evaluation are two different concepts. Both assessment and evaluation are functions carried out by teachers to gather information needed to make wise instructional decisions. "whereas assessment focuses on gathering and synthesizing information, the term evaluation usually refers to the process of making judgements, assigning value, or deciding on worth"(Arends, 2001, p.195). So, assessment is gathering information about students and evaluation is the process of making judgements and deciding on the worth of a particular programme or approach or of student's work.

Based on the procedure used and formality practised in the administration of assessment tools and criteria used by teachers and administrators, assessment can be classified into two types viz. external and internal. If a testee's ability is tested against the externally established criteria it is known as external assessment. The annual examination of Tribhuvan University, S.L.C. examination etc. are related to external assessment in which, efforts are made to gather information about students or programmes after a set of instructional activities have occurred. In external assessment, information about students and programmes is collected for the purpose of making summative evaluation. On the other hand, internal assessment refers to the way of measuring students'

success before or during instruction and used to make adaptations to instruction. In internal assessment, simple tools are used (often informally) to collect information about students' work and the information is generally used for the purpose of making formative evaluation. Internal assessment is a continuous process that goes along with the instructional programme. So the term assessment is generally understood as a continuous process which is a blend of internal and external criteria.

Internal assessment contributes to encourage and promote successful learning. So that continuous internal assessment is to be provisioned by the respective teacher in order to assess the progress and speed of learning as well as to assess learners' strengths and weaknesses in learning. It is also required to provide feedback to both teacher and learner and decide remedial teaching.

1.1.1 Internal Assessment in Tribhuvan University

Tribhuvan University, the oldest and largest university in Nepal is the place for many students to complete their higher education. Like all the universities, it also has its own system of assessing students' level of academic achievement. Among the different ways of assessing students' performance, examination is one that has been used widely to measure students' level of achievement in order to provide certificate.

From 2019 B.S., Tribhuvan University began conducting examination on its own curriculum. Since then, it has used annual examination very widely. "With the implementation of New Education System in 2028, it also practiced the semester system. In this system there was provision of internal assessment weighting 20% of the total marks in all the departments at all levels and 80% of the total marks was allocated for semester assessment system. But the system became very expensive in terms of time, money and labour, and also, the quality of education

did not improve as it had been expected. Due to these problems, university could not persist with it. So, from 2038 B.S. university dismissed the semester system and restored the annual examination system in which still there was the provision of internal assessment weighing 20% of the total marks. But even this internal assessment system could not last very long since discipline among teachers and students could not be maintained and there came practical problems in handling the system" (Madhusudhan Dhungana, Chief of the records Division, TU, OCE Balkhu, personal communication, June 29, 2008). So, ultimately annual examination system became the only tool to assess students' level of academic achievement.

Tribhuvan University has now again introduced the internal assessment system weighing 20% of the total marks in each subject in all departments at Master's level in addition to the annual examination that takes place at the end of every academic year. Tribhuvan University, office of the Academic Council decided on 22 Kartik, 2062 to implement this system to be effective from the academic year 2062/63. Till that time, internal assessment system was used only in the faculty of management and in the institute of science as practical facet. But now, this system is implemented in the faculty of humanities and social science and education also. In this type of assessment system, Professors have been given full authority to make judgement of the achievement level of students.

Different faculties and departments have their own systems and techniques of assessing students. The curriculum has not specified the contents, techniques, tools and procedures to conduct the system. So, Professors are using their own ideas to assess the learners.

Clarifying the concept of internal assessment system implemented in TU, Khaniya (2006, p.4) says, "Like external examinations internal assessment is also a competent academic process of educational delivery. It is based on the

academic principle that it is the teachers who keep close look on students' learning, and thereby could best judge their performance". Internal assessment is integrated with the day-to-day teaching and learning directly involving teacher and it reduces the risk of overdependence on one-shot exam. The degree of proportion of the scores of students in the internal and external examination is a credible property of internal assessment. Therefore, university must do everything possible to improve the teaching and learning and avoid the impression that it adopted the system simply to please the students. However, the validity of internal assessment lies on how successfully, effectively and objectively it is handled.

1.1.2. Issues Related to Assessment

Although examination is taken as unfair means of assessment, we still have final examinations in each subject area and decide the final result of our evaluation according to the marks the students have secured in the exams. The total progress of students during the entire period of instruction is assessed through a few hour written examinations. Due to such kind of unfair means of assessment tool, students show unusual behaviours during the exam. Because of the total dependences in examination, our whole system of education has become examination oriented. Even the talented students may fail in this type of examination either because of their illness or some other problems which are not accounted in the system.

According to Chauhan (1993, p.131), "End product of learning which is evaluated by written examination tells what the learner has memorised. It does not tell how the examinee achieved the particular score." This also indicates that some kind of monitoring during the course of instruction is essential. All these issues and problems related to one-shot-exam force us to think about the alternative way of student assessment.

To avoid the weaknesses of assessments based on the final examination, and to assess the total aspects of the students' learning as well as to provide feedback even to the teachers and students during the instruction internal assessment can be appropriate tool. It can promote the meaningful involvement of the students with materials that is central to the teaching objectives of the given course. There are several world class universities which argue that evaluation of students without internal assessment can never be complete. They put more emphasis on internal rather than external evaluation. So, it needless to argue that internal assessment is a proven and widely used system of evaluation.

On the other hand, even the internal assessment system is not free from criticisms. There are a number of issues and problems related to the internal assessment system. The issues are related to the procedure to be adopted, tools and techniques to be used, experts to be involved, duration of assessment, ways to be followed in the process of assessing students' work and so on.

The most dangerous aspect of internal assessment is related to the unfairness and partiality in the process of assessing and grading students. As Arends (2001, p.198) has stated "In any assessment situation, the possibility of teacher bias is always a worry. ...you undoubtedly know how important it is for teachers to be perceived as fair and impartial in their treatment of young people. Being free from bias is particularly important when judging student work and assigning grades." Teachers hold different criteria and strategies for judging student work and that they are influenced by numerous subjective factors, such as the expectations teachers have for a particular student, his/her handwriting, his/her relation with the teacher and so on. So, reliable and valid strategies must be devised to reduce bias and subjectivity in assessment and grading.

In the name of objectivity and in the rush to reform to education, the classroom teacher's expertise and common sense also should not be pushed aside. The

professionals with extensive knowledge of assessment and learning can determine what is best for their students. On this issue, Angaran (1999, p.72) states, "Teachers are not provided enough time to analyse the test data and to translate the information into meaningful goals for their students. At times we seem to be putting the cart before the horse: Shouldn't we change the way we teach and then assess students? It doesn't make a lot of sense to continue to assess students when we are not given the time to modify our methodology. Without ample time to reflect and change, aren't we simply assessing either old practices or underfunded undervalued, half-hearted attempts at educational change? Without the time to examine our teaching practices and to work with our colleagues we will continue to teach in the ways we were taught 20 or more years ago."

To avoid these problems, strategies should be devised in such a way that ensures reliability and validity of assessment and reduce subjectivity and bias in it. The correspondence between the two assessments (i.e., internal and external) will indicate the validity of it and a high degree of discrepancy will reduce faith in the university. So, it is necessary to establish the modes of assessments and evaluation that encourages, rather than frustrates the desire and ability to learn for the learners and that also fosters teachers' expertise and creative involvement with the students.

1.1.3 Uses of Assessment

Assessment results taken from different types of tests and inventories provide useful information in the process of making academic decision. Based on the results obtained from student assessment, we can evaluate the whole educational programme. It also serves the functions of both formative and summative assessments. Linn and Gronlund (2003) have mentioned three uses of assessment. The first one is *pretesting*, which is used at the beginning of an

instructional segment. The second use is the *formative use* of assessment which is used during instruction to monitor learning progress. The third use is the *summative use* of assessment which is known as end-of-instruction assessment and is used to measure the extent to which the intended learning outcomes and performance standards have been achieved.

According to Zunker (1982, pp.2-4), mainly there are three uses of assessment which are:

-) **Diagnostic uses of assessment:** In diagnostic uses, assessment results are often used to evaluate individual strengths and weaknesses in order to detect it and provide appropriate feedback to the learners as well as teachers. In the same sense assessment results may help in the identification of the need for remedial training, or skill development.
-) **Predictive uses of assessment:** Assessment results may also be used to predict future academic performance. The probability of performing well in an educational programme is relevant information to base further exploration on it. When assessment results are used to predict subsequent performance, it should be ensured that relevant predictive validity has been established for the tests that are used.
-) **Comparative uses of assessment:** Comparing one's personal characteristics, abilities, interests, values etc. with those of criterion groups is the comparative use of assessment. Comparing one's performance with average performance of the group in which he/she belongs to shows how different his/her performance is from the group performance. The similarities and differences found can help us to decide on any special treatment to provide him/her if necessary.

Among the three uses of assessment, the second one mentioned above (i.e. predictive use) is the concern of the present study. As a predictive tool,

assessment results forecast the probability of performing well in a training or educational programme. More specifically, this research work tries to find out whether or not the internal assessment of M. Ed. students works as a predictive device to forecast the performance of the final examination.

1.1.4. Validity of Assessment

The central concept in all assessment is that of validity. A given mark or quantification guarantee a certain standard of knowledge only if it is valid. “When constructing or selecting assessments, the most important questions are: (1) To what extent will the interpretation of the scores be appropriate, meaningful, and useful for the intended application of the results? And (2) what are the consequences of the particular uses and interpretations that are made of the results?” (Linn and Gronlund 2003, p. 73). Validity is one of the very important qualities of a good assessment system or test. Validity of a test ensures its meaningfulness.

According to Hughes (1995, p.22), “A test is said to be valid if it measures accurately what it is intended to measure”. Therefore, to what extent the measurement of a test is suitable, accurate, relevance; concrete, valid and representative in light of the purpose for which it is administrated indicates its validity. The examiner should prepare items for the examination keeping in mind what he/she intends to obtain. There should be a good correlation between items asked in the exam and behavioural change expected after teaching certain teaching items. The same text may be used for different purposes and its validity may be high for one, moderate for another and low for the next one. So, the essential question of test validity is how good a test does the job it is employed to do.

There are different types of validity. Among them the major types are briefly introduced below (Heaton, 1975, Davies, 1977, Hughes, 1989):

1. Content validity

A test is said to have content validity if its content constitutes a representative sample of course content given. The sample of assessment tasks should well represent the domain of tasks to be measured. The content validity being a conceptual validity requires no statistical procedure but a rigorous analysis by the expert of the material which it desires to test. It is concerned with what goes into the test.

2. Construct Validity

Construct validity is also a conceptual validity of a test which measures just the ability which it is expected to measure. It is related to how well performance on the assessment can be interpreted as a meaningful measure of some characteristics of quality. Construct validity is discussed as the centre of all methods of validating a test. Accepted the inclusiveness of the construct validity, it can be argued that if a test provides evidence that it has content and criterion – related validity, the exam also provides some evidences that it has construct validity as well.

3. Face Validity

A test is said to have face validity if it looks as if it measures what it is supposed to measure. Face validity is “what it appears superficially to measure” (Anastasi 1982, p. 136, as cited in Khaniya, 2005). It is pseudo validity. Hughes (1989) also describes it as hardly a scientific but very important concept because if a test does not look right to other testers, teachers, moderators and testees, it can not be described as having validity.

4. Criterion Related or Empirical Validity

The criterion related validity which is described as statistical validity can further be divided into two types:

a. Concurrent validity

b. Predictive validity

According to Heaton (1975), empirical validity is obtained as a result of comparing the results of the test with the results of some criterion measure such as:-

1. An existing test, known or believed to be valid and given at the same time; or
2. The teacher's ratings on any other such form of independent assessment given at the same time; or
3. The subsequent performance of the testees on a certain task measured by some valid test; or
4. The teacher's rating or any other such form of independent assessment given later.

Results obtained by either of the first two methods above are measure of the test's concurrent validity in respect of the particular criterion used. The third and fourth methods estimate the predictive validity of a test which is used to predict future success. He further says, "The test situation or the technique used is always an important factor in determining the overall validity of any test.

Although an ideal test situation will by no means guarantee validity, a poor test situation will certainly detract from it" (Heaton 1975,p.161).

A test to have predictive validity should be positively correlated with another future criterion. There is a need to show that a positive relationship exists between scores on the test (the predictor) and scores on some acceptable measure of future performance (the criterion). As Ebel & Frisbie (2004, p.106)

state the criterion measure is “a accepted standard against which some test is compared to validate the use of the test as a predictor.”

The proposed study is concerned only with the predictive validity of internal assessment system conducted at M. Ed. level in TU. To establish the predictive validity of any test or exam, some standard measure of performance must be pre-specified. It is this pre-specified performance which serves as a criterion. In this study students' marks obtained in internal assessment will be taken as the predictor and validated against the marks obtained in the final/external examination (i.e., the performance criterion) by calculating the correlation of the two assessment systems.

1.1.5. Correlation- Coefficient as a Statistical Tool

Correlation-Coefficient is a statistical tool, a quantitative measure that represents the degree of relationship between two (or more) variables. It expresses how much two series of numerical observations have in common. There may exist two or more variables in a co relational study but as Ingram (1977,p.25) claims ‘correlations are appropriate when observations come in pairs’. For example, when a group of student have taken two tests, the scores for each individual form a pair and two sets of scores for a group of individuals can be correlated to see how much agreement there is between them.

In testing, correlations are used to assess reliability and validity and also to assess the degree of correspondence which exists among subtests. If one is going to predict from one set of scores to another, the correlation values need to be high or very high. The measurement of correlation between two variables results in the value that ranges from +1 to -1 which indicates the degree to which two variables are related or vary. That means to say, in estimating predictive validity, the nearer the correlation approaches +1, the more closely the scores on one test

predict the scores on the other test. The higher the correlation between criterion scores and test scores, the more accurately we can predict from scores on one to scores on the other. So, positive correlation between two sets of scores always indicates high predictive validity and negative correlation indicates low predictive validity.

For the purpose of comparison and further analysis, a relative numerical measure has been developed by Karls Pearson called Pearson's co-efficient of correlation or Product movement correlation coefficient which is "the most often used and most precise coefficient of correlation" (Best and Kahn,1996,p.302) and is denoted by 'r'. The relationship between two variables as defined by Pearson's correlation coefficient is measured by:

$$r = \frac{N \sum XY - \sum X \sum Y}{\sqrt{N \sum X^2 - (\sum X)^2} \sqrt{N \sum Y^2 - (\sum Y)^2}}$$

Where, $\sum X$ = Sum of the x score

$\sum Y$ = Sum of the y scores

$\sum X^2$ = Sum of the squared x scores

$\sum Y^2$ = Sum of the squared y scores

$\sum XY$ = Sum of the products of paired x and y scores

N = Number of paired scores

1.2. Literature Review

Batala (2004) carried out a research work on 'Validation of the SLC English Examination' and attempted to find out the predictive and content validity of English question paper of SLC examination. He used both primary and secondary sources of data. For primary data, the sample consisted of 10 secondary level English teachers, 20 students and 2 question setters who were selected by using judgemental random sampling procedure. For secondary data, the sample consisted of 100 students' marks that they had secured in the SLC

and grade xi English examination. The study revealed that the predictive validity of SLC English examination was very low whereas its content validity was found to be quite satisfactory.

Khaniya (2005) in his book 'Examination for Enhanced Learning' writes about wash back validity of an examination. Based on his Ph. D. research findings, he states that wash back is an inherent quality of an exam. As he mentions, for his study he had collected data from SLC students in Nepal with the objective to examine the extent to which preparation for an exam like SLC English exam had an effect on learning.

Thapa (2005) conducted a research entitled 'Predictive Validity of Grade xi Annual English Test' in order to find out the predictive validity and uniformity especially that of all students passed HSEB exam from Biratnagar during the period of 2054-2060 B.S. For this the researcher mainly analyzed the secondary data as census population (i. e. 4264 students' scores that of grade xi and xii English from HSEB). The researcher concluded that predictive validity of grade xi English test was significant to predict the future performance, though the predictive validity was not uniformed.

Timsina (2006) in his M. Ed. thesis attempted to find out whether the Language Testing exam paper had content validity or not. For this he used only secondary source of data. He analysed five years question papers from 2058 to 2062 B.S. of Language Testing paper and found that language testing tests had low content validity in terms of both coverage and weighting.

Bhandari (2007) conducted another research on 'Content Validity of Grammar Theory and Practice'. He also used similar methodology as Timsina (2006) and analysed the question papers of the concerned subject of the years from 2057 to 2063. He concluded his study with the finding that Grammar Theory and

Practice test had high content validity in terms of coverage but low content validity in terms of relevance.

Similar type of study using the same methodology was conducted by **Khatri (2007)** on 'The Content Validity of Translation Theory and Practice Exam at M.Ed. Level'. He analysed question papers of the years from 2058 to 2063 and found the average content validity in terms of content coverage and low content validity in terms of content weighting of the subject. **Hamal (2007)** also conducted a research on 'Content Validity of General English Test Items at B. Ed. Level: A case of Reading and Writing'. Using similar methodology, he found that Reading and Writing test items had low content validity in terms of content coverage and high content validity in terms of course content weightage.

Bhattarai (2008) attempted to find out the attitude of teachers and students towards the internal assessment system in his M. Ed. thesis 'Attitude of Teachers and Students towards Internal Assessment'. For this he collected information from five teachers teaching in the Central Department of English Language Education and 30 students of M. Ed. first and second year. His work revealed that teachers had negative attitude towards the internal assessment whereas students had positive attitude towards it.

The present study is different from the above mentioned studies in the sense that it is related to the predictive validity of *Internal Assessment* of M. Ed. level of the academic year 2062- 63. No study has been conducted on this topic till now.

1.3. Objectives of the Study

The proposed study had the following objectives:

1. To compare the degree of predictive validity of different departments within the faculty of education.
2. To find out the predictive validity of internal assessment of M. Ed. level.
3. To suggest some pedagogical implications.

1.4. Significance of the Study

The study mainly tries to find out the predictive validity of internal assessment system implemented at Master's level in TU. More specifically, the study is related to the internal assessment of education faculty. So, the findings of this study will be beneficial to the concerned institution as the study is concerned with how valid the newly implemented internal assessment system is. The study will be equally useful to the question setters, evaluators, and all the professors who are working at M.Ed. level with this system under various departments as this research work will provide some insight to evaluate their own work. Furthermore, the policy makers can move forward in the light of such research to make future educational policy.

1.5. Definition of the Specific Terms

Assessment: Process of the collecting a full range of information about students and classroom for the purpose of making instructional decisions

Classroom procedures: Interpersonal and group processes that help classroom participants' deals with issues of expectations, leadership, attraction, norms, communication and cohesiveness.

Correlation: A term used to express how or more variables are related.

Correlational Research: A type of research that investigates relationship between variables which exists naturally

Correlation and Coefficient: Number ranging from +1.00 to -1.00 that describes the numerical relationship between variables.

Formative Evaluation: Evaluation that occurs before or during instruction and is used to assist with planning or making adaptation.

Instructional Objectives and goals: Statements that describes a teacher's instructional intents

Performance Assessment: Assessment procedures that have students demonstration their abilities to perform particular tasks in testing situation.

Summative Evaluation: Evaluation done after instruction to determine programme effectiveness or the worth of students' work.

Reliability: The degree to which a test procedures consistent results over several administrations.

Validity: The degree to which a test measures what it claims to measure.

CHAPTER- TWO

METHODOLOGY

The followings methodological strategies were adopted to achieve the desired objectives:

2.1 Sources of Data

2.1.1 Secondary Sources of Data

I used only secondary source of data. The data were the raw marks of M.Ed. students obtained in internal assessment and external examination. The mark-sheets of the concerned students were collected and their marks were analyzed. In addition to this the researcher consulted various books, journals, and articles etc. which were found to be useful to the study. Some of them are Heaton (1975), Davies (1977), Hughes (1989), Linn & Gronlund (2003), Best and Kahn (2004), Khaniya (2005), Khaniya (2006).

2.2 Population of the Study

The population of the study was the raw score of M. Ed. students of Central Department of Education who appeared in both internal assessment and final examination in the academic year 2062/63. The total data consisted of the raw scores of 80 students sampled proportionately from the selected departments of the Faculty of Education, University Campus, Kirtipur.

2.3 Sampling Procedure

In this study, the sample size consisted of 80 students. On the basis of the list of admitted students, the departments holding more than 20 students were selected for study. The departments which came under this criterion were seven

departments of the Faculty of Education running M. Ed. programme. The total number of students admitted in these departments in the academic year 2062/63 was 1255. Out of this number, the required number of population was selected by using stratified proportionate random sampling procedure. For this, seven departments were considered as seven strata and 80 students were selected proportionately based on the number of admitted students. The followings were the number of students sampled from the respective departments:

Table- 1: No of Students Selected from Different Departments

S.N.	Department.	No. of Selected Students
1	English Language Education	32
2	Mathematics Education	11
3	Curriculum and Evaluation	10
4	Nepali Language Education	10
5	Health and Physical Education	9
6	Educational Planning and Management	5
7	Economics Education	3
	Total	80

2.4 Process of Data Collection

For data collection, I visited Central Department of Education and the concerned departments under it. For this, I visited the departments, explained the purpose of the visit and asked for their permission to take data. With their permission, I selected and recorded the necessary information.

2.5 Limitations of the Study

The proposed study had the following limitations:

- a. The study was limited to the predictive validity of internal assessment

- b. The study was limited to the internal assessment system of Tribhuvan University which is conducted at Master's level.
- c. The study was limited to the data of the Central Department of Education.
- d. The study was limited to the departments holding equal or more than 20 students.
- e. The study was limited to the internal assessment of the academic year 2062-63
- f. The study was limited to the sample size of 80 students of the Faculty of Education.
- g. The study was limited to the raw scores of those students who appeared in both internal assessment and external examinations

CHAPTER - THREE

ANALYSIS AND INTERPRETATION

After having the data gathered from different sources, the data (i.e., the raw scores) were refined and finalized keeping in mind the limitations of the study. To analyze the finalized data, I used correlation coefficient as the statistical tool.

In the present study, I have used Pearson's coefficient of correlation to compare and analyze the numerical values of the two variables where, the variable 'x' represents the percentage of the marks obtained in internal assessment and variable 'y' represents the percentage of the marks obtained in external examination. The correlation measure obtained using Pearson's coefficient of correlation formula may result in different values and the values are interpreted accordingly.

Connolly and Sluckin (1971, p. 154) give what they call 'a rough but useful guide to the degree of relationship indicated by the size of the coefficient which provides a useful guideline to find out how positively or negatively two sets of scores are related. According to them:

0.90-1.00 very high correlation; very strong relationship

0.70-.90 high correlation; marked relationship

0.40-0.70 moderate correlation; substantial relationship

0.20-0.40 low correlation; a definite relationship but a small one

0.20 and less slight correlation; relationship so small as to be negligible.

Best and Kahn (1996, P. 308) also provide a measure having five degrees of correlation which is slightly different from that of Connolly and Sluckin. They give the following measure.

.00 to .20- Negligible

.20 to .40- Low

.40 to .60 - Moderate

.60 to .80 - Substantial

.80 to 1.00- High to very high

Similarly, Sthapit et al. (2004) have also given some relevant criteria for the interpretation of correlation coefficient which I have used for my convenience in this research work as follows:

Table -2, Interpretation of Correlation-coefficient

Degree	Direction	
	Positive	Negative
perfect	+1	-1
Very high	+0.75 to +1	-0.75 to -1
High	+0.50 to +0.75	-0.50 to -0.75
Low	+0.25 to +0.50	-0.25 to -0.50
Very low	0 to +0.25	0 to -0.25
Absent	0	0

For the purpose of presenting the relationship or magnitude between the two sets of scores, I have used the two easiest methods in this study: Scatter Diagram (also Scattergram) and Bar Diagram

Scatter diagram is a graphical method of studying the correlation coefficient between two variables. Though the designation of the two variables in a scatter diagram is arbitrary, the left axis by convention is the independent variable and the right is dependent variable. In the present study, marks obtained in the internal assessment have been described as independent variable and the marks obtained in the final examination the dependent variable. The sets of points of the two scores (variables) have been plotted along the x axis y-axis of

rectangular co-ordinates. Each point in the scatter diagram has represented a student's scores. Observing the closeness and dispersion of the coordinates, the degree and direction of correlation between the two sets of scores (variables) can be interpreted as follows:

- i. The correlation coefficient is positive if the direction of the dots or points moves from lower left to upper right corner.
- ii. The correlation coefficient is negative if the direction of the dots or points moves from lower right corner to the upper left.
- iii. The correlation coefficient is high if the dots or points take the shape of narrow band.
- iv. The correlation coefficient is low if the dots or points take the shape of scattered band.
- v. The correlation coefficient does not exist at all if the dots or points form a circle.

Similarly, bar-diagram has been used in this research in order to display the magnitude or pattern of relationship between the two variables. Bar diagram is an effective way of displaying the relationship between variables. By convention, frequency is presented on x-axis and scores on y- axis in a bar diagram. Here bar-diagram has been used to present how the two sets of the scores are related.

3.1 Position of Education in the Examination Result

Department of Education is a single faculty department under the faculty of education Tribhuvan University. Under the faculty of education, there are twelve constituent departments running M.Ed. programme. Besides central departmental of education, there are other campuses in the country running M.Ed. program which together form the faculty of education.

The internal assessment system was introduced in the faculty of education as well as in other faculties in Tribhuvan University from the academic year 2062/063. In the first year of its implementation, the faculty of education and other three faculties got the following results.

Faculty	No. of exam appeared students	Passed	Percent
Law (L.L.M)	89	34	38.20
Humanities (M.A)	12965	4936	38.07
Education (M.Ed.)	5043	1799	35.67
Management (MBS)	9523	2467	25.91

As it has been presented above, in comparison to three other faculties the position of M.Ed. in its examination result is in the third position in terms of the percentage/number of students who successfully completed the first year examination. But the result of M.Ed. in the year 2063 seems satisfactory when it is compared with that of its results of the pervious years. The following table presents the result of M.Ed. from the year 2059 to 2063.

Year	No. of Exam Appeared Students	Passed	Percent
2059	2012	342	17.00
2060	2616	476	18.20
2061	3164	695	21.97
2062	3789	896	23.65
2063	5043	1799	35.67

As the above table shows, the percentage of students who successfully completed M.Ed. first year examination had increased highly in the year 2063; the first year of the implementation of internal assessment system. From this we can hypothesize that the internal assessment also contributed a lot for the qualitative improvement of examination result of M.Ed. However, in order to find out the quality of the result, it is necessary to find out the quality of internal assessment itself. Validity is one of the most important qualities of any

assessment. In the following sub-sections, predictive validity of internal assessment has been analyzed considering the limitations of the present study.

3.2 Degree of Predictive Validity of Different Departments in the Faculty of Education.

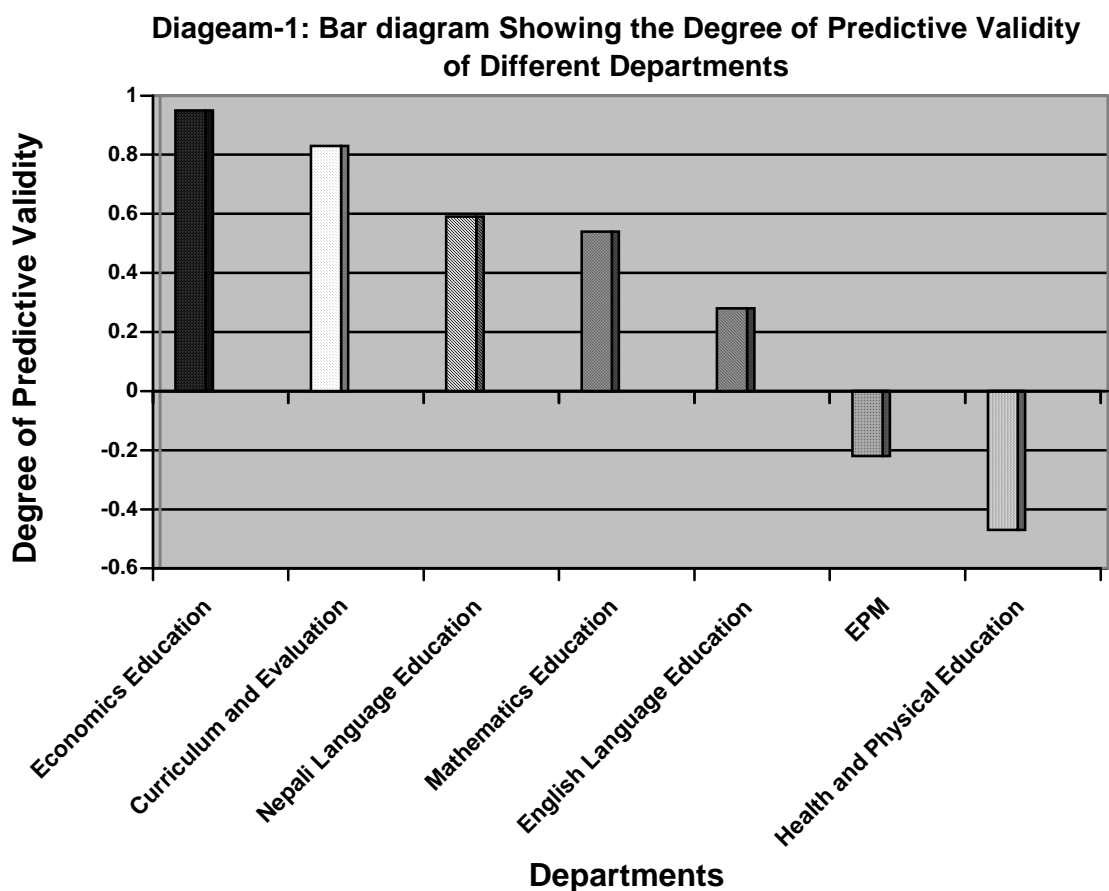
From the analysis, it has been found that there is not uniformity in the correlation between the two sets of scores of the students from different departments. In the study of predictive validity, correlation coefficient is taken to be same as the predictive validity. High degree of correlation indicates high degree of predictive validity and low correlation indicates low degree of predictive validity. So, it can be said that there is no uniformity in the degree of predictive validity of internal assessment of different departments in the faculty of education.

Although, the magnitude of the correlation coefficient vary from -1 to 0 to +1, no department has been found to have perfectly positive (i.e. +1 and perfectly negative (i.e. -1) correlation. In the following table the seven departments of the faculty of education included in the study have been presented in the order of predictive validity of internal assessment from high to low, the respective departments have.

Table -3 Degree of Predictive Validity of Different Departments

S.N.	Departments	Correlation	Predictive validity
1	Economics Education	+0.95	Very high, positive
2	Curriculum and Evaluation	+0.83	Very high, positive
3	Nepali Language Education	+0.59	High, positive
4	Mathematics Education	+0.54	High, positive
5	English Language Education	+0.28	low, positive
6	Educational Planning and Management	-0.22	very low, negative
7	Health and Physical Education	-0.47	Low, Negative

As shown in the above table, in the Department of Economics Education, internal assessment scores can predict the final examination scores since it has very high predictive validity. Similarly, Department of Curriculum and Evaluation also has very high predictive validity. In the Departments of Nepali Language Education and Mathematics Education the internal assessment has high predictive validity and in the Departmental of English Language Education it has low predictive validity. But in the Departments of Educational Planning and Management and Health and Physical Education, the internal assessment scores negatively predict the final examination scores where, the former has very low negative predictive validity whereas the later has low negative predictive validity. The degree of predictive validity of the seven departments can be shown more vividly in the bar diagram as follows:



As the above bar diagram shows, in the Department of Economics Education, the internal assessment has the highest predictive validity and the Department of English Language Education has the lowest predictive whereas in the departments of Educational Planning and Management and Health and Physical Education the internal assessment has negative predictive validity which negatively predict the final examination scores.

3.3 Department-wise Analysis of Correlation-coefficient

In this section of the research report, the selected departments and correlation analysis of the two sets of scores of the students from the respective departments has been presented.

3.3.1 Department of English Language Education

Department of English Language Education is the largest department in Central Department of Education in terms of student enrollment. In the academic year 2062, 474 students were admitted in 1st year in this department out of which, 32 students have been selected for the study. Marks (in percentage) obtained by the selected students in both internal assessment (i.e. variable 'x') and final examination (i.e. variable 'y') and their correlation analysis has been presented below.

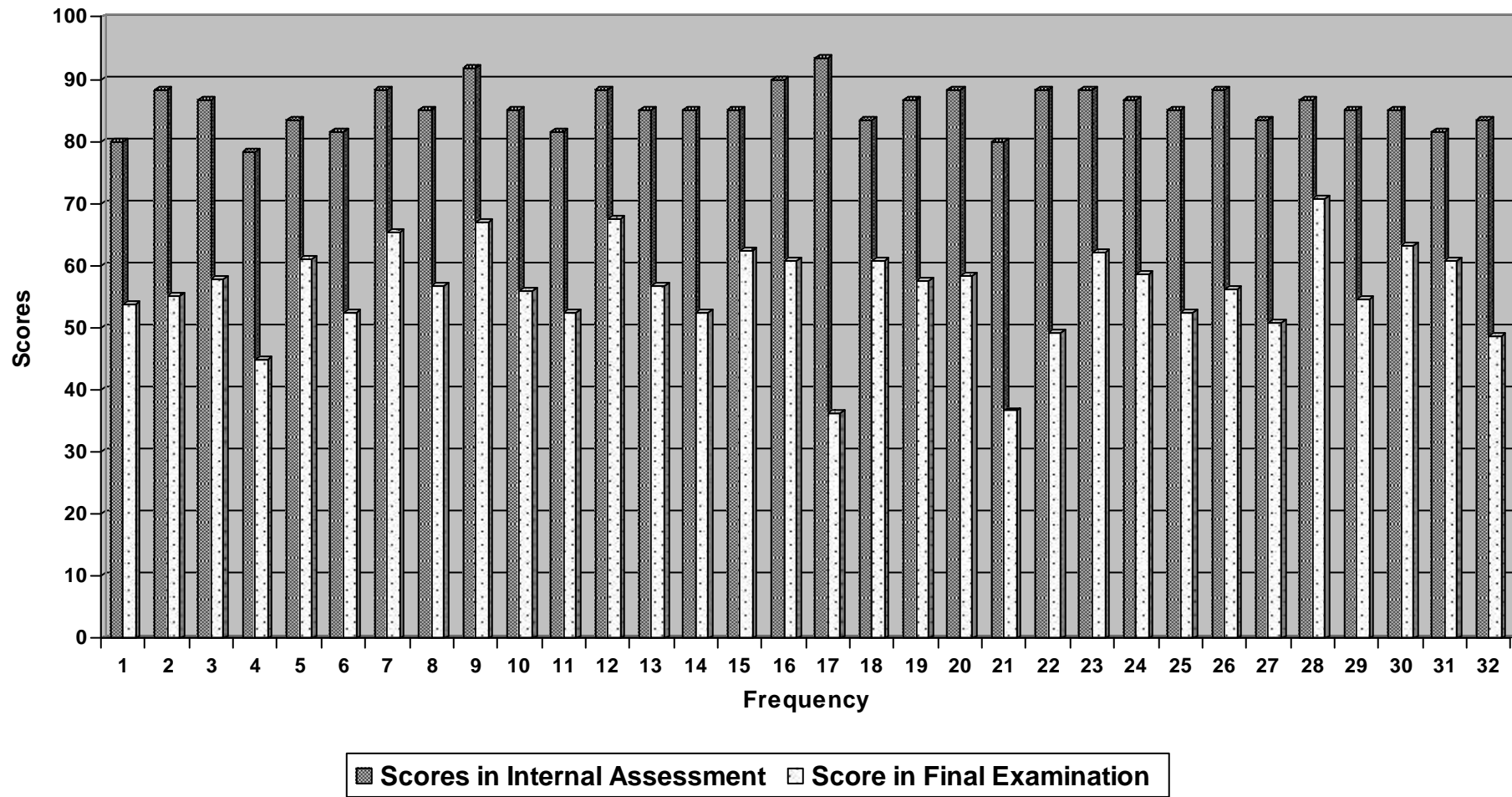
Table- 4: Analysis of Scores of Students from DELE

S. N.	x	y	x ²	y ²	xy	r
1	80	53.75	6400.00	2889.06	4300.00	
2	88.33	55	7802.19	3025.00	4858.15	
3	86.66	57.91	7509.96	3353.57	5018.48	
4	78.33	45	6135.59	2025.00	3524.85	
5	83.33	61.12	6943.89	3735.65	5093.13	
6	81.66	52.50	6668.36	2756.25	4287.15	
7	88.33	65.41	7802.19	4278.47	5777.67	
8	85	56.66	7225.00	3210.36	4816.10	
9	91.66	67.08	8401.56	4499.73	6148.55	
10	85	55.83	7225.00	3116.99	4745.55	
11	81.66	52.50	6668.36	2756.25	4287.15	
12	88.33	67.50	7802.19	4556.25	5962.28	
13	85	56.66	7225.00	3210.36	4816.10	
14	85	52.50	7225.00	2756.25	4462.50	
15	85	62.50	7225.00	3906.25	5312.50	
16	90	60.83	8100.00	3700.29	5474.70	
17	93.33	36.25	8710.49	1314.06	3383.21	
18	83.33	60.83	6943.89	3700.29	5068.96	
19	86.66	57.50	7509.96	3306.25	4982.95	
20	88.33	58.33	7802.19	3402.39	5152.29	
21	80	36.66	6400.00	1343.96	2932.80	
22	88.33	49.16	7802.19	2416.71	4342.30	
23	88.33	62.08	7802.19	3853.93	5483.53	
24	86.66	58.75	7509.96	3451.56	5091.28	
25	85	52.50	7225.00	2756.25	4462.50	
26	88.33	56.25	7802.19	3164.06	4968.56	
27	83.33	50.83	6943.88	2583.69	4235.66	
28	86.66	70.83	7509.96	5016.89	6138.13	

29	85	54.58	7225.00	2978.98	4639.30	
30	85	63.33	7225.00	4010.69	5383.05	
31	81.66	60.83	6668.36	3700.29	4967.38	
32	83.33	48.75	6943.89	2376.56	4062.34	
	$\phi X =$ 2736.57	$\phi Y =$ 1800.21	$\phi X^2 =$ 234383.44	$\phi Y^2 =$ 103152.29	$\phi XY =$ 154179.10	+0.28

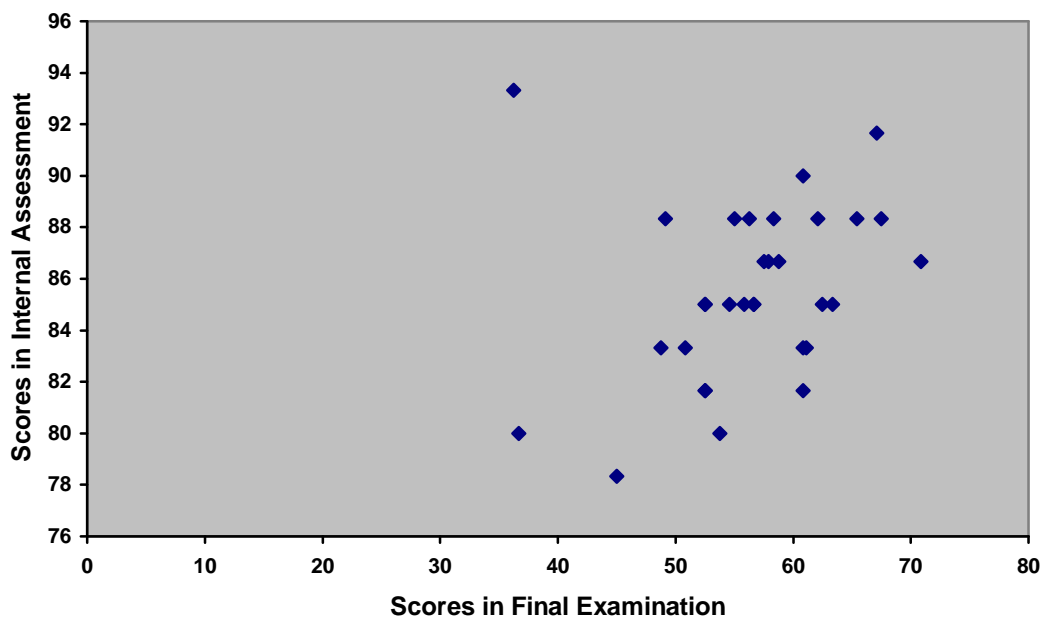
The bar diagram given below is based on the above table and represents the degree of relationship shared by the variables 'x' and 'y'

Diagram- 2: Bar Diagram Showing the Two Sets of Score of the Students from DELE



The above bar diagram indicates that all the students in the Department of English Language Education obtained higher marks in internal assessment than in final examination. Not even a single student got even equal marks in final examination as he/she got in internal assessment. Although the marks obtained by all students in internal assessment is higher than the marks they obtained in final examination, the relationship shared by the two sets of scores may have different interpretation. To observe this relationship the two sets of scores have been presented in the following scatter gram.

Diagram- 3: Correlation Co-efficient of the two sets of scores of the students from DELE



As the above diagram presents the correlation coefficient between two sets of scores of the thirty- two students from the department of English Language Education is low since the dots in the figure take the shape of scattered band. The direction of the dots in the figure has moved slightly to the upper right corner which indicates the positive correlation coefficient.

The calculation of correlation coefficient based on the table no: 3 above has also resulted that the coefficient of correlation between the two sets of scores of the students from the Department of English Education is only +0.28 (See appendix-II) This makes obvious that the degree of correlation between the two sets of scores is found to be low although they are positively related.

3.3.2 Department of Mathematics Education

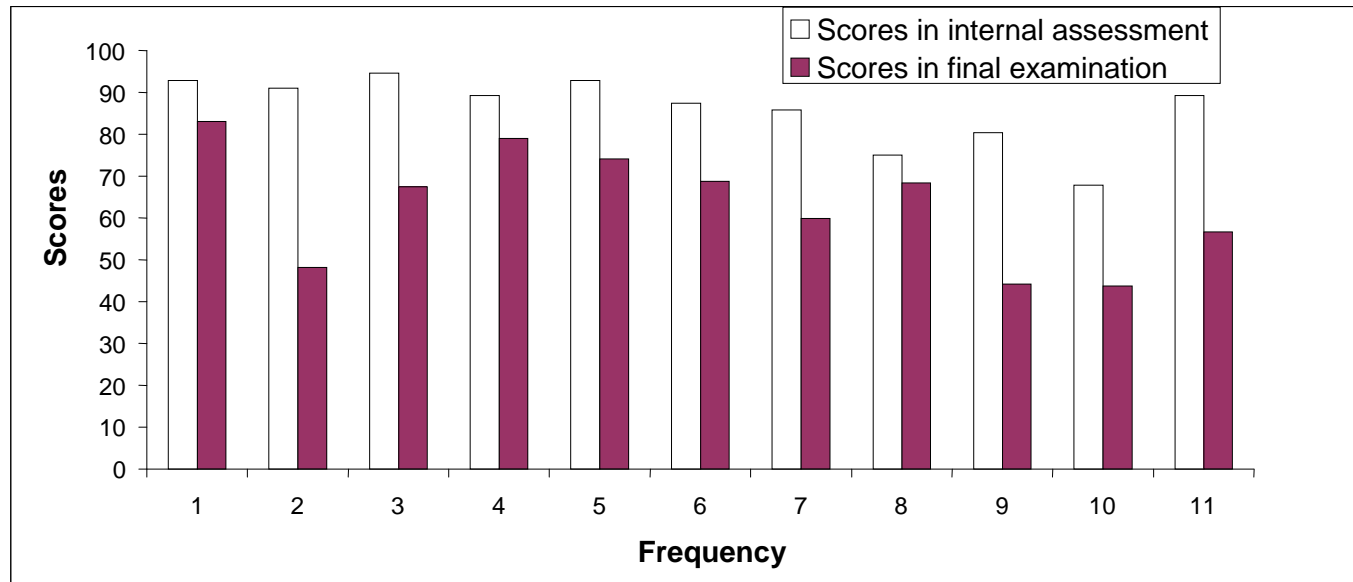
In terms of student enrolment in M.Ed. first year, the Department of Mathematics was the second largest department in Central Department of Education in the academic year 2062. In this academic year, 183 students were admitted in first year in this department from which 11 students have been selected for the study. Marks (in percentage) obtained by the selected students in both internal assessment (i.e. variable 'x') and final examination (i.e. variable 'y') and their analysis has been presented below:

Table- 5: Analysis of Scores of Students from DME

S. N.	X	y	x ²	y ²	xy	r
1	92.85	83.03	8621.12	6893.98	7709.34	
2	91.07	48.21	8293.74	2324.20	4390.48	
3	94.64	67.41	8956.73	4544.11	6379.68	
4	89.28	79.01	7970.92	6242.58	7054.01	
5	92.85	74.10	8621.12	5490.81	6880.19	
6	87.50	68.75	7656.25	4726.56	6015.63	
7	85.71	59.82	7346.20	3578.43	5127.17	
8	75.00	68.30	5625.00	4664.89	5122.50	
9	80.35	44.19	6456.12	1952.76	3550.67	
10	67.85	43.75	4603.62	1914.06	2968.44	
11	89.28	56.69	7970.92	3213.76	5061.28	
	ϕx =946.38	ϕy =693.26	ϕx^2 =82121.74	ϕy^2 =45546.14	ϕxy =60259.39	+0.59

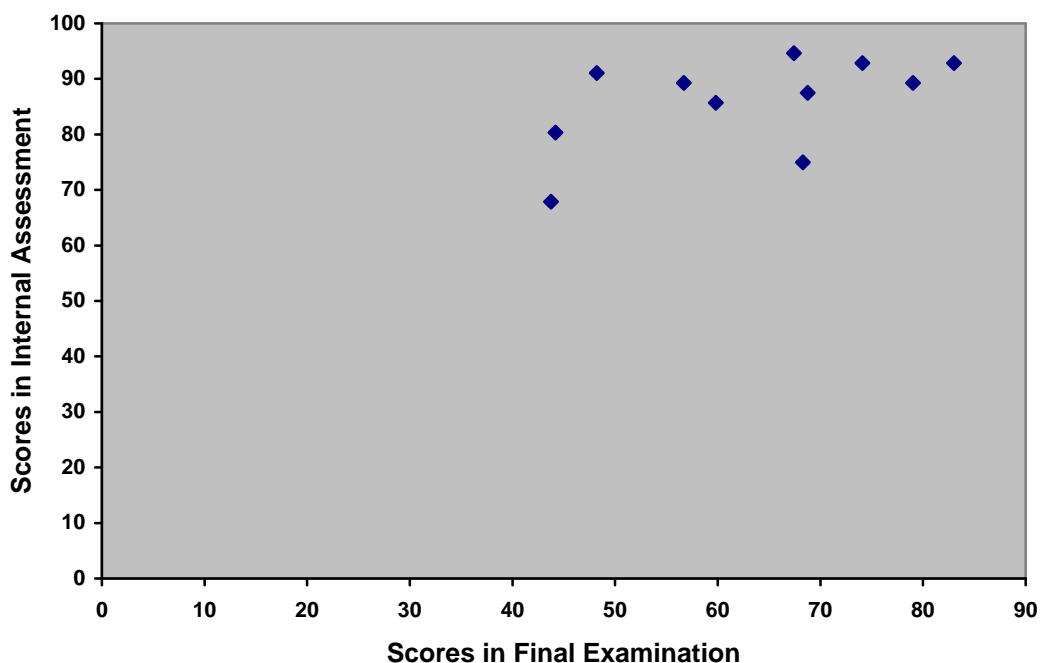
The numerical values represented by the two variables x and y above can be presented in a bar diagram to show in what way the two variables are related:

Diagram -4: Bar-diagram showing the two Sets of Scores of Students from DME



The above bar diagram shows that on the whole students obtained better marks in internal assessment than in final examination. None of the students got equal mark in both the examinations. The correlation ship between the two variables can be observed in a scatter gram which can be presented as follows:

Diagram -5: Correlation Co-efficient of the two Sets of Scores of Students from DME



From the diagram above it is clear that the correlation coefficient between the two sets of scores of the students from the Department of Mathematics Education is high since the dots in the figure take the shape of narrow band moving towards the upper right corner. So the two sets of scores are positively related.

The calculation also has revealed the fact that the correlation value of the two sets of score given in table-4 is +0.54 (see appendix II). This indicates that the degree of correlation between the two sets of scores of the students from the department of mathematics education is high.

3.3.3 Department of Curriculum and Evaluation

In the Department of Curriculum and Evaluation, Central Department of Education, 165 students were admitted in the academic year 2062. Out of 165

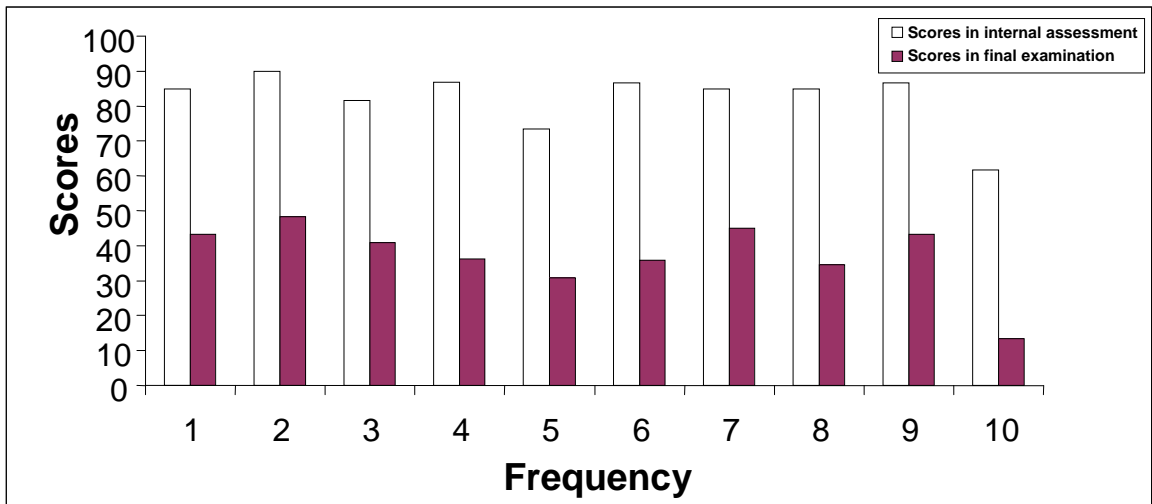
students, 10 students have been selected for the study. Marks (in percentage) obtained by the selected students in both internal assessment (ie. variable 'x') and final examination (i.e., variable 'y') and their analysis has been presented below:

Table - 6: Analysis of Scores of Students from DCE

S.N.	x	y	x^2	y^2	xy	r
1	85.00	43.33	7225.00	1877.49	3683.05	
2	90.00	48.33	8100.00	2335.79	4349.70	
3	81.66	40.83	6668.36	1667.09	3334.18	
4	86.88	36.25	7509.96	1314.06	3141.43	
5	73.33	30.83	5377.29	950.49	2260.76	
6	86.66	35.83	7509.96	1283.79	3105.03	
7	85.00	45.00	7225.00	2025.00	3825.00	
8	85.00	34.58	7225.00	1195.78	2939.30	
9	86.66	43.33	7509.96	1877.49	3754.98	
10	61.66	23.33	3801.96	544.29	1438.53	
	$x \Sigma$ 821.63	$y \Sigma$ 381.64	$x^2 \Sigma$ 68152.49	$y^2 \Sigma$ 15071.27	$xy \Sigma$ 31831.96	+0.83

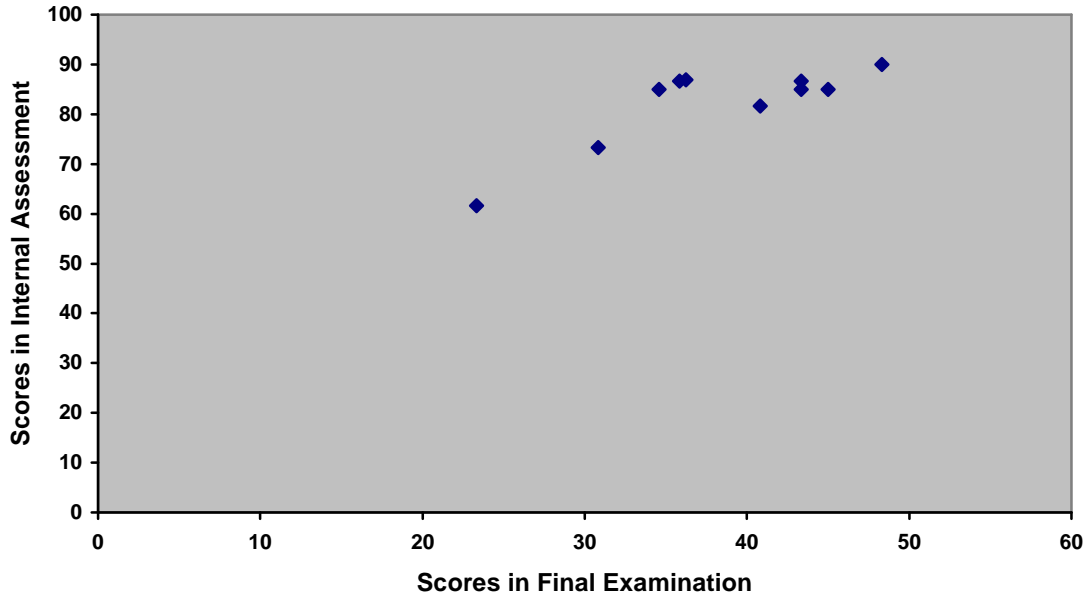
To show the relationship between two sets of scores represented by the two variables x and y above, the bar diagram is given below presenting which variable presents the higher scores.

Diagram -6: Bar-diagram Showing Two Sets of Scores of Students from DCE



As the above diagram shows, all the students from the Department of Curriculum and Evaluation secured higher scores in internal assessment than in final examination. The internal assessment score of most of the students is more than twice as much as the final examination marks. Although there is a marked difference between the two sets of scores, there exists a positive correlation between the two sets of scores. This can be observed in the following scatter gram.

Diagram - 7: Correlation Co-efficient of the two Sets of Scores of Students from DCE



Since the direction of the dots moves from lower left to upper right corner forming a narrow band in the above diagram, the correlation coefficient between the two sets of scores is high. The calculation also shows that the correlation value of the two sets of scores is +0.83 (see appendix II) Therefore, the correlation coefficient between the two sets of scores of the students from the Department of Curriculum and Evaluation is found to be very high.

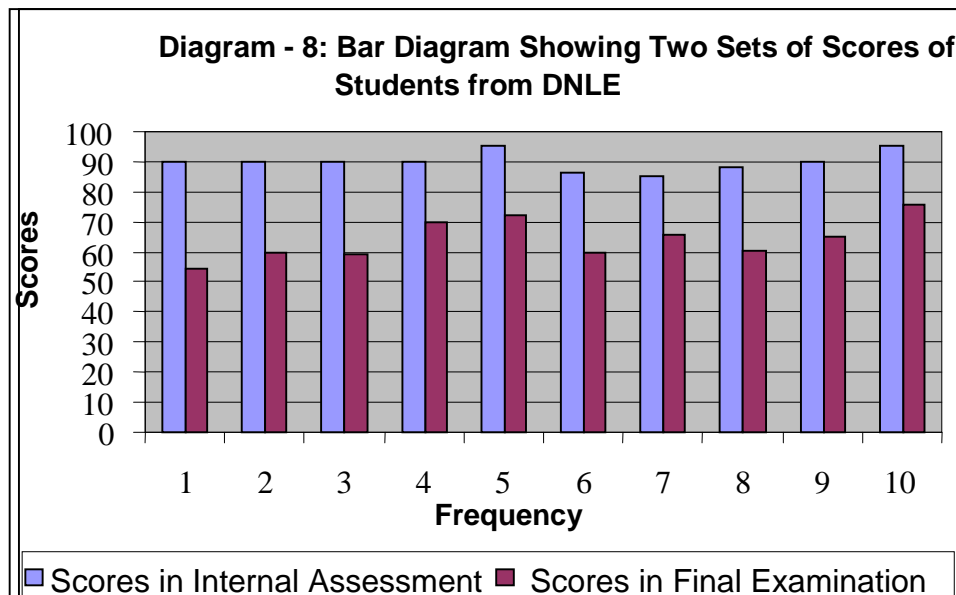
3.3.4 Department of Nepali Language Education

In the Department of Nepali Language Education, Central Department of Education, 163 students were admitted in 1st year in the academic year 2062. Out of the total number of admitted students, 10 students have been selected for the study. Marks (in percentage) obtained by the selected students in both internal assessment (i.e. variable x) and final examination (i.e. variable y) and its analysis has been presented below:

Table -7: Analysis of Scores of Students from DNLE

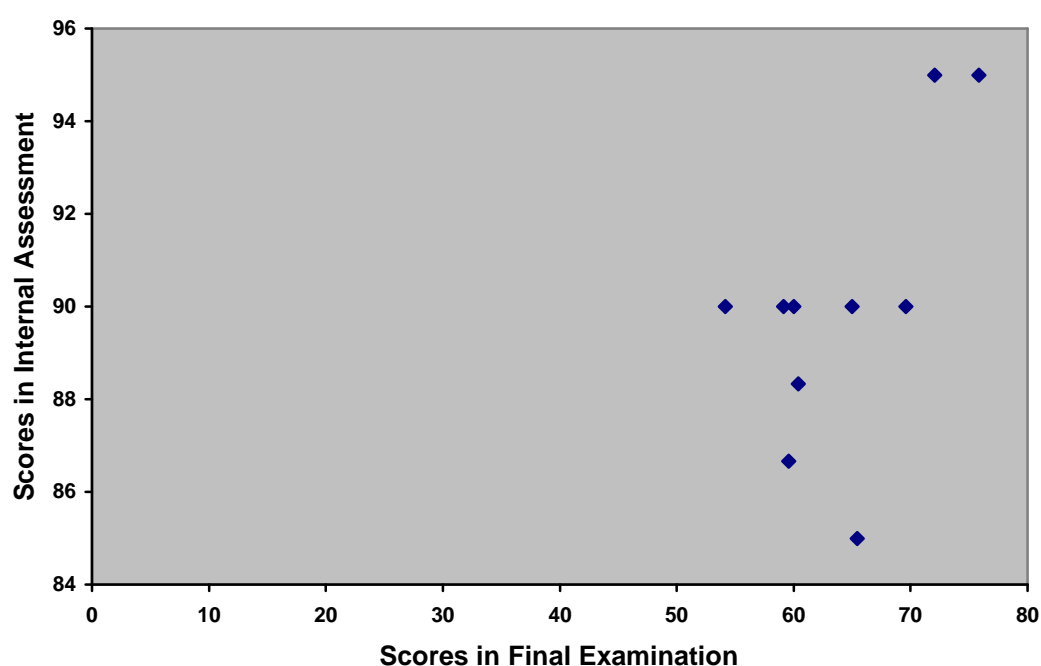
S. N.	x	Y	x^2	y^2	xy	r
1	90.00	54.16	8100.00	2933.30	4874.40	
2	90.00	60.00	8100.00	3600.00	5400.00	
3	90.00	59.16	8100.00	3499.90	5324.40	
4	90.00	69.58	8100.00	4841.37	6262.20	
5	95.00	72.08	9025.00	5195.52	6847.60	
6	86.66	59.58	7509.95	3549.77	5163.20	
7	85.00	65.41	7225.00	4278.46	5559.85	
8	88.33	60.41	7802.18	3649.36	5336.02	
9	90.00	65.00	8100.00	4225.00	5850.00	
10	95.00	75.83	9025.00	5750.18	7203.85	
	$x \text{ X}$ 899.99	$y \text{ X}$ 641.21	$x^2 \text{ X}$ 81087.13	$y^2 \text{ X}$ 41522.86	$xy \text{ X}$ 57821.52	+0.59

The table- 6 can be better explained with the help of the following bar diagram from which two variables can be compared.



The bar diagram shows that all the students from the Department of Nepali Language Education secured higher scores in internal assessment than in final examination.. There is much difference between the scores the student secured in internal assessment and in final examination. But this does not mean there is negative correlation between the two sets of scores. The correlation coefficient between the two variables can be observed in the following scatter gram.

Diagram - 9: Correlation- Coefficient of the Two Sets of Scores of Students from DNLE



In the above scatter gram, the dots take the shape of narrow band which indicates high correlation between two variables. When calculated the correlation coefficient of the two variables of table.6, it resulted the value of +0.59 (see appendix II) which indicates high positive correlation. Therefore, the coefficient of correlation between the two sets of scores of the students from the Department of Nepali Language Education is high.

3.3.5 Department of Health and Physical Education

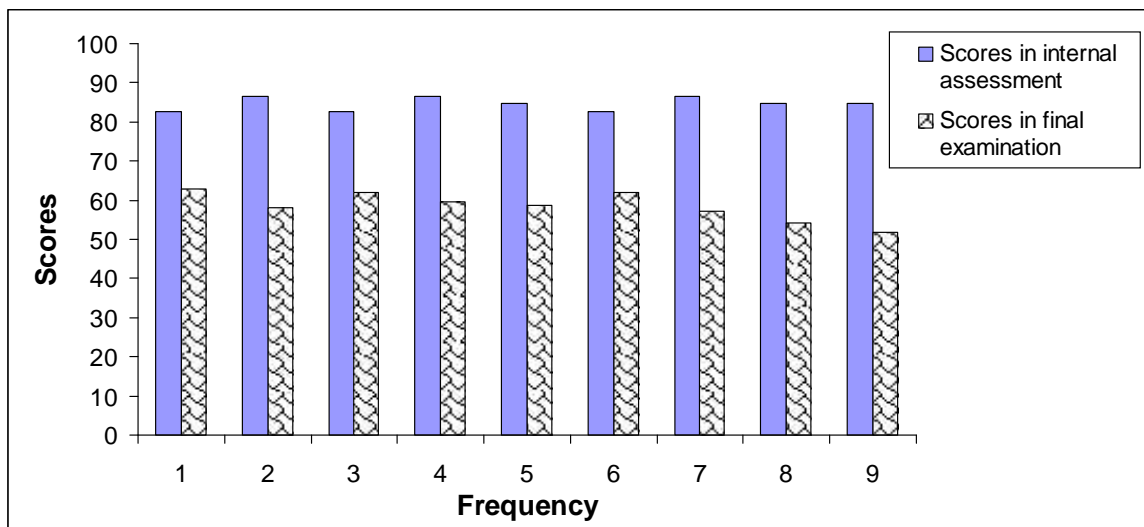
Department of Health and Physical Education was the 5th largest department in the Central Department of Education in the academic year 2062 in terms of student enrolment in M. Ed first year. In the academic year 2062, 162 students were admitted in this department out of which, 9 students were randomly selected for the present study. Analysis of the scores of the selected students is given below where variable 'x' represents the internal assessment scores and variable y represents the final examination scores.

Table -8 Analysis of the Scores of Students from DHPE

S.N.	x	y	x^2	y^2	xy	r
1	82.69	62.98	6837.64	3966.48	5207.82	
2	86.53	58.17	7487.44	3383.75	5033.45	
3	82.69	62.01	3837.64	3845.24	5127.61	
4	86.53	59.61	7478.85	3553.35	2158.05	
5	84.61	58.65	7158.85	3439.82	4962.38	
6	82.69	62.01	6837.64	3845.24	5127.61	
7	86.53	57.21	7478.44	3272.98	4950.38	
8	84.61	54.32	7158.85	2950.66	4596.02	
9	84.61	51.92	7158.85	2695.69	4392.95	
	$x X$ 761.49	$y X$ 526.88	$x^2 X$ 64451.79	$y^2 X$ 30953.21	$xy X$ 44556.27	+0.47

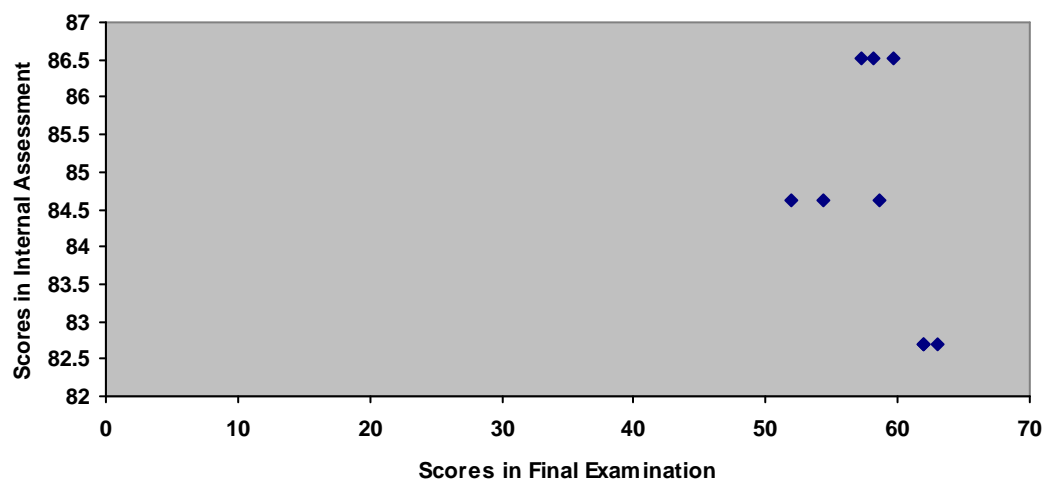
In the above table, the variable x and y represent the two sets of scores of the students from the department of health and physical education. These two sets of scores can be presented more vividly using the following bar diagram.

Diagram - 10 Bar diagram Showing Two Sets of Scores of Students from DHPE



The bar diagram shows that all the students from the Department of Health and Physical Education obtained better marks in internal assessment than in final examination. The diagram also shows that there is much difference between the two scores. But the correlation between the two variables can be observed only in the scatter gram which is given below:

Diagram-11: Correlation Co-efficient of the Two Sets of Scores of Students from DHPE



The shape of dots in the diagram indicates that the two variables are negatively related. From the calculation also the correlation value of the variables x and y

given in table -7 is determined to be -0.47 (see appendix II). This value indicates that there exists low negative correlation between the two sets of scores of the students from the Department of Health and Physical Education.

3.3.6 Department of Educational Planning and Management

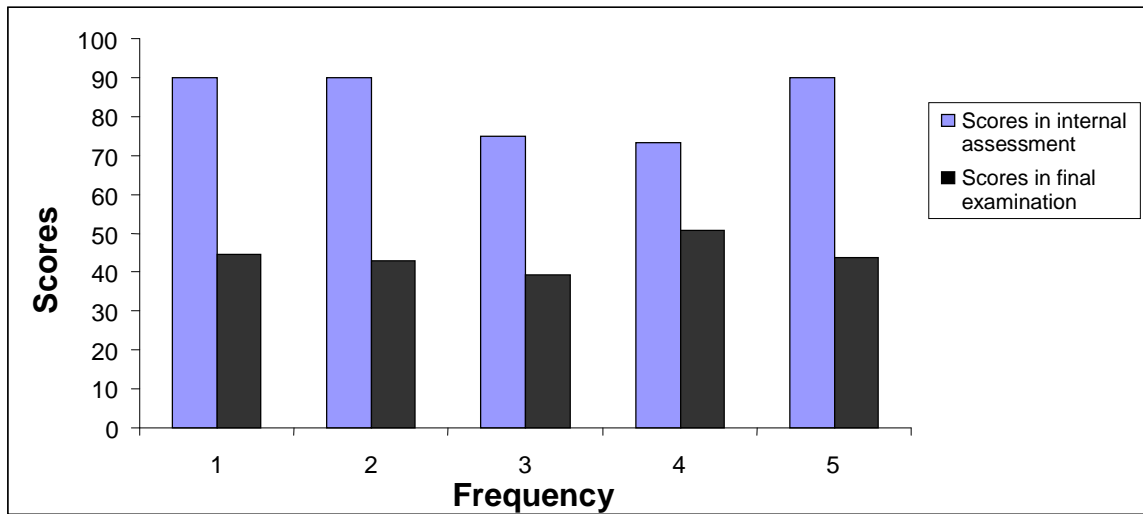
In the Department of Educational Planning and Management, Central Department of Education, 74 students were admitted in M.Ed. first year in the academic year 2062. Out of this number 5 students have been randomly selected for the purpose of the present study. Scores of the selected students in both internal assessment (i.e. represented by variable x) and final examination (i.e. represented by variable y) and its analysis has been presented below:

Table -9 Analysis of Scores of Students from DEPM

S.N.	x	y	x^2	y^2	xy	r
1	90.00	44.58	8100.00	1987.37	4012.20	
2	90.00	42.91	8100.00	1841.26	3861.90	
3	75.00	39.16	5625.00	1533.50	2937.00	
4	73.33	50.83	5377.28	2583.68	3727.36	
5	90.00	43.75	8100.00	1914.06	3937.50	
	$x \times 418.33$	$y \times 221.23$	$x^2 \times 35302.28$	$y^2 \times 9859.87$	$xy \times 18475.96$	-0.22

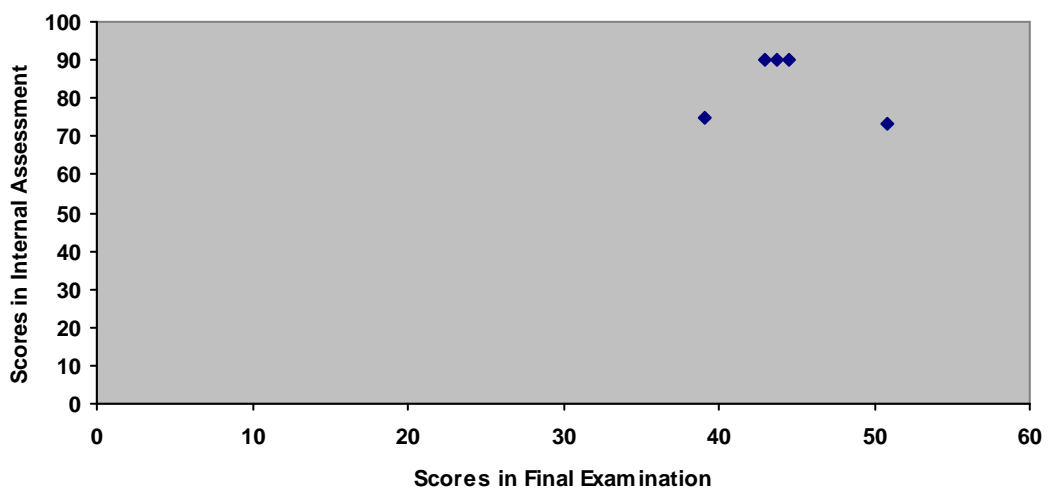
The variable x and y in the above table can be better explained and can be compared with the help of the following bar diagram.

Diagram - 12 Bar diagram Showing Two Sets of Scores of Students from DEPM



The above diagram shows that all the students from the Department of Educational Planning and Management secured higher scores in internal assessment than in final examination. Out of five students, three students' internal assessment score is more than twice as much as the final examination score. There is a marked difference between the two sets of scores. Now, the coefficient of correlation between the two sets of scores can be observed in the following scatter gram:

Diagram- 13 Correlation Co-efficient of the Two Sets of Scores of Students from DEPM



In the above scatter gram, the dots take the shape of scattered band moving slightly towards the upper left corner. This indicates that the two variables are negatively related. The calculation also shows that the correlation vale of the two sets of scores is -0.22 (see appendix II). So the correlation coefficient between the two sets of scores of the students from the Department of Educational Planning and Management is very low and the direction of their relationship is negative.

3.3.7 Department of Economics Education

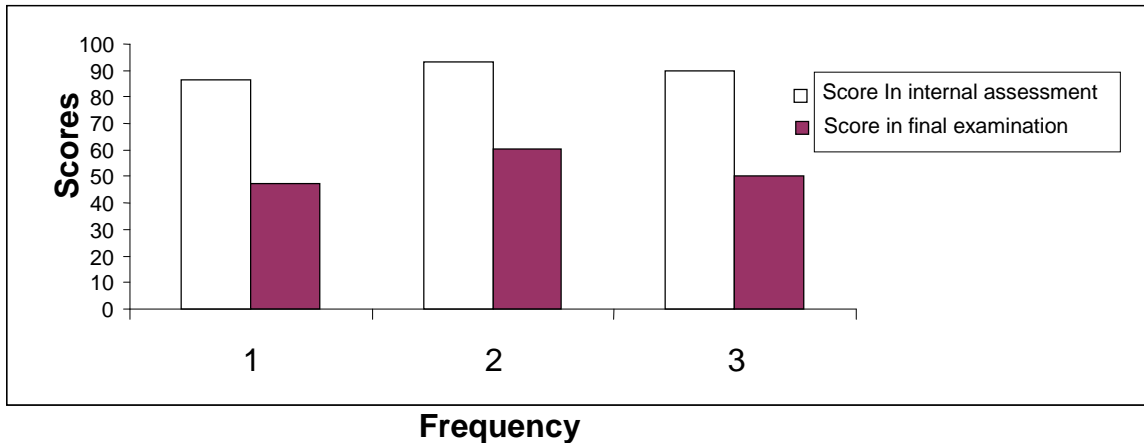
Department of Economics Education is the smallest department among the departments selected for the study. But this department is the largest department among the departments of social science education in terms of student enrolment. In the academic year 2062, 34 students were admitted in this department out of which, three students have been selected for the study. Analysis of the scores of the selected students is given below where variable x represents the internal assessment scores and variable 'y' represents the final examination scores.

Table -10 Analysis of Scores of Students from DEE

S.N.	x	y	x^2	y^2	xy	r
1	86.66	47.50	7509.95	2256.25	4116.35	
2	93.33	60.41	8710.48	3649.36	5638.06	
3	90.00	50.41	8100.00	2541.16	4536.90	
	$x X$ 269.99	$y X$ 158.32	$x^2 X$ 24320.43	$y^2 X$ 8446.77	$xy X$ 14291.31	+0.95

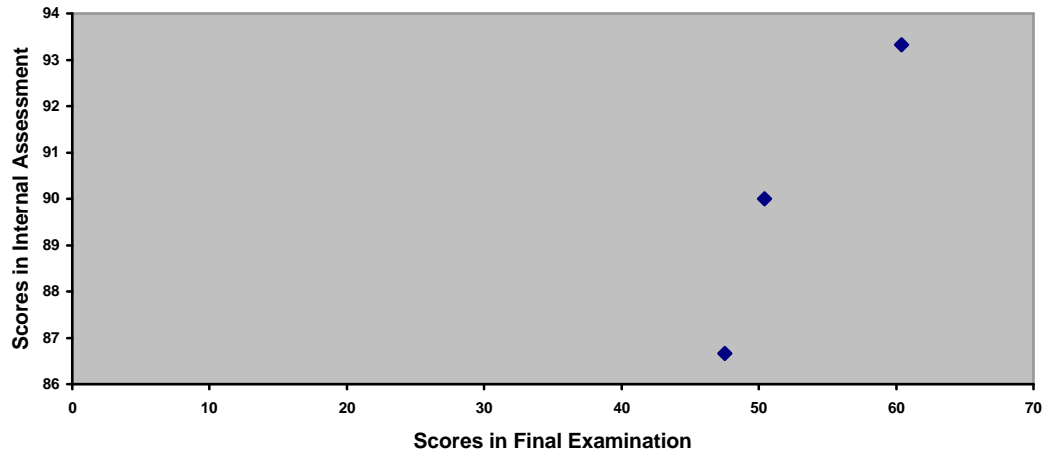
In the above table the variables x and y represent the two sets of scores of the students from the department of economics education. These two sets of scores can be presented more vividly using the following bar diagram.

Diagram -14: Bar Diagram Showing Two Sets of Scores of Students from DEE



The bar diagram shows that all the students from the Department of Economics Education obtained better marks in internal assessment than in final examination. The diagram also shows that there is much difference between the two sets of scores. The coefficient of correlation between the two sets of scores can be observed in the scatter gram below:

Diagram- 15: Correlation-Coefficient of the Two Sets of Scores of Students from DEE



The scatter gram shows that the variables are positively related since the direction of he dots moves from lower left to upper right corner. The calculation of these two variables has also resulted the value +0.95 (see appendix II) Therefore, the correlation between the two sets of scores of the students from the Department of Economics Education is very high and the direction of their relationship is positive.

CHAPTER - FOUR

FINDINGS AND RECOMMENDATIONS

This chapter deals with the major findings of the study. On the basis of the analysis of collected data, findings have been drawn and presented in this chapter. The chapter also includes some recommendations and pedagogical implications which are suggested on the basis of the findings

4.1. Findings of the Study

After analyzing the collected data, the following findings have been derived:

1. When compared the degree of predictive validity across departments, the following findings have been derived.
 - i. Department of Economics Education is in the top position in terms of predictive validity of internal assessment having very high predictive validity.
 - ii. Department of Curriculum and Evaluation is in the second position which also has very high predictive validity and the Departments of Nepali Language Education and Mathematics Education are in third and fourth position respectively having high predictive validity of internal assessment. Similarly, the Department of English Language Education is in fifth position having low predictive validity.
 - iii. The Departments of Educational Planning and Management and Health and Physical Education are in sixth and seventh position in terms of the degree of predictive validity of internal assessment. In these two departments, internal assessment scores negatively predict the final examination scores. In the department of EPM, internal assessment has very low negative predictive validity and in the Department of Health and Physical Education it has low negative

predictive validity and it is in the bottom position among the seven departments in the order of the degree of predictive validity.

2. From the department-wise analysis of the two sets of scores, it has been found that there is no uniformity in the degree of predictive validity of internal assessment.
 - i. In the Department of English Language Education, the predictive validity of internal assessment is low since the correlation value of two sets of scores is +0.28
 - ii. In the Department of Mathematics Education the predictive validity of internal assessment is high where the coefficient of correlation between the two sets of scores is found to be +0.54
 - iii. In the Department of Curriculum and Evaluation, the predictive validity of internal assessment is very high since the correlation analysis resulted the value +0.83.
 - iv. In the Department of Nepali Language Education, the predictive validity of internal assessment is high where +0.59 is the value of correlation between the two sets of scores.
 - v. In the Department of Health and Physical Education the predictive validity of internal assessment is low negative since the correlation between two sets of scores is -0.47. This indicates that internal assessment scores negatively predict the final examination scores.
 - vi. In the Department of Educational Planning and Management, the internal assessment has very low negative predictive validity which indicates that the internal assessment scores negatively predict the final examination scores and the correlation value of the two sets of scores is -0.22
 - vii. In the Department of Economics Education, the predictive validity of internal assessment is very high since the correlation value of the two sets of scores is +0.95.

4.2 Recommendations and Pedagogical Implication

Based on the findings derived from the analysis and interpretation of the collected data, some pedagogical implications with some recommendations have been suggested below:

1. The departments which have low predictive validity or negative predictive validity of internal assessment should try to improve their assessment systems in order to achieve high degree of predictive validity.
2. These departments should make their assessment system more objective and accurate.
3. To achieve higher degree of predictive validity, tools of assessment should be designed in such a way that assesses actual performance of students.
4. The departments which have high or very high degree of predictive validity of their internal assessment should also make efforts to achieve even higher degree or perfect predictive validity since the goal of every test is to achieve the perfect predictive validity.

References

- Allen, J.P.B. & Davies, A. (Eds.) (1977). *Testing and experimental methods* (EC in AL Vol. 4). London: OUP.
- Arends, R.I. (2001). *Learning to teach*. New York: McGraw Hill Higher Education.
- Batala, K.P. (2004). *Validation of SLC examination*. An Unpublished Thesis of M. Ed., TU, Kathmandu.
- Best, W.J. & Kahn, J.R. (2004). *Research in education*. New Jersey: Prentice Hall.
- Bhandari, M.P. (2007). *Content validity of grammar theory and practice (Eng. Ed. 512) Tests*. An Unpublished Thesis of M. Ed., TU, Kathmandu.
- Bhattarai, P.P. (2008). *Attitude of teachers and students towards internal assessment: a case of M. Ed.* An Unpublished Thesis of M.Ed., TU, Kathmandu.
- Chauhan, C.P.S. (1993). *Emerging trends and educational evaluation*. Delhi: Common Wealth Publisher.
- Connolly, T.G. & Sluckin, W. (1971). *An introduction to statistics for the social sciences*. London: Macmillan.
- Davies, A. (1977). *The construction of language tests*. In Allen and Davies (Eds.) (1977). London: OUP.
- Ebel, R.L. & Frisbie, D.A. (2004). *Essentials of educational measurement*. New Delhi: Prentice Hall of India Pvt. Ltd.
- Faculty of Education Central Department of Education (2005). *Department profile 2005*. Kathmandu: FOE,TU.
- Hamal, R. (2007). *Content validity of general English test items at B.Ed.* An Unpublished Thesis of M.Ed., TU, Kathmandu.
- Hughes, A.(1989). *Testing for language teachers*. London: Cambridge

- Heaton, J.B. (1975). *Writing english language tests*. London: Longman.
- Ingram, E. (1977). *Basic concepts in testing*. London: OUP.
- Kandel, R.P. (2006). *A study on the correlationship between Sent up and SLC examination results*. An Unpublished Thesis of M.Ed., TU, Kathmandu.
- Khaniya, T.R. (2006 May 7). Internal Assessment in TU. *The Kathmandu post*, p.4.
- Khaniya, T. R. (2005) *Examination for enhanced learning*. Kathmandu: Millennium Publication Pvt. Ltd.
- Khatri, R. (2007). *The content validity of translation theory and practice language testing test at M. Ed. level*, An Unpublished Thesis of M. Ed., TU, Kathmandu.
- Kumar, R. (2006). *Research methodology*. London: Sage Publication Ltd.
- Linn, R.L. & Gronlund, N.E. (2003). *Measurement and assessment in teaching*. Delhi: Pearson Education Pte. Ltd.
- Sthapit, A.B. et al. (2004). *Statistical Methods*. Kahmandu: BAE Pvt. Ltd.
- Timsina, H.L. (2006). *Testing the test investigating the content validity of language testing test at M.Ed. level*. An Unpublished Thesis of M.Ed., T.U., Kathmandu.
- Thapa, T.B. (2005). *Predictive validity of grade XI annual English test*. An Unpublished Thesis of M.Ed., T.U., Kathmandu.
- Zunker, V.G. (1982). *Using assessment results in career counselling*. USA: Brooks/Cole Publishing Company Monterey California.

Department of English Language Education

Department of Mathematics Education

Department of Curriculum and Evaluation

Department of Nepali Language Education

Department of Health and Physical Education

Department of Educational Planning and Management

Department of Economics Education