

**A STUDY ON RELIABILITY USING TEST-RETEST METHOD**

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

**By  
Ishwori Shrestha**

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

**2007**

# **A STUDY ON RELIABILITY USING TEST-RETEST METHOD**

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

**By**

**Ishwori Shrestha**

**M.Ed. (Second Year)**

**Faculty of Education**

**Tribhuvan University**

**Kirtipur, Kathmandu, Nepal**

**2007**

**T.U. Regd. No.: 17495-94**

**Campus Roll No.: 691/059**

**Exam Roll No. (2<sup>nd</sup> year): 28189/060**

**Date of Approval of the Thesis**

**Proposal: 2063/09/02**

**Date of Submission: 2064-3-25**

## RECOMMENDATION FOR ACCEPTANCE

This is to certify that **Mrs. Ishwori Shrestha** has prepared this dissertation entitled "**A Study on Reliability Using Test-Retest Method**" under my guidance and supervision.

I recommend this dissertation for acceptance.

Date: \_\_\_\_\_

\_\_\_\_\_

**Dr. Anjana Bhattarai (Guide)**

Lecturer

Department of English Language Education

Faculty of Education

University Campus, T.U.

Kirtipur, Kathmandu, Nepal

## RECOMMENDATION FOR EVALUATION

This dissertation has been recommended for evaluation by the following research guidance committee:

**Dr. Chandreshwar Mishra**

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

---

Chairperson

**Dr. Anjana Bhattarai (Guide)**

Lecturer,  
Department of English Language Education  
T.U., Kirtipur, Kathmandu, Nepal

---

Member

**Mr. Prem Phyak**

Lecturer,  
Department of English Language Education  
T.U., Kirtipur, Kathmandu, Nepal

---

Member

Date: \_\_\_\_\_

## EVALUATION AND APPROVAL

This dissertation has been evaluated and approved by the following thesis evaluation committee:

**Dr. Chandreshwar Mishra**

Reader and Head,

Department of English Language Education

Tribhuvan University

Kirtipur, Kathmandu, Nepal

---

Chairperson

**Dr. Jai Raj Awasthi**

Professor,

Department of English Language Education

Tribhuvan University

Kirtipur, Kathmandu, Nepal

---

Member

**Dr. Anjana Bhattarai (Guide)**

Lecturer,

Department of English Language Education

Tribhuvan University, Kirtipur

Kathmandu, Nepal

---

Member

Date: \_\_\_\_\_

# DEDICATION

*I dedicate this dissertation to my parents who helped me a lot to introduce me as a person what I deserve these days.*

## ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. Anjana Bhattarai, Lecturer of the Department of English Education for her regular guidance, invaluable suggestions, cooperation and encouragement from the very beginning to the end of the thesis.

My sincere gratitude goes to Prof. Dr. Shishir Kumar Sthapit, Department of English Education who provided valuable knowledge during my academic year. I would like to express due respect to Prof. Dr. Shanti Basnyat for her suggestions and encouragement to carry out research. Similarly, I would like to express my sincere gratitude to Dr. Chandsreshwar Mishra, Head of the Department of English Education who provided invaluable suggestions for this research.

I would like to express my respect to Prof. Dr. Jai Raj Awasthi, Prof. Dr. Govinda Raj Bhattarai, Assistant Dean, FOE, T.U., Prof. Dr. Tirth Raj Khaniya for their kind suggestions during this study. Likewise, I am equally grateful to Mr. Ram Ekwel Singh, Mr. Bishnu Singh Rai, Dr. Bal Mukunda Bhattarai, Mr. Leknanath Pathak and other teachers of the Department of English Education.

I am gratefully acknowledge to Mr. Ganga Ram Gautam, Mr. Rishi Ram Rijal and Mr. Netra Prasad Paudel, lecturers of Mahendra Ratna Campus, Tahachal for their kind support, encouragement and inspiration. Similarly, my indebtedness goes to Mr. Lekhanath Sharma, Lecturer of Department of Mathematics Education, T.U. Kirtipur for his suggestions to carry out statistical analysis in this research work.

Likewise, my special thanks go to Mrs. Nanda Kumari Maharjan, Mr Ganesh Gurung, Mr. Kamal Raj Devkota, Mr. Govinda Bhatta and Hom Kumar Shrestha for their kind co-operation.

My special thanks go to my husband Mr. Amrit Shrestha, brother Mr. Prakash Shrestha and sister Miss Sangita Shrestha for their co-operation for this work. Finally, I express my thanks to Resunga Computer Service, for attractive typing.

**Ishwori Shrestha**

## **ABSTRACT**

This research work attempts to find out the reliability of test-retest method in order to measure the writing skill of secondary level students' (i.e. grade nine). In order to do so, the researcher collected data from students of secondary level schools in Kathmandu district. The sample population consisted of 56 students who were selected by random sampling procedure. Only subjective type of test items were used to find out the reliability of tests by using test-retest method. From the analysis, she found that the reliability of test-retest method in assessing of writing skill is high.

The study consists of four chapters which are given below:

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

Chapter two deals with methodology. It includes sources of data, population of the study, sample population and sampling procedure, tools for data collection, process of data collection and limitations of the study.

Chapter three consists of analysis and interpretation of the data. the analysis of data has been carried out interms of school, sex and item wise. For that different statistical tools like mean, standard deviation, reliability etc. have been used.

Chapter four deals with the findings and recommendations of the study. On the basis of the analysis and interpretation, some significance, findings are enlisted and on the experience of the researcher while conducting the research, some recommendations are made. This chapter follows references and appendices.

## TABLE OF CONTENTS

	<b><u>Page No.</u></b>
Recommendation for Acceptance	i
Recommendation for Evaluation	ii
Evaluation and approval	iii
Dedication	iv
Acknowledgements	v
Abstract	vii
Table of Contents	ix
List of Tables	xi
Abbreviations	xii

### CHAPTER –I : INTRODUCTION

1.1. General Background	14
1.1.1. Language Testing	15
1.1.2. Qualities of a Good Test	16
1.1.3. Reliability	18
1.1.4. Test-Retest Method	20
1.2. Literature Review	21
1.3. Objectives of the Study	23
1.4. Significance of the Study	23

### CHAPTER –II : METHODOLOGY

2.1. Sources of Data	24
2.1.1. Primary Source of Data	24
2.1.2. Secondary Sources of Data	24
2.2. Population of the Study	24

2.3 Sampling Procedure	24
2.4. Tools for Data Collection	16
2.5. Process of Data Collection	25
2.6. Limitations of the Study	26

### **CHAPTER – III : ANALYSIS AND INTERPRETATION**

3.1 School-wise description of reliability of test-retest method	28
3.1.1 Nepal Yubak Secondary School (NYSS)	28
3.1.2 Shree Gita Mata School (SGMS)	29
3.1.3 Shree Siddhi Ganesh School (SSGS)	30
3.1.4 Shree Man Singh school (SMSS)	31
3.2 Sex-wise description of reliability of test-retest method	32
3.2.1 Female Students of NYSS	32
3.2.2. Male Students of NYSS	33
3.2.3. Female Students of SGMS	34
3.2.4 Male students of SGMS	35
3.2.5 Female Students of S.S.G.S	35
3.2.6. Male Students of SSGS	36
3.2.7 Female Students of S.M.S.S	37
3.2.8 Male Students of S.M.S.S.	38
3.3 Item-wise description of reliability of test-retest method	40
3.3.1 Controlled Writing	40
3.3.2 Sentence- Matching	41
3.3.3 Re-arrange	42
3.3.4 Guided Writing	44
3.3.5 Free Writing	45
3.4. Holistic description of reliability of test-retest method	46

### **CHAPTER – IV : FINDINGS AND RECOMMENDATIONS**

4.1 Findings	48
4.1.1 As a Whole Finding	48
4.1.2 School-Wise Findings	48
4.1.3 Sex-Wise Findings	48
4.1.4 Item-Wise Findings	49
4.2 Recommendations	49
<b>REFERENCES</b>	<b>38-39</b>
<b>APPENDICES</b>	
APPENDIX-1	
APPENDIX-2	
APPENDIX-3	

## LIST OF TABLES

	<b>Page No.</b>
Table No. 1: Students' scores in 2 Tests	28
Table No. 2: Students' Scores in 2 Tests	29
Table No. 3: Students' Scores in 2 Tests	30
Table No. 4: Students' Scores in 2 Tests	31
Table No. 5: Students' Scores in 2 Tests	32
Table No. 6: Students' Scores in 2 Tests	33
Table No. 7: Students' Scores in 2 Tests	34
Table No. 8: Students' Scores in 2 Tests	35
Table No. 9: Students' Scores in 2 Tests	36
Table No.10: Students' Scores in 2 Tests	37
Table No. 11: Students' Scores in 2 Tests	38
Table No.12: Students' Scores in 2 tests	39
Table No. 13: Mean, S.D. and Reliability of Four Schools	40
Table No. 14: Mean, S.D. and Reliability of Four School	41
Table No. 15: Mean, S.D. and Reliability of Four Schools	43
Table No. 16: Mean, S.D. and Reliability of Four Schools	44
Table No. 17: Mean, S.D. and Reliability of Four School	45
Table No. 18: Holistic Description of Reliability (Male + Female)	38

## **ABBREVIATIONS**

F.M.	Full Mark
i.e.	That is
N.Y.S.S.	Nepal Yubak Secondary School
O.M.	Obtained Mark
P.M.	Pass Mark
S.G.M.S.	Shree Gita Mata School
S.M.S.S.	Shree Man Singh School
S.N.	Serial Number
S.S.G.S.	Shree Siddhi Ganesh School
St.Dev.	Standard Deviation

## CHAPTER –I

### INTRODUCTION

#### 1.1. General Background

Language is not an end in itself . . . It is a way of connection between souls, a means of communication."

- Otto Jespersen (1904: 4)

Language is the most widely used means of communication for human beings. Human beings express their feelings, desires, thoughts and emotions through language. Because of the access of language, they have become able to reach the summit of this advanced civilization. According to Robins (1999:135) "Language is a symbol system based on purely arbitrary convention infinitely extendable and modifiable according to changing needs and conditions of speakers". Here, this definition clearly mentions that language is a symbolic system. It is extendable, modifiable infinitely according to the intention, desire, need of the speakers.

The aforementioned definitions clearly mention that language is not a haphazard compilation of smaller units. It is a system of systems. Therefore, sound units are systematically structured to form words and phrases and so on. In fact, language is a cover term that can be studied at four levels viz. phonological level, morphological level, syntactic level and semantic level. There are four skills. They are listening, speaking, reading and writing. The writing skills are complex and sometimes difficult to teach, requiring mastery not only of grammatical and rhetorical devices but also of conceptual and judgemental elements. The writing skill in general has various sub-skills. Heaton (1975) has

attempted to group many and varied skills necessary for writing good prose into four main areas. (i) Grammatical skills: the ability to write correct sentences, (ii) stylistic skills: the ability to manipulate sentences and use language effectively, (iii) mechanical skills: the ability to use correctly those convention peculiar to the written language. e.g. punctuation, spelling, (iv) judgement skills: the ability to write in an appropriate manner for a particular purpose with an ability to select, organize and order relevant information. Therefore, testing writing indicates testing the skills mentioned above. The writing ability is tested usually in the form of essays, letters and testing writing vary according to whether they deal separate factors such as punctuation, spelling, structure or vocabulary. However in every case, we have to set up a stimulus to obtain a controlled response in written form containing the problem we wish to test. The type of stimulus to be used depends upon testing of writing at different levels. e.g. sentence level, supra-sentence level and creative writing.

### **1.1.1. Language Testing**

Testing is the cover term to refer to any means of checking what students can do and cannot do with the language taught. Testing is an essential and integral part of education. Education involves teaching and testing both. It is said that there is whole part relationship between language education and testing. To show the clear-cut relationship between language education and testing, we need one more term language teaching. Language teaching and testing can function like the combination of a pick and a shovel to dig deep into the language education.

According to David Baker (1989: 3) "Language testing is a complicated subject and much of this complication stems from problems of description and measurement which are particularly acute in linguistic

and psychological investigation. It can be instructive therefore to look at other kinds of tests which do not share those particular difficulties". Life is full of tests of varying degrees of formality and important principles can often be seen operating more clearly in non-linguistic tests, where issues are simpler. Extending these principles to language testing can help to think clearly about what tests do and what they are for. There are two fundamental principles which provide a starting point for thinking about the goals of any kind of testing. A test is a way of arriving at a meaningful decisions and language tests also lead to decision.

A similar observation can be made about the so-called progress test. If we decide that we need a real test, identifying the decision that needs to be made is an important first step in constructing or choosing an appropriate instrument. If we discover that we don't need a real test, the operation of this criterion may save a lot of time and expense. Appreciation of the close link between testing and decision making enables the test user or writer to approach the task of evaluating a group of learners with a much clear ideas of what kind of test is needed if indeed a test is needed at all.

The goal of language testing is to check what students can do and cannot do with the language taught.

### **1.1.2. Qualities of a Good Test**

The specific qualities that determine the overall usefulness of a given test. There are six test qualities which are including in test usefulness: reliability, construct validity authenticity, interactiveness, impact and practicality.

Reliability is defined as consistency of measurement. A reliable test score will be consistent across different characteristics of the testing

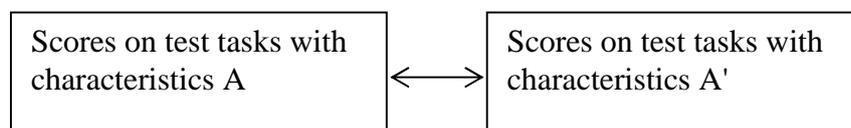
situation. Thus, reliability can be considered to be a function of the consistency of scores from one set of test and test tasks to another. Construct validity pertains to the meaningfulness and appropriateness of the interpretations that we make on the basis of test scores. When we interpret scores from language tests as indicate of test takers' language ability, a crucial question is, 'To what extent can we justify these interpretations?' the clear implication of this question is that as test developers and test users we must be able to provide adequate justification for any interpretation we make of a given test scores. Authenticity is the degree of correspondence of the characteristics of a given language test-task to the features of a TLU task authenticity to be an important test quality because it relates the test task to the domain of generalization to which we want our scores interpretations to generalize. Authenticity thus provides a means for investigating the extent to which score interpretation generalized beyond performance on the test to language use in the TLU domain. The define interactiveness as the extent and type of involvement of the test taker's individual characteristics in accomplishing a test task. The individuals characteristics that are most relevant for language testing are the test taker's language ability. The interactiveness of a given language test task can thus be characterized in terms of the ways in which the test taker's area of language knowledge; meta cognitive, strategies, topical knowledge, and effective schemata are engaged by the test task. A variety of individuals will be affected by and thus have an interest, or hold a 'stake' in the use of a given test in any particular situation. 'Stakeholders' that are directly affected include the test takers and the test users or decision makers. In addition, a large number of individuals will be indirectly affected. Finally, to the extent that it has an impact on the societal or educational system is indirectly affected by the use of the test. Rather than attempting to discuss the

general systematic effect of test use or the potential indirect impact on individuals, we will focus our attention here on the impact on these individuals who are most directly affected by test use: test takers and teachers. The last test quality that needs to be considered in practicability, which is different its nature from the other five qualities, while those qualities pertain to the uses that are made of test scores, practicability pertains primarily to the ways in which the test will be implemented, and, to a large degree whether it will be developed and used at all. Practicality is a matter of the extent to which the demands of the particular test specifications can be met within the limits of existing resources.

As mentioned above, reliability is one of the most important qualities of a good test.

### **1.1.3. Reliability**

Reliability is defined as consistency of measurement. A reliable test scores will be consistent across different characteristics of the testing situation. Reliability can be considered to be a function of the consistency of scores from one set of tests and test tasks to another. It can be said that reliability can be considered to be a function of consistencies across different sets of test task characteristics. This can be represented as follows:



In this figure, the double-headed arrow is used to indicate a correspondence between two sets of task characteristics (A and A') which differently in incidental ways. For example, if the same tests were to be administered to the same group of individuals on two different occasions,

in two different settings, it shouldn't make any difference to a particular test-taker whether she takes that test on one occasion and setting or the other. Ebel (1965) clarifies this concept saying "The reliability coefficient for a set of scores from a group of examinees is the coefficient of correlation between that set of scores and another set of scores on an equivalent test obtained independently from the members of the same group. So, we can say 'reliability means the consistency with which set of test scores measure whatever they do measure. But there are some factors which influence reliability. They are length of the test, spread of scores, difficulty level of item, objectivity, administration of the test and ability level of examinee. There are four methods to find out reliability of test. They are as follows:

- a. Test-retest method
- b. Equivalent form method
- c. Split-halves method and
- d. Kuder-Richardson method

Test-retest is an estimate to determine the reliability of a test in which we get a group of subjects to take the same test twice. In other words, the same test is administered to the same group of learners twice keeping the considerable interval between the two administrations. The term 'considerable' refers to not too short because candidates are likely to recall items and their responses to them and not too long because learning or forgetting will have taken place. After administering the same test twice to the same group of learners the two sets of scores are compared. In alternative forms method, two parallel or equivalent forms of a test which try to measure exactly the same skill or ability, which use the same method of testing and which are of equal length and difficulty, are

administered to the same group of subjects. In this method, the reliability of a test is determined by the extent to which the learners get similar scores on parallel forms of a test. Split-half method involves only one administration of one test. In this method, the subjects take the test in the usual way, but each subject is given two scores so as to obtain two sets of scores. One score is for on half of the test and the other score is for the other half. Then, the two sets of scores are used to obtain the reliability coefficient as if the whole test had been taken twice. In this method, it is necessary to split the test into two really equivalent halves through the careful matching of items. In fact, where items in the test have been ordered in terms of difficulty, a split to odd-numbered items and even-numbered items may be adequate. This method resembles the alternative forms method except that the two 'forms' are only half the length. In Kudar-Richardson method, a measure of the degree to which the items or parts of a test are homogeneous or consistent with each-other. It is often estimated by comparing the two halves of a test or by a Kuder-Richardson reliability coefficient. Among them, test-retest method is the most important and easiest one in which the researcher wants to write and tries to find out something.

#### **1.1.4. Test-Retest Method**

According to Heaton (1968: 163), "Test-retest method of reliability is the method of measuring reliability of a test is to re-administer the same test after a lapse of time. It is assumed that all candidates have been treated in the same way in the interval that they have either all been taught or that none of them have provided that such assumptions can be made, comparison of the two results would then show how reliable the test has proved. It is an estimate to determine the reliability of a test in which we get a group of subjects to take the same test twice. In other

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

- to provide the interesting materials which are needed to the students for information, knowledge and pleasures.
- to develop the techniques to remove the factors which affected to the reliability.
- to find out the consistency with the sets of test scores and tests.
- to be able to allocate the time to the test.

## **1.2. Literature Review**

Some studies have been carried out to test the proficiency of different aspect of language in the Department of English Education.

Khanal (1997) conducted a research entitled "A study on the effectiveness of the close test over conventional objectives test in testing reading comprehension in English." He found that the private schools performed better than the public schools in both objectives and the close tests, it is not surprising that the private school performed better in non-textual materials and it was found more effective and reliable. So, it was recommended for the school of Nepal.

Baral (1999) had done a study on "Measurement of readability of a course in general English." He found the difficulty level of 46 passages in general English and ranks them in terms of their readability level.

Kshetree (2000) did a study on the "Washback effect of the SLC examination." He found that the washback effect is the real hindrance in achieving the goal of teaching English and he found the washback effect of exams in teaching and learning English also.

Ghimire (2002) tried to find out the Pronunciation Proficiency of the students of lower secondary level." He found that the proficiency of pronunciation of the boys in some selected words from their textbooks are better than girls as well as private students performed better than government school counterparts.

Similarly, Bhattarai (2002) carried out a research entitled "A comparative study of the writing proficiency of Bachelors Level Students". He found that the students of institutes have greater proficiency in writing than the students of faculties, within institutes, students of medical science are more proficient than those of the others.

"The reliability of a test suffers to the consistency of scores obtained of the same individuals in different occasions or with different sets of equivalent" (Anastosi: 133). In this study, the researcher will attempt to find out the reliability of a written test with the reference of test-retest method in grade nine. So far, the research has been on test of proficiency in the Department of English. The present study is a new endeavour and different from other studies in the sense that it is concerned with the reliability and test-retest method to find out the

reliability with the help of writing skill, using test-retest method in grade nine.

### **1.3. Objectives of the Study**

- a. To find out reliability of test-retest method in order to measure the writing skill.
- b. To compare the reliability of the test-retest method interms of following variables:
  - o School-wise
  - o Sex-wise
  - o Item-wise
- c. To suggest some pedagogical application and implication.

### **1.4. Significance of the Study**

As the proposed study is a new field in our country. It will find out how test-retest method shows the reliability of a test on the basis of tests from the English textbook of class nine. Anyone who wants to get in this field will be expected to be benefited with it. It is significance for the authors and the examiners. It is equally significance for the teachers who check the answer sheets and correct the mistakes. It provides the guidelines for the future researchers interested in the stated area.

## **CHAPTER –II**

### **METHODOLOGY**

Methodology is very important to carry out any research work. This chapter deals with the source of data, process of data collection and the limitations of the study that the researcher applied.

#### **2.1. Sources of Data**

This study was carried out on the basis of both primary and secondary sources of data.

##### **2.1.1. Primary Sources of Data**

The primary sources of data were the ninth graders studying in Kathmandu.

##### **2.1.2. Secondary Sources of Data**

Apart from the primary sources, the researcher studied different related books, journals, articles, reports, previous research work etc. in order to facilitate the study. For example, Ur (1992), Heaton (1975), Bachman (1989).

#### **2.2. Population of the Study**

The population of the study consisted of 56 students of grade nine of 4 schools of Kathmandu district.

#### **2.3 Sampling Procedure**

For this study, the researcher randomly selected fifty-six students from four secondary schools. The sample population can be shown clearly in the table below.

## **Sample Population of the Study**

Schools	Students
N.Y.S.S.	14
S.G.M.S.	14
S.M.S.S.	14
S.S.G.S.	14
Total population	56

### **2.4. Tools for Data Collection**

The researcher used the test-items to collect the data. (Appendix-1) There are five test items. They are controlled writing, sentence matching, re-arranging, guided writing and free writing, which are given in the table below clearly.

#### **Description of the Research Tools**

S.N.	Assessing quality	Mode of items	FM
1	Controlled writing	Subjective	10
2	Sentence matching	Subjective	10
3	Re-arranging	Subjective	10
4	Guided writing	Subjective	10
5	Free writing	Subjective	10
Total			50

### **2.5. Process of Data Collection**

After preparing the test items for data collection, the researcher visited the selected schools and explained the purpose of her visit to the schools authorities. She requested for a noiseless class and visited the schools in the given time. She went to the class-room and told her purpose, introducing herself and administered the test. The researcher personally visited the schools and administered the test. Instructions were clearly explained for each test items and no individual treatment was

given to the students. They were given only one hour. Most of them finished within the given time. After taking the exam, she assigned marks 3 for grammar, 5 for subject matter and 2 for punctuation marks of test A. In test B, the process was also the same which was taking after a lapse of time i.e. 10 days. In this way, the data for the study were obtained.

## **2.6. Limitations of the Study**

The study has following limitations:

- a. This research was limited to the secondary schools students in Kathmandu district.
- b. Only fifty-six students of the schools under study were included as the population of this study.
- c. The primary data for this study was collected only from the written test items.
- d. This study was further limited to the test-retest method of reliability.
- e. The test items were extracted only from the textbook of grade nine.
- f. Same test items were administered twice within 10 days in the same class.
- g. Students' homely environment and students' proficiency level of English are not accounted in this study.
- h. The analysis was limited to the writing skill only.

## CHAPTER – III

### ANALYSIS AND INTERPRETATION

This chapter is concerned with the analysis and interpretation of the data obtained from the students. For this research work, the researcher selected five questions to the students. She administered the test twice and calculated the total scores achieved by students. Here, the researcher has used different statistical tools to analyze and interpret the data. She has used mean, standard deviation and reliability to analyze the data.

The researcher has analyzed the data in terms of the following variables:

- a) School-wise
- b) Sex –wise
- c) Item-wise
- d) To analyze data, the researcher has used the following formula:

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

Where,

$r_{xy}$  = correlation between x and y

N = Number of given scores

x and y = Individual scores

– Pearson

#### **B. Correlation coefficient**

$r_{xy} < 0.3 \rightarrow$  poor correlatives

$0.3 < r_{xy} < 0.7 \rightarrow$  moderate/Satisfactory

$0.7 < r_{xy} \rightarrow$  highly correlated

### 3.1 School-wise Description of Relativity of Test-Retest Method

Four schools were used to obtain this information. School- wise description is presented below.

#### 3.1.1 Nepal Yubak Secondary School (N-Y-S-S)

Fourteen students of this school were given the same test twice to find out the reliability of test-retest method in assessing writing skill of those students. The description is presented below.

**Table No. 1: Students' Scores in 2 Tests**

S.N.	Testees	Test A	Test B
1	Suman Thapa	13	14
2	Dhiraj Thakur	18	22.5
3	Sudhir Malla	17	18
4	Suv Raj Khatri	22	18
5	Moti Gurung	11.5	21.5
6	Suman Rai	10	8
7	Raj Kumar R. Magar	24	26
8	Rabina Shrestha	17	18
9	Nabina Shakya	9	12
10	Susmita Shrestha	23.5	23.5
11	Rabita Gurung	15	14
12	Hasina Banu	30	30
13	Pramila Shrestha	16.5	23.5
14	Chandika Gurung	19	19

Mean: Test A: 4.464286

Test B: 4.464286

Standard dev. Test A: 1.11742

Test B: 0.77 220

Reliability: 0.8431

The above table shows that there were 14 students in a school who attended test twice. With the scores they got, the mean of test A is 4.46428, test B is 4.46428, standard deviation of test A is 1.11742. test B is 0.77120 and the reliability is 0.8431 which means the reliability of test retest method is high in this school.

### 3.1.2 Shree Gita Mata School (SGMS)

Fourteen students of this school were given the same test twice to find out the reliability of test-retest method in assessing writing skill of those students. The description is presented below.

**Table No. 2: Students' Scores in 2 Tests**

S.N.	Testees	Test A	Test B
1	Sabitri Ghale	36	32
2	Sarita Bhandari	22	22
3	Binita Shrestha	27	28
4	Rajan Maharjan	32.5	33
5	Kos Lal Gurung	31	31
6	Shital Sharma	23.5	24
7	Amita Sharma	22	20
8	Manisha Ghimire	35	36
9	Susmita Rai	34	33
10	Sharmila Pantha	28	29
11	Bishnu Bhandari	14	15
12	Madhu S. Giri	9	10
13	Basanta Shrestha	22.5	22
14	Basanta Sharma	25	24

Mean: Test A: 4.03714

Test B: 4.25000

Standard Deviation: Test A: 0.88718

Test B: 1.03310

Reliability: 0.8199

The above table shows that there were 14 students in a school who attended test twice. With the scores they got, the mean of test A is 4.03571, test B is 4.46428 and standard deviation of test A is 0.887180 and test B is 1.03310. The reliability is 0.8431 which means the reliability of test-retest method is high in this school.

### 3.1.3 Shree Siddhi Ganesh School (SSGS)

Fourteen students of this school were given the same test-twice to find out the reliability of test-retest method in assessing writing skill of those students. The description is presented below.

**Table No. 3: Students' Scores in 2 Tests**

S.N.	Testees	Test A	Test B
1	Arati Rai	32	32
2	Maya Tamang	19	18
3	Sharada Pandey	24	25
4	Sujan Pudasaini	20	20
5	Rita Thapa	10	12
6	Pratima Pariyar	15	14
7	DilBahadur Gurung	13	12
8	Krishna Khadka	27	22.5
9	Rahul K.C.	16	17
10	Sabin Shahi	15	15
11	Rabita Shahi	32.5	32.5
12	Ramita Karki	13	10.5
13	Sapana Thapa	23	23.10
14	Prakash Shrestha	24	22

Mean: Test A: 4.53571

Test B: 4.71428

Standard Deviation: Test A: 0.86523

Test B: 0.64194

Reliability: 0.5543

The above table shows that there were 14 students in a school who attended test twice. With the scores they got, the mean of test A is 4.53571, test B is 4.71428 and standard deviation of test A is 0.86523, test B is 0.64194 and the reliability is 0.5543 which means the reliability of test-retest method is moderate or the result is satisfactory.

### 3.1.4 Shree Man Singh school (SMSS)

Fourteen students of this school were given the same test twice to find out the reliability of test-retest method in assessing writing skill of those students. The description is presented below:

**Table No. 4: Students' Scores in 2 Tests**

S.N.	Testees	Test A	Test B
1	Hikmat Gurung	9.5	14.5
2	Rupa Gurung	30	23
3	Saurav Shrestha	37	38
4	Dev Raj Gurung	26.5	27.5
5	Padam Kumari Thapa	29	27
6	Ratna Tamang	12	14
7	Niraj Kattel	17	12
8	Kalpana Gurung	22.5	15.5
9	Chandra Tamang	28	25
10	Dil Bahadur Gurung	20	21
11	Kalpana Lama	12	13
12	Ang Babu Sherpa	14.5	17.5
13	Nirmala K.C.	32	33
14	Sabina Shrestha	15	15

Mean Test A: 4.321429

Test B: 4.250000

Standard deviation: Test A: 1.10256

Test B: 0.99518

Reliability: 0.8156

The above table shows that there were 14 students in a school who attended test twice. With the scores they got, the mean of test A is 4.32142, test B is 4.2500 and the standard deviation of test A is 1.102546, test B is 0.99518 and the reliability is 0.8156 which means the reliability of test-retest method is high in this school.

### **Conclusion**

Four schools were used to obtain the information. Among them, NYSS has done well because it has the highest scores i.e. 0.8431 which means the reliability of test-retest method in writing test is high in the contest of school.

### **3.2 Sex-Wise Description of Reliability**

Four schools were used to obtain this information. Sex-wise description is presented below:

#### **3.2.1 Female Students of NYSS**

Seven female students of this school were given the same test twice to find out the reliability of test-retest method in assessing writing skill of those students. The description is presented below:

**Table No. 5: Students' Scores in 2 Tests**

S.N.	Testees	Test A	Test B
1	Rabina Shrestha	17	18
2	Nabina Shakya	9	12
3	Susmita Shrestha	23.5	23.5
4	Rabita Gurung	15	14
5	Hasina Banu	36	30
6	Pramila Shrestha	16.5	23.5
7	Chandika Gurung	19	19

Mean: Test A: 19.14293

Test B: 21.28572

Standard Deviation: Test A: 8.71736

Test B: 7.76673

Reliability: 0.9028

The above table shows that seven testees were there who attended test twice. With the scores they got, the mean of test A is 19.14293, test B is 21.28572 and standard deviation of test A is 8.71776, test B is 7.76673 and the reliability of the test is 0.9028 which means the reliability of test-retest method is high in female students.

### 3.2.2. Male Students of NYSS

Seven male students of this school were given the same test twice to find out the reliability of test-retest method in assessing writing skill of those students. The description is presented below:

**Table No. 6: Students' Scores in 2 Tests**

S.N.	Testees	Test A	Test B
1	Suman Thapa	13	14
2	Dhiraj Thakur	18	22.5
3	Sudhir Malla	17	18
4	Suv Raj Khatri	22	18
5	Moti Gurung	11.5	21.5
6	Suman Rai	10	8
7	Raj Kumari R. Magar	24	26

Mean: Test: 16.78572

Test B: 18.28572

Standard Deviation: Test A: 5.13054

Test B: 5.94326

Reliability: 0.7858

The above table shows that seven testees were there who attended test twice. With the scores they got, the mean of test A is 16.78572 and test b is 18.285726, standard deviation of test A is 5.13054 and test B is 5.9432. The reliability is 0.7858 which means the reliability of test-retest method is high in male students.

### 3.2.3. Female Students of SGMS

Seven female students of this school were given the same test twice to find out the reliability of test-retest method in assessing writing skill of those students. The description is presented below:

**Table No. 7: Students' Scores in 2 Tests**

S.N.	Testees	Test A	Test B
1	Sabitri Ghale	30	32
2	Sarita Bhandari	22	22
3	Binita Shrestha	27	28
4	Amita Sharma	22	20
5	Manisha Ghimire	35	36
6	Sushma Rai	34	33
7	Sharmila Bhandari	28	29

Mean: Test A: 28.28574

Test B: 28.57144

Standard Deviation : Test A: 5.18702

Test B: 5.82692

Reliability: 0.9841

The above table shows that seven testees were there who attended test twice. With the scores they got, the mean of the test A is 28.28574, test B is 28.57144 and standard deviation of test A is 5.18702, test B is 5.82692. The reliability of these scores is 0.9841 which means the reliability of test-retest method is high in female students.

### 3.2.4 Male students of SGMS

Seven male students of this school were given the same test-twice to find out the reliability of test-retest method in assessing writing skill of those students. The description is presented below:

**Table No. 8: Students' Scores in 2 Tests**

S.N.	Testees	Test A	Test B
1	Rajan Maharjan	32.5	33
2	Kos Lal Gurung	31	31
3	Shital Sharma	23.5	24
4	Bishnu Bhandari	14	18
5	Madhu S. Giri	3	10
6	Basanta Shrestha	22.5	22
7	Basanta Sharma	25	24

Mean: Test A: 22.28573

Test B: 21.42864

Standard Deviation: Test A: 10.35163

Test B: 8.96025

Reliability: 0.9672

The above table shows that seven testees were there who attended the test twice. With the scores they got, the mean of test A is 22.28573, test B is 21.42864 and the standard deviation of test A is 10.35163, test B is 8.96025. The reliability of those scores is 0.9672 which means the reliability of test re-test method is high in male students.

### 3.2.5 Female Students of S.S.G.S

Seven female students of this school were given same test twice to find out the reliability of test re- test method in assessing writing skill of those students. The description is presented below.

**Table No. 9: Students' Scores in 2 Tests**

S.N.	Testees	Test A	Test B
1	Aarati Rai	32	32
2	Maya Tamang	19	18
3	Sharada Pandey	24	25
4	Sapana Thapa	22	24
5	Rita Thapa	10	12
6	Pratima Pariyar	15	14
7	Ramita Karki	13	10.5

Mean: Test: A : 25.50000

Test B: 23.00000

Standard deviation: Test A: 7.54980

Test B: 6.51280

Reliability: 0.8959

The above table shows that there were seven testees who attended test twice. with the scores they got, the mean of test A is 25.00000, test B is 23.00000 and the standard deviation of test A is 7.54980, test B is 6.51280. The reliability of these scores is 0.8959 which means the reliability of test-retest method is high in female students.

### **3.2.6. Male Students of SSGS**

Seven male students of this school were given the same test twice to find out the reliability of test-retest method in assessing writing skill of those students. The description is presented below.

**Table No.10: Students' Scores in 2 Tests**

S.N.	Testess	Test A	Test B
1	Sagan Paudasaini	20	20
2	Dil Bahadur Gurung	20	21
3	Krishna Khadka	27	22.5
4	Rahul K.C.	10	17
5	Sabin Shahi	15	16
6	Rabin Thapa	21	20
7	Prakash Shrestha	24	22

Mean : Test A: 22.14292

Test B: 22.14292

Standard Deviation : Test A :8.81623

Test B: 8.94763

Reliability: 0.9729

The above table shows that seven students were there who attended the same test twice. With the scores they got, the mean of test A is 22.14292, test B is 22.142923 and the standard deviation of the test A is 8.3162, test B is 8.94763. The reliability of these scores is 0.9729 which shows the reliability of test re-test method is high in male students.

### **3.2.7 Female Students of S.M.S.S**

Seven female students of this school were given the same test twice to find out the reliability of test re-test method in assessing writing skill of those students. The description is presented below.

**Table No. 11: Student Scores in 2 Tests**

S.N	Testees	Test A	Test B
1	Rupa Grrung	30	23
2	Kalpana Gurung	22.5	15.5
3	Chandra Tamang	28	25
4	Bipana Thapa	22	25
5	Kalpana Lama	12	13
6	Angila Sherpa	17.5	14
7	Sabina Shrestha	15	16

Mean: Test A: 20.78575

Test B: 20.57145

Standard Deviation:

Test A: 9.00863

Test B: 9.28453

Reliability: 0.9935

The above table shows that seven testees were three who attended the same test twice. With the scores they got, the mean of test A is 20.7857 and test B is 20.5714 and the standard deviation of test A is 9.00863 and test B is 9.28453. The reliability of these scores is 0.9935 which means the reliability of test re-test method is high in female students.

### **3.2.8 Male Students of S.M.S.S.**

Seven students of this school were given the same test twice to find out the reliability of test re-test method in assessing writing skill of those students. The description is presented below:

**Table No.12: Students' Scores in 2 tests**

S.N.	Testees	Test A	Test B
1	Hikmat Gurung	9.5	14.5
2	Saurav Shrestha	37	38
3	Devi Raj Gurung	26.5	27.5
4	Niraj Kattel	17	12
5	Basanta Shrestha	22.5	22
6	Ratna Tamang	12	14
7	Ang Babu Sherpa	14.5	17.5

Mean: Test A: 21.00000

Test B: 18.78570

Standard Deviation: Test A: 6.97610

Test B: 4.22150

Reliability: 0.8994

The above table shows that seven testees were there who attended the same test twice. With the scores they got, the mean of test A is 21.00000, test B is 18.78570 and the standard deviation of test A is 6.9761 and test B is 4.22150. The reliability of these scores is 0.8994 which means the reliability of test-retest-method is high in male students

### **Conclusion:**

Students of four schools were used to obtain the information. Among them students of SMSS have done well because they have the highest score i.e. 0.9935 (Female) and 0.8994 (Male) in sexwise description. They mean the reliability of test re-test method in assessing writing which tests are high.

### 3.3 Item-wise Description of Reliability

Four schools were used to obtain this information. Item-wise description is presented below.

#### 3.3.1 Controlled Writing

Four schools were given the same test twice to find out the reliability of test re-test method in controlled writing of the students. The description is presented below.

**Table No. 13: Mean, S.D. and Reliability of Four Schools**

N.Y.S.S	Mean: Test A : 4.46420 Test B: 4.46420	St.dev: Test A : 1.11742 Test B: .77122	Reliability: .8431
S.G.M.S	Mean: Test A: 4.03570 Test B :4.25000	St.dev:Test A : .8871 Test B:1.10331	Reliability: .8199
S.S.GS.	Mean: Test A: 4.53572 Test B:4.71422	St.dev: Test A : .86523 Test B.64193	Reliability: .5543
S.M.S.S	Mean: Test A: 4.32144 Test B: 4.25000	St.dev: Test A: 1.10253 Test B:.99516	Reliability: .8156

The above table shows that N.Y.S.S has the different mean and standard deviation in controlled writing. The reliability of N.Y.S.S in controlled writing is .8431 which means the reliability of test re-test is high in controlled writing. The mean of test A is 4.46420, Test B is 4.46420 and standard deviation in test A is 1.11742 and test B is 0.77122. S.G.M.S has also different mean and standard deviation in controlled writing. The reliability in controlled writing of S.G.M.S is .8199. The mean of test A is 4.0357, Test B is 4.25000 and the standard deviation of test A is .88710 and test B is 1.1331. The reliability .8199 shows that the reliability of test re-test method is high in controlled writing. Like wise, S.S.G.S has different mean and standard deviation in controlled writing.

The mean of test A is 4.53572, test B is 4.71422 and standard deviation of test A is .8652 and test B is .64193 in controlled writing. The reliability in controlled writing of S.S.M.S is .5543 only which means there reliability of test-retest method is moderate or satisfactory. At last, S.M.S.S has also different mean and standard deviation in controlled writing the mean of test A is 4.32144 and test B is 4.25000 and the standard deviation in test A is 1.10253 and test B is .99516. The reliability of the test in controlled writing of S.M.S.S is .8150 which means the reliability of test-retest method is high

### **Conclusion:**

Students of four schools were used to obtain the information. Among them, students of NYSS have the highest reliability i.e. 0.8431 in controlled writing. It means the reliability of test-retest method in assessing writing skill is high.

### **3.3.2 Sentence- Matching**

Four schools were given the same test twice to find out the reliability of test re-test method in sentence-matching of the students. The description is presented below.

**Table No. 14: Mean, S.D. and Reliability of Four School**

N.Y.S.S	Mean: Test A : 6.42863 Test B: 6.96423	St.dev: Test A : 2.35220 Test B: 2.21410	Reliability: 0.9780
S.G.M.S	Mean: Test A: 5.64282 Test B: 5.75002	St.dev: Test A : 1.88546 Test B: 1.67276	Reliability: 0.9315
S.S.GS.	Mean: Test A: 6.07143 Test B: 6.46423	St.dev: Test A : 1.83832 Test B: 1.77006	Reliability: 0.9400
S.M.S.S	Mean: Test A: 6.50000 Test B: 7.35716	St.dev: Test A: 1.96113 Test B: 1.74782	Reliability: 0.9551

The above table shows that N.Y.S.S has the different mean and deviation in sentence matching. The researcher has found that the mean of test A is 6.42853, test B is 6.96423 and the standard deviation of test A is 2.35220 and test B is 2.21410 and reliability of tests is .9780 which means the reliability of test-retest method is high in sentence matching. S.G.M.S has the different mean and standard deviation in sentence matching. The mean of test A is 5.64252, test B is 5.75002 and the standard deviation of the test A is 1.88546, test B is 1.67276. The reliability of the test is 0.9315 which means the reliability of test re-test method is high of S.G.M.S. in sentence matching. Likewise SSGS has different mean and standard deviation in sentences matching. The mean of test A is 6.07143 and test B is 6.40423 and the standard deviation of test A is 1.83823, test B 1.77006 and reliability is 0.9400. The reliability is 0.9400 which means the reliability of test re-test method is high in sentence matching. At last, S.M.S.S. has also the different mean and standard deviation. The mean of test A is 6.5000, Test B is 7.3571 and the standard deviation of test A is 1.9611, test B is 1.7478 in sentence matching. The reliability of test in sentence matching is 0.9551 which means the reliability of test retest method is high in sentence matching.

### **Conclusion:**

Students of four schools were used to obtain the information. Among them the students of N.Y.S.S have the highest reliability i.e. .9780 in sentence matching. It means the reliability of test re-test method in assessing writing skill is high.

### **3.3.3 Re-arrange**

Four schools were given the same test twice to find out the reliability of test re-test method in re-arrange of the students. the description is presented below.

**Table No. 15: Mean, S.D. and Reliability of Four Schools**

N.Y.S.S	Mean: Test A : 5.07146 Test B: 5.46425	St.dev: Test A : 2.19134 Test B: 2.17913	Reliability: 0.9791
S.G.M.S	Mean: Test A: 4.25002 Test B : 4.17853	St.dev:Test A : 0.99512 Test B: 0.97284	Reliability: 0.8131
S.S.GS.	Mean: Test A: 3.60712 Test B: 4.07123	St.dev: Test A : 1.09500 Test B: 1.12410	Reliability: 0.9554
S.M.S.S	Mean: Test A: 4.21423 Test B: 4.28574	St.dev: Test A: 0.82540 Test B: 0.75220	Reliability: 0.8418

The above table shows that N.Y.S.S has the different mean and standard deviation in re-arrange. The researcher found that the mean of test A is 5.07146, Test B is 3.85715 and the standard deviation of test A is 2.19134, test B is 2.17913 and the reliability of the test is 0.9791 which means the reliability of test re-test method is high in re-arrange of N.Y.S.S. Likewise, S.G.M.S has also the different mean and standard deviation in re-arrange. The re-researcher has found that the mean of Test A is 4.25002 and test B is 4.17853. The standard deviation of test A is 0.99512 and test B is 0.97284. The reliability is .8131 which means the reliability of test re-test method is high in re-arrange. S.S.G.S has different mean and standard deviation. The mean of the test A is 3.60 712 ,test B is 4.07148 in arrange and the standard deviation of test A is 1.09500 and test B is 1.12410. The reliability in re-arrange of S.S.G.S is 0.9554 which means the reliability of test re-test method is high. At last, S.M.S.S has also the different mean and standard deviation in re-arrange. The mean of test A is 4.21423, Test B is 4.28574 and the standard deviation of test A is 0.82540 and test B .75220 in re-arrange. The reliability in re-arrange. of S.M.S.S is .8418 which means the reliability of test-retest method is high.

## Conclusion:

The students of four schools were used to obtain the information. Among them, the students of school N.Y.S.S have the highest reliability i.e. 9791 in re-arrange. It means the reliability of test-retest method in assessing writing skill is high.

### 3.3.4 Guided Writing

Four schools were given the same test twice to find out the reliability of test re-test method in guided writing of the students. The description is presented below.

**Table No. 16: Mean, S.D. and Reliability of Four Schools**

N.Y.S.S	Mean: Test A : 5.71423 Test B: 5.82140	St.dev: Test A : 1.63740 Test B: 1.46243	Reliability: 0.9310
S.G.M.S	Mean: Test A: 5.50000 Test B : 5.96420	St.dev:Test A : 1.67563 Test B: 01.48663	Reliability: 0.9456
S.S.G.S.	Mean: Test A: 6.25000 Test B: 6.67850	St.dev: Test A : 1.85793 Test B: 1.50133	Reliability: 0.9165
S.M.S.S	Mean: Test A: 5.00000 Test B: 5.78570	St.dev: Test A: 1.93153 Test B: 1.81562	Reliability: 0.8389

The above table shows that the N.Y.S.S has the different mean and standard deviation .The mean of Test A is 5.71423, test B is 5.82140 and the standard deviation of test A is 1.63720 and test B is 1.46243. The reliability of N.Y.S.S in guided writing is 0.9310 which means the reliability of test-retest method is high in guided writing. S.G.M.S has also has the different mean and standard deviation in guided writing. The mean of the test A is 5.5000, test B is 5.96420 and the standard deviation of the test A is 1. 6756, test B is 1.48663 in guided writing. The reliability of the test in guided writing of S.G.M.S is .9456 which means the

reliability of test re-test method is high in guided writing in guided writing. Likewise, S.S.G.S has also the different mean and standard deviation in guided writing. The mean of test A is 6.2500, test B is 6.67850 and standard deviation of test A is 1.85793, test B is 1.50133 in guided writing of S.S.G.S. The reliability of tests is 0.9165 which means the reliability of test-retest method is high. At last, S.M.S.S has also the different mean and standard deviation. The mean of test A is 5.00000, test B is 5.78570 and the standard deviation of test A is 1.93153 and test B is 1.81562 in guided writing of S.M .S.S. The reliability of test is .8389 which means the reliability of test re-test method is high in guided writing of S.M.S.S.

### **Conclusion:**

Students of four schools were used to obtain the information. Among them the students of S.G.M.S have the highest reliability i.e. 0.9456 in guided writing. It means the reliability of test re-test method in assessing writing test is high.

### **3.3.5 Free Writing**

Four schools were given the same test twice to find out the reliability of test re-test method in free writing of the students. The description is presented below.

**Table No. 17: Mean, S.D. and Reliability of Four School**

N.U.S.S	Mean: Test A : 3.85710 Test B: 3.85710	St.dev: Test A : 0.94922 Test B: 1.08182	Reliability: 0.6818
S.G.M.S	Mean: Test A: 4.03572 Test B : 3.82140	St.dev:Test A : 0.94990 Test B: 0.77470	Reliability: 0.7871
S.S.GS.	Mean: Test A: 4.10710 Test B: 4.00000	St.dev: Test A : 0.81208 Test B: 0.83203	Reliability: 0.7014
S.S.M.S	Mean: Test A: 4.03573 Test B: 4.00000	St.dev: Test A: 0.94990 Test B: 0.87713	Reliability: 0.6083

The above table shows that N.Y.S.S has different mean and the standard deviation in free writing . The mean of test A is 3.85710 test B is 3.85710 and the standard deviation of test A is .94922 and test B is 1.08182. The reliability of test is .6818 which means the reliability of test re-test method is satisfactory. S.G.M.S has also the different mean and standard deviation in free writing. The mean of the test A is 4.03572, test B is 3.82140 and standard deviation of test A is .94990 and test B is 0.77470. The reliability is .7871 which means the reliability of test re-test method is moderate. Next, S.S.G.S also has the different mean and standard deviation. The mean of test A is 4.10710, test B is 4.00000 and deviation of test A is 0.81280 and test B is .8320 in free writing. The reliability in free writing of S.S.G.S is .7014 which means the reliability of test-retest method is moderate. At last, S.M.S.S has also the different mean and standard deviation in free writing. The mean of test A is 4.03573, test B is 4.00000 and standard deviation of test A is .94990 and test B is .87713 in free writing. The reliability in free writing of S.M.S.S is .6803 which means the reliability of test-retest method is moderate.

### **Conclusion**

Students of four schools were used to obtain the information. Among them, the students of S.G.M.S have the highest reliability i.e. .7817 in free-writing. It means the reliability of test-retest method in assessing to writing test is high .

### **3.4. Holistic Description of Reliability**

Fifty-six students of grade nine of four schools in Kathmandu district were used to obtain the information. Only subjective test of 50 marks was used twice in schools.

**Table 18: Holistic Description of Reliability (Male + Female)**

	Test A	Test B
Numbers of testees	56	56
F.M.	50	50
Mean	21.81253	21.75890
Standard deviation	8.96020	8.75033
Reliability	0.9441	

The above table shows that there were 56 students as a whole and the full mark was 50 in both tests. The mean of the test A is 21.81253 and test B is 21.75890. The standard deviation of test A is 8.96020 and test B is 8.75033. The reliability of the test is 0.9441 which means the reliability of test retest method is high. It means as a whole, all testees of nine grades and all schools have done well because the reliability is more than 0.7 i.e. 0.9441 which shows the tests are highly reliable.

### **Conclusion**

As a whole, the reliability of test-retest method in assessing writing skill is high. The tests are highly reliable.

## **CHAPTER – IV**

### **FINDINGS AND RECOMMENDATIONS**

#### **4.1 Findings**

The purpose of the study was to determine the reliability of test using test-retest method in assessing of writing skill of grade nine exercise book. In the process of research, the researcher herself selected the questionnaire based on the objectives of secondary level English book of grade nine. After the analysis and interpretation of the data obtained from the subjects, the researcher drew the following findings.

The main finding is the reliability of test-retest method in assessing writing skill is high.

##### **4.1.1 As a whole Finding**

The reliability of test-retest method in assessing writing skill is high

##### **4.1.2 School-Wise Findings**

- i. The reliability of the tests (test-retest method in writing skill) of NYSS is the highest.
- ii. The reliability of the tests of S.G.M.S is the second highest.
- iii. The reliability of tests of S.M.S.S is the third highest.
- iv. The reliability of tests of S.S.G.S. is the last.

##### **4.1.3 Sex-Wise Findings**

- i. The reliability of the tests (sex-wise) of S.G.M.S is the highest.
- ii. The reliability of the tests (Sex-wise) of S.M.S.S in the second highest.

- iii. The reliability of the tests (sex-wise) of S.S.G.S is the third highest.
- iv. The reliability of the tests (sex-wise) of N.Y.S.S is the last.

#### **4.1.4 Item-Wise Findings**

- i. The reliability of test –retest method in assessing writing tests of students of four schools in controlled writing is high.
- ii. The reliability of test-retest method in assessing writing tests of students of four schools in sentence matching is very high.
- iii. The reliability of test-retest method in assessing writing tests of students of four schools in guided writing is also high.
- iv. The reliability of test-retest method in assessing writing tests of students of four schools in free writing is high.

#### **4.2 Recommendations**

On the basis of the finding above the following recommendations have been made:

- i. Test-retest method should be used to ensure the reliability of test items in assessing writing skill.
- ii. Test-retest method should be used in schools to ensure the reliability of test items.
- iii. In sex-wise also, test-retest method should be used to find out the reliability assessing writing skill. The research findings show the female students need more practice in every exercise.
- iv. Test-retest method should be used in different test items to ensure the reliability in assessing the writing skill.

## REFERENCES

- Bachman, L.F. (1989). *Fundamental Consideration in Language Testing*. London: OUP.
- Baral , R.K(1999). **Measurement of Readability of a Course in General English:** An unpublished M. Ed. Theses T.U. Kathmandu.
- Bhattarai, G.R. (2005). *A Sample of Research Proposal. Young Voices in ELT*, Vol. 4.
- Bhattra, I.B.M. (2002). **A Comparative Study of the Writing Proficiency of Bachelors Level Students.** An unpublished M. Ed Thesis, T.U Kathmandu.
- Brad Field, J.M, Moredock, H.S. (1952). *Measurement and Evaluation in Education*. New York: The Mac Millan Company.
- Brown, G. and Yule, G. (1993). *Discourse Analysis*. New York: Cambridge University Press.
- Brown, J. and Rodgers S. (2002). *Doing Second Language Research*. New York: Cambridge University Press.
- Cook, G. (1995). *Discourse*. New York: Oxford University Press.
- Ebel, R.L. (1965). *Measuring Educational Achievement*. Prentice Hall. Engle Wood Cliffs New Jersey, USA.
- Ghimire, L.M. (2002). **Pronunciation Proficiency of the Students of Lower Secondary Level.** An Unpublished M. Ed Thesis T.U. Kathmandu.
- Hancock, Charles R. (1994). *Teaching, Testing and Assessment*. USA.: Lincoln Wood, Illinois.
- Harmer, J. (1983). *The Practice of English Language Teaching*. Longman Group UK Ltd.

- Harrison, A. (1983). *Language Testing Handbook*. London: ELTS, Mac Millan Publishers.
- Heaton, J.B. (1988). *Writing English Language Tests*. London: Longman.
- Hughes, A. (1989). *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- Khanal, U.R. (1997). **A Study on the Effectiveness of the Close Test Over Conventional Objectives Test in Testing Reading Comprehension in English**. An unpublished M. Ed. Theses, T.U. Kathmandu.
- Kshetree, G.L. (2000). **Washback Effect of the SLC Examination: An Unpublished M. Ed. Thesis, T.U. Kathmandu.**
- Kumar, R. (1976). *Research Methodology*. London: Sage Publication.
- Narayanswami, V.R. (1994). *Strengthen Your Writing*. Madras: Orient Longman Ltd.
- Nunan, D. (1992). *Research Methodology in Language Learning*. Cambridge: Cambridge University Press.
- Otto, J. (1996). *A Dictionary of Linguistics and Phonetics*. Blackwell.
- Rivers, Willga, M. (1968). *Teaching Foreign Language Skills*. Chicago: University of Chicago Press.
- Robin, J.C. (1999). *Longman Dictionary of Language Teaching and Applied Linguistics*. London: Longman.
- Tribble, C. (1996). *Writing*. Oxford: Oxford University Press.
- Ur, P. (1992). *A Course in Language Teaching*. Practice and Theory, Cambridge University Press.

## APPENDICES

### Appendix-1

#### Test Items

Name: F.M.: 50  
School: P.M.: 20  
Subject: Time: 1 hr.

Attempt all questions.

1. What is your partner planning to do during the holidays? Write down the plans using v + to inf. ..., am/is/are + going to , will + v. or will be + v. ing.

You can start like this: (Laxmi is my best friend. She is going to .) 10

2. Make sensible sentences by re-arranging the jumbled words. Add a capital letter and a full stop. 10

a. are/to eat/ students/not allowed/ in the class

b. allowed/tourists/here/photos/are/to take/not

c. here/strictly/is prohibited/smoking

d. not/be here/rubbish/to/allowed/thrown/is

e. during/cheating/prohibited/examination/is strictly/the

3. Imagine that you are staying in a town from your home. Write a postcard to your friend describing what you do every day. Use the simple present tense and the clues in the box. 10

- stay with	- school finish at 5.00 am
- eat lunch at school	- get up at 7.00 am
- return and do home	- school starts at 9.00 am
- go to marries or write letters	- miss home
- have breakfast at 8am	- very much

4. Make possible five sentences from the following table.

Arice cooker	in order to	drive a bus
		be a lecture
A washing mach	to	be an engineer
ine		drive a car
she got a license	is meant	washing clothes
	for	cooking rice

5. Imagine that one of your friends had a motorcycle accident and died in hospital. Write a letter of condolence to his parents using the clues in the box in about 100-150 words. 10

Sukhodeo, your best friend, studied together for 8 years, very helpful, always cracked jokes, popular with both teachers and students, miss him very much, never forget

**Appendix-3**  
**Results of Test-Grade 9**  
**(Holistic)**

S.N.	Name of the testees	Name of school	Sex	Test A	Test B
1	Suman Thapa	N.Y.S.S.	M	13	14
2	Dhiraj Thakur	N.Y.S.S.	M	8	22.5
3	Sudhir Malla	N.Y.S.S.	M	1.7	18
4	Suv Raj Khatri	N.Y.S.S.	M	22	18
5	Moti Gurung	N.Y.S.S.	M	11.5	21.5
6	Suman Rai	N.Y.S.S.	M	10	8
7	Raj Kumar R. Magar	N.Y.S.S.	M	24	26
8	Rabina Shrestha	N.Y.S.S.	F	17	18
9	Nabina Shakya	N.Y.S.S.	F	9	12
10	Susmita Shrestaha	N.Y.S.S.	F	23.5	23.5
11	Rabita Gurung	N.Y.S.S.	F	15	14
12	Hasina Banu	N.Y.S.S.	F	36	30
13	Pramila Shrestha	N.Y.S.S.	F	16.5	23.5
14	Chandika Gurung	N.Y.S.S.	F	19	19
15	Sabitri Ghale	S.G.M.S.	F	30	32
16	Sarita Bhandari	S.G.M.S.	F	22	22
17	Binita Shrestha	S.G.M.S.	F	27	28
18	Rajan Maharjan	S.G.M.S.	M	32.5	33
19	Kos Lal Gurung	S.G.M.S.	M	31	31
20	Shital Sharma	S.G.M.S.	M	23.5	24
21	Amita Sharma	S.G.M.S.	F	22	20
22	Manisha Ghimire	S.G.M.S.	F	35	36
23	Susmita Rai	S.G.M.S.	F	34	33
24	Sharmila Pantha	S.G.M.S.	F	28	29
25	Bishnu Bhandari	S.G.M.S.	M	14	15
26	Madhu S. Giri	S.G.M.S.	M	9	10

27	Basanta Shrestha	S.G.M.S.	M	22.5	22
28	Basanta Sharma	S.G.M.S.	M	25	24
29	Arati Rai	S.S.G.S.	F	32	32
30	Maya Tamang	S.S.G.S.	F	19	18
31	Sharada Pandey	S.S.G.S.	F	24	25
32	Rita Thapa	S.S.G.S.	F	20	20
33	Pratima Pariyar	S.S.G.S.	F	15	14
34	Dil Bahadur Gurung	S.S.G.S.	M	13	12
35	Rahul K.C.	S.S.G.S.	M	16	17
36	Sabin Shahi	S.S.G.S.	M	15	15
37	Rabita Shahi	S.S.G.S.	F	32.5	30.5
38	Ramita Karki	S.S.G.S.	F	13	10.5
39	Sapana Thapa	S.S.G.S.	F	23	23
40	Prakash Shrestha	S.S.G.S.	M	24	22
41	Krishna Khadka	S.S.G.S.	M	27	22.5
42	Sujan Pudasaini	S.S.G.S.	M	20	20
43	Hikmat Gurung	S.M.S.S.	M	9.65	14.5
44	Rupa Gurung	S.M.S.S.	F	30	23
45	Saurav Shrestha	S.M.S.S.	M	37	38
46	Dev Raj Gurung	S.M.S.S.	M	26.5	27.5
47	Padam Kumari Thapa	S.M.S.S.	M	29	27
48	Ratna Tamang	S.M.S.S.	M	12	14
49	Niraj Kattel	S.M.S.S.	M	17	12
50	Kalpana Gurung	S.M.S.S.	F	22.5	15.5
51	Chandra Tamang	S.M.S.S.	F	28	25
52	Dil Bahadur Gurung	S.M.S.S.	M	20	21
53	Kalpana Lama	S.M.S.S.	F	12	13
54	Ang Babu Sherpa	S.M.S.S.	M	14.5	17.5
55	Nirmala K.C.	S.M.S.S.	F	32	33
56	Sabina Shrestha	S.M.S.S.	F	15	15

