CHAPTER THREE

METHODOLOGY

In the previous chapter, the researcher explained the concept of Critical Thinking and English Reading Comprehension as well as their nature, role and importance in learning with reference to English as a Second language (L2). In this chapter, the research method implied to explore the research problem, namely to determine the probability of the relationship between Critical Thinking skills and English Reading Comprehension ability in second language learners at undergraduate level with science students in focus. It is pertinent to get familiarized with the theoretical background considering research framework, quantitative research, the construction, measuring, testing and administering of questionnaires to collect data, analyzing data by processes of descriptive and inferential statistical procedures and the ethical procedures that are required, when conducting research of this kind. It covers methodology employed to test the research questions. The description of the empirical research as well as literature relevant to research tools for data collection adopted for the study is specified. The main purpose of this chapter is to specify a comprehensive description of instruments selected to measure Critical Thinking skills and English Reading Comprehension ability were used. It also emphasis aims and objectives of the study, Empirical Research, Population and Sample, Instruments, Design and Procedure of data collection, techniques Data Analysis used and Ethical considerations.

3.1. Aims and Objectives of the Study

As suggested in the opening chapter, the primary objective of this study is to determine the probability of the relationship between Critical Thinking skills and English Reading Comprehension in second language learners at undergraduate level. So, the overall aims and objectives are stated as follows:

- To discover if there is any meaningful relationship between Critical Thinking skills (Inference, Recognition of Assumptions, Deduction, Interpretation, Evaluation of Arguments) and English Reading Comprehension ability in the second language learners at undergraduate level.
• To determine if there is any impact of Critical Thinking skills on English Reading Comprehension ability in the second language learners at undergraduate level.

• To see if gender moderate in relationship Critical Thinking skills and English Reading Comprehension ability in the second language learners at undergraduate level.

• To determine if there is gender and department wise significant difference in Critical Thinking skills and English Reading Comprehension ability in the second language learners at undergraduate level.

Keeping in mind the above objectives, the design of the research study is systematically given hereunder:

3.2. Empirical Research

Research is about (i) understanding the world/phenomena and (ii) how one sees the world/phenomena (Cohen et in Maree & Van der Westhuizen, 2007). These observations implied that the researcher should:

a) Determine whether phenomena will be understood from an external, objective (realist) view, or through the words created by individuals (nominalist view) Cohen et al in Manion & Van der Westhuizen, 2007).

b) Determine whether knowledge gathered can be viewed as objective (positivist stance) or interpretive (anti-positivist stance).

c) Adopt a stance regarding human nature, that is whether humans respond mechanically to their environment, where they initiate their own actions or whether they fall somewhere between the two mentioned extremes.

In the context of this study, the researcher wanted to understand the two phenomena: (i) Critical Thinking; and (ii) English Reading Comprehension, from an external and objective stance. The researcher thus operated from a positivistic paradigm and envisaged objectively obtaining information to achieve the objectives of the study.

The following sections will highlight the choice of the research method, research design, data collection instruments and the tools, and procedure for data analysis as related to the positivistic research paradigm.
3.3. **Research Method**

Choosing research methodology had to be relevant to the data needs for predetermined objectives of the study with all its dimensions and scope; considered carefully the needs of the present research, and a view taken as to what kind of data was needed to be gathered. In support of the positivistic paradigm, the researcher decided to make use of data in the form of numbers (quantitative) and to be an external observer, whose intent was to establish and confirm relationships and to develop generalizations on the basis of analysis that would contribute to a theory. Keeping in view the research problem, a quantitative approach was chosen for the study. The researcher undertook the responsibility to find out and establish how and to what extent Critical Thinking skills and Reading Comprehension ability are correlated, in order to provide opportunities for the development of Critical Thinking skills and English Reading Comprehension ability. By exploring of appropriate methods for evolving better approaches to teaching and learning second language and acquiring proficiency in this regard as few authors are mentioned in earlier studies.

Quantitative research can be defined as a process that is systematic and objective in using numerical data from only a selected sub-group from the relative universe (or population) to generalize the findings applicable to the universe that is being studied. According to Eldabi, Irani, Paul and love (2002), quantitative research is “a logical linear structure, in which hypotheses take the form of expectations about likely causal links between the constituent concepts identified in the hypotheses”. It is used to answer questions about relationship among measured variables with the purpose of explaining, predicting and controlling phenomena (Leedy & Ormrod, 2005).

As it was the intention of the researcher to examine relationship between level of Critical Thinking skills and English Reading Comprehension in science students, quantitative research was chosen. Quantitative research relies on the measurement and analysis of statistical data that determines the relationship between one set of data to another. It is important for the researcher to establish whether quantitative research could be regarded as a valid way to conduct research that set out to determine the possibility of the correlation between critical thinking and English Reading Comprehension in science students.
3.4 Research Design

For the purpose of the quantitative study, as in this case required, a non-experimental, descriptive survey research design was chosen. Descriptive research provides a summary of an existing phenomenon by using numbers to characterize individuals or a group. Leedy and Ormrod (2005) indicate that descriptive research involves either identifying the characteristics of an observed phenomenon or exploring possible correlations between two or among more phenomena. Descriptive research assesses the nature of existing conditions and aims to characterize something as it is (McMillan & Schumacher, 2006). Survey research, refers to almost any form of descriptive, quantitative research (Gay and Airasian in Leedy and Ormrod, 2005). Survey research involves acquiring information about one or more groups of people about their characteristics, opinions, attitudes or previous experiences (McMillan and Schumacher, 2006). The ultimate goal is to learn about a large population by surveying a sample of the population. Inferences about a particular population are drawn from the responses obtained from the sample. As it was the intention of the researcher to explore the opinions, attitudes and experiences of the research participants with regard to the possibility for the correlation between Critical Thinking skills and Reading Comprehension skills in science students for whom English was a second language (ESL), survey research was deemed suitable and hence adopted.

Survey research poses a series of questions to the willing participants, summarizes their responses with percentages, frequency counts, or by more sophisticated indexes, and then draws inferences about a particular population from the responses of the sample (Leedy and Ormrod, 2005). In the context of the current study, the participants’ responses were reported as frequencies. Means and percentages were worked out before inferences were drawn.

3.5. Research Instruments:

Due to the nature of the present study, two different instruments were chosen to aid the gathering of required data pertaining to critical thinking and reading comprehension.
3.5.1 Critical Thinking (Questionnaire)

3.5.1.1 The Watson Glaser Critical Thinking Appraisal (WGCTA) Form A

To evaluate students’ Critical Thinking ability, the “Watson Glaser Critical Thinking Appraisal” (CTA) was selected and employed. It is mentioned in the test’s booklet, in the early part of 20th century, Watson became interested in to find out how it might be possible to measure fair mindedness. In 1937, Glaser modified and revised Watson’s idea for use in ‘An Experiment in the Development of Critical Thinking’. Since then, the Watson Glaser Critical Thinking Appraisal (WGCTA) has undergone many modifications and developments, including extensive internationalization, to take account of the now global market for such type of testing. This test defines Critical Thinking (CT) as a composite of Attitude, Skill and Knowledge (Suliman and Halabi, 2007).

This test is an intellectually challenging-tool for addressing critical thinker’s cognitive ability. According to Watson and Glaser (1980), the five subscales of the WGCTA are “each designed to tap a somewhat different aspect” of Critical Thinking (CT) skills (cited in Bernard et al, 2008). There are four standardized versions and one experimental edition of the Watson Glaser measure. The latest, the WGCTA, was last revised in 1994. The number of items varies across versions but the subscales and their descriptions have remained consistent over time. Practitioners now acknowledge WGCTA as the most widely used psychometric test of Critical Thinking (CT) ability. Its international reputation, and its continued and widespread use, means that WGCTA is very robust in its standardization, reliability and validity, making it one of the most trusted testing tools available (Watson-Glaser Critical Thinking booklet, 2002).

The version used in this study is the form A of the WGCTA, which was standardized and applied in 1980 by the Psychological Corporation in the United States of America. The test was fully explained to the test-participants before its administration. This test comprises of 80 items and consists of 5 sub-tests as follows (Watson – Glaser Critical Thinking booklet, 2002):

Sub-test 1: Inference: Discriminating among degrees of truth or falsity of inference drawn from given data. It consists of 16 Likert scale questions with
alternatives ranging from: T as true, PT as probably true, ID as insufficient data, PF as probably false and F as false.

Sub-test 2: Recognizing Unstated Assumptions: Recognizing unstated assumptions or presuppositions in given statements of assertions. This sub-test consists of 16 two-scaled questions with alternatives ranging from MADE and NOT MADE.

Sub-test 3: Deduction: Determining whether certain conclusions necessarily follow from information in given statements or premises. It consisted of another 16 two-scaled questions with alternatives ranging from FOLLOW and does not FOLLOW.

Sub-test 4: Interpretation: Weighing evidence and deciding if generalizations or conclusions based on the given data are warranted. This part consisted of 17 two–scaled questions with the alternatives ranging FOLLOW and does not FOLLOW.

Subtest 5: Evaluation of Arguments: Distinguishing between arguments that are strong and relevant and those that are weak or irrelevant to a particular question at issue. It consisted of 15 two-scaled questions with the choice of STRONG and WEAK.

The test items include problems, statement, arguments, and interpretation of data similar to those that are encountered on a daily basis in the classroom, at work and in newspaper and magazine articles.

The instrument has two distinct sets of items i.e. “Neutral set” items deal with the weather, scientific facts or experiments, and other subject matter about which people generally do not have strong feelings or prejudices. “Controversial set” items denote to logical structure to neutral items, refer to political, economic, and social issues that frequently provoke very strong feelings. As has often been shown, strong attitudes, opinions, and biases affect the ability of some people to think critically (Watson and Glaser, 1980).

3.5.2. English Reading Comprehension Test

In order to evaluate students’ Reading Comprehension, the researcher conducted a Reading Comprehension test prepared by English Language Teaching
(ELT) experts specific to reading scientific texts of different kinds. This instrument contains 40 different items in all on reading comprehension. Difficult vocabulary items were avoided, and great care was taken to ensure simplicity and comprehensibility of the items. Special attention was given to ensure that the phrases in passages contain high frequency words most likely known to all participants. Two TEFL experts also checked the draft instrument with all the items, and their comments were considered. This test included 40 multiple-choice items measuring the participants’ comprehending. The allotted time for the test was 40 minutes. The test consisted of four Reading Comprehension passages; each passage included ten multiple-choice tests and participants were asked to select the correct option out of four choices (i.e., correct option, three distracters. Participates were given complete and clear instructions as to what to do.

3.6. Validity and Reliability of quantitative research for the present study

3.6.1. Validity

Validity is a measure or instrument that is valid if it measures what it is supposed to measure. To Leedy and Ormrod (2005), validity means accuracy, meaningfulness and credibility of a study as a whole. McMillan and Schumacher (2006) referred to validity as the truthfulness of findings and conclusions. For the results of a study to be trustworthy, the research must have a high degree of both internal as well as external validity. It was necessary for the researcher to determine whether the present research conducted by his/her complied with the following validity criteria identified by Leedy and Ormrod (2005) and McMillan and Schumacher (2006).

A high degree of internal validity of the research design means that there was sufficient control of the variables and that the design allowed the researcher to draw accurate conclusions (Leedy & Ormrod, 2005). Researchers often use triangulation, which is the cross-validation measure among data sources, data collection strategies, time periods and theoretical schemes to increase validity (McMillan & Schumacher, 2006). In this regard, Maree and Van der Westhuizen (2007) refer to crystallization. Lincoln and Guba (in Maree & Van der Westhuizen, 2007) refer to crystallization as “attending to voices that differ from your own view enable you to study multiple constructed realities”. The researcher thus gathered data from science students to find
regularities in the data and to check whether the same students kept recurring regularities in order to strengthen the conclusion made. In this way, the study could be regarded as compliant with criteria for internal validity.

External validity, on the other hand, refers to the degree to which results can be generalized to the entire population and to other contexts (Leedy & Ormrod, 2005; McMillan and Schumacher, 2006; Maree & Petersen, 2007). Although the sample used in the study was small and geographically bound to small area, and results could not be generalized to be applicable to the population, the external validity of the study was enhanced by the fact that the study was conducted in a real life setting to construct validity of data concerning the efficiency of using a particular data collection instrument for the research purpose (McMillan and Schumacher, 2006).

Statistical conclusion validity refers to the appropriate use of statistical tests and procedures to determine whether purported relationships are a reflection of actual relationships (McMillan & Schumacher, 2006). The construct validity of the Critical Thinking Test has already been calculated by many researchers according to its manual, which has been reported here as well. The researcher checked also the validity of the two instruments as ascertained through panel of experts.

3.6.2. Reliability

Reliability is the degree of consistency that the instrument or procedure demonstrates. The reliability of a research instrument concerns the extent to which the instrument yields the same results on repeated trials. In other words, a test is reliable to the extent that it measures whatever it is measuring consistently, i.e., in case they are repeated, almost the same results are obtained. In tests that have a high coefficient of reliability, errors of measurements have been reduced to a minimum.

The reliability or stability or stability of a test is usually expressed as a correlation coefficient. There are several types of reliability:

- Stability over time (test-retest)
- Stability over items sample (equivalent or parallel forms)
- Stability of items (internal consistency) which is of two types:
a. Split halves

b. Formula

- Stability over scorers (inter scorer reliability)
- Standard error of measurement

Reliability is a necessary but not a sufficient condition for validity. That is, a test must be reliable for it to be valid, but a test can reliable and still not be valid.

In the present study, the reliability of all tests was calculated by the use of the formula.

The reliability of the Watson-Glaser test has been determined in three ways. (i) Estimates of the test’s internal consistency (ii) stability of the test scores over time, and (iii) the correlation between scores on alternate forms. Internal consistency has been measured using split-half reliability coefficients using the Spearman–Brown Formula. Testing stability over time, by administrating the test to the same group with an internal difference, the correlation between scores has indicated an acceptable level at 0.73. Regarding validity, the Watson–Glaser test enjoys all areas of fact, content, criterion and constructs validity (Watson–Glaser Critical Thinking booklet, 2002).

3.7. Ethical Consideration

Prior to conducting the study, ethics of research, especially when human participants are involved, were considered and followed. Subjects’ anonymity, the right to withdraw from the study at any time, and the offer to access the findings of the study if interested, were taken into account. Therefore it is important that a researcher protects the rights of the participants of research study and of the departments in which the study was conducted. The research process may influence individuals or groups of people directly or indirectly. Therefore, it is important that the researcher should consider the ethical considerations while designing or structuring the research process. Leedy and Ormrod (2005), McMillan and Schumacher (2006) and Van Deventer (2007) identified ethical principles that researchers have to observe, particularly where human participants are involved. In
the context of this study, the researcher complied with ethical principles in the following ways:

- Permission was obtained from the Heads of different Departments of Science and the teachers of those science classes, which were concerned in this study.

- The purpose of the research was explained properly to all participants. Participants were provided with description of what their participation would involve, as well as statement including that participation was voluntary and could be terminated at any time.

- Participants were assured that responses would be treated confidentially as only the researcher, the study leader, would possess and store them.

- Participants were informed that the results would be made known to them after the research.

3.8. Participants of the current study

Participants in this study were a total number of 208 under-graduates students from different science departments at Panjab University, Chandigarh, India. The participants who took part in this study were both male and female of almost the same age group. Gender based differences are considered as a possible moderating variable and some differences have been observed in terms of background knowledge of topics of more interest to one gender than the other (Bugel and Buunk, 1996).

There were about 350 students enrolled in science courses at undergraduate level at Panjab University, Chandigarh. The researcher selected 220 students randomly, who volunteered to take part in the study. The participants were explained the purpose of the study and could exercise their option to take part in it or not. The researcher expressed to participants how much their correct and honest responses to the questionnaires would be vital for the result of the research so they were encouraged to ask any questions to eliminate any ambiguity. Thus the subjects cooperated with extensive care and they were rewarded for their efficient participation. From 220 students, 12 students failed to complete all the requirements for the study that includes completing a questionnaire of reading comprehension skills and the test of critical thinking skills. Therefore, their data were removed from the
study and we were left with final 208 students. Thus the sample response rate was 94.5%.

### 3.9. Demographic Information

Table No. 1 below contains data in respect of the number of Female and Male learners that participated in the present research from different departments of science.

**Table No. 1**

**Number of total (Female and Male) Second Language learners that participated in Research**

<table>
<thead>
<tr>
<th>Department</th>
<th>Females (%)</th>
<th>Males (%)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anthropology</td>
<td>5 (50%)</td>
<td>5 (50%)</td>
<td>10 (100.0%)</td>
</tr>
<tr>
<td>Biochemist</td>
<td>16</td>
<td>12</td>
<td>28 (100.0%)</td>
</tr>
<tr>
<td>Biophysics</td>
<td>10</td>
<td>3</td>
<td>13 (100.0%)</td>
</tr>
<tr>
<td>Biotech</td>
<td>5</td>
<td>4</td>
<td>9 (100.0%)</td>
</tr>
<tr>
<td>Botany</td>
<td>14</td>
<td>2</td>
<td>16 (100.0%)</td>
</tr>
<tr>
<td>Chemistry</td>
<td>15</td>
<td>7</td>
<td>22 (100.0%)</td>
</tr>
<tr>
<td>Electro</td>
<td>11</td>
<td>6</td>
<td>17 (100.0%)</td>
</tr>
<tr>
<td>Geology</td>
<td>4</td>
<td>7</td>
<td>11 (100.0%)</td>
</tr>
<tr>
<td>Mathematics</td>
<td>13</td>
<td>12</td>
<td>25 (100.0%)</td>
</tr>
<tr>
<td>Microbiology</td>
<td>10</td>
<td>12</td>
<td>22 (100.0%)</td>
</tr>
<tr>
<td>Physics</td>
<td>10</td>
<td>16</td>
<td>26 (100.0%)</td>
</tr>
<tr>
<td>Zoology</td>
<td>7</td>
<td>2</td>
<td>9 (100.0%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>120 (57.7%)</strong></td>
<td><strong>88 (42.3%)</strong></td>
<td><strong>208 (100.0%)</strong></td>
</tr>
</tbody>
</table>
3.10. Data Collection Procedure

This study comprised two phases. In first phase administration of the WGCTA to collect data on Critical thinking, and in the second phase Reading Comprehension test was done. Both the tests were self-administered. Before data collection began, consent from the various science departments of Panjab University was acquired to conduct the study. In pre-administration tryouts of the WGCTA test, it was acknowledged that the recommended time of 60 minutes was a sufficient amount of working time for the large majority of examinees to finish the test. The procedure of data collection began from the month of November 2014 till the end of February 2015.

3.10.1 Phase I: Administration of WGCTA

The first phase of data collection included the administration of the Watson – Glaser Critical Thinking Appraisal (WGCTA) that was given to the participants. This was not a concurrent test for all the participants since it was not possible for the researcher to gather all the sampled 208 participants at one place at the same time. Therefore, the test was given in separate sessions. The students of 12 departments of science took part for the study, which included departments of Physics, Chemistry, Anthropology, Biotechnology, Biochemistry, Biophysics, Microbiology, Mathematics, Botany, Zoology, Geology and Electronics. The total number of participants was 120 Female and 88 Male science students studying in the first year of their study.

The researcher explained the procedure to the teachers fully and completely whose services were requisitioned to supervise and guide students so that they could in turn help their own students to understand what they were supposed to do. To make sure that the test was employed properly, the researcher was present when the teachers explained the procedure for their students and answered in case there were any questions or doubts. Due to the time limitation of each class, which was just one hour, the researcher explains the instructions one day before administering the test for each class. Participants were given 1 hour to complete the test, which is the standard time. The gathering of the WGCTA from all classes covering all the sample-participants took about 2 months for completion. No one took the test home, or left the class midway while taking the test. The participants cooperated fully.
It was illuminated that they had to choose the closest option, which described them and their feelings or reactions best. Given the intimate and intrusive nature of multitude of questions in most psychological test including this one, the participants were reassured that the results were, obviously, to be held in confidence and not revealed under any circumstances to the third party or used for any purposes other than research. It was explained that the results of this questionnaire had to be compared with other ones, which had to come later in the semester, and hence they were required to mark their identity by providing either their real names or pseudonyms. It was clarified that using nicknames was acceptable as long as they chose a name were sure would be unique and made sure they would not forget the name for the coming test sessions.

3.10.2 Phase II: Administration of the Reading Comprehension Test

Two weeks after the first test session, the prepared Reading Comprehension Test was administration to the classes of different departments representing the population targeted by the study. Ideally, it was reasonable to allow a few-month interval between the two sessions of Critical Thinking Test and Reading Comprehension Test administrations, so that the sample was minimally influenced by its answers on the written version. However, due to time restrictions before the end of the semester and virtually impossible access to the participants after the current term at the university the second questionnaire was given after two weeks. Furthermore, they were once more urged to use the identities they had previously adopted.

The second phase of the study, like the previous phase, was conducted on different days, one class at a time. The participants were given 40 minutes, which was the standard time limit to complete the test. The participants were given separate answer sheets to give their answers. No one left the test session midway and no one was allowed to take the test home. Explanations were given to the teachers as what to do and how to administer the test. The researcher was also present when the test was conducted. Naturally, they were further reminded that they were supposed to use that same identity as the last time – either their real names or pseudonyms. As stated earlier in this chapter, a number of the students from the initial sample did not show up for the second round of administration and the sample size shrank. It took one and half months to gather reading comprehension test results for the second phase. The
WGCTA answer sheets and Reading Comprehension test were objectively assessed and scores obtained by the participant-students. As a token of appreciation, all students who had contributed to two tests were distributed gifts from the researcher at the end of this session.

3.11. Design of the Study

Since the nature of the present study was aimed at correlation analysis, the design of the current research was ex-post facto research. In this design, the researcher appeared on the scene after all the events had occurred. In other words, the researcher had no control over the events. According to Hatch and Farhady (1981), in ex-post facto designs, what is important for the researcher is the degree of relationship between two variables and not the cause and effect relationship.

The variables were not manipulated, controlled, or modified as recommended by Mackey and Gass (2005). In the present study, the researcher had no control over whatever the students had learned and they answered the Critical Thinking tests based on their prior knowledge. The application of an ex-post facto design was vindicated on the ground that the distinction between the dependent and independent variables appeared to be arbitrary rather than any rule governed. Since the events or conditions had already occurred or existed, the researcher merely selected the relevant variables for an analysis of their relationship. Any relationship between the scores of the groups would not be related to any instructional programme. The researcher took the effect (or dependent variable) and examined that data retrospectively to establish causes, relationships or associations and their meanings. The researcher had no control over what had already happened to the subjects. There were no cause and effect concerns between variables, only the type and or degree of their relationship were under question. Ex-post facto design just sought to find answers to questions through an analysis of variable relationships. Hatch and Farhady (1981) support this by stating, “ex-post facto design is often used when the researcher does not have control over the selection and manipulation of the independent variables”.

Cohen, Manion and Morrison (2003) believe that it should not be concluded that ex-post facto studies are of little value; many important investigations in education and psychology are ex-post facto designs. Most of the time, there is no choice in the matter; a researcher cannot cause one group to become failure,
delinquent, etc. Research necessarily should rely on existing groups. Alternatively, the inability of ex-post facto design to incorporate the basic need for control makes them vulnerable from a scientific point of view and the possibility of their being misleading should be clearly acknowledged. Ex-post facto designs are probably better realized not as experiments with the greater certainty but more as surveys, useful as sources of hypotheses, to be tested by more conventional experimental means of a later date. This research is particularly suitable in social, educational and psychological contexts where the independent variable or variables lie outside the researcher’s control.

Spector (1981) identifies the advantages of ex post-facto research as follows:

- Ex-post facto research meets an important need of the researcher where more rigorous experimental approach is not possible.

- The method yields useful information concerning the nature of phenomenon what goes with what and under what conditions. In this way, ex post-facto research is a valuable exploratory tool.

- Improvements in statistical techniques and general methodology have made ex-post facto designs more defensible.

- In some ways and in certain situations, the method is more useful than the experimental methods, especially where the setting up of the latter would introduce a note of artificiality into research proceedings.

- Ex-post facto research is particularly appropriate when simple cause and effect relationships are being explored.

- The method can give a sense of direction and provide a fruitful source of hypotheses that can subsequently be tested by more rigorous experimental methods.

Among the limitations and weaknesses of ex-post facto design, Specter (1981) explained the following:

- There is problem of lack of control in that the researcher is unable to manipulate the independent variable or to randomize her/his subjects.
• One cannot know for certain whether the causative factor has been included or even identified.

• It may be that no single factor is the cause.

• A particular outcome may result from different causes on different occasions.

• When a relationship has been discovered, there is problem of deciding which the cause is and which is the effect; the possibility or reverse causation has to be considered.

• The relationship of two factors does not establish cause and effect.

• Classifying into dichotomous groups can be problematic.

• There is difficulty of interpretation and the danger of the ex-post facto assumption being made; that, believes that because X precedes O, X causes O.

• It often bases conclusions on too limited sample or number of occurrences.

• It frequently fails to single out the really significant factor or factors, and fails to recognize that events have multiple rather than single causes.

• As a method, it is regarded by some as too flexible.

• It lacks nullifiability and confirmation.

• The sample size might shrink massively with multiple matching.

Thus the researcher implied ex post facto research design as per objectives of the study to find the degree of relationship between the independent variable that is Critical Thinking English Reading Comprehension as dependent variable.

3.12. Data Analysis

To find the answers to the research questions the collected data was captured, coded, analyzed and interpreted using descriptive statistics. Descriptive statistic was used to organize and summarize data meaningfully in order to endorse an understanding of the data characteristics. The researcher conducted two sessions for
data collection relating to critical thinking and then reading comprehension variables. The raw scores were tabulated and data of each instrument i.e. items 1-80 on Critical Thinking Test and numerical ratings obtained from and 1-40 on Reading Comprehension Test were entered from response sheets to compute scores of critical thinking into five dimensions and reading comprehension scores. The data was fed in and analyzed by using Statistical Package for the Social Sciences (SPSS 22) for each of the 208 respondents along with their demographic data. The analysis was performed to test hypothesis by applying following tests:

To answer the first research question, Coefficient of Correlation inferential statistics was used to examine the relationship between learners’ critical thinking skills and reading comprehension ability use by computing Pearson correlation (r), to allow us to look at the two variables and evaluate the strength and direction of their relationship or association with each other (Dornyei, 2007). The strength of the relationship is expressed as a correlation coefficient between -1 and +1. Correlations with positive values close to 1.0 imply strong relationships whereas values close to 0.0 imply weak (or no) relationships. It is necessary to note that a correlational study is different from a cause-effect, as the latter, examines a causal relationship between two or more variables under study.

To answer the second research question, regression analysis was applied. Regression analysis helps to find best regression model without testing all possible regressions. The result gives the best fitting model with only statistically significant variables.

To answer the third and fourth research questions for differential statistics, Student T test was applied on mean score of the CTA and RC of each of the male and female groups of subjects to find significant difference. In other words, to investigate the possible significant difference of critical thinking score and vocabulary learning strategy use, between two groups of proficient and less proficient learners. As stated by Dornyei (2007), in order to find that the difference in any two sets of scores reflects any ‘real’ difference, a t-test is needed, to check whether the results are generalizable or the scores are likely to be an artifact of random variation.

One-way analysis of variance (ANOVA) is based on the mean comparisons to examine a ratio (F-distribution) of more than two groups in the independent variable,
which is applied to the test when there is one interval-scale dependent variable and one nominal-scale independent variable (Brown, 2001; Walsh, 1990).

To conclude, this chapter restated the purpose of this research and presented the research questions. The participants were chosen through a random sample of various science departments of Panjab University, Chandigarh, India. The selection of the 220 participant samples from the target population was discussed. In addition, the validity and reliability of the instruments was presented, including the tests of reliability of the revised instrument. The data collection procedures were also discussed in this chapter. The response rate of the respondents was 94.5%, as data of 208 was found usable for the study. Finally, the methods of data analysis for each of the research questions were presented, followed by a discussion of statistical power analysis. The details of the data analysis depicting the answers to the research questions would appear in full in the next chapter.