# CHAPTER - III

# CHAPTER – III
# METHODOLOGY AND PROCEDURES

## 3.0 INTRODUCTION

This Chapter – III is developed to present complete information on the plots made to experiment and the processes followed by the researcher during the entire research study conducted on the 'Development And Implementation Of Instructional Strategy Based On The Structure Of Observed Learning Outcomes (SOLO) Taxonomy For Mathematics Of Class–IX'. This chapter embraces the descriptions about the research method and design selected for this research study; the population and sample; sources and instruments used for the data collection; various threats and preventive measures as well as about the techniques used for the data analysis, which are elaborated in further sections and sub-sections. This Chapter-III is organized with following sections.

- ► Methodology
- ► Population and sample
- ► Tools and Techniques of Data collection
- ► Techniques of data analysis

## 3.1 METHODOLOGY

The Methodology, according to **Howell (2013)** is the general research strategy that outlines the way in which research is to be undertaken and among other things, identifies the methods to be used in it. The methods described in the methodology, define the means or modes of data collection or sometimes, how a specific result is to be calculated. Methodology does not define specific methods, even though much attention is given to the nature and kinds of processes to be followed in a particular procedure or to attain an objective. Here, several methods and processes conducted for this research study are constituted with this section of methodology and expounded step by step as below.

### 3.1.1 Research Method And Design Of The Study

This research study is the Experimental research study. The research design for a study is the 'Non-Equivalent Control Group Posttest Only', tabulated (table-3.1) below is one of the 'Quasi-Experimental-Designs'. Quasi-experimental designs are commonly employed in the evaluation of educational programs when random assignment is not possible or practical.

**Table–3.1:**
**Notation for 'Non-Equivalent Control Group Posttest Only' research design**

| Group | Design | Pre-Test | Treatment / Intervention | Post-Test |
|---|---|---|---|---|
| **Experimental Group** | **N X $O_1$** | No | Yes | Yes |
| **Control Group** | **N C $O_2$** | No | No | Yes |

Where,

**N** - Non-equivalent groups.

**X** – Experimentation - a group With Treatment of SOLO based Instructional Strategy

**C**– Conventional/Control-a group Without Treatment of SOLO based Instructional Strategy

**$O_1$ and $O_2$** - Posttest for X and C respectively.

According to **Green et al. (2006),** 'Non-Equivalent Control Group Posttest-Only' design involves an experiment in which a control group is utilized as well as a posttest. It is important to note however, that control group and experimental group used by the researchers in this design are groups that have not been created through random assignment.

The **Non-Equivalent Groups Design** is probably the most frequently used design in social research. It is structured like a pretest-posttest randomized experiment, but it lacks the key feature of the randomized designs -- random assignment. In the Non-Equivalent Groups Design, we most often use intact groups that we think are similar as the treatment and control groups. In education, we might pick two comparable classrooms or schools. In community-based research, we might use two similar communities. We try to select groups that are as similar as possible so we can fairly compare the treated one with the comparison one. But we can never be sure the groups are comparable. Or, put another way, it's unlikely that the two groups would be as similar as they would if we assigned them through a random lottery. Because it's often likely that the groups are not equivalent, this design was named the non-equivalent group design.

This design is also termed as "After-Only with Control Design" as mentioned by **Kothari & Garg (2014)**. In this design two groups or areas (test area and control area) are selected and the treatment introduced into the test area only. The dependent variable is then measured in both the areas at the same time at post level only.

In this design, instead of using random assignment of subjects to experimental and control group, uses the technique of matching. In this technique, the subjects are paired so that their scores on matching variable/s means the extraneous variable/s that the experimenter wants to control are as close as possible. One subject of each pair is randomly assigned to one group and the other to the second group. The groups are designated as experimental and control by random assignment (tossing a coin). The major limitation of the design is that it is very difficult to use matching as a method of controlling extraneous variables because in some situations it is not possible to locate a match and some subjects are excluded from the experiment. For this research study also, 'Matching' technique was employed which is elaborated ahead in a Section-3.2.

### 3.1.2 Variables For The Experimental Study

As the Experimental research design consists of a comparison about the effectiveness of particular intervention program or treatment with traditional (or conventional) program based on assumed dependent variable which is observed for the variations or the effects. And the independent variable is the variable that is assumed to be the cause of the effect.

For the present research study, the independent as well the dependent variables are determined in a way as:

- **Independent Variable:** The SOLO Taxonomy based Instructional Strategy developed by the researcher

- **Dependent Variables:** The (i) Achievements as well as the (ii) Reactions of the group studied through the SOLO based developed instructional strategy.

### 3.1.3 Threats And Validity For The Experimental Study

The present experimental research study was carried in the actual context, so that all the variables which are generally functioning in the teaching-learning situations were considered as functioning during this experimental research study also. But some variables were identified as the threats for experimental validity as given below, where some were controlled by the researcher in a fair manner with the precautionary measures.

For the present study also, Experimental validity are divided mainly in two parts as: (i) Internal Validity as well (ii) External Validity and both are elaborated as below.

#### 3.1.3.1 Internal Validity

Internal validity refers to the extent to which the manipulated or independent variables actually have a genuine effect on the observed results or dependent variable. This validity is affected by the lack of control of extraneous variables. As **Campbell & Stanley (1963), i**n the research literature, the extraneous variables also referred to as intervening variables, directly affect the action of the independent variable on the dependent variables. Intervening variables are those variables that occur in the study setting. They include economic, physical, and psychological variables. Therefore, it is important to control extraneous variables to study the effect of independent variable on dependent variable. We must be very careful to control all possible extraneous variables that might intervene the dependant variable.

Extraneous Variables are the threats for the Internal Validity. Similar variables were detected for this experimental research study also and the researcher had put reasonable, unbiased, positive and possible efforts to control the extraneous variables in order to avoid or minimize the influences on the Post-treatment measurement/s.

Following table-3.2 represents a list of the Extraneous Variables which were the threats to the internal validity along with a brief description and the preventive measures attempted by the researcher. The following list is from **Campbell & Stanley (1963)** as interpreted by **Kirk (1995)**:

**Table-3.2:**
**Threats to Internal Validity observed for the present research study**

| No. | Extraneous Variables | Description / Explanation | Precautionary measures taken by the Researcher |
|---|---|---|---|
| **Threats Associated with Participants** | | | |
| 1 | **Differential Selection** | If we select groups for treatment and control differently than the results may be due to the differences between groups before treatment. Randomization solves this problem by statistically equating groups. | For the present study also, two groups (two classes) of class-IX were selected randomly with the criteria of same class/grade-level and of the same average age. And academic equity level of both the groups was measured by the previous academic records of class-VIII of both the groups |
| 2 | **Maturation** | The natural process of human growth can result in changes in post-test scores quite apart from the treatment. | It was natural to accept. The experimentation period was of six months. And the Post-tests were managed by the researcher to conduct along-with or during or immediate to the treatment. |
| 3 | **Selection-Maturation Interaction of Subjects** | Refers to a mixing of selection and maturation factors to compound the extraneous influence on measurements. | For the present study also, two groups(two classes) of class-IX were selected randomly with the criteria of same class/grade-level and of the same average age. And academic equity level of both the groups was measured by the previous academic records of class-VIII of both the groups |
| 4 | **Experiment al Mortality** | Also called "attrition," refers to the loss of subjects from the experiment. | To deal with such situation, the researcher kept the records for the attendance of both the selected groups for the entire period of the |

| | | | |
|---|---|---|---|
| | | | experimentation. And appropriate criterions were employed to retained with optimum sample size. |
| **5** | **History** | Refers to events other than the treatment that occur during the experiment which may influence the post-treatment measure of treatment effect. | A note was made on influences noticed by the researcher as well suitable precautionary measures were made to minimize the effect of such events that affected the Post-treatment measurement. |
| **6** | **The John Henry Effect** | If subjects in a control group find out they are in competition with those in an experimental treatment, they tend to work harder. When this occurs, differences between control and treatment groups are decreased, minimizing the perceived treatment effect. | For the remedy, the researcher had selected two groups from two different schools which were geographically located at distance at two different places to remove this influence and to isolate the treatment to the particular experimental group. |
| **Threats Associated with Measurement** | | | |
| **7** | **Testing** | Refers to the Pretest-Posttest research design. If the same test is given both times, the group may show an improvement simply because of their experience with the test. This is especially true when the treatment period is short and the tests are given within a short time. It is better to only give a post-test as randomly assign subjects to groups to render the dependent variable statistically equal at the beginning of the study. | Hence, for the present research study, only the Post-tests with the reasonable time-span were conducted for both the groups. |
| **8** | **Instrumentation** | Refers to the problem of using the same test twice in pre- and | Hence, for the present research study, only the Post-tests with the |

| | | | |
|---|---|---|---|
| | | post-measurements. But if you use different tests for pre- and post-measurements, then the change in pre- and post-scores may be due to differences between the tests rather than the treatment. The best remedy is to use randomization and a post-test only design. | reasonable time-span were conducted for both the groups. |
| **9** | **Statistical Regression** | Refers to the tendency of extreme scores, whether low or high, to move toward the average on a second testing. Subjects who score very high or very low on one test will probably score less high or less low when they take the test again. That is, they regress toward the mean. | For the present research study, investigator will study the full range of the scores. |
| **10** | **Treatment diffusion** | Similar to the John Henry effect is treatment diffusion. Over the course of the experiment, some of the materials of the treatment group may be borrowed by the control group members. Over time, the treatment "diffuses" to the control group, minimizing the treatment effect. This often happens when the groups are in close proximity. | As aforesaid, For the remedy, the researcher had selected two groups from two different schools which were geographically located at distance at two different places to remove this influence and to isolate the treatment to the particular experimental group. |

For the present research study, by conducting only the posttests on two distinct and distanced groups were helped to minimize the influence of some threats. Other concern was with the non-randomization, so the researcher used the 'Matching' criterion for the subjects mean to study the actual effect of the treatment. Further, the researcher kept the records for other non-controlled extraneous variables which might have influenced the effect of the treatment in any manner and also stated in a report.

#### 3.1.3.2 External Validity

Sources of external invalidity cause changes in the experimental groups so that they no longer reflect the population from which they were drawn. Variables those were the threats to the External validity were observed for the present research study. Thus, the researcher had intentionally choose the aforesaid research design in order to put possible and reasonable efforts to follow and maintain the natural phenomenon of normal functioning of the education as well an examination system.

Following table-3.3 is a list of the variables as threats to the external validity observed for the present experimental research study. Also, brief descriptions about the threats and the preventive measures taken by the researcher in term to minimize the effects of threats are included in the following table.

**Table-3.3:**
**Threats to External Validity observed for the present research study**

| No. | Variables | Description / Explanation | Precautionary measures taken by the Researcher |
|---|---|---|---|
| **1** | **Reactive effects of testing** | Subjects in samples may respond differently to experimental treatments merely because they are being tested. Since the population at large is not tested, experimental effects may be due to the testing procedures rather than the treatment itself. This reduces generalizability. One type of reactive effect is *pre-test sensitization*. Subjects who take a pre-test are sensitized to the treatment. Another is *post-test sensitization*. The post-test can be, in itself, a learning experience that | Researcher had design the post-test with the care to measure the treatment effect only with the unbiased manner and with the complete justifications. |

| | | | |
|---|---|---|---|
| | | helps subjects to "put all the pieces together." | |
| **2** | **Treatment & Subject Interaction** | *Subjects* in a sample may react to the *experimental treatment* in ways that are hard to predict. This limits the ability of the researcher to generalize findings outside the experiment itself. | Researcher had put effort to make possible prediction/s with the help of the reactions of the subjects and by the observation. |
| **3** | **Testing & Subject Interaction** | *Subjects* in a sample may react to the process of ***testing*** in ways that are hard to predict. This limits the ability of the researcher to generalize findings outside the experiment itself. If there is a systematic bias of test anxiety or "test-wiseness" in a sample, then treatment effects will be different when applied to a different sample. | Researcher had put efforts to make possible prediction/s with the help of MLL test, post-test, reactions of the subjects and by the observation. |
| **4** | **Multiple Treatment Effect** | If an experiment exposes subjects to three treatments (A, B, and C) and test scores show that treatment C produced the best results, one cannot declare treatment C the best. It may have been the *combination of the treatments* that led to the results. | In the present study, investigator will be conducting only single-single post-test to each of the concepts. |

The Researcher had put reasonable, unbiased, positive and possible efforts to control the other threats to external invalidity in order to secure representative samples to study so that inferences can be made back to the population from which the samples were drawn.

### 3.2 POPULATION AND SAMPLE

The students of the thirty CBSE English medium secondary schools of the Vadodara city were constituted as the Population for the present research study, (source: CBSE website on 10/01/2013-some were affiliated and some were in pipeline to affiliation). List of such schools is enclosed in an Appendix - .G.

Two schools out of thirty CBSE English medium secondary schools were selected **random-purposively** with respect to the permissions of the schools in terms to be a sample of the said research study. Those two schools were selected as sample for the present experimental research study based on the following criterions as the researcher had set according to the need of the study: **(a) Selected CBSE schools having secondary section that is at least having Class-IX. (b) A School whoever ready and permit to conduct experiments for the said research study for the longer duration as well to provide flexibility, liberty and necessary facilities as per the requirements of the research study. (c) After matching process, at least to get or have 25-40 students for a sample of the study.**

Though, two schools with Class-IX were selected for the study where one school was consisting of only one section of Class-IX with 43 students considered as one group while other school was consisting of two sections of Class-IX with total of 74 students considered as another group. In terms to make two comparable groups that is probably two equivalent groups, 'Matching' technique was employed to equate both the groups as well to randomize the groups and to make homogeneous groups of a sample before the experimentation of the study. Further analytical interpretations were made based on the performances of these two matched groups.

### 3.2.1 Matching For The Equivalence Of Two Groups

Randomization or in randomly assigning individuals as according to **Creswell (2012)** is mean to controlling for extraneous variables that might influence the relationship between the new practice and the outcome. Extraneous factors are any influences in the selection of participants, the procedures, the statistics, or the design likely to affect outcome and provide an alternative explanation for such results than what is expected. All experiments have some random error (where the scores do not reflect the true scores of the population) that one cannot control, but one can try to control extraneous factors as much as possible. If Random assignment is a decision made by the investigator before the experiment begins, than other control procedures one can use both before and during the experiment are pretests, covariates, matching of participants, homogeneous samples and blocking variables.

Two distanced schools were randomly selected for the present experimental research study. All the students studying in class-IX of these two schools were determined as a sample but to equalize these two groups with appropriate subjects (students), the researcher used the procedures of 'Matching' the participants in the beginning that was before to begin an actual experimental study. The criterion for Matching the participants were based on the overall performance of the respective participants in Mathematics during the previous academic year (year 2013-2014) that is based on the achievement scores were gained for Class-VIII Mathematics. In such a way, two equal randomized groups were formed as an actual sample for the present research study and the detailed procedure of Matching the groups is mentioned below in a figure-3.1 and table-3.4. The list of the subjects or students of a sample with the scores (in percentages) of class-VIII Mathematics is enclosed in Appendix- G.

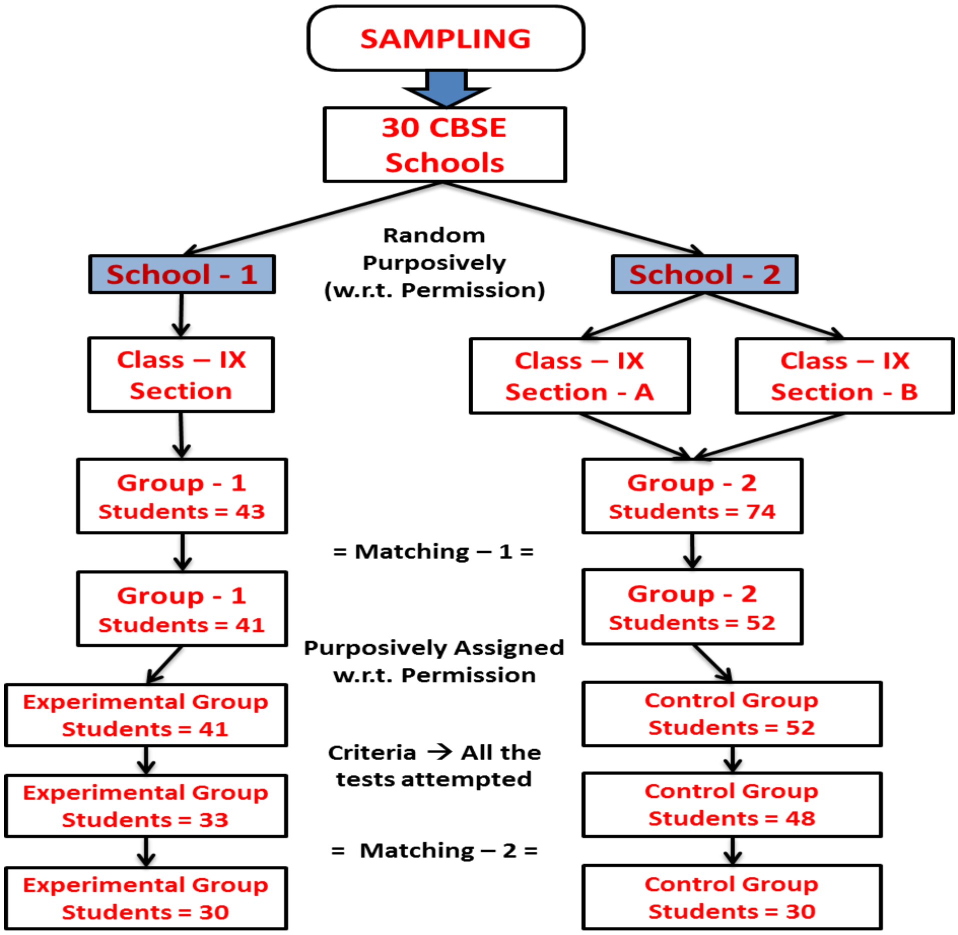


**Figure – 3.1: Matching and Sampling**

The table-3.4 is showing the filtration of participants using Matching process in terms to form two academically homogeneous groups for a sample of an experimental research study. Initially, there were 43 students in one group and 74 students in other group. While after matching the scores (in percentages) of class-VIII Mathematics, 41 students were found in one group and 52 students were found in other group. The equality of both the groups was measured through Measure of Central Tendency. Then, purposively the groups were assigned (with reference to the permission from the school to conduct an actual experiment) as an Experimental group (with 41 students) and other as Control group (with 52 students).

But finally, one more criterion was employed after the completion of an intervention program was 'only the participants of the matched groups who were attempted all the achievement post-tests'. Then it was found 33 students in experimental group and 48 students in control group which is shown in the table-3.4, so again same Matching technique was repeated between a group of 33 students and a group of 48 students in terms to have two probably equivalent groups for the fair comparisons between them. Finally two groups as sample determined as shown in the said table, where effort was made for each group to consist of 30 students.

**Table: 3.4**
**Matching-2: Frequency distribution on percentages of VIII Mathematics scores**

| Sr. No. | Class-interval of the Percentage (%) | Frequencies based on the Results of Class-VIII Mathematics | | Selected minimum frequency from A & B |
|---|---|---|---|---|
| | | Experimental Group A | Control Group B | |
| 1 | **40% to 49%** | **1** | 8 | **1** |
| 2 | **50% to 59%** | **5** | 13 | **5** |
| 3 | **60% to 69%** | **3** | 6 | **3** |
| 4 | **70% to 79%** | 7 | **6** | **6** |
| 5 | **80% to 89%** | 7 | **6** | **6** |
| 6 | **90% to 99%** | 10 | **9** | **9** |
| **TOTAL** | | 33 | 48 | **30** |

The Means of both the finally matched groups were derived from their scores/results in Mathematics of their previous academic year (2013-2014) that is based on the records of Class -VIII are shown in the table – 3.5.

**Table-3.5:**
**Means and SD of the scores gained in Class -VIII Mathematics by both the groups**

| Sample | Size | Mean | SD | t-test |
|---|---|---|---|---|
| **Experimental Group (E)** | **30** | **76.73** | **16.56** | **0.093** |
| **Control Group (C)** | **30** | **76.34** | **16.08** | |

As the critical value of is t = 2.663 for df = 58 at significance level of 0.01 is greater than a calculated value of t = 0.093. So it accepts a Null hypothesis means that both the groups are equal.

Thus, the two groups formed or matched through Matching process, are assumed as probably equivalent groups with reference to the academic caliber, maturity, level of understanding and other possible characteristics or attributes. With such assumptions, further analyses conducted are reported in the chapter of Data Analysis. Also, all achievement tests attempted by this sample were taken for the analysis and interpreted about the intervention program that is SOLO based instructional strategy. For the analysis based on the Reaction tools, the whole intake classes of both the schools were taken into consideration in terms to get more responses and views about the intervention programs as well as about various components included in it. Next section is explaining about research instruments developed to collect several data needed for the present study.

## 3.3 TOOLS FOR DATA COLLECTION

Research Tools or instruments are the means to collect the data for the study and are either readily available or need to develop according to the requirements of the study. For the present experimental research study, the researcher had developed many tools to collect the data and study the desired variables of the research study. Following are the details about these research instruments which were mainly divided in two ways as: (a) tools to measure the components of the study and (b) tools to support the study.

### 3.3.1 Tools To Measure The Components Of The Study

Data collected through such tools were considered for an actual analysis and the findings derived from this analysis were taken into account for making the predications about the overall research study in terms of its impact or effectiveness as well to provide appropriate justifications for this research study. Mainly two tools (i) Achievement tests and (ii) Reaction Scale were designed and developed by the researcher as well taken for the actual analysis in terms to draw out the findings and to make conclusions from it for this research study.

**3.3.1.1 Achievement Tests:** Achievement Tests are the ways to collect the quantitative data which meant to signpost about the learning of the participants. For the present experimental research study, several Achievement tests (only Posttest) were designed and developed by the researcher based on all the levels of SOLO Taxonomy in order to measure the progressive learning outcome of the sample. The central focus was to frame 'Understanding' based questions for all the achievement tests and to address all the levels of SOLO Taxonomy.

The Achievement tests (only Posttest) were prepared for each of the five selected chapters (mentioned in a table-3.6 below) as well as for overall (final) test that is total six Achievement tests were developed. And all achievement tests were administered at Post level for both the groups – an experimental and control group. All the chapter-wise achievement tests were based on the content of respective chapters and were conducted immediately after the completion of the respective chapters. Chapter-wise Achievement Tests are enclosed in the Appendices - A to E. Final or Overall achievement test was examined at the end of the intervention program and was consisting of cumulative content from all the five chapters. Overall Achievement test is enclosed in Appendix – F. Here, following table-3.6 is detailing about the Achievement tests developed by the researcher of this research study.

**Table-3.6:**
**Details about the developed Achievement tests for Post-tests**

| Sr. No. | Achievement Test For | | No. of Questions Consist | Time Duration | Maximum Mark |
|---|---|---|---|---|---|
| 1 | Chapter – 12 | Heron' Formula | 07 | 30 minutes | 25 |
| 2 | Chapter – 4 | L. E. In Two Variables | 10 | 30 minutes | 25 |
| 3 | Chapter – 8 | Quadrilaterals | 08 | 30 minutes | 25 |

| | | | | | |
|---|---|---|---|---|---|
| 4 | Chapter – 14 | Statistics | 06 | 30 minutes | 25 |
| 5 | Chapter – 15 | Probability | 08 | 30 minutes | 25 |
| 6 | Final / Overall Achievement Test | | 25 | 2½hours | 75 |

All the six achievement tests were validated by the experts and necessary changes recommended by the experts were incorporated before the actual implementation of the Pots-tests.

**3.3.1.2 Reaction Tools:** The Reaction Tools developed by the researcher for this research study were aimed to collect the opinions from the participants in terms to know about the learning experiences of the participants with the new developed instructional strategy as well about the other components of the teaching-learning processes like activities, home-tasks and the achievement tests. The tools to collect the responses in terms of reactions were divided into two parts as: (a) SOLO Reflective Reaction Sheet and (b) Reaction Scale are detailed as below.

**[A] SOLO Reflective-Reaction Sheet:** It was designed and developed chapter-wise by the researcher based on the SOLO levels in terms to collect the reactions only from an experimental group that is the group studied trough the developed Instructional Strategy. These Chapter-wise Reactions sheets were administered after the completion of respective chapters in terms to collect the reactions on the learning, activities, post-tests and overall experiences gained for each of the chapter studied through developed instructional strategy.

The number of items in each of the reaction sheets varies with respect to the chapters and the SOLO level-wise as it depends on the number of activities or tasks were exercised through the teaching-learning processes of an intervention program. Each item of the reaction sheets was consists of limited options to choose. All the five reaction sheets were validated by the experts and necessary suggestions were incorporated. All the chapter-wise SOLO reaction sheets are enclosed in the Appendices – A to E. The details of these chapter-wise reaction tools are tabulated in a table-3.7 as given below.

**Table-3.7:**
**Details about the developed Reaction tools for chapter-wise reactions**

| Sr. No. | Reaction Tool | | Respondents | No. of Item-statements | |
|---|---|---|---|---|---|
| | | | | Close-ended | Open-ended |
| 1 | Chapter – 12 | Heron's Formula | Experimental Gr. | 16 | 05 |
| 2 | Chapter – 4 | L.E. In Two Variables | Experimental Gr. | 14 | 04 |
| 3 | Chapter – 8 | Quadrilaterals | Experimental Gr. | 16 | 03 |
| 4 | Chapter – 14 | Statistics | Experimental Gr. | 16 | 03 |
| 5 | Chapter – 15 | Probability | Experimental Gr. | 16 | 03 |

**[B] Reaction Scale:** It was prepared by the researcher to collect the responses from both the groups for being a part of this research study. Two separate reaction sheets were prepared, one for control group consists of two sections (I & II) and another for an experimental group consists of three sections (I to III). Section-I of the Reaction Scale was designed with intension to collect the opinions from both the groups about their thoughts for the Mathematics. Section-II was designed to collects the responses about the achievement test (posttest). And Section – III was designed to collect the responses about the Mathematics teaching-learning processes conducted with developed instructional strategy. All three Sections are briefed here in a table-3.8 as:

**Table-3.8:**
**Details of the developed overall reaction tool – The Reaction Scale**

| Reaction Scale | | | | | |
|---|---|---|---|---|---|
| Sections | Details | Type for Response | No. of Items-statements | | Respondents |
| | | | Close ended | Open-ended | |
| I | About You in Mathematics | Five statements as options | 11 | - | Experimental Gr. & Control Gr. |
| II | About The Achievement Tests | Five point Likert scale | 30 | - | Experimental Gr. & Control Gr. |
| III | Your Learning Experience Throughout New Intervention Program | Five point Likert scale | 25 | 01 | Experimental Gr. |

A Likert scale, the scale is named after its inventor and psychologist **Likert (1932)** is a psychometric scale commonly involved in research employing questionnaires. When responding to a Likert questionnaire item, respondents specify their level of agreement or disagreement on a symmetric agree to dis-agree scale for a series of statements. Thus, the range captures the intensity of their feelings for a given item. For the present experimental research study also, based on the idea of five points of the Likert scale following five points or scales of reactions taken as the options for the responses.

(i)  = Very Unhappy / Very Upset
(ii)  = Unhappy / Upset
(iii)  = Normal / Neutral
(iv)  = Happy
(v)  = Very Happy

The first two sections were kept common in both the reaction sheets, while third section was kept only in a reaction sheet given to the group studied through the developed instructional strategy as it was about the participants' overall learning experiences through the newly developed instructional strategy. These tools were validated by the experts. And it is enclosed in an Appendix - F.

### 3.3.2 Tools To Support The Study

Data were collected through these instruments were not considered to measure any component of the study or for the actual analysis. But these tools were used as evidences to support the research study in all possible dimensions and in terms to identify the desired elements needed for the research in terms to make it more appropriate or effective or with deeper aspects. Both kinds of tools are elaborated below.

**3.3.2.1 Field Notes:** As per the Emerson, **Emerson et al. (1995),** Field notes are one means employed by qualitative researchers whose main objective of any research is to try to understand the true perspectives of the subject being studied. Field notes allow the researcher to access the subject and record what they observe in an unobtrusive manner.

For the present study also, Field notes were prepared based on the class observations conducted in the first phase of the study. This instrument was used to collect the data from the actual situations in terms to study the current status of the teaching-learning processes exercised for Mathematics of Class-IX in various CBSE schools of Vadodara city. Notes were made on 'as it is whatever seen or observed on the spot (fields) in fair or unbiased manner' by the researcher. The mentioned points were kept in mind by the researcher while observing the classes as: (a) Infrastructural aspects of a classroom; (b) The Mathematics Teacher/s; (c) The Students of Class-IX; (d) Mathematic teaching-learning process; (e) Innovative approach; (f) Level of Thinking, Learning or Understanding and (g) Interactions with Mathematics teachers and the students inside and outside the class.

The field notes were prepared on the situational presentations and prepared before the planning as well development of the instructional strategy. These notes helped in planning, designing, developing and implementation of the instructional strategy. The findings were derived from these notes on class-observations (situational observations) were reported in the last chapter of this thesis.

**3.3.2.2 Observation Notes:** It was unstructured in order to record the on-field situations had taken place during the implementation of the developed Instructional Strategy in terms to observe and to note down as it was. The relevant points kept in mind by the researcher to observe and to note down were: (a) the interest, behaviors, attitudes, readiness for the learning of the participants; (b) The disciplinary aspects as well the challenges came across and (c) The unusual incidents or moments took place during the implementation of an instructional strategy with reference to various threats which influences the validities.

The observations noted down by the researcher helped to discuss on the interpretations derived from the various analyses of this research study and also incorporated in the discussion part of the Chapter-VI. These notes on observations supported the researcher in planning as well making necessary modifications in terms to improve the teaching-learning processes conducted during the implementation of the developed instructional strategy. Also helped to explain and elaborate the numerous experiences learnt during the study and to draw pin-point suggestions or recommendations to incorporate in the chapter of summary of this thesis.

**3.3.2.3 Reflection Notes:** It was prepared by the researcher in order to note down the self-reflections of the researcher being as a teacher and to note down about all the experiences gained related with the actual implementation of the developed Instructional Strategy. These notes were prepared after the completion of each of the components like after each class, each activity as well as each chapter of the experimental study and overall completion of the experimental study. Reflection notes were prepared based on the points as: (a) The teaching-learning challenges and difficulties; (b) Time constraints; (c) Several classroom situations; (d) The academic-administrative processes and (e) Other disciplinary manners.

This practice was helped the researcher to make necessary behavioral changes and other modifications required in any manner for the better next step during the implementation of the experimental study. Some important points and suggestions outlined from these notes are incorporated in the part of discussions and further improvements or recommendation in the last chapter for this thesis.

## 3.4 TECHNIQUES OF DATA ANALYSIS

Data were collected personally by the researcher with the help of the research instruments. Both the Qualitative and Quantitative type of the data were collected which are shown in the following table - 3.9.

**Table – 3.9:**
**Overview on the type of collected data and methods for data-analysis**

| Sr. No. | Name of the Tools | Type of Data | With Reference to the Objectives | Statistical Method |
|---|---|---|---|---|
| **Tools Considered To Measure The Components Of The Study** | | | | |
| 1 | Achievement Tests | Quantitative (Ordinal data) | Objective – 3 to 5; Hypotheses–1 to 6; Hypotheses–7 to 11 | Descriptive Statistics; Graphs; Non-Parametric Method & Rasch Model/Analysis |

| | | | | |
|---|---|---|---|---|
| 2 | SOLO Reflective Reaction Sheet | Qualitative | Objective – 6 Hypotheses–12 to 16 | Frequency and Percentage; Graphs; Chi-Square Test |
| 3 | Reaction Scale | Quantitative Qualitative | Objective – 7; Hypothesis - 17 | Frequency and Percentage; Graphs; Chi-Square Test |
| **Tools To Support The Study** | | | | |
| 1 | Field Notes | Qualitative | Objective – 1, 2 | Content Analysis |
| 2 | Observation Notes | Qualitative | Objective – 3 to 7 | Content Analysis |
| 3 | Reflection Notes | Qualitative | Objective – 3 to 7 | Content Analysis |

In terms of the Data Analysis, the data collected through two main instruments, the Achievement Tests and the Reaction Tools were taken into account to measure the variables of the study. Detailed data analysis is given in a Chapter-V.

For the analysis of the data collected for the present research study, the researcher had employed Non-Parametric methods to analyze the data and to draw the interpretations. Following are the criteria which referred to use Non-Parametric methods for the data analysis as: **(i) Purposive Sampling; (ii) Size of the sample was 30; (iii) Data to be considered as the Ordinal data or fall in the Ordinal scale; (iv) Data were not normally distributed.**

**Ordinal scale or Data**

As the quantitative data collected through the achievement tests were considered in the 'Ordinal scale' or as 'Ordinal data' for the present experimental research study. As according to **Creswell (2012),** researchers use ordinal scales (or ranking scales or categorical scales) to provide response options where participants rank from best or most important to worst or least important some trait, attribute or characteristics. These scales have an implied intrinsic order. For example, a researcher might record individual performance in a race for each runner from first to last place. Many attitudinal measures imply an ordinal scale because they ask participants to rank order the importance (as highly important to of no importance) or the extent (to a great extent to a little extent) of topics. As this example illustrates the information is categorical in a ranked order. According to **Wu & Adams (2007)**, when numbers are assigned to objects to indicate ordering among the objects, the numbers are said to be ordinal. Ordinal measurements are often used, such as for ranking students, or for ranking candidates in an election or for arranging a list of objects in order of preference.

Further, the test for the Normal Distribution of the quantitative data collected through the present research study was carried out by using the Measures of Central Tendency and checked whether three M that is Mean, Mode and Median are probably equal or not. Following Table-3.10 is showing the values of three M as well as Standard Deviation (SD) and Variance calculated on the Post Achievement Test scores gained by the sample that is by both the control (C) and experimental (E) groups for the five selected chapters of Class-IX Mathematics.

**Table – 3.10:**
**Measures of Central Tendency on Post-test Scores of sample for Class-IX Mathematics**

| Measures | Group | MEAN | MODE | MEDIAN | SD | VARIANCE |
|---|---|---|---|---|---|---|
| Chapter – 12 | C | 8.23 | 6 | 7.25 | 4.92 | 24.22 |
| | E | 18.34 | 18 | 18.50 | 3.80 | 14.43 |
| Chapter – 4 | C | 13.47 | 20 | 12.38 | 4.70 | 22.14 |
| | E | 17.96 | 16 | 17.25 | 2.33 | 5.45 |
| Chapter – 8 | C | 11.99 | 8.5 | 12.38 | 4.78 | 22.88 |
| | E | 18.01 | 18.5 | 18.50 | 1.87 | 3.49 |
| Chapter – 14 | C | 8.92 | 8 | 8.75 | 2.50 | 6.23 |
| | E | 18.84 | 18 | 18.88 | 1.86 | 3.44 |
| Chapter – 15 | C | 12.33 | 16 | 12.50 | 2.58 | 6.66 |
| | E | 20.67 | 22.5 | 21.00 | 2.20 | 4.85 |
| Overall (Final) | C | 19.71 | 23.5 | 19.25 | 5.90 | 34.85 |
| | E | 53.80 | 50 | 52.25 | 6.26 | 39.16 |

Following graphs are another way of showing whether the data in the form of test-scores is normally distributed or not. Chapter-wise graphs are showing the checks for normal distribution of test-scores depicted along with the standard values. The standard values for the Normal Distribution were generated in MS-Excel 2007 by using the Statistical Function NORMDIST (X, Mean, SD, Cumulative_distribution). Bell curves were formed through this standard values with reference to the scores gained by a control group and by an experimental group in the chapter-wise as well final achievement post-tests are shown in the following graphs-3.2 to 3.7. The group of following six figures presented is consists of graphical presentation on the checks for Normal Distribution on each chapter's scores as well final scores achieved by a control group in the chapter-wise post-tests.

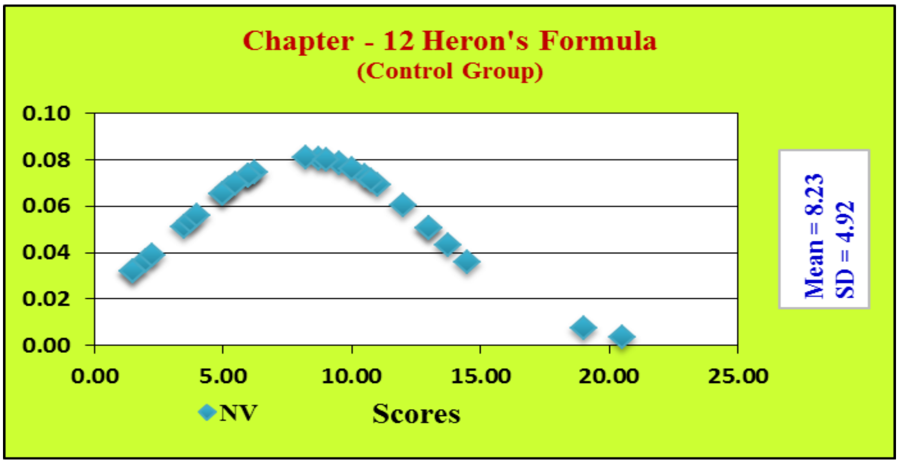


**Figure – 3.2: Check for Normal distribution on scores of Control group for Chapter-12**

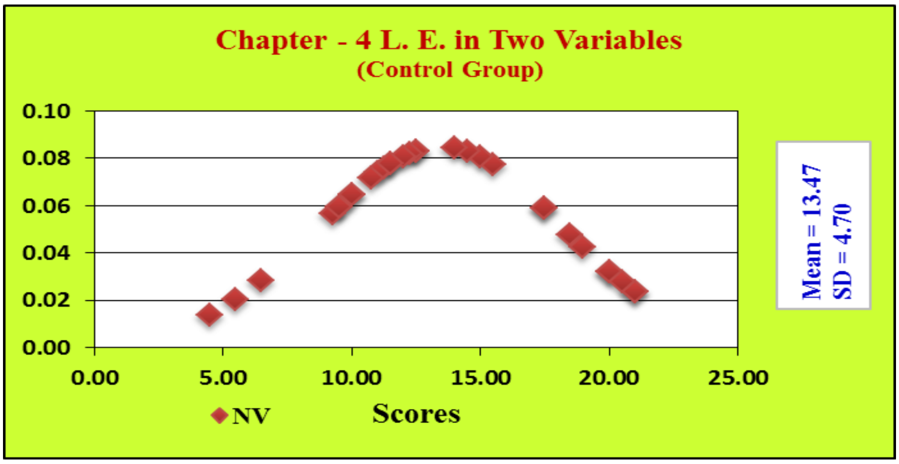


**Figure – 3.3: Check for Normal distribution on scores of Control group for Chapter-4**

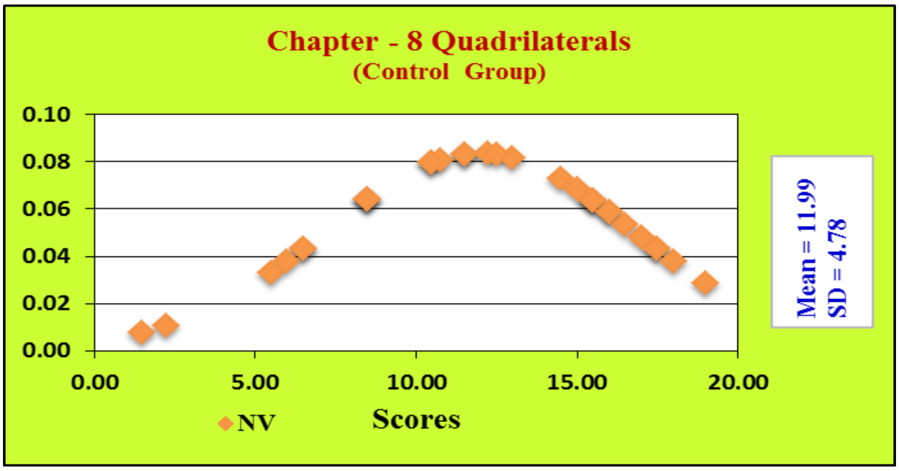


**Figure – 3.4: Check for Normal distribution on scores of Control group for Chapter-8**

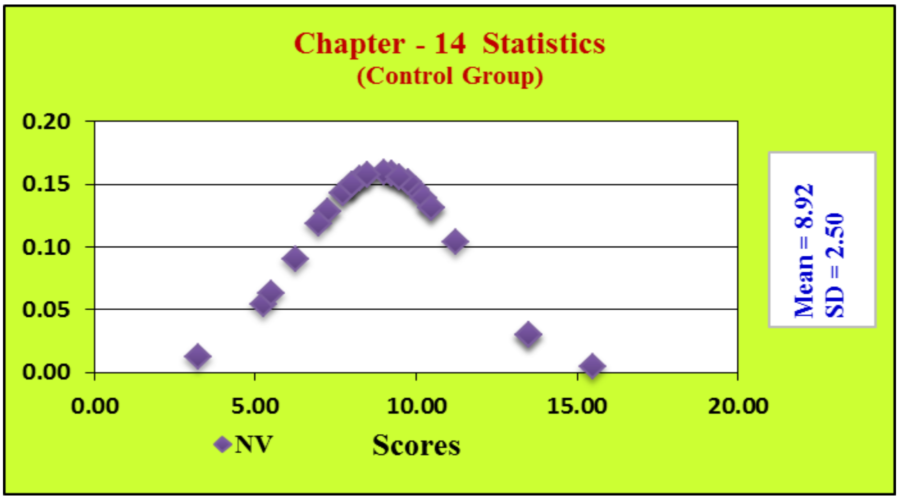


**Figure – 3.5: Check for Normal distribution on scores of Control group for Chapter-14**

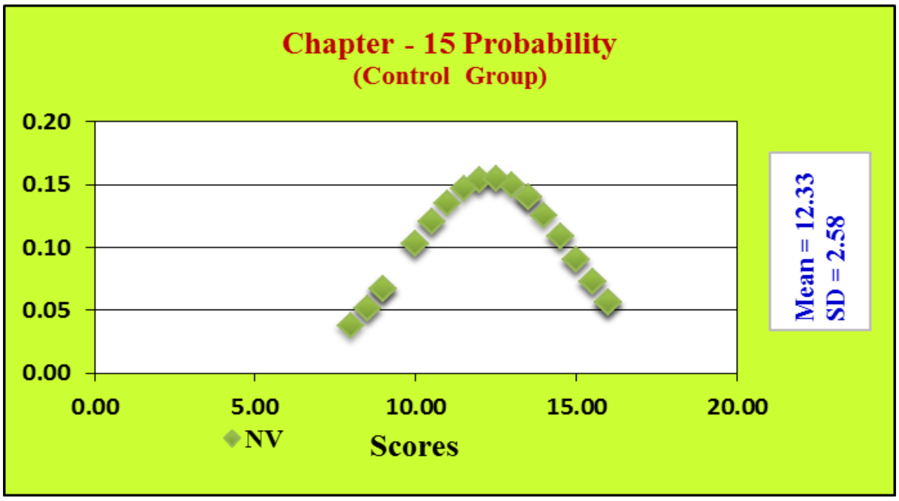


**Figure – 3.6: Check for Normal distribution on scores of Control group for Chapter-15**

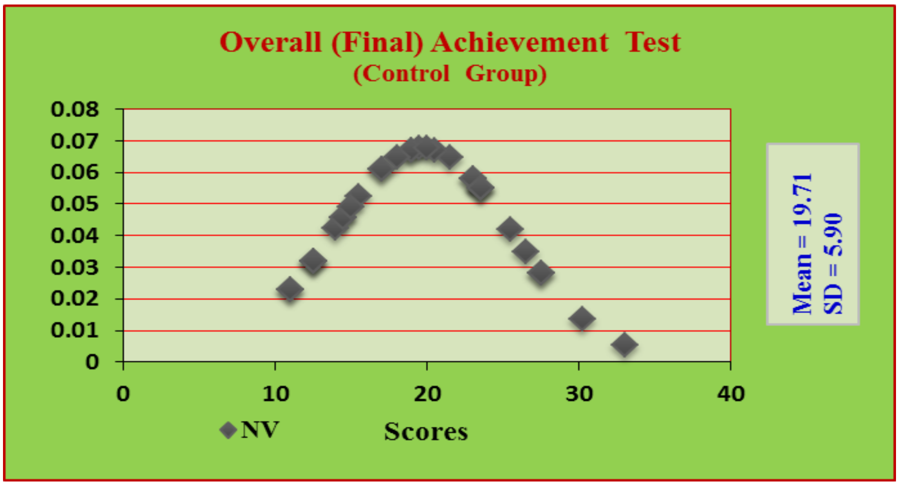


**Figure – 3.7: Check for Normal distribution on scores of Control group in Overall Test**

The group of following six figures-3.8 to 3.13 presented is consists of graphical presentation on the checks for Normal Distribution on each chapter's scores as well final scores achieved by an experimental group in the chapter-wise post-tests.

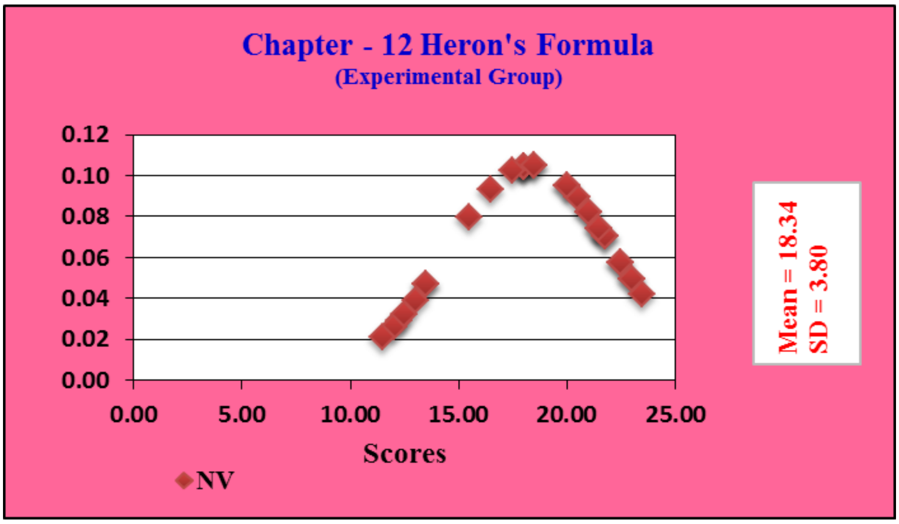


**Figure – 3.8: Check for Normal distribution on scores of Experimental group for Chapter-12**

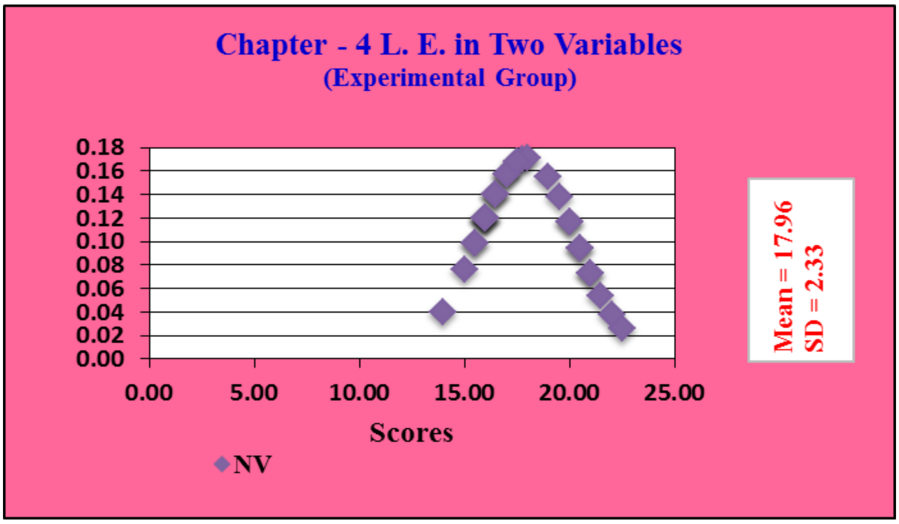


**Figure – 3.9: Check for Normal distribution on scores of Experimental group for Chapter-4**

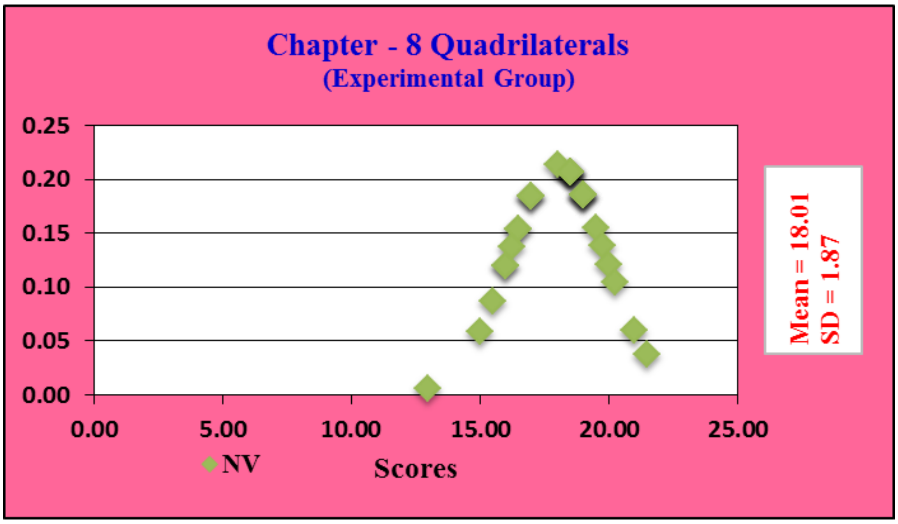


**Figure – 3.10: Check for Normal distribution on scores of Experimental group for Chapter-8**

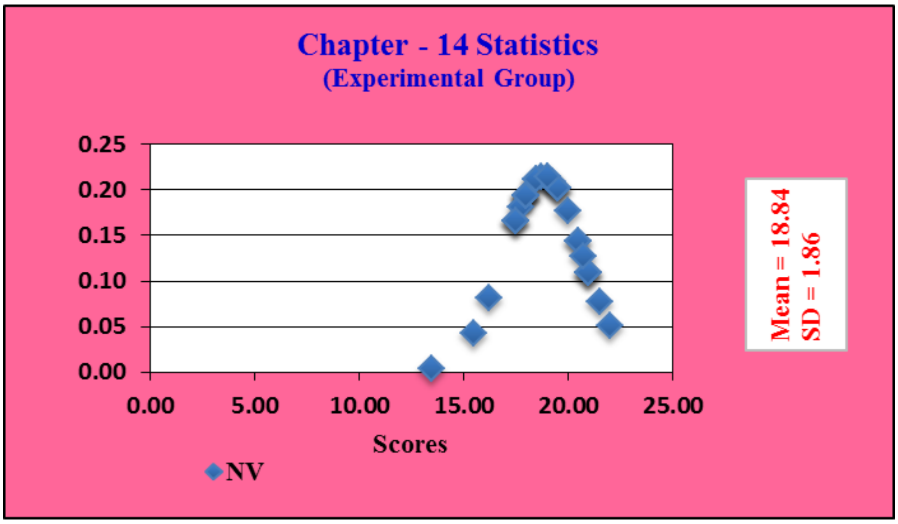


**Figure – 3.11: Check for Normal distribution on scores of Experimental group for Chapter-14**

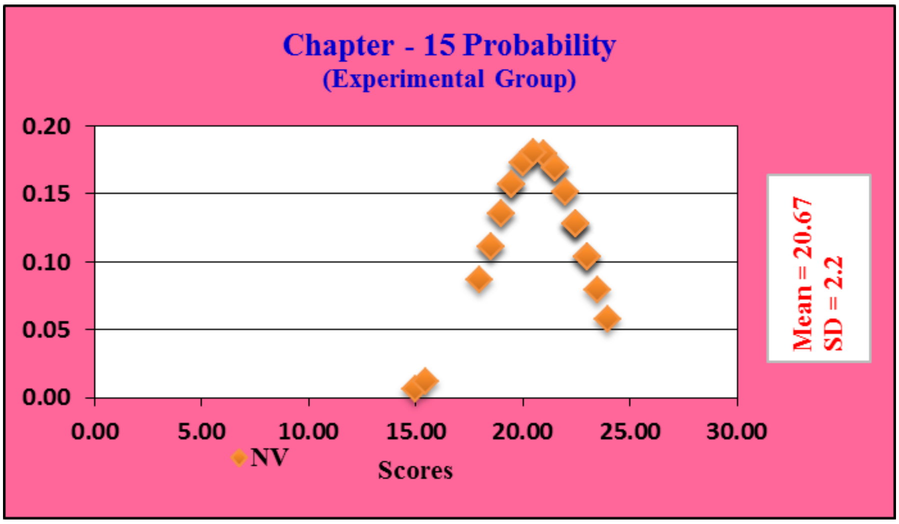


**Figure – 3.12: Check for Normal distribution on scores of Experimental group for Chapter-15**

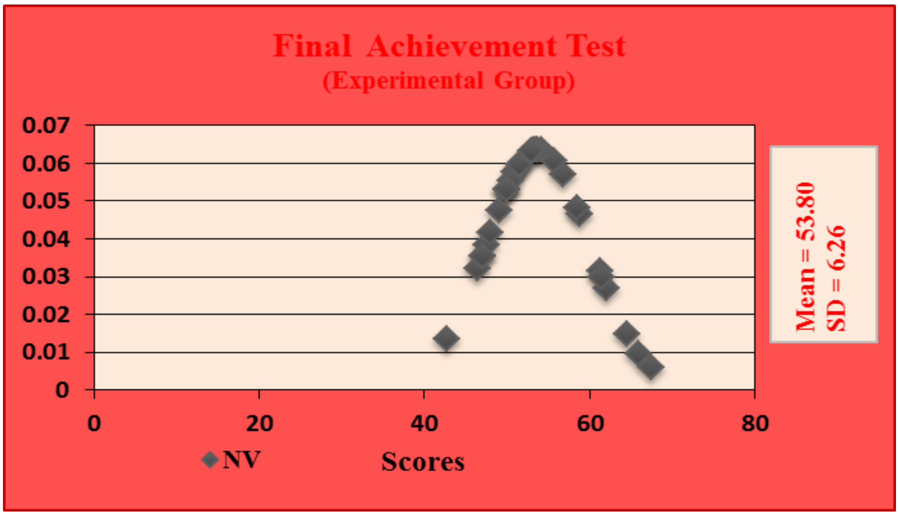


**Figure – 3.13: Check for Normal distribution on scores of Experimental group in Overall Test**

As looking to the several graphs, it is clearly seen whether the data have distributed normally or not. For the further checks, the Measure of Skewness calculated for the test-scores of both the experimental and control group which is shown in the following table-3.11.

**Table – 3.11:**
**Measures For Skewness on test-scores of both the groups**

| Sr. No. | Details | Measures For Skewness ($S_k$) | |
|---|---|---|---|
| | | Experimental Group | Control Group |
| 1 | Chapter – 12 | **0.09** | **0.45** |
| 2 | Chapter – 4 | **0.84** | **-1.39** |
| 3 | Chapter – 8 | **-0.79** | **0.73** |
| 4 | Chapter – 14 | **0.45** | **0.20** |
| 5 | Chapter – 15 | **-0.45** | **-0.19** |
| 6 | Overall / Final | **0.61** | **0.23** |

Looking to the Normality Tests as shown graphically (graph – 3.2 to 3.12) as well as in tabular form (table-3.10 and table-3.11), it is perceived that all the data are not normally distributed. Also the data are of Ordinal type. Also, there is a quite differences between the Variances calculated on the scores (as shown in the table-3.10) of both the groups. With these references, Non-Parametric Tests have been professed to analyze the quantitative data collected from the present research study.

**Non-parametric methods**

The **Non-parametric methods** according to **Wikipedia (2016)** are widely used to studying populations that take on a ranked order. The use of Non-parametric methods may be necessary when data have a ranking but no clear numerical interpretation, such as when assessing preferences. In terms to the level of measurement, Non-parametric methods results in “Ordinal Data”.

Following table-3.12 is showing the Parametric Test Methods and their alternative Non-Parametric Methods.

**Table – 3.12:**
**Overview on Parametric Test vs. Non-Parametric Tests**

| Sr. No. | Parametric Test | Alternative Non-Parametric Test |
|---|---|---|
| 1 | 1-Sample Z-test<br>1-Sample t-test | 1-Sample Sign Test |
| 2 | 1-Sample Z-test<br>1-Sample t-test | 1-Sample Wilcoxon Test |
| 3 | 2-Sample t-test | Mann-Whitney U Test |
| 4 | One-way ANOVA | Kruskal-Wallis Test |
| 5 | One-way ANOVA | Mood's Median Test |
| 6 | Two-way ANOVA | Friedman Test |

(Source: http://support.minitab.com/Understandingnonparametrictests, 2016)

Moreover, Non-parametric methods make fewer assumptions, their applicability is much wider than the corresponding Parametric methods. In particular, they may be applied in situations where less is known about the application in question. Also, due to the reliance on fewer assumptions, Non-parametric methods are more robust. Another justification for the use of Non-parametric method is simplicity. Due both to this simplicity and to their greater robustness, Non-parametric methods are seen by some Statisticians as leaving less room for improper use and misunderstanding. For the present research study, data were analyzed using **Mann-Whitney U Test** in terms to test hypotheses framed for the scores of all the achievement tests administered at post level and to study about the efficacy of the experimentation.

**Mann-Whitney U Test**

According to **Kothari & Garg (2014) and Wikipedia (2016),** Mann-Whitney U Test is also called as Mann-Whitney-Wilcoxon (MWW) or Wilcoxon Rank-Sum Test or Wilcoxon-Mann-Whitney Test. Mann-Whitney U Test is a Non-parametric test of the null hypothesis that two samples come from the same population against an alternative hypothesis, especially that a particular population tends to have large values than the other. Unlike the t-test it does not require the assumption of Normal Distribution. It is nearly as efficient as the t-test on Normal Distribution. Hence, in this study Mann-Whitney Test is applied to the Post-test scores. And following are the steps followed as the procedure of a said method.

1. Rank all scores together, ignoring which group they belong to
2. Add up the Ranks separately for both the groups.
3. Select the larger Rank-sum as Tx
4. Derive n1=number of participants in first group; n2=number of participants in second group and nx= size of group whose Rank-sum is larger
5. Calculate U using a formula : U = n1* n2 + nx * (nx + 1)/2 - Tx
6. To be significant, calculated U has to be equal or less than this critical value.
7. The Critical (Tabled) Values for U are available only for small sample size that is up to n1=n2=20. For larger samples, U is considered approximately distributed and in such cases 'z' could be helpful to find out the significance level/s. Another, logic is made out looking to the table of Critical values of U that it increases with the increase in sample sizes. Thus, the highest Critical values of a table for U have been taken into consideration to check the fulfillment of the conditions. If not, then further z will be calculated for the same. The highest Critical values are as follows.

**Table – 3.13:**
**Highest Critical value of U in a U-Table For n1=n2=20**

| Critical U Values For sample size → n1 = n2 = 20 | | |
|---|---|---|
| **Significance Level** | **0.05** | **0.01** |
| **Values** | **127** | **105** |

8. As the critical values for the U in a table are limited up to the sample size n1 = n2 = 20. For the larger sample sizes, U is approximately normally distributed. In that case, the standardized value is,

$$\mathbf{z = (U - Mean_u) / SD_u}$$

Where, $Mean_u = n1*n2 / 2$

$SD_u = sqrt [ n1*n2*(n1+n2+1) / 12]$

OR

Calculating new Critical value U for large sample as below:

$\mathbf{U_{critical} = Mean_u - z * SD_u - 0.5}$ (here, 0.5 is the Correction factor)

Where, z values are as,

For Significance level = 0.05

z = 1.96 for two tailed test

z = 1.64 for one tailed test

For Significance level = 0.01

z = 2.58 for two tailed test

z = 2.33 for one tailed test

For the present research study, The MWU test has applied to test the null hypothesis framed for the chapter-wise achievement test-scores, overall test scores and SOLO level-wise score gained by both the groups conducted at post-level. Detailed analysis presented in a Chapter-V of this thesis.

**Rasch Analysis**

Moreover, further analysis was carried with partial use of **Rasch Model** or **Rasch Analysis** in terms to measure and study the achievements of the sample through all the five levels of SOLO Taxonomy. Many reviews found as mentioned in Chapter-II of this thesis had been experimented the SOLO Taxonomy and analyzed the data using Rasch Analysis in terms to study the outcomes. The Rasch Model is based on the Item Response Theory (IRT).

As Item Response Theory (IRT) as well as about the comparisons between IRT and Classical Test Theory (CTT) elaborated and reported by **Le (2013)** that how IRT is more beneficial over CTT within learning environments too and in terms to measures the person's responses with respect to the abilities. As reported, Item response theory (IRT) is a set of latent variable techniques especially designed to model the interaction between a subject's "ability" and the item level stimuli (difficulty, guessing, etc.). The focus is on the pattern of responses rather than on composite or total score variables and linear regression theory. In IRT the item responses are considered the outcome (dependent) variables, and the examinee's ability and the items' characteristics are the latent predictor (independent) variables. Classical test theory (CTT) was the dominant approach until 1953 when Frederic Lord published his doctoral dissertation on Latent Trait Theory. While CTT models test outcomes based on the linear relationship between true and observed score (Observed score =

True Score + Error), IRT models the probability of a response pattern of an examinee as a function of the person's ability and the characteristics of the items in a test or survey.

In the 2000s as per Le (2013), the IRT field was promoted by a new wave of researchers who not only expanded the technical aspects of the framework (estimation, model identification, and goodness of fit), but also advanced its computational aspects. IRT has been an extremely active area of research for more than half a century. The combination of methodological advances and increasingly powerful software has increased applicability and interest. IRT is widely used in assessment and evaluation research to describe the interaction between examinees and test questions. For many years CTT remained the dominant framework used in education despite the development and progress of IRT. Currently IRT is finding widespread application in the engineering of large-scale assessments as well as on a smaller scale in sociological and psychological assessments.

Furthermore, both CTT and IRT have advantages as well limitations respectively. But looking to the focus of the present research study, focusing on what benefits an IRT has over CTT. In the IRT framework, item characteristics are sample-independent and a person's latent scores are test-independent provided that the selected models fit the data well. Thus, scores that describe examinee proficiency are not dependent on test difficulty. Their scores may be lower on more difficult tests and higher on easier tests, but their ability scores remain constant over any test at the time of testing or surveying. IRT also permits calculation of the probability of a particular respondent selecting a category on a test item. IRT models are more complex and the parameter estimation methods often involve complicated numerical methods. Latent traits as well as item parameters can also be difficult to interpret both graphically and numerically. Rasch models (one-parameter models of IRT) are more straightforward to apply than other IRT models.

Hence, Le (2013) conclude as the IRT is a general framework for specifying mathematical functions that characterize the relationship between a person's ability or trait as measured by an instrument and the person's responses to the separate items in the instrument. In educational testing, IRT offers an alternative to classical test theory, which depends on total scores or number correct as outcome variables. IRT models have become a popular framework in many fields including psychology, nursing, and public health. But IRT-based analyses are still scarce in the social sciences and in the educational research literature.

Also, stated by **Cavanagh & Waugh (2011)** about Rasch model as, since 1960, when Georg Rasch (1901–1981) produced his now well-accepted measurement model published as the "Probabilistic Models for Some Intelligence and Attainment Tests", there has been a quiet revolution in measuring variables in education, psychology, business and medicine. Rasch's initial measurement model, now called the Simple Logistic Model of Rasch, only applied to items in a dichotomous format, such as no/yes; disagree/agree, wrong/right. That measurement model has now been extended to more than two scoring categories and is called the Partial Credit Model of Rasch or the Polytomous Logistic Model or the Rating Scale Model. Rasch measurement has also been extended to include a judges model and the Rasch Pair-Wise Comparison Model (where many people compare many essays, projects or assignments in pairs as worse/better).

Further, mentioned by **Wu & Adams (2007)** as, in the case of the Rasch model, the mathematical function of the item characteristic curve for a dichotomous item (A dichotomous item is one where there are only two response categories as correct and incorrect) is given by:

$$p = P(X = 1) = \frac{\exp(\theta - \delta)}{1 + \exp(\theta - \delta)}$$

Where,

- → X is a random variable indicating success or failure on the item. X=1 indicates success (correct response) on the item, and X=0 indicates failure (incorrect response) on the item.
- → θ is a person-parameter denoting the person's ability on the latent variable scale, and
- → δ is an item-parameter, generally called the item difficulty, on the same latent variable scale.

The above equation shows that the probability of success is a function of the difference between a person's ability and the item difficulty. When the ability equals the item difficulty, the probability of success is 0.5.

Moreover, looking to the Properties of the Rasch Model outlined by Wu & Adams (2007) are as: (a) the aforesaid equation has a special property called specific objectivity; (b) Indeterminacy of An Absolute Location of Ability; (c) Equal Discrimination; (d) Indeterminacy of An Absolute Discrimination; (e) Length of a logit; (f) Raw scores as sufficient statistics; (g) Fit of Data to the Rasch Model.

Thus, as learnt about the Rasch Model or Analysis, the researcher of the present research study was intended to analyse the data using Rasch Model in terms to have deeper insight or interpretations of the collected data and to examine the proceedings of the 'Progressive Understanding' through the levels of SOLO Taxonomy. The researcher had applied the Rasch Model in terms to do SOLO Level-wise analysis on the scores of Overall Achievement test. Detailed analysis as mentioned above are explained in a chapter on data analysis that is Chapter-V.

●●●●●●●●●●