CHAPTER III
CHAPTER III

DESIGN AND METHODOLOGY OF THE STUDY

3.1 INTRODUCTION

The methodology for this study was designed to determine the effectiveness of performance assessment in metacognition and academic achievement among preuniversity chemistry students. The population and sample, data collection procedures, instrumentation, and data analysis techniques are presented in this chapter.

3.2 LOCALE OF STUDY

3.2.1 City of Malayer in IRAN

Malayer is a city in western Iran, in the province of Hamedan. Malayer is a historical and cultural city. Malayer has existed at least since Sassanid times (3rd-7th century CE). It is 282 kms from the Iranian capital, Tehran. The second largest city of the province, Malayer, is located between Hamedan and Borujerd and South West of Tehran. The capital of Malayer, Hamedan, is 80 kms north of Malayer. Malayer has an estimated population of 450,000, in the Zagros mountains. Malayer is the center of a rich agricultural region. The weather in summer is warm and dry (max 28°C on Sun afternoon, min 20°C on Tue night) and in winter it is too cold. Malayer is a city with a rich culture in Iran. Most of the people in Malayer are educated.

3.3. VARIABLE SELECTED FOR THE STUDY

This study was undertaken to examine the impact of performance assessment on metacognitive skills and academic achievement among preuniversity chemistry students. The study has the following three-fold purposes: (I) to determine effect of performance assessment on metacognitive skills in preuniversities chemistry students, (II) to study the effect of performance assessment on academic achievement in preuniversities chemistry students, (III) To study the difference between boys and girls in metacognitive skills and academic achievement in chemistry. In this research, metacognitive skills and academic achievement are considered as dependent variables and performance assessment and gender are considered as independent variables. An attempt is made to survey the possible successful, significant and enlightening point
in this research to make recommendations for their applications to bring improvement in the assessment system in education.

**3.4 DESIGN AND METHODOLOGY OF THE STUDY:**

As mentioned in previous chapter the main object of present research is recognize effectiveness of performance assessment in metacognitive skills and academic achievement comparative with traditional assessment in pre university students in Malayer city- Iran.

To gain to this goal a plan with following characteristics would be suitable:

1. Sampling should be random.
2. A control group and an experimental group should exist.
3. In a specific period the experimental group should be put under independent variable.
4. A suitable tool measures and analyzes the changes made in the experimental group.
5. The difference between two groups should be compared with a proper statistic tool for the independent variable.

Based upon the factors mentioned above the quasi experimental method has been determined as a suitable method.

In this research with a design including two patterns of performance assessment and traditional assessment tries to do a quasi experimental research to determine the effect of these patterns on the students’ meta cognitive skills and academic achievement. From the variant quasi-experimental designs, non-equivalent pretest-posttest controls design seems very appropriate. The most common quasi-experimental research design consists of two groups of experimental and control. The proposed design is a two-factor design consisting of the independent variables of assessment patterns and gender as its factors. Given the design, the selected classes are randomized into two experimental and control groups. The meta-cognitive skills and chemistry academic achievement pretest was taken by both groups. After performing the pre-test of the academic achievement and meta-cognitive skills Inventory, the experimental group will be evaluated by performance assessment pattern. The students in the control group had their normal Chemistry classes with
their regular class teachers who were only requested to encourage the students to be serious in studying Chemistry for better achievement. Again, after performing the program, both groups will take academic achievement post test and meta-cognitive skills Inventory.

Table 3.1: Design of the study

<table>
<thead>
<tr>
<th>Randomly assigned</th>
<th>Pre-Test</th>
<th>Independent variable</th>
<th>Post-Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental group(e)</td>
<td>O1</td>
<td>Performance assessment</td>
<td>O1</td>
</tr>
<tr>
<td>Control group (c)</td>
<td>O2</td>
<td>Traditional assessment</td>
<td>O2</td>
</tr>
</tbody>
</table>

In this research, the researcher must make sure that these groups are homogeneous before the experiment. In this design, equal conditions for the variables of age, academic grade and course of study should be controlled.

Research design is a bi–factor design in which independent variables are the assessment patterns and the gender.

Table 3.2: Design of the study by group, gender and intervention

<table>
<thead>
<tr>
<th>Group</th>
<th>Gender</th>
<th>Pre test</th>
<th>Treatment</th>
<th>Post test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>Boy</td>
<td>Chemistry Academic Achievement</td>
<td>Performance Assessment Pattern</td>
<td>Chemistry Academic Achievement</td>
</tr>
<tr>
<td>Group</td>
<td>Girl</td>
<td>Meta-cognitive skills</td>
<td></td>
<td>Meta-cognitive skills</td>
</tr>
<tr>
<td>Control</td>
<td>Boy</td>
<td>Chemistry Academic Achievement</td>
<td>Traditional Assessment Pattern</td>
<td>Chemistry Academic Achievement</td>
</tr>
<tr>
<td>Group</td>
<td>Girl</td>
<td>Meta-cognitive skills</td>
<td></td>
<td>Meta-cognitive skills</td>
</tr>
</tbody>
</table>
In fact classes were used for the study in order not to disrupt administrative arrangement of the school. This became necessary as the study lasted for 4 months. The researcher will select sample subjects from population by random method and then randomly assign patterns to experimental group and control group.

**Figure 3.3: Design of the study**

3.5 TOOLS USED FOR THE STUDY:

The following tools were used for the collection of data to measure different variables of study in this research. Also, researcher did translation for utilize and easy understanding of each item of Meta-cognitive skills Inventory by teachers. Meta-cognitive skills questionnaires were translated from English language to Persian language.
Table 3.4: Tools used for the study

<table>
<thead>
<tr>
<th>TOOLS USED</th>
<th>VARIABLE MEASURED</th>
<th>CONSTRUCTED BY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Achievement Pre-Test</td>
<td>Academic Achievement</td>
<td>Were prepared by investigator</td>
</tr>
<tr>
<td>Academic Achievement Post-Test</td>
<td>Academic Achievement</td>
<td>Were prepared by investigator</td>
</tr>
</tbody>
</table>

3.5.1 ACADEMIC ACHIEVEMENT PRE-TEST IN CHEMISTRY (AAPC)

The academic Achievement pre-test in chemistry was a paper-pencil multiple-choice test with 30 questions prepared by the researcher and teacher and was used to measure the academic achievement of pre university students on the subject chemistry. In this study purpose of academic achievement pre-test was assessing learners’ mastering amount of learning objects at the first of program in pre university chemistry (Information, concepts, principals, rules and skills). Pre-test of chemistry was standardized by researcher. Context validity of the academic achievement pre-test was investigated by teachers who were professional in chemistry. The results have provided sufficient evidence for the context validity of this instrument. The reliability of this scale was investigated by researcher through statistical methods. Cronbach coefficient alpha reliability of chemistry academic achievement pre test was .83. Data analysis was performed by t test, ANCOVA test, and descriptive analysis of mean and standard deviation.

3.5.2 ACADEMIC ACHIEVEMENT POST-TEST IN CHEMISTRY (AAPC)

The academic Achievement post-test in chemistry was a paper-pencil multiple-choice test with 30 questions prepared by the researcher and teacher and was used to measure the academic achievement of pre university students on the subject chemistry. In this study purpose of academic achievement post-test was assessing
learners’ mastering amount of learning objects at the end of program in pre university chemistry (Information, concepts, principals, rules and skills) by standardized academic achievement test for chemistry. Context validity of the academic achievement post-test was investigated by teachers who were professional in chemistry. The results have provided sufficient evidence for the context validity of this instrument. Cronbach coefficient alpha reliability of chemistry academic achievement pre test was .81.

3.5.3 METACOGNITIVE SKILLS SCALE (MSS)

Even though it has been decades that the professionals of education have introduced the effects of Meta-cognitive skills on the academic achievement of the students, there haven’t been many standard tools to measure these skills. One of the useful tools to do this is questionnaire of analyzing Meta-cognitive skills of ONils and Abedi in 1996.

This instrument was designed by O’Nils & Abedi (1996) to measure four dimensions of meta-cognitive skills. The instrument includes 5 statements to measure planning (4,8,12,16,20), 5 statements to measure monitoring or self-checking (2,6,10,14,18), 5 statements to measure cognitive strategy (3,7,11,15,19) and 5 statements measure awareness (1,5,9,13,17). Students were asked to rate the statements on a 4-point scale. The rating was scored on Likert-type scale. The scale ranges from 1 for “Not at all”, 2 for “some what”, 3 for “moderately so” and 4 “very much”. The obtained score was then divided by the number of items on the subscale to obtain a mean score that reflected the original unit of measurement. This procedure allowed the researcher to make comparisons between the subscales. The possible scores for forth the subscales could range from 5 to 20, with higher scores indicating more efficacious in each subscale. The reliability (alpha above .70) of the instrument was initially established by the authors. Construct validity of meta-cognitive inventory was acceptable, but for performing in the Iranian context Cronbach coefficient alpha reliability of meta-cognitive skills questionnaire was investigated and reported .85. Construct validity of the instrument was investigated by factor analysis. The results have provided sufficient evidence for the construct validity of this instrument.
3.6 SAMPLE AND SAMPLING

Sampling is the selection of subjects from the population in such way that they are representative of the population. A sample is a smaller representation of the population it is small group, which representative all the traits and characteristic of the population under study. A good representative sample allows us to generalize our results to the population (Korn & Graubard, 2000). There are different types of sampling methods and most of them can be categorized into two:

1) Probability sampling methods
2) Non probability sampling

Probability sampling methods are likelihood of inclusion of each element or individual in the sample. Major probability sampling methods are; - simple random sampling – stratified random sampling- Area or cluster sampling and multi stage cluster random sampling. The sample numbers of experimental research have been suggested 15 people for each control and experimental groups (cohen & Manion 2000).

Non probability sampling is any sampling method where some elements of the population have no chance of selection (these are sometimes referred to as 'out of coverage'/under covered ’), or where the probability of selection can't be accurately determined. It involves the selection of elements based on assumptions regarding the population of interest, which forms the criteria for selection. Hence, because the selection of elements is nonrandom, non probability sampling does not allow the estimation of sampling errors. These conditions give rise to exclusion bias placing limits on how much information a sample can provide about the population. Information about the relationship between sample and population is limited, making it difficult to extrapolate from the sample to the population.

Important techniques of Non probability sampling methods are: - Quota sampling – Accidental sampling – purposive sampling – systematic sampling – snowball sampling – statistical sampling.

In the present study multi stage cluster random sampling was used. Multi-stage sampling involves selecting a sample in at least two stages. Multi-stage sampling is generally used when it is costly or impossible to form a list of all the units
in the target population. The other reason for choosing multi-stage cluster sampling in this study was doing the study in the second semester of the year so choosing the students in order to match them was not possible. Therefore the classes were selected as the unit of sampling. The population of this study was all of the students (1050 students) who were studying in chemistry in Malayer pre-universities (20 pre universities) in 2011-2012. The sample included for this research was 92 students (23 boys in experimental group and 23 boys in control group, 23 girls in experimental group and 23 girls in control group) from pre university students who were studying in 4 schools in Malayer city. At the first stage from among this 20 schools 4 schools were selected randomly (boys and girls pre university schools). At the second stage from each school 1 classes were selected and assigned to experimental and control group randomly. In addition, the participants were homogeneous in a number of controllable features, such as age, academic grade, field of study, intelligence, and, place of study. Number of participants in each group (experimental and control) was recommended to be 15 at least (Cohen & Manion, 2000), following other previously-conducted researches which have utilized the same sample size. The sample breakup is given below:

Table 3.5: Break up of the sample on the basis of gender

<table>
<thead>
<tr>
<th>Group</th>
<th>Gender</th>
<th>Pre test</th>
<th>Treatment</th>
<th>Post test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>Boy</td>
<td>Academic Achievement</td>
<td>Performance Assessment Pattern</td>
<td>Academic Achievement</td>
</tr>
<tr>
<td>Group</td>
<td>Girl</td>
<td>Meta-cognitive skills</td>
<td>Meta-cognitive skills</td>
<td>Meta-cognitive skills</td>
</tr>
<tr>
<td>Control</td>
<td>Boy</td>
<td>Academic Achievement</td>
<td>Traditional Assessment Pattern</td>
<td>Academic Achievement</td>
</tr>
<tr>
<td>Group</td>
<td>Girl</td>
<td>Meta-cognitive skills</td>
<td>Meta-cognitive skills</td>
<td>Meta-cognitive skills</td>
</tr>
</tbody>
</table>

3.7 TREATMENT PROCEDURES

The main goal of this study is to determine the effects of performance-based assessment model on developing Meta-cognitive skills and academic achievement of the students. In order to do this, traditional and performance-based assessment
methods have been compared with each other. Then the data was gathered by questionnaires and tools that have been explained previously, in different stages.

After choosing the sample, in the first step the teachers were acquainted to the performance-based assessment method and the experimental group students have also been completely justified on the new method and their participation. Before using the method two pretests were administrated on the students: I) Meta-cognitive skills and II) Achievement test in chemistry. Because the test was done in the second half of the academic year, the chemistry pretest was only from the first half of the book. The tests were same in this stage for both groups.

In the second step, in both of the classes of experimental groups (male and female), the teaching and the other activities of the teachers by performance-based continued for 4 months (each week two hours) to assess level of learning of the students and planning on reactions to improve their learning.

The intervention program of the experimentation consists of assessment items, measuring components of meta-cognitive skills:

1) Planning, example: *For doing an experiment in chemical reaction which instrument do you need.*

2) Monitoring, example: *Explain and experiment that when a candle burns, it reacts with the oxygen in the air.*

3) Cognitive strategy, example: *Explain another way to say that no atoms are created or destroyed in a chemical reaction.*

4) Awareness, example: *Are atoms created or destroyed in a chemical reaction?*

   *Student: No*

   *Teacher: How do you know?*

In the third step at the end of the semester, post-test in meta-cognitive skills and academic achievement were done on both groups in the same condition. The chemistry post-test was contained the second half of the book. Table 3.6 shows an example of intervention in experimental groups.
Table 3.6 An example of intervention in experimental groups

<table>
<thead>
<tr>
<th>Objectives</th>
<th>Content</th>
<th>Pre test and formative performance assessment</th>
<th>Experiment and Activities</th>
<th>Summative performance assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Students will be able to:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Explore chemical energy</td>
<td>• Storing energy in physical and chemical changes</td>
<td>• Implementation of meta-cognitive skills and chemistry achievement pre test in the first session of study design (In experimental and control groups)</td>
<td>- <strong>Storing energy in physical changes:</strong></td>
<td>- <strong>Final assessment in classroom</strong></td>
</tr>
<tr>
<td>• Write an energy equation</td>
<td><strong>Key concepts:</strong></td>
<td>• Implementation of performance assessment in the experimental groups.</td>
<td>• The rubber band experiment</td>
<td>• Teacher asks the students with doing an experiment write a response to this question and report the results to the class: “Where is the energy involved in physical and chemical changes stored?”</td>
</tr>
<tr>
<td>• Recognize that energy is chemical changes</td>
<td>• Energy definition</td>
<td>Questions related to storing energy in physical changes:</td>
<td>- <strong>Storing energy in chemical changes:</strong></td>
<td><strong>Project Beyond School</strong></td>
</tr>
<tr>
<td></td>
<td>• Chemical changes</td>
<td>- Question 1: “Stretch and relax a rubber band without letting it go. What differences, if any, do you feel between its stretched and relaxed states?”</td>
<td>• Chemical Reactions and Stored Energy experiment</td>
<td>• Teacher asks the students to contact the local power company to determine where and how the electrical</td>
</tr>
<tr>
<td></td>
<td>• Physical changes</td>
<td>- Question 2: “Hold the rubber band without stretching it, then let it go. What happens?”</td>
<td>Activity:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Energy saving</td>
<td></td>
<td>Students have to write word equations for the energy changes in a chemical reaction.</td>
<td></td>
</tr>
</tbody>
</table>

- **Storing energy in form of electromagnetic”**
- **Question 3:** “Stretch the rubber band and release it towards a wall (not towards any person). What happens now?”

- **Question 4:** “What difference is there?”

*Questions related to storing energy in chemical changes:*

- **Question 5:** “What was common to all four reactions?” [Light, heat, energy released.]

- **Question 6:** “Where was the energy stored?” [Match—sulfur; Lighter—butane; Paper—Carbon bonds; Wood—carbon bonds.]

*Questions related to storing energy in form of electromagnetic:*

- **Question 7:**
  “Does energy have mass?”

- **Question 8:**
  “What did you observe as you placed the magnets farther and farther apart?”

<p>| •The magnet experiment | energy you use in your home is harnessed. |</p>
<table>
<thead>
<tr>
<th>Question 9:</th>
<th>“What can you conclude about the amount of energy stored in the magnets when they are pushed close together compared to when they are held not so close?”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 10:</td>
<td>“Did the magnets store energy?”</td>
</tr>
<tr>
<td>Question 11:</td>
<td>“What evidence do you have of this?”</td>
</tr>
</tbody>
</table>
The stages of performance assessment in the class were done based on the meta-cognitive Choys' model (2006).

The learning assessment model that was used in this study is contained 7 stages. These 7 stages were expanded by an example in chemistry to be used in classes of experimental groups as follows:

An example of teaching and assessment in experimental groups

**Subject: Energy in chemical and physical changes**

1. Determining the learning objectives.

**Objectives**

- Students will be able to explore chemical energy.
- Students will be able to write an energy equation.
- Students will be able to recognize that energy is conserved during chemical changes.

2. Determining experiments and activities in the class.

**Experiments**

- The rubber band experiment
- Chemical Reactions and Stored Energy experiment
- The magnet experiment

**Activity**

*Students have to write word equations for the energy changes in a chemical reaction.*

3. Introducing the learning subject in form of a problem.

*Teacher tells the students* “Today we are going to explore energy. We are going to try to answer this question:

**Problem:** Where is the energy involved in physical and chemical changes stored?”

4. Implementation of meta-cognitive skills and chemistry achievement pre test in the first session of study design (In experimental and control groups).

*To assess students’ prior knowledge, teacher asks them to answer the initial questions about energy.*
5. Implementation of performance assessment in the experimental groups.

5.1 Questions related to storing energy in physical changes:

In order to implement of curriculum, teacher divides students into groups and asks them to perform the rubber band activities in their Student Pages. Remind students to read the questions thoroughly and answer them completely. Question #1: “Stretch and relax a rubber band without letting it go. What differences, if any, do you feel between its stretched and relaxed states?” Question #2: “Hold the rubber band without stretching it and then let it go. What happens?” Question #3: “Stretch the rubber band and release it towards a wall (not towards any person). What happens now?” Question #4: “What difference is there?” Teacher asks the students to fill in chart.

5.1.1 Teacher guidance and students evaluation

When students have completed the rubber band experiment, teacher says: “If you had a hard time answering this last question, you are not alone. Here is how scientists explain rubber bands flying across the room: When you stretch a rubber band you give it energy. When you let the rubber band go, that energy is responsible for it flying across the room and perhaps hurting someone or possibly moving something small that it hits. But you can feel that it is stored in the rubber band when you hold it apart, because it takes some work for you to stretch it. Stored energy is converted into energy of motion when the rubber band leaves your hand and begins to fly across the room. When it hits something, its energy of motion can be transferred to another object to make it move (or it can sting someone if it hits them).”

5.2 Questions related to storing energy in chemical changes:

Teacher ask students to do the “Chemical Reactions and Stored Energy” experiment. Discuss the answers to questions: Question #5: “What was common to all four reactions?” [Light, heat, energy released.] Question #6: “Where was the energy stored?” [Match—sulfur; Lighter—butane; Paper—Carbon bonds; Wood—carbon bonds.]

5.2.1 Teacher guidance and students evaluation

When students have completed the chemical reactions experiment, teacher says: “The idea of storing energy is very important in explaining where the heat and light of fires
comes from. Much more heat and light is given off by any burning reaction than is used to get it started.”

Teacher asks: “But what does it mean to say that energy is stored in butane or in wood, or in the match compound? Those chemicals (butane, wood, paper) are not like a rubber band. You cannot feel the stored energy in wood! The wood, paper, match, or butane were not hot or glowing before they burned. Is there a clue as to where the energy came from?” [Students should be able to conclude that the end produces are different from the reactants. For example, they get black carbon instead of light colored paper. The products are different, so a chemical change has occurred. The bonds between the atoms have been changed, and the atoms rearranged into new substances. From this students can conclude that the energy may be in the bonds.]

5.3 Questions related to storing energy in form of electromagnetic

Teacher asks students to do the magnet experiment and use the following questions to generate a class discussion: Question #1: “Does energy have mass?” Question #2: “How do you know?” Question #3: “Can energy be created?” Question #4: “What happens?” [They repel.] Question #5: “What happens?” [The repelling force becomes weaker.] Question #6: “What did you observe as you placed the magnets farther and farther apart?” Question #7: “What can you conclude about the amount of energy stored in the magnets when they are pushed close together compared to when they are held not so close?” [Force is weaker.] Question #8: “Did the magnets store energy?” [Yes.] “What evidence do you have of this?”

5.3.1 Teacher guidance and students evaluation

When students have completed the magnet experiment, teacher says: “Magnets have stored energy, too. This energy is in the structure of their molecules, just as there is energy in the structure of wood, paper, or butane. But in the case of the magnet, the energy is electromagnetic. It is the result of the movement of electrons, and the substance does not have to be changed or rearranged for us to see the result of the energy.”
6. Discovering and analyzing the problems of learning.

6.1 Problem of learning

A common student misconception is that energy is made of matter and therefore that it has mass. Mass can be converted to energy in nuclear reactions, but this conversion is never seen in normal life on Earth.

6.1.1 Developing a plan for learning problems

To help students construct new knowledge, teacher asks them to think about types of energy (heat, light, sound, energy of motion, stored energy). Then ask them: “How could you find the mass of heat, light, or sound?” Students should realize that it is impossible to do so and that energy is not made of matter.

6.2 Problem of learning

Another common student misconception is that energy is created; for instance, many students believe that power plants create energy. Students need to come to the scientific conception that energy can only be converted from one form to another. Often they fail to recognize that kinetic energy harnessed from a moving river or from the wind is converted into electrical energy via turbine-driven generators, or that stored energy in coal or uranium is released to heat water, which moves turbine-driven generators.

6.2.1 Developing a plan for learning problems

The application beyond school portion of this lesson may help students overcome this misconception.

Activity: Teacher asks the students to write word equations for the energy changes in a chemical reaction. These are qualitative statements. After students express the combustion of paper in a word equation, teacher provides a few more examples orally. [Energy in food = Heat energy + Muscle Movement Energy].

At the end of the class teacher asks the students to fill in the summary at the end of the Student Pages.

**7.1 Final assessment in classroom**

- *Teacher asks the students with doing an experiment write a response to the question in Step 1 and report the results to the class: “Where is the energy involved in physical and chemical changes stored?”*

**7.2 Project Beyond School**

*Contact your local power company to determine where and how the electrical energy you use in your home is harnessed.*

These steps were followed in all of the sessions in experimental groups and diagnostic and formative performance-based assessment has been going on all the way, also performance assignments were given to students as individual and grouping activities.

**Scoring plan:**

In this study the teacher used analytical assessment method to give marks to students. In this method the teacher indicates the components of the answer and gives each part a different point. For example, a written report of an experiment, teacher can use accuracy of data, quality of the results analysis and the method of results justification as points to give mark. For example to score summative performance assessment can use the following frame:

<table>
<thead>
<tr>
<th>Components</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy of response</td>
<td>2</td>
</tr>
<tr>
<td>Quality of result analysis</td>
<td>1</td>
</tr>
<tr>
<td>Result justification</td>
<td>1</td>
</tr>
<tr>
<td>Report frame</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table 3.7: Scoring plan**

<table>
<thead>
<tr>
<th></th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>5</td>
</tr>
</tbody>
</table>
3.8 STATISTICAL TECHNIQUES EMPLOYED FOR ANALYSIS OF THE DATA

The data collected was analyzed using the following statistical techniques:

1. Descriptive statistics like mean and standard deviation have been employing to analyze the data for the descriptive part of the study.

2. Analysis of Covariance (ANCOVA), and dependent samples t-test that brief description of these techniques are given below:

1. **Analysis of covariance (ANCOVA):** Covariance is a measure of how much two variables change together and how strong the relationship is between them (Howell, 2009). ANCOVA is a general linear model which blends ANOVA and regression. ANCOVA evaluates whether population means of a dependent variable (DV) are equal across levels of a categorical independent variable (IV), while statistically controlling for the effects of other continuous variables that are not of primary interest, known as covariates (CV). Therefore, when performing ANCOVA, we are adjusting the DV means to what they would be if all groups were equal on the CV (Keppel, 1991).

   ANCOVA can be used to increase statistical power (Tabachnick, 2007), (the ability to find a significant difference between groups when one exists) by reducing the within-group error variance. In order to understand this, it is necessary to understand the test used to evaluate differences between groups, the F-test. The F-test is computed by dividing the explained variance between groups (e.g., gender difference) by the unexplained variance within the groups. Thus, if this value is larger than a critical value, we conclude that there is a significant difference between groups. Unexplained variance includes error variance (e.g., individual differences), as well as the influence of other factors. Therefore, the influence of CVs is grouped in the denominator. When we control for the effect of CVs on the DV, we remove it from the denominator making F larger, thereby increasing your power to find a significant effect if one exists.

   Another use of ANCOVA is to adjust for preexisting differences in nonequivalent (intact) groups. This controversial application aims at correcting for initial group differences (prior to group assignment) that exists on DV among several intact groups. In this situation, participants cannot be made equal through random
assignment, so CVs are used to adjust scores and make participants more similar than without the CV. However, even with the use of covariates, there are no statistical techniques that can equate unequal groups. Furthermore, the CV may be so intimately related to the IV that removing the variance on the DV associated with the CV would remove considerable variance on the DV, rendering the results meaningless (Miller, 2001).

**Assumptions for ANCOVA:**

There are four assumptions that underlie the use of ANCOVA and affect interpretation of the results:

**Assumption 1: Randomness and Independent Sampling**

Observations must be randomly sampled from the population and independent from each other. If this assumption is violated, the test will produce inaccurate results.

**Assumption 2: Normality**

There must be a normal distribution of the DV in the population. In the event that a distribution that is non-normal (e.g., skewed or kurtotic) and sample sizes are small, p-values may be invalid (Green, 2011).

**Assumption 3: Homogeneity of Variances**

The variances of the DV must be equal for all levels of the IV and the CV.

**Assumption 4: Homogeneity of Regression Slopes**

The slope of the line predicting the DV from the CV must be equal for each level of the IV. That is, the CV must not have differential effects on the DV at different levels of the IV. This assumption is violated when there is a significant interaction between the IV and the CV. If this assumption is violated, ANCOVA should not be performed (Engqvist, 2005). If the correlations of the covariates with the DV are very different in different cells of the design, gross misinterpretations of results may occur. In ANCOVA, we basically perform a regression analysis within each cell to partition out the variance component due to the CV. The homogeneity of slopes assumption implies that we perform this regression analysis subject to the constraint that all regression equations (slopes) across the cells of the design are the same. If this is not the case, serious biases may occur.
Covariate analysis is similar to variance analysis; however, in covariate analysis any other variables that may have relationship with dependent variables (except of independent variables) are considered and controlled. In this study, IQ and prior learning's of students were controlled by covariate analysis. In the present study ANCOVA was employed to compare differences between performance assessment group and traditional assessment group based on the gender with controlling effect of pretest.

2. T-Test

The T-Test assesses whether the means of two groups are statistically different from each other. This analysis is appropriate whenever you want to compare the means of two groups, and especially appropriate as the analysis for the posted – only two – group randomized experimental design. There are two main types of T-test: Independent sample T-test and paired sample T-test. An independent sample T-test is used when we want to compare the mean score, on some continues variable for two different groups of subjects. Paired sample T-test is used when we have only one group of people and we collect data from them under two different conditions. This approach is also used when we have matched pairs of subjects (Pallant, 2001). The T-test assesses the significance of the difference between the means of two groups or two sets (Before and after) of scores (Somer & Somer, 1986). If the T-value exceed a cut off point (depending on degrees of freedom) the difference in the mean is significant. When the T-test value is below to cut off point the difference is said to be not significant. In this research dependent sample t-test is used to compare the scores before and after implementation of performance assessment.
<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Statistical analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>There is a significant difference between the performance assessment group and traditional group in meta-cognitive skills.</td>
<td>ANCOVA</td>
</tr>
<tr>
<td>There is a significant difference between boys and girls in meta-cognitive skills.</td>
<td>ANCOVA</td>
</tr>
<tr>
<td>There is a significant difference between the performance assessment group and the traditional assessment group in monitoring sub scale.</td>
<td>ANCOVA</td>
</tr>
<tr>
<td>There is a significant difference between boys and girls in monitoring sub scale.</td>
<td>ANCOVA</td>
</tr>
<tr>
<td>There is a significant difference between the performance assessment group and the traditional assessment group in planning sub scale.</td>
<td>ANCOVA</td>
</tr>
<tr>
<td>There is a significant difference between boys and girls in planning sub scale.</td>
<td>ANCOVA</td>
</tr>
<tr>
<td>There is a significant difference between the performance assessment group and the traditional assessment group in cognitive strategy sub scale.</td>
<td>ANCOVA</td>
</tr>
<tr>
<td>There is a significant difference between boys and girls in cognitive strategy sub scales.</td>
<td>ANCOVA</td>
</tr>
<tr>
<td>There is a significant difference between the performance assessment group and the traditional assessment group in awareness sub scale.</td>
<td>ANCOVA</td>
</tr>
<tr>
<td>There is a significant difference between boys and girls in awareness sub scales.</td>
<td>ANCOVA</td>
</tr>
<tr>
<td>There is a significant difference between the experimental group and control group in academic achievement in chemistry.</td>
<td>ANCOVA</td>
</tr>
<tr>
<td>There is a significant difference between boys and girls in academic achievement in chemistry.</td>
<td>ANCOVA</td>
</tr>
</tbody>
</table>