IMPACT OF FORMATIVE EVALUATION AND FIXED INTERVAL SCHEDULE REINFORCEMENT ON ACADEMIC ACHIEVEMENTS OF SECONDARY SCHOOL STUDENTS

A thesis submitted to the
NATIONAL UNIVERSITY OF MODERN LANGUAGES
Islamabad, Pakistan

By
TAHIR MEHMOOD

in partial fulfillment of the requirement for the degree of
Doctor of Philosophy in Education
Faculty of Advanced Integrated Studies and Research

NATIONAL UNIVERSITY OF MODERN LANGUAGES
Islamabad, Pakistan
December 2007
THESIS/DISSERTATION AND DEFENCE APPROVAL FORM

The undersigned certify that they have read the following thesis, examined the defence, are satisfied with the overall exam performance, and recommend the thesis to the Faculty of Advanced Integrated Studies & Research for acceptance:


Submitted by: Tahir Mehmood
Name of Student

Registration No: 080-PhD/Edn/2002

Doctor of Philosophy
Degree Name in Full (e.g Master of Philosophy, Doctor of Philosophy)

Education
Name of Discipline

Professor Brig ® Dr. Allah Bakhsh
Name of Research Supervisor

Signature of Research Supervisor

Professor Dr. Shazra Munnawer
Name of Dean (FASIR)

Signature of Dean (FASIR)

Brig ® Dr. Aziz Ahmad
Name of Rector

Signature of Rector

Date
Title “Impact of formative evaluation and fixed interval schedule reinforcement on academic achievements of secondary school students”

The study was designed to find out the impact of formative evaluation and fixed interval schedule reinforcement on academic achievements of secondary schools students. It was experimental design and completed in two phases with the duration of 120 days, 60 days for each phase. In first phase experimental group was evaluated by formative evaluation and in second phase experimental group was evaluated formatively and also was given fixed interval schedule reinforcement. In first phase 60 students were randomly selected from 9th class of Govt. High School Fateh Jang. This sample was divided into two equal groups. One served as control group and other as experimental group. A pretest was conducted to both the groups in the subjects of Math and English. Both groups were taught by subject specialist of Math and English. Control group was only taught and was not given any treatment whereas experimental group was taught and evaluated by teacher made test during session. After completion of 60 days teaching, a post test was conducted in Math and English.  

In second phase 60 students were randomly selected from 10th class of Govt. High School Fateh Jang. This sample was divided into two equal groups. One served as control group and other as experimental group. A pretest was conducted to both the groups in the subjects of Math and English. Both groups were taught by subject specialist of Math and English. Control group was only taught and was not given any treatment whereas experimental group was taught and evaluated by teacher made test and reinforcement was given with a fixed interval schedule, after every six days, during session. After completion of 60 days teaching, a post test was conducted in Math and English. It was found that formative evaluation and fixed interval schedule reinforcement has significant relationship with learning and academic achievements in the subjects of Math and English at secondary school level. It is recommended that educational institutional management may arrange capacity building programs for teachers to familiarize them with evaluation, its importance and its process. The
managers of educational institutions may monitor teaching learning process with a view to implementing formative evaluations policy. Curriculum designers and developers may design and develop formative evaluation tools within curriculum documents for each subject keeping in view the requirements of fixed interval schedule reinforcement in the manuals for textbook instruction.
CANDIDATE DECLARATION FORM

(Declaration Form to be filled in by the Candidate at the time of Submission of Thesis to the Supervisor for Internal Evaluation. Follow this pattern strictly, and also let the dotted lines appear on the page)

I ______________ Tahir Mehmood

Daughter/Son of ___ Abdul Rehman

Registration # ___080-PhD/Edn/2002

Discipline ________ Education

Candidate of Degree of Doctor of Philosophy at the National University of Modern Languages do hereby declare that the thesis titled “Impact of Formative Evaluation and Fixed Interval Schedule Reinforcement on Academic Achievements of Secondary School Students” submitted by me in partial fulfillment of PhD degree is my original work, and has not been submitted or published earlier. I also solemnly declare that it shall not, in future, be submitted by me for obtaining any other degree from this or any other university or institution.

I also understand that if evidence of plagiarism is found in my thesis/dissertation at any stage, even after the award of a degree, the work may be cancelled and the degree revoked.

________________________                     ______________________________________
Date                                                                                      Signature of Candidate

Tahir Mehmood  .
Name of Candidate
<table>
<thead>
<tr>
<th>Chapter</th>
<th>INTRODUCTION</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.2 Statement of the problem</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>1.3 Rationale of the problem</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>1.4 Theoretical Framework</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>1.5 Objectives of the study</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>1.6 Significance of the study</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>1.7 Limitations</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>1.8 Hypotheses</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>1.9 Procedure of the study</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>1.10 Population</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>1.11 Sampling techniques</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>1.12 Instrument development</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>1.13 Instrument validation</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>1.14 Data analysis</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>2 REVIEW OF THE RELATED LITERATURE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.1 Evaluation</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>2.2 Cognitive and Educational Theories</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>2.3 Learning in complex domain</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>2.4 Educational evaluation</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>2.5 Formative evaluation</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>2.6 Learning through formative assessment</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>2.7 Student attitudes towards formative Evaluation</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>2.8 Training students as formative evaluators</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>2.9 Over view of cognitive and learning science theories</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>2.10 Fostering and assessing competence</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>2.11 Teachers and teaching</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>2.12 Scope of Modern evaluation</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>2.13 Recent trends in evaluation</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td>2.14 Guiding as a trend in evaluation</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>2.15 Motivation of learning by examinations</td>
<td>116</td>
</tr>
<tr>
<td></td>
<td>2.16 Evaluation as appraisal</td>
<td>118</td>
</tr>
<tr>
<td></td>
<td>2.17 Simulation-based learning</td>
<td>121</td>
</tr>
<tr>
<td></td>
<td>2.18 Evaluation of performance</td>
<td>123</td>
</tr>
<tr>
<td></td>
<td>2.19 Simulators</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td>2.20 Cognitive load theory</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>2.21 Learning layers of complexity</td>
<td>130</td>
</tr>
</tbody>
</table>
2.22 Evaluation for conflict resolution in curriculum……. 132
2.23 Evaluation and emotional intelligence…………… 134
2.24 Evaluation and empirical research……………….. 137
2.25 Process evaluation……………………………. 139
2.26 Evaluation and communication…………………… 141
2.27 Evaluation and speaking competencies………… 144
2.28 Criterion-referenced evaluation…………………… 146
2.29 Evaluation and students intellectual development… 148
2.30 Self-reactions and adaptations in learning………2 152
2.31 Evaluation and teacher qualification……………… 154
2.32 Instruction and assessment……………………… 158
2.33 Behavioral model of reinforcement……………… 159
2.34 Reinforcement……………………………………. 159
2.35 Schedules of reinforcement……………………… 162
2.36 Research and reinforcement……………………. 171
2.37 Experiments in reinforcement…………………… 177
2.38 Dynamic reinforcement schedules………………. 179
2.39 Principles of reinforcement……………………… 187
2.40 Process of reinforcement……………………….. 188
2.41 Challenges in reinforcement learning……………. 193
2.42 Reinforcement learning methods…………………. 197

3 METHOD AND PROCEDURE
3.1 Research design……………………………………. 206
3.2 Population…………………………………………. 207
3.3 Sample ……………………………………………. 208
3.4 Instrument development………………………….. 209
3.5 Instrument validation……………………………… 213
3.6 Experimentation…………………………………2 214
3.7 Data analysis………………………………………. 215

4 ANALYSIS OF DATA
4.1 Descriptive statistics for phase I…………………. 226
4.2 Descriptive statistics for phase I I……………….. 254

5 SUMMARY,FINDINGS,CONCLUSIONS AND RECOMMENDATIONS
5.1 Summary…………………………………………. 282
5.2 Findings…………………………………………. 284
5.3 Discussion………………………………………. 293
5.4 Conclusion………………………………………. 295
5.5 Recommendations ………………………………. 296

BIBLIOGRAPHY 299
ACKNOWLEDGMENTS

It is all glory and praise to Almighty Allah who bestowed his blessings upon me in every moment of life. He also enabled me to complete my thesis.

I am greatly indebted to my respected and honorable supervisor, Brig ® Dr. Allah Bakhsh Malik for his full time valuable guidance. His professional attitude became the constant source of inspiration throughout my research work. His advice strengthened me to overcome all the obstacles that came in my way. He is born researcher and has the capability to infuse the spirit of research in his students. He has dedicated himself for the promotion of educational research.

I feel honor to express gratitude to worthy Rector, Brig ® Dr. Aziz Ahmad and Dean FAIS, Professor Dr. Shazra Munnawer for their kind guidance and encouragement during research process.

I am really indebted to the love and kindness of my mother, brothers and sisters who wished and prayed for me and remained a constant source of support and encouragement to improve my education.

Tahir Mehmood
CHAPTER 1

1.1 Introduction

Education has an enduring impact on human life. It is both innate and acquired that gradually brings changes in the personality of the learner. It means differently to different people and is defined according to the needs and goals set by an individual or a society or a nation. It is a life long process which helps in the complete growth of an individual. Educational system has some defined goals to achieve; it may be social efficiency, good citizenship or development of desired personality.

In Pakistan, various commissions, committees and study groups were constituted from time to time to improve the standard of educational system at all levels. The reports, so produced have tried to analyze the factors which influence the standard of education, e.g., physical facilities, teachers attitude, student attendance, socio-economic background and the capabilities of the students, age group, teaching methods and aids, curriculum, college and home environment, methods of evaluation and examination system.

Evaluation in all forms is important because every thing else orbits around it. Evaluation is significant in any system of education for determining the individual differences in attainment. Economic Survey of Pakistan (2001-02) stated that in the past three decades a gradual improvement in the examination system corresponding to different stages of education has taken place. Teaching and evaluation are interdependent as such one can not think of one without the other. Evaluation is helpful to both teachers and the students.
Norman (1990) defined evaluation as, “Evaluation is a systematic process of collecting analyzing and interpreting information to determine the extent the pupils are achieving instructional objectives.(p.21). Evaluation is considered as a powerful educational tool. It can provide information on student academic achievements with which students can glimpse their own progress. It can allows teachers to monitor the progress of individual student and also to obtain evidence about the efficacy of their own teaching.

Cizek (1997) portrayed, “evaluation programs for whole education systems specify what the system takes to be important for students to learn. That specification can provide a very salient indication of where schools and teachers should direct major effort”(p.67). Evaluation in schools is basically concerned with two major approaches for making decisions, formative evaluation and summative evaluation.

Formative evaluation aims at ensuring a healthy acquisition and development of knowledge and skills by students. Formative evaluation is also used to identify student needs in order to guide them towards desired goals. As student needs and difficulties are identified, appropriate remedial measures are taken to solve such problem. The purpose is to find out whether after learning experience students are able to do what they were previously unable to do. Impara (1996) placed it as, “the primary purpose of formative evaluation is to help as much as possible to ensure that summative evaluation comes out positive. It is a process of channeling input variables through a process that will yield expected outputs” (p.23).

The classroom teacher plays role as the best formative evaluator. Formative evaluation is type of assessment for which serves the purpose of promoting pupils’
learning. It is different from assessment designed primarily to serve the purposes of accountability, or of ranking, or of certifying competence.

Black (1998) elucidated, “an assessment activity can help learning if it provides information to be used as feedback, by teachers, and by their pupils in assessing themselves and each other, to modify the teaching and learning activities in which they are engaged. Such assessment becomes ‘formative assessment’ when the evidence is actually used to adapt the teaching work to meet learning needs”. (Pp.2-3)

Stiggins (2000) explained the relation between evaluation and learning, “Evaluation depends on learning and learning may be enhanced by reinforcement. In operant conditioning, reinforcement is an increase in the strength of a response following the presentation of a stimulus contingent on that response”. The relationship between response and frequency is also important in this whole process. Response strength can be assessed by measures such as the frequency with which the response is made or the speed with which it is made. The stimulus and response both also play important role in reinforcement process. Stiggins (2000) also defined the reinforcer as, “The stimulus contingent on a response is called a reinforcer. Reinforcement can only be confirmed retrospectively, as objects, items, food or other potential 'reinforcers' can only be called such by demonstrating increases in behavior after their administration. It is the strength of the response that is reinforced, not the organism”.

Schedules of reinforcement are the specific rules that are used to present or to remove reinforcers subsequent to a particular operant behavior. These rules are elucidated in terms of the time or the number of responses required in order to present or to remove a reinforcer. Different schedules of reinforcement create individual effects on operant
behavior. Interval schedules require a minimum amount of time that may pass between consecutive reinforced responses. Responses which are made before this time has elapsed are not reinforced. Interval schedules may specify a fixed time period between reinforcers or a variable time period between reinforcers.

Fixed interval schedules create an increased rate of response as the time of reinforcement approaches. Reinforcement plays very important role in learning. The study made attempt to unearth the liaison of formative evaluation and fixed interval schedule reinforcement with learning.

1.2 Statement of the Problem

The problem under investigation was to find out the impact of formative evaluation and fixed interval schedule reinforcement on academic achievements of secondary school students.

1.3 Rationale of the Problem

The Government’s Education Policies 1992, 1998-2010 and the Education Sector Reforms (ESR) Action Plan 2001-2005 emphasize the quality of education as an important factor in national development. The monitoring of academic achievements of students is one of the key components to evaluate and to improve the quality of education. National Education Assessment System (NEAS) is one of the key programs of the ESR under Quality Assurance innovative programme agenda of the Government. The program would enlarge national capacity for monitoring the academic achievements of elementary students in order to advance the quality of services i.e., curriculum, textual material, teachers’ delivery, policy formulation, etc in the education sector.
One of the objectives of the National Education Policy 1998-2010 and the Education Sector Reform was to build assessment capacity at the school, provincial and federal level to better measure learning outcomes, improve the quality and effectiveness of program interventions. The Ministry of Education played a leading role in this scheme to incorporate Provincial/Area assessment activities to direct them to a common purpose and standard of measurement so that national level assessments of student achievement could be made on a periodic basis.

Specifically, the objectives of the project are:

1. to institutionalize a monitoring system which permits and encourages continuous educational improvement at the elementary level;

2. to develop institutional capacity for test development, test administration, statistical analysis and report writing targeted at stakeholders, capacity developers;

3. to establish a baseline of student achievement and develop institutional capacity to conduct periodic assessments for obtaining monitoring indicators;

4. to analyze student performance with reference to variation in instructional context, student background and other factors affecting student achievement in order to identify the effectiveness of educational inputs and interventions;

5. to identify strong and weak areas of student learning with reference to the curriculum and target competencies, for appropriate action;

6. to assess performance of teachers and educational institutions by relating it to the learning achievements of students;
7. to provide feedback to policy makers and frontline implementers to enable them to develop effective strategies and action plans for improving the quality of education;
8. to inform parents, community members and other stakeholders about the quality of education;
9. to establish professional linkages with relevant national and international institutions;
10. to promote uniformity of academic standards in accordance with national and international institutions.

The study will provide a baseline for accomplishment of objective # 3 of above mentioned project NEAS.

1.4 Theoretical Framework

Evaluation is the mode used to establish the worth or value of an instructional program. It is also a practice to develop the teaching-learning process or to fix on whether or not to continue it. Educational evaluation can also be defined as the process of delineating, collecting, and providing information useful for judging teaching-learning decision alternatives (Stiggins, 1992).

Evaluation involves making judgments about the exactness of the strategy and mainly about the outcomes or explicit benefits that the approach delivers. "Worth" is more than just fiscal worth. Rather, it answers the questions of "Is training achieving the results that it was set up to achieve?", "Are the actual results worth having?", and
"Were the results achieved by the most cost-effective methods?" (Newby, 1991). Educational evaluation has two types; formative and summative evaluation

Formative Evaluation occasionally known as interim evaluation, it is carried out during implementation phase of projects or programmes for improving their performance. Formative evaluations are also conducted for some other reasons such as compliance or legal requirements. It is anticipated for managers and direct supporters of a project. Formative evaluation, including pre-testing, is designed to evaluate the strengths and weaknesses of materials or operational strategies before execution. It permits necessary reviews before the full effort goes forward. Its basic function is to maximize the chance for program success prior to the communication activity starts. Evaluation of a program in its developmental stages is also called formative evaluation. In the instructional design process, formative evaluation takes place earlier than the final product completion.

Most frequently outcome of formative evaluation results in changes to the instructional program for making it more effective. Formative evaluation is also an assessment of efforts earlier to their completion for the rationale of improving the efforts. It is a practice that has happened to well developed in the education and training evaluation literature.

Summative evaluation comprised of all those activities which judge the worth of a completed program. In the instructional design procedure, summative evaluation is often considered as the final stage. Summative evaluation generally does not result in changes to the instructional program being evaluated; it most often guides the trainer and organization about students learned behaviors from the training. It provides information
on the product's efficiency, its capability to do what it was designed to do. For example, did the learners become skilled at what they were expected to learn after using the instructional module. Logically, it helps the learners know "how they did," but more significantly, by looking at how the learner's did, it helps you know whether the artifact teaches what it is supposed to teach. Summative evaluation characteristically is quantitative, using numeric scores or letter grades to evaluate learner attainment.

1.4.1 Behavioral Model of Reinforcement

Learning can be defined as the practice leading to relatively stable behavioral change or probable behavioral change. Watson (1958) was the first to study how the process of learning affects our behavior. The central idea behind behaviorism is that only observable behaviors are worthy of research since other abstraction such as a person’s mood or thoughts are too subjective. Skinner (1938) believed that internal situation could influence behavior just as external stimuli. Behavioral Psychology is mostly interested in how behavior results from the stimuli both in the environment and within organism. Behavioral learning may be divided in classical and operant conditioning.

In classical conditioning learning can be associated with stimulus-response condition. Unconditioned stimulus (UCS) may produce unconditioned response (UCR) and conditioned stimulus (CS) may generate the conditioned response (CR) respectively. Operant conditioning is called a type of learning. The term "Operant" refers to how an organism operates on the environment, and hence, operant conditioning comes from how organism responds to what is presented to in
environment. It can also be defined as learning occurs as a result of the natural consequences of actions.

The term reinforce can be defined as means to strengthen, and refers to anything stimulus which strengthens or increases the probability of a particular response. Skinner (1938) defined reinforcer as the treat, which increases the response. Reinforcement can be categorize in four categories: positive, negative, punishment, and extinction. In operant conditioning, reinforcement takes place when an event following a response creates an increase in the chance of that response occurring in the future. Response strength can be judged by measures such as the frequency with which the response is made or the speed with which it is made. The environment change contingent upon the response is called a reinforcer (Skinner 1938).

Skinner (1938) argued that reinforcers are defined by a change in response strength, that is functionally rather than causally, and that which is a reinforcer to one person may not be to another. Hence, activities, foods or items which are usually considered pleasing or pleasurable may not essentially be reinforcing; they can only be considered so if the behavior that instantly leads the potential reinforcer increases in alike future situations.

Positive Reinforcement is a boost in the future occurrence of a behavior due to the addition of a stimulus instantly following a response. Giving or adding food to a dog conditional on its sitting is an example of positive reinforcement if this results in an increase in the future behavior of the dog sitting. Positive reinforcement refers to
adding something in environment in order to boost a response. For example, adding a treat will increase the response of sitting;

Negative reinforcement is an addition in the future occurrence of a behavior when the consequence is the taking away of an aversive stimulus. Avoidance conditioning is a form of negative reinforcement that happens when a behavior prevents an aversive stimulus from starting or being applied. Negative reinforcement refers as to taking something away in order to increase a response. The removal of negative stimulus is reinforcer and will likely increase the chances of occurrence of desired behavior. Skinner explained that while it may appear so, Punishment is not the opposite of reinforcement. Somewhat, it has some other effects as well as decreasing undesired behavior. Reinforcement can be concurrently positive and negative, such as a drug addict taking drugs for the added euphoria and eliminate removal symptoms. Punishment refers to adding something aversive in order to decrease a behavior. The most common instance of this is disciplining a child for disobedience. The reason for it, is because the child starts to correlate being punished with the negative behavior. The punishment is not liked and therefore to avoid it, he or she will stop behaving in that manner.

Extinction refers to a situation when something is removed in order to decrease a behavior. Research has declared positive reinforcement as the most powerful amongst all types of reinforcement. Adding a positive reinforcer for increasing a response not only works better, but allows both parties to focus on the positive aspects of the situation. Punishment, when applied instantly following the negative behavior
can work efficiently, but results in extinction when it is not applied consistently. 
Punishment can also calls upon other negative responses such as anger and resentment. 
By controlling organism's surroundings, its behavior patterns after reinforcement 
become predictable, even for very complex behavior patterns. The schedule of 
reinforcement is the procedure for determining when responses or behaviors will be 
reinforced, ranging from continuous reinforcement, in which every response is 
reinforced, and extinction, in which no response is reinforced.

Partial reinforcement lies between these extremes is intermittent or where only 
some responses are reinforced. Continuous Schedule can be defines as applying one of 
the four types of reinforcement every time the behavior occurs, getting a raise after 
every successful project or getting spanked after every negative behavior.
Shepard (2000) explained different response rate of the four simple schedules of reinforcement by above chart, each hatch mark designates a reinforcer being given. Simple schedules have a single rule to determine when a single type of reinforcer is delivered for specific response.

Fixed ratio schedules refer to a situation where a response is reinforced only after a specified number of responses. This schedule constructs a high, sturdy rate of responding with only a succinct pause after the delivery of the reinforcer. Fixed ratio schedule can be defined as a situation in which reinforcement is applied after a specific number of behaviors. In the present study FIR schedule is used because it can conveniently be fitted in the formative assessment system for regularity and consistency. The problem is that the child or anyone for that matter, will begin to realize that he can get away with two requests before he has to act. Therefore, the behavior does not tend to change until right before the preset number. The remedy lies in decreasing the time interval so that response is strengthened nearly all the time. In present study the interval is kept at 5 days.

In fixed interval schedule reinforcement, the first response is rewarded only after a specified amount of time has elapsed. This schedule causes high amounts of responding near the end of the interval, but a lot slower responding instantly after the delivery of the reinforcer. Fixed interval schedule is referred to a situation where the reinforcer is applied after a specific amount of time. A major problem with this schedule is that people tend to improve their performance right before the time period expires so as to "look good" when the review comes around.
Variable ratio schedule refers to a situation when a response is reinforced after an unpredictable number of responses. This schedule produces a high stable rate of responding. Variable ratio schedules are considered to work best in many situations.

Variable Interval schedule occurs when a response is rewarded after an unpredictable amount of time has passed. This schedule turns out a leisurely, steady rate of response. Reinforcing someone after a variable amount of time is the final schedule. The variable schedules are considered as more powerful and result in more consistent behaviors. This may not be work for punishment since consistency in the application is so important, but for all other types of reinforcement they tend to result in stronger responses.

1.5 Objectives of the Study

The following were the objectives of the study:

1. To find out the nature of the relationship of formative evaluation with academic achievements of secondary school students.

2. To find out the impact of formative evaluation on academic achievements of secondary school students.

3. To find out the nature of the relationship of fixed interval schedule reinforcement with academic achievements of secondary school students.

4. To find out the impact of fixed interval schedule reinforcement on academic achievements of secondary school students.

5. To compare the performance of the control group at the time of pre-test and post-test.
6. To compare the performance of experimental group at the time of pre-test and post-test.

7. To compare the performance of control group and experimental group at the time of pre-test.

8. To compare the performance of control group and experimental group at the time of post-test.

1.6 Significance of the study

The study will prove beneficial for teachers, administrators, parents, curriculum developers, policy and decision makers. Formative evaluation’s role in teaching-learning process will provide feedback to teachers. It will be helpful in improving their methodology. Parents will be able to get benefit from its results and they may be able to manage said activity at home. The curriculum developers will be able to develop the curriculum, keeping in view the results of the study, based on formative evaluation strategy. The administrators will, also, be able to guide teachers about usage of this technique for improvement of teaching-learning process. It will also prove beneficial for designing the syllabi into sub units for whole academic session.

1.7 Limitations

Following were the limitations of the study;

1. The teachers, already teaching the subject of Math and English to secondary classes, were selected as subject specialist.

2. Teacher made test were used during experiments.
1.8 Hypotheses

Following were hypotheses of the study:

1. There is no significant relationship of formative evaluation with academic achievements of secondary school students.

2. There is no significant impact of formative evaluation on academic achievements of secondary school students.

3. There is no significant relationship of fixed interval schedule reinforcement with academic achievements of secondary school students.

4. There is no significant impact of fixed interval schedule reinforcement on academic achievements of secondary school students.

Following were sub-hypotheses of the study:

1. There is no significant difference between the performance of control group and experimental group in pre-test in the subject of Math.

2. There is no significant difference between the performance of control group and experimental group in pre-test in the subject of English.

3. There is no significant difference between the performance of control group and experimental group in post-test in the subject of Math.

4. There is no significant difference between the performance of control group and experimental group in post-test in the subject of English.
The study was undertaken to unearth the impact of formative evaluation and fixed interval schedule reinforcement on academic achievements of secondary school students. The study was completed into two phases. In first phase 60 students were randomly selected from 200 students of 9th class of government high school Fateh Jang. This sample was divided into two equal groups. One served as control group and other as experimental group. The experiment was completed by administering the achievement tests for pre-test and post-test to control and experimental group. During the experiment, teacher made test were used to measure the learning outcome as formative evaluation.

A pretest was conducted to both the groups in the subjects of Math and English. Both groups were taught by subject specialist of Math and English. Control group was only taught and was not given any treatment whereas experimental group was taught and evaluated by teacher made tests during session. After administering the teacher made tests, relevant teachers marked the tests and gave these back to students to observe what were their mistakes and how to remove those. After completion of 60 days teaching, a post test was conducted in Math and English. In second phase 60 students were randomly selected from 212 students of 10th class of government high school Fateh Jang. In this phase, same exercise was repeated with addition of fixed interval schedule reinforcement which was provided to experimental group after very 6th day.

In phase I, first part of experiment was completed and sketched out the relationship and impact of formative evaluation with academic achievements. Next step was an effort to unearth the relationship and impact of fixed interval schedule reinforcement with academic achievement of students. For the purpose a separate group
was selected for experiment. This strategy helped out in verifying results of phase first along with completion of second part of experiment.

The instructors were trained in the procedures for the study by the researcher. During the experiment the instructors were provided with a detailed syllabus for the lessons and an instructor guide, which they were directed to follow closely.

1.10 Population

All secondary schools of district Attock were dealt as population of the study.

<table>
<thead>
<tr>
<th>Table 1.1 Secondary Schools of district Attock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sector</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Government</td>
</tr>
<tr>
<td>Private</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

1.11 Sampling Technique

As the study was experimental in nature, the researcher preferred convenient sampling technique. A high school named Govt. High School Fateh Jang was the venue of the experiment. Experiment was completed in two phases. A sample of 60 students was randomly selected, for phase I, out of 200 students studying in 9th class. For second phase a sample of 60 students was randomly selected out of 212 students studying in 10th class. Two groups were formed for each phase of the experiment. One was called the control group and the other was called the experimental group.
Table 1.2 Strength of selected sample

<table>
<thead>
<tr>
<th></th>
<th>Group</th>
<th>Phase I</th>
<th>Phase II</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Control</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Experimental</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>60</td>
<td>60</td>
</tr>
</tbody>
</table>

1.12 Instrument development

Separate achievement tests were developed for pre-test and post-test by consultation of 6 experts, selected from elementary teachers training college and secondary schools. During the experiment teacher made tests were used for formative evaluation of experimental group.

1.13 Instrument validation

The achievement tests were administered to 20 students for each phase, members of same population but excluded from the sample. After data collection, by item analysis, its reliability was found. A group of experts (annex D) testified the validity of the achievement test.

1.14 Data collection

Data were collected by administering the achievement tests for pre-test and post-test to control and experimental group separately. During the experiment, teacher made test were used to measure the learning outcome as formative evaluation. The study was completed in two phases. In first phase 60 students were randomly selected from 9th class...
of Govt. High School Fateh Jang. This sample was divided into two equal groups. One served as control group and other as experimental group. A pretest was conducted to both the groups in the subjects of Math and English. Both groups were taught by subject specialists of Math and English. Control group was only taught and was not given any treatment whereas experimental group was taught and evaluated by teacher made test during session at the discretion of teachers. 10 tests were given to experimental group of phase I for formative evaluation and 10 tests were given to experimental group of phase II for proving reinforcement. After completion of 60 days teaching, a post test was conducted in the subjects of Math and English.

In second phase 60 students were randomly selected from 10th class of Govt. High School Fateh Jang. This sample was divided into two equal groups. One served as control group and other as experimental group. A pretest was conducted to both the groups in the subjects of Math and English. Both groups were taught by subject specialists of Math and English. Control group was only taught and was not given any treatment whereas experimental group was taught and evaluated by teacher made test and reinforcement was given with a fixed interval schedule during session. After completion of 60 days teaching, a post-test was conducted in the subjects of Math and English.

1.15 Data analysis

Data was analyzed by using SPSS, statistical package for social sciences. There were 8 null hypotheses. Each hypothesis was analyzed at 0.05α or 0.01α. Pearson correlation was applied on data and with the help of table value, correlation between variables was found. For finding out the impact of formative evaluation and fixed interval schedule
reinforcement on academic achievements, t test was applied. By using this strategy data analysis was completed.
CHAPTER 2

REVIEW OF RELATED LITERATURE

The academic achievement of secondary school students has strong relationship with the factors including prior knowledge, study habits, confidence, and motivation (Covington, 1992; Kohn, 1993; Stiggins, 2001; Tuckman, 1998). Teachers use a variety of strategies to have an effect on these factors, hoping to attain positive results in both studying and learning. Conceivably the most ingrained strategy is that of tests and grades, which function in a punishment–reward fashion. While the appropriate use of standardized admission tests is debated and controversial (Linn, 2001). The suitable form and use of tests within the school classroom have received far less concentration and research.

2.1 Evaluation

Evaluation is the practice of determining the value, significance, or worth of e.g., a program generally by vigilant appraisal and study. The primary goal is to conclude the effectiveness of a program for achieving pre-set targets and objectives, and recognizing strengths and weaknesses and points for revision (Airasian, 1994). There are manifold approaches to evaluation; that is, viewpoints and organizing scaffolds that guide the development and conduct of the evaluation. A number of the major evaluation approaches include systems analysis, behavioral objectives, decision-making, goal-free, art criticism and case study. Approaches differ in their target techniques, models, and tools used.
Evaluation design mostly use the methods of either quantitative, qualitative, or both. Qualitative designs center on description and explanation in evaluation, whereas quantitative designs focus on validation (Hawkes and Santhya, 2002). One of the famous mixed-methods design is the extended-term mixed-method (ETMM), which is usually used to collect evidence on the efficacy of educational programs, rather than using the results of randomized field trials, which is indicated by the What Works Clearinghouse’s (WWC) standards (Chatterji, 2005). ETMM involves qualitative and quantitative methods for programs evaluation from infancy throughout their lifespan to establish effectiveness and what works. Generally, descriptive research is conducted on the program at the former stages during program adoption and implementation in the specific environment. This is followed by experimental studies of the program at a afterward stage.

Generally education targets at two different levels; (1) acquisition of knowledge, skills, attitudes, and behaviors and (2) self-directed learning, where learners take responsibility of their own learning. Learning occurs informally through everyday experiences, but at the same time formal education and training serve as complementary component. During last three decades, focus of learning has moved from repetition and practice to understanding and application of knowledge.

The goals of education have direct link with the strategies used in the educational process and in curriculum design. These goals and strategies may be based on educational theory for bridging up the gap between educational theory and practice, eventually leading to better outcomes and performance (Kaufman 2003). As an
example, self-directed learning can be a teaching strategy as well as a goal of the learner. As a teaching policy, learning and instruction may be organized so that the learning tasks are typically under the learner’s control. Learners can strive to take responsibility of their own learning (Bruner, 1960). Learners should practice certain skills to achieve self-directed learning which comprises asking questions, perilously examining new information, identifying one’s own holes in knowledge and skills, and by reflecting one’s own learning.

Self-directed learning is also a theme of education, students’ self-assessments of their skills and deficiencies, especially in formal education, have largely been poor indicators of authentic competence and performance (Eva et al. 2004). Though, teachers can also add to effective learning by promoting self-efficacy among learners by modeling, labeling a clear picture of the desired upshot, providing the obligatory basic knowledge and skills for the assignment, providing guidance and remedial advice, and giving learners chances for reflecting on their learning. Constructivism is the base of all these ideas about the teacher’s role is as a guide and facilitator of learning, not merely a transmitter of knowledge (Kerins, 1984). This instructional method is known as scaffolding. Primarily, learning involves a important degree of outer support, through early on environmental structuring, particularly for beginning students. More and more, as capability is attained, there is a reduction in environmental support leading to opportunities for internalized self-regulation, throughout guided apprenticeship. In this way, as the through external support fades away, learning happens increasingly under the control of the learner.
For achieving the desired goals of education, students can be engaged in reflection on their own learning experiences in two different ways: “reflection in action,” which occurs throughout the experience and engrosses actively engaging in the new experience by applying precedent experiences and reasoning to this new situation; and “reflection on action,” which follows the experience and involves thinking about past experiences and how they might persuade future experiences and practice (Kaufman 2003).

2.1.1 Purposes of Evaluation

Educational evaluation is the evaluation procedure of characterizing and appraising some aspects of an educational process or programme Explicit principle of evaluation in education comprises of documenting learning performance and achievement, making decisions and strategies for educational programs, and promoting faculty for recompense of teaching brilliance. The major components of evaluation in education contain an assessment of learners and learning, and specified teaching standards of teachers and teaching content (Chapman, 1989) . For the purpose of accurately capturing student achievement and forecast performance, assessment of students and teachers needs to take place either within the actual performance circumstance where the tasks are ecologically valid (Baker 2007), or in the laboratory where there is increased psychological strength but a certain lack of ecological validity (e.g., Norcini et al. 2002; Norman 2005a).

Some common purposes of educational evaluation which educational institutions generally use to demonstrate their effectiveness to donors and other stakeholders, and
also for marketing purposes. Educational evaluation is also considered as a professional activity that individual educators need to carry out if they aim to continuously analyze and enhance the learning they are endeavoring to make possible.

The evaluation and assessment methods which apply to program evaluation and evaluation of education generally based on same domain. Cognitive and learning theories provide the necessary rationale for educational evaluation and reforms.

The teachers’ use midterm or hourly exams as a testing strategy and also as a general norm in school classrooms, particularly in lecture-style courses. This strategy has disadvantages, namely, the inclination for students to cram to get ready for midterms, which is in-proficient for learning and retention (Donovan & Radojevich, 1999; Willingham, 2002).

2.2 Cognitive and Educational Theories

Cognitive theory is a learning theory which aims at explaining human behavior by considering the thought processes Donovan and Radojevich (1999) concluded that spaced or distributed practice conditions yield significantly superior results when compared to massed or “cramming” practice conditions (p-25). This reveals that students achieve better when they spend a longer time for study for studying a smaller amounts of information.

By analyzing more than 100 years of cognitive research, Willingham (2002) concluded that massed practice or “cramming” is the least effective learning and studying strategy (p. 37). Cramming can be useful if testing takes place 1 or 2 days after
the activity, but is not valuable for “the long haul.” Willingham also reported on research revealing that spaced studying promotes learning and memory over a period or delay of many years.

Social Cognitive Theory (SCT) illustrates learning as the interrelationship between behavior, environmental factors, and personal factors. It also gives the theoretical scaffold for interactive learning normally used to develop both Constructivism and Cooperative Learning.

Social Cognitive Theory assumes that the learners acquire knowledge as environment converges with personal characteristics and personal experience. It is nearly combination of present and past experiences. New experiences are appraised vis-a-vis the past; past experiences help to consequently guide and inform the learner as to how the present should be investigated.

Social cognitive theorists also recommend avoiding traditional assessment practice. They assume that the traditional hour-exam strategy may not be efficient. Self-efficacy or the belief about capability of producing positive results can motivate students for study and learning. (Pajares, 1996). The cognition processes involved in self-efficacy are called “conative.” Conation is considered as the connection between behavior and belief. At the same time it is the intentional, goal-oriented, and striving aspects of motivation that guides to proactive behaviors (Baumeister, Bratslavsky, Muraven, & Tice, 1998).

Based upon educational theory, Stiggins (2001) argued that student-centered assessment is imperative to the maturity and confidence. Tests may be used to motivate
students, then they should be reinforced by continuous formative classroom assessment. The utilization of frequent tests is associated with the implications of cognitive studies that students may not cram or use massed learning practices to study content, as much as they would in midterm testing arrangement. There are many other educational benefits of using recurrent tests which merge with social cognitive processes associated with effective learning. Similar to Stiggins (2001); Huba and Freed (2000) quoted that multiple tests provide a probable opportunity of communication between the instructor and students debate of the tests is a apparatus to provide feedback more continuously and permit school students to regulate their study habits accordingly. Student-centered approach has been initiated by these lines of communication and it helps with overall student development and achievement. Research has shown faculty–student interaction as one of the most consistent predictors of student learning and achievement (Kuh & Hu, 2001).

Overall, this type of continuous assessment encourages faculty to concentrate on the student-learning element of teaching, whereas faculty facilitates learning and motivation by providing students with key feedback on their learning progress and helping them to identify and solve learning problems (Angelo & Cross, 1993). Certainly, faculty requires constant feedback on student performance and learning for understanding student’s academic achievements, to identify barriers to learning, and to elucidate any misconceptions and miscommunication. Hakel (1997) claimed that classroom learning is often too far from what was intended or wanted, and this is repeatedly different from what is actually taught.
Multiple tests are often used to motivate students. Critics argue that tests are a form of extrinsic motivation and that students perform best when motivated intrinsically. Kohn (1993) quoted that extrinsic motivation can undermine a student’s natural inquisitiveness. Orpen (1998) analyzed the relationship between motivation and procrastination and found greater procrastination among school students who are primarily extrinsically motivated. Extrinsic motivation and poorer academic performance are associated with each other. Hancock (2001) found that students performed more poorly in classrooms with high evaluative conditions compared to those with low evaluative conditions. It is also found that negative effects of extrinsic motivation were greater for students with test anxiety.

Assessment of students’ academic achievement is a crucial step in any educational plan since it provides information about the accomplishment of specific teaching objectives (Guilbert 1992; Wass et al. 2001). It is well recognized that students deal with learning processes along with the particular type of assessment used (Rolfe and McPherson 1995; Cohen-Schotanus 1999; Wass et al. 2001), and accordingly assessment should closely match the learning objectives.

Educational experts advocate intently the use of formative assessment in addition to summative assessment (Rolfe and McPherson 1995; Friedman 1999; Nendaz and Tekian 1999; Wass et al. 2001; Liaison Committee on Medical Education 2002; Rushton 2005). The function of formative assessment is not authorization; it provides direct advice about the learning and teaching processes and may have useful effects for both students and teachers (Rolfe and McPherson 1995; Wass et al. 2001). There are
some formal schools where formative assessments are institutionalised and carried out scientifically (Neufeld et al. 1989; Rolfe and McPherson 1995; Nendaz and Tekian 1999; Nennin and Kalishman 1999).

The concepts of formative and summative roles of evaluation were primarily proposed by Scriven (1967). He explained formative evaluation as being done throughout program development to develop or make changes while the program is proceeding in accomplishment, and summative evaluation to be assessment of effects and results which is completed at the end of a program and works as a decision-making instrument about program continuation or change. The evaluation paradigm of summative and formative assessment has the prospective to put into current conceptualizations of program evaluation and psychotherapy upshots research and practice. Actually, some contemporary theory-driven evaluations have continued to accentuate the significance of both summative and formative evaluation techniques regardless of the considerable implementation challenges that arise during practice (Donaldson 2003). The use of this combined approach can have far reaching insinuations both for the action procedure as well as for program efficacy. The noticeable benefit from using this model in psychotherapy outcomes research is that it concurrently allows for instant advice within the curative relationship to help determine treatment change variables while providing for treatment revision were attractive. Ongoing instantaneous feedback leads to quick generation of hypothesis about the cause or exacerbation of symptoms and treatment changes can be made quickly. In program evaluation, the use of this model allows for cultured decisions about program
modification, thus capturing the sought after concept of continuous quality expansion along with final outcomes at the end of the evaluation procedure.

While the purpose of formative and summative assessment is widely documented, they have conventionally been considered as detach processes (Prescott et al. 2002). Mackler (1974) had previously clarified dissimilarity between formative and summative assessment in the framework of service-oriented evaluation. He reported that formative assessment addresses enhancements in programs or answers process questions about adjustment of programs, while summative assessment normally leads to universal or general judgments and conclusions based on program impact.

With the initiation of this distinction in evaluation roles, the serious task of evaluation was border lined among experts in the field. Cronbach (1966) observed that formative evaluation was much more imperative than summative evaluation whereas Scriven (1967) argued that the eventual worth of any evaluation lies in the potency of summative findings aimed at making end-point rulings (Stufflebeam 2000). Supporting for a new conceptualization of program evaluation and treatment effect, Beer and Bloomer (1986) noticed that the formative-summative terminology carried on a false severance of process and impact. As an alternative, they proposed that the center of any evaluation must be to show the program’s outcomes as a result of the program’s processes. Conversely, even though it is normally recognized that summative assessment can be used in a formative manner to develop performance, using formative assessment for summative functions has been disputed because of potential bewilders such as damaging the open relationship between trainer and trainee (Prescott et al.
Moreover, while much concentration has been focused on formative and summative assessment terminology in the educational field, little concentration has been paid to the use of formative and summative techniques in mental health programming, predominantly programs that work with youth with serious emotional commotion in school-based mental health programs (Algozzine et al. 2001; Sugai and Horner 2002).

2.3 Learning in complex domains

Complexity can be defined in three different ways. One can understand it as being mainly subjective, that is, on psychological aspect such as perceived task importance and identity, as principally objective, that is, as a function of objective task characteristics, or as an interface of subjective and objective characteristics (Campbell 1988). The objective explanation is based on the degree of orderedness of a task or the chance of multiple solution paths (Byström and Järvelin 1995; Campbell 1988). When the procedure of task concert in a domain can be explained in detail a priori i.e., is well-structured, it is considered less multipart, in dissimilarity, when it is less structured or ill-structured it is highly complex. In amorphous or ill-structured domains, such as teaching, medicine, or engineering, it is repeatedly the case that manifold paths can lead to a suitable solution for tasks, or even that multiple solutions are probable. A teacher, for example, needs to utilize his or her theoretical knowledge structures of classroom management, of learning material, and of pedagogical and instructional strategies concurrently (Ebel, 1962). Furthermore, this needs to be done in continually varying situations, in which there is always a possibility of involvement of another party, mostly
students, but occasionally also parents or colleagues that is reacting on the teacher’s concert and thereby influencing the outcome.

2.4 Educational Evaluation

Evaluation is the process of determining significance or worth, usually by careful appraisal and study. Basically it deals with the analysis and comparison of actual progress vs. former plans, oriented headed for improving plans for future execution. It is also a part of a continuing management progression comprised of planning, implementation, and evaluation; preferably with each following the other in a unremitting cycle until flourishing completion of the activity.

Gullickson (1996) delineated evaluation as "the systematic application of social research procedures for assessing the conceptualization, design, implementation, and utility of the programs." (p-41)

Evaluation is the methodical attainment and assessment of information to afford valuable feedback about some object. Eventually it is a procedural area that is strongly related to, but discernible from more conventional social research. Evaluation uses many of the identical techniques used in conventional social research, but because evaluation takes place within a orderly and organizational circumstance, it requires group skills, management capability, sensitivity to manifold stakeholders and other skills that social research in broad-spectrum does not rely on as much.

Norman (1990) delineates evaluation as, “evaluation is a systematic process of collecting analyzing and interpreting information to determined extent the pupils are achieving instructional objectives. (p.21)
2.4.1 Types of Evaluation

Mostly, evaluation can be carried out at two levels; program level and student level.

2.4.1.1 Program Evaluation

Program evaluation is carried out for determining the successfully implementation of some program. Stiggins (2000) has raised some salient questions which may be posted and answered. These questions include the following, among others.

- Is the content of the program of desirable quality?
- Is there a positive relationship between actual learning outcomes and intended learning?
- Are intended learning outcomes achieved?
- Are unintended learning outcomes identified and corrections made for them?
- Is the content relevant to the needs of the students?
- Is the content relatively simple and comprehensible, and is it able to be extended and generalized to situations within and outside the school. (p-37)

2.4.1.2 Student Evaluation

Testing is also an integral part of student evaluation. The basic purpose of student evaluation is to find out the extent how well a student is performing in a program. A series of oral questions is often used for it. Different techniques like paper-pencil tests, manipulative skill tests, discussions, tutorials, individualized instruction, assignments, projects and so on, are also used in this type of evaluation. The student is gradually guided towards a desired goal by evaluation under this technique. Tow types of students’ evaluation are common in practice: formative
and summative. The former is guidance-oriented, while the latter is judgmental in nature.

2.5 Formative Evaluation

The basic goal of formative evaluation is to ensure a healthy acquisition and development of knowledge and skills by students. Formative evaluation is also used for sorting out student needs in order to guide them towards desired objectives. When student needs and difficulties are identified, it is practical to take appropriate remedial measures to solve such problem. The purpose behind this effort is to find out the extent whether after learning experience students are able to perform what they were previously unable to do.

As Shepard (2000) reported it, “the primary purpose of formative evaluation is to help as much as possible to ensure that summative evaluation comes out positive. It is a process of channeling input variables through a process that will yield expected outputs”(p-13). The classroom teacher is considered as the best formative evaluator. His in-depth involvement in teaching learning process enables him to draw a more reliable conjecture about his students than an external evaluator. He can identify the levels of cognitive process of his students and also can choose the appropriate teaching techniques and materials for them. The teacher can effectively appraise the level of cognitive abilities of students such as memorization, classification, comparison, analysis, explanation, quantification, application and so on.

Formative evaluation serves as a tool for providing the evaluator helpful information about the strength or weakness of the student within an instructional
context. Bhola (1990) described formative evaluation as, “It is a method of judging the worth of a program while the program activities are forming or happening. Formative evaluation focuses on the process (p-34).

Formative evaluation also serves the purpose of assessing the ongoing activities. It starts operation at program start-up and continues all over the life of the program. It is intended to provide information for improving the program. According to evaluation theorist Airasian (1994), “formative evaluation is a bit more complex than summative evaluation. It is done with a small group of people to "test run" various aspects of instructional materials"(p-64).

Scriven (1994) describes formative evaluation as

Formative evaluation is typically conducted during the development or improvement of a program or product and it is conducted, often more than once, for in-house staff of the program with the intent to improve. The reports normally remain in-house; but serious formative evaluation may be done by an internal or an external evaluator or preferably, a combination; of course, many program staff are, in an informal sense, constantly doing formative evaluation. (p-73)

Airasian (1994) give purpose of formative evaluation as,

“The purpose of formative evaluation is to validate or ensure that the goals of the instruction are being achieved and to improve the instruction, if necessary, by means of identification and subsequent remediation of problematic aspects.”(p-32)

Cizek (1997) is of the view that;

Formative evaluation is conducted to provide program staff evaluative information useful in improving the program. The formative classroom assessment using cooperative groups has four purposes:
- to increase students' understanding of concepts through verbal interaction with peers,
- to provide feedback to the instructor on the cognitive processes students use to answer questions,
• to reinforce the classroom learning environment, and
• to model a variety of assessment methods. (p-89)

Classroom assessment serves some additional purposes in secondary education. Midterm evaluation also falls in category of formative evaluation and often is used to provide feedback to the students and instructors; summative assessment is normally used for finding out the extent whether the student will pass or fail the course. The formative classroom assessment has four additional purposes:

• To increase students' understanding of concepts through verbal interaction with peers (Johnson, 1994);
• To provide feedback to the instructor on the cognitive processes students use to answer questions (Nemer, 1995).
• To reinforce the classroom learning environment (Brookhart, 2000)
• To model a variety of assessment methods (Hartog, 1993)

Impara (1996) depicted,

Teacher made test based assessment is the traditional form of assessment within classrooms. Teachers construct questions, evaluate student responses, assign and check homework, monitor projects, and informally assess student progress hundreds of times a day. These assessments may be accurate or they may be faulty, depending upon the teacher's skill as a judge of various indicators and their applicability to the question at hand. However, unrelated factors can affect this type of assessment. For instance, research has shown that good handwriting can influence grading on essays. (p-29)

In educational process, learning can be assessed generally by the demonstration of competence in predefined domains. Usually these domains address them as codified in curricular frameworks and course outlines. Teacher chalk out a course of instruction
based upon predefined guidelines and consequently educational assessment also follows those guidelines.

Scriven (1994) states, “The modern concept of evaluation which has evolved largely, though gradually, in recent decades, stemmed from a newer philosophy of education which called for the development of more adequate techniques of assessing pupil growth and development” (p-71). This philosophy of education has accentuated the assignments of the teacher for the development of concepts, information skills, and habits, and also for the motivation of learner growth in attitudes, admirations, interests’ powers of thinking, and personal-social compliance. During past these objectives have become clear and included in instructional practices, suitable techniques of assessment, both formal and informal, have been worked out to judge the adequacy of the schools' programs.

Modern evaluation has changed the concept of appraisal in several ways. It tries to measure a widespread range of objectives addressed by the modern school curriculum in-spite of subject-matter achievement only. It also employs a number of techniques of appraisal, such as achievement, attitude, personality, and character tests. It also use a variety of tools like rating scales, questionnaires, judgment scales of products, interviews, controlled-observation techniques, socio-metric techniques, and anecdotal records. Modern evaluation also includes integrated approach for interpreting these various behaviors of an individual or an educational situation.
2.6 Learning through formative assessment

An important element of formative assessment is to have an insight into how targets need to be or can be successfully completed. Summative assessment is conducted after completion of a learning phase and it ranges from a single course to whole curriculum and serves liability or certification purposes. Formative assessment is conducted during a learning phase with the objective learning promotion (Birenbaum 1996; Sadler 1989, 1998). Both formative and summative assessment can be intended for the learners’ level of acquired knowledge. During past decades preference has started to shift towards alternative assessment forms which helps out the measurement of acquired knowledge as well as cognitive and physical skills through performance on relevant whole tasks (Segers . 2003).

Learning can be reinforced through formative assessment if it succeeds in helping out learners for identify their weaker and stronger points. It also supports them in guiding to overcome the weaker points during the learning process (Boud 1990; Dierick and Dochy 2001; Topping 2003). Learners need to develop an understanding of the different performance criteria and standards and it is considered the crucial aspect of formative assessment. Learners should be able to know the aspects of performance which needs to be assessed (criteria) and also about the indicators which constitute poor, average, good or excellent performance on those aspects (standards; cf. Arter and Spandel 1992).

Learners may be able to practice the kinds of assignments and tasks which will also be assessed by summative evaluation. Merely understanding about performance
criteria will not prove helpful if associated tasks will not be performed. In formative assessment procedure, instructional tasks and formative assessment tasks can be matching. Learners can work on a assignment, evaluate their performance on that assignment, decide about the aspects of their performance which need improvement, and can move for another task to work on that will prove helpful in improving the aspects of performance. It can also be done by external source as they can also have their performance assessed, diagnosed and have the task selected by external source or person. (Sluijsmans et al. 1998, 2006).

According to Scriven’s formulation (2003), the participative model applied to evaluation is characterized by a mutual aspect which eventually supports the incorporation of the examiner and examined points of view (Semeraro 2006c, d; Cousins 2003). Some studies based on a participative model, show students’ and teachers’ conceptions about learning as an important indicator for evaluating teaching (Lecouter and Del Fabbro 2001). Furthermore, the studies indicate an absolute definition of “good teaching” as it doesn’t exist since it varies depending on what kind of teaching conception exists in learner’s mind. The identical participation of the students have been found in Giles’ s research (2004), which present a participative evaluation model where teachers and students worked collectively for building a teaching evaluation program, and assigning a strong and active role to students.

Saroyan and Amundsen (2001), presented teaching concept as an articulated process derived from the interface of three elements; learning and teaching conceptions, subject matter knowledge and action. Consequently didactical planning, endorses the
use of manifold evaluation mechanism for an enhanced understanding of the teaching-learning process complexities.

2.6.1 Learning by formative assessment

Becher and Trowler (2001) analyzed epistemological and social communities on the basis of their structures, and they found that the disciplines vary also in the different models of communication and publishing research outcomes. Kekäle (2000) suggests to make a distinction among each disciplinary context, and to keep in consideration these differences when quality is focused and evaluated, since every academic discipline has diverse research topics, objectives and goals, perspectives, social values and behavior models.

Palmer and Marra (2004) focused their concentration on differences in epistemological conceptions of students from diverse disciplinary fields. The results reveal that activities proposed during the course; teaching strategies, and students’ diverse experiences, persuade the changing of one’s ideas of teaching, creating new conceptions regarding the meaning of teaching-learning process.

Self-directed learning is considered important in on-demand education. Though concept of SDL formerly emerged from the area of adult education, with specific relevance to workplace learning. Students in secondary vocational institutions are also considered to direct their own learning processes, as well as assessing their own performance, deducing their learning needs, and selecting appropriate learning sources for meeting desired needs (Knowles 1975).
While several theorists in adult education endorse the benefits of SDL (Brookfield 1986; Tough 1979), students in secondary vocational education frequently face problems with it, leading to adaptation difficulties or yet open rejection (Nolan and Nolan 1997a, b; Slevin and Lavery 1991; Williams 1996). The majorities of students who enter vocational education are familiar with a learning environment with a strong tradition of teacher-center learning and are not well prepared for SDL. Many times, teachers imperfectly presume that students already possess SDL skills, or students will develop those skills by participating in an on-demand learning environment which requires them to direct their own learning (Levett-Jones 2005). It is misconception among teachers which effects the learning of students. Consequently, the prospective benefits of on-demand education are easily undermined by both the lack of SDL skills of students who enter vocational education and the lack of support for learning SDL skills.

2.6.3 Learning enhancement formative assessment

The advantages of formative evaluation for enhancing student learning and motivation have been of interest to several researchers (Black and William 1998; Sadler 1989; Wolf et al. 1991). Formative evaluation, as initially described by Scriven (1967), is evaluation conducted for the reason of improvement in educational programs that are still under development process. In recent times, researchers have expanded the concept of formative evaluation by using it for describing evaluation of student work that is in process and has not yet reached in its final form (Weston et al. 1997).
Mainly three types of individuals have been involved in evaluation process and served as the evaluators in formative evaluation of student work in the classroom: the teacher, the student him/herself, and student peers. Teachers are considered as the most common evaluators of their students’ classroom work because they have very direct interaction with the students. Research has proved that formative evaluation has a positive effect on student learning and attitudes. Formative evaluation has strong linkage with academic achievement of students. Olina and Sullivan (2002) found that secondary school students who received formative evaluation during classroom produced higher quality reports of their experiments and scored significantly higher on a knowledge-based posttest than students who were not evaluated formatively by the teacher. Elawar and Corno (1985) quoted about sixth-grade students’ performance and attitudes toward mathematics were improved when they received constructive written feedback on their homework on a weekly basis. Ozogul et al. (2007) found during their research work that pre-service teachers who were given formative feedback by teacher on their draft lesson plans wrote appreciably better final plans than students who were given feedback from self or peer evaluators.

Literature shows that many teachers are reluctant or unwilling for providing formative evaluation because of time constraints to do the evaluation and provide useful feedback from it. Though formative evaluation by the teachers has showed encouraging results in several studies. (Elawar and Corno 1985; Olina and Sullivan 2002). Alternative strategies include self-evaluation and peer evaluation which have received increased attention because they provide the chance to students for active involvement in their own learning (Gopinath 1999; Orsmond et al. 1997). In this practice,
responsibility of formative evaluation is shifted to the students and considerably reduces the teacher’s workload in the formative-evaluation process (Ballantyne et al. 2002; Fallows and Chandramohan 2001).

During last three decades several areas of research provided the support for use of student self-evaluation in instruction process, including metacognition (Smith 2000; Bransford et al. 1999; Flavell 1976; Winne and Hadwin 1998), self-regulation (Kitsantas et al. 2004; Ley and Young 2001) and classroom assessment (Gipps 1994; Sadler 1989; Stiggins 2001). Metacognitive engagement of students in appraisal of their own work may push them to monitor their performance and to perceive shortcomings in it. Basically it is an attempt of involvement of students in teaching-learning process more actively and effectively. Students can monitor their own progress through formal self-evaluation which may improve their self-regulation skills and performance. Self-evaluation as a classroom assessment practice encourages students to participate as active role player in their own learning and may facilitate them in acquiring a longer-term self-evaluation skill (Davies 2002; Fallows and Chandramohan 2001; Wiggins 1998).

Despite the positive claims for self-evaluation, experiential findings from different studies about its effects on learning have been proved. Fontana and Fernandes (1994) reported the performance of elementary-school students who were trained in self-evaluation and evaluated themselves in math scored considerably higher on a mathematics test than a control group which did not use self-evaluation. On other hand, Andrade and Boulay (2003) found during their research process that there is no

Peer evaluation has the probability to be useful to both the teacher and the student. Peer evaluation, like self-evaluation, can have an effect on the amount of time students consumed on an assignment and on their level of rendezvous in it (Anderson; 2001, Topping; 2000) also supported the fact that peer evaluation gives students a chance to evaluate their own work with the work of a peer, a procedure that may result in better metacognitive awareness and add to the development of self-evaluation skills. Topping (1998) concluded from his meta-analysis of 145 studies of peer evaluation in higher education that peer evaluation of writing could yield comparable outcomes to teacher evaluation. Smith. (2002) described his findings about peer evaluation which augmented university undergraduate students’ consciousness of the evaluation process and their utilization of the scoring criteria on their own work.

2.7 Student attitudes toward formative evaluation

Student attitudes are usually more positive toward teacher evaluation as compared to self-evaluation or peer evaluation. Zhang (1995) stated that a best part of his subjects chosen teacher advice over self and peer response, and a number of students chose peer evaluation over self-evaluation when comparing only those two forms. Olina and Sullivan (2004) quoted that students liked teacher assessment more than self-evaluation. Nevertheless, training of self- and peer evaluators on such aspects like
objective scoring and providing productive feedback may result in more positive attitudes toward this student-based evaluation process.

Researchers have also revealed the student’s trend as more favorable toward evaluating peers’ work especially when the peer assessor is unknown to the student being evaluated than when they are known (Davies 2000, Miller and Ng 1994). Miller and Ng (1994) reported this effect to the peer evaluator not wanting to make inauspicious judgments concerning the person being evaluated. Brindley and Scoffield (1998) reported that when the student being evaluated knew who peer assessor was, the peer assessor often did not assign a low grade even if it were suitable.

2.8 Training students as formative evaluators

Numerous researchers have mentioned the significance of providing training for student assessors on conducting formative evaluation and on using an evaluation rubric (Hanrahan and Isaacs 2001; Smith 2000; Williams 1992). Ozogul (2007) found from his study with pre-service teachers that suitable training material should be considered in the significance of objective scoring to evade inflated scores, the value of remedial feedback, and teaching practice on use of the evaluation rubric. Ballantyne (2002) recommended the incorporating practice sessions that may comprise prototypes of good, average and poor work and response on students’ performance as evaluators.

During the past few years, audience response technology (ART) has been extensively adopted in school environment. This sort of instructional technology, also known to as “audience feedback,” or “clicker” technology, as well as by a diversity of
brand names e.g., the Classroom Performance System has become particularly popular amongst instructors of bulky lecture classes. Existing ART packages with coordinated hardware and software facilitate instructors to ask variety of questions, obtain instant responses from students by their response devices “clickers” or “remote controls”, and show the model of answers in a tabular or graphic format that preserves individual anonymity. Practically any class size can be accommodated.

Supporters of ART have declared that the technology improves student engagement and learning (Ward; 2003). Such findings have played an important role and prompted many school instructors to slot in ART into their classes, particularly for bulky lectures in which student participation and learning outcomes can be less than desired. So far, many studies on ART recommended that students respond favorably to the technology. Conversely, the number of studies is still comparatively small. Additionally, these studies can’t be generalized due to their limitations for the undergraduate student in large-lecture classes. Such limitations comprise the studies of atypical students and classes, assessments conducted after limited use of the technology, and evaluation on a limited number of evaluative dimensions (Rice and Bunz 2006; Fitch 2004; Latessa and Mouw; 2005). Therefore, it seems imperative that student experiences with ART receive more experiential evaluation.

Major roles of educational institutions are to rationalize professional process in the social division of labor and also to provide professional education, to elevate professionals skilled with modern techniques and to create scientific knowledge. Science plays a vital role in meeting all these anticipations imposed by a positivist
world view. The porch of scientific thinking into an extended sphere of culture brings architecture and its education itself close to mathematics and also to the stream of productive thinking (Vesely 2004, p. 23).

Scriven (1967) characterized formative evaluation as evaluation conducted with the rationale of improving educational programs that are still in the process and also can be more developed. Formative evaluation techniques are normally used to identify potential improvements for the duration of pilot testing or field testing of new educational programs. These improvements are then integrated into the program before it is made accessible for general users.

2.8.1 Observing students as formative evaluators

Researchers have broadened the conception of formative evaluation ahead of the evaluation of instructional programs in current years by applying for the evaluating the student work that is still under development (Fallows & Chandramohan, 2001). In usual procedure the students have to use given guidelines or specifications to develop a product e.g., a research statement, a business diagram in draft or near-final form. The student products are then evaluated, often by using the similar guidelines, and response based on the evaluation is provided to every student. The students make use of this feedback to modify their products into final form. Such type of formative evaluation has been extensively discussed as an approach with great potential to improve student learning and amplify learner motivation (Black & William, 1998; Sadler, 1989; Wolf, Bixby, Glenn, & Gardner, 1991). Black and William (1998) concluded after review of
250 articles on classroom assessment that formative evaluation has positive effects on performance of students of all ages and ability levels.

Research has generally yielded positive results of the effects of teacher evaluation on student performance especially when it is used for formative evaluation purposes (Black & William, 1998). Olina and Sullivan (2002) found that secondary school students who were reinforced formatively by teachers have produced higher quality reports of their experiments and scored considerably higher on a knowledge-based posttest than students who were not formatively evaluated by the teacher. Elawar and Corno (1985) explained the performance and attitudes of sixth-grade students’ were improved toward mathematics when teachers provided positive written feedback on their homework on a weekly basis.

A dilemma correlated with formative evaluation of student work by the teacher is the time of teacher required for the evaluation and for providing useful feedback to the students. Responsibility is shifted to the students in both student self-evaluation and peer-evaluation in recent days and may significantly reduce the teacher’s workload in the formative evaluation procedure (Ballantyne, Hughes, & Mylonas, 2002; Fallows & Chandramohan, 2001).

More significantly, student self-evaluation and peer-evaluation are supposed to have vital learning benefits for students in addition to potentially dropping teacher workload. Several areas of research support the use of student self-evaluation in instruction, including metacognition and classroom assessment (Bransford, Brown, & Cocking, 1999; Flavell, 1976; Winne & Hadwin, 1998). From a metacognition
viewpoint, formal rendezvous in assessment of student work prompt students to watch their levels of mastery and perceptiveness and to identify shortcomings. If classroom assessment perspective is considered then self-evaluation helps students to take a livelier role in their own learning and can assist them gaining an important long-term skill that may improve the excellence of their work without relying heavily on others (Davies, 2002; Fallows & Chandramohan, 2001; Wiggins, 1998; Wolf et al., 1991). According to Chi, Bassok, Lewis, Reimann, and Glaser, (1989), good students employ in self-explanations and explanation of their actions to a greater degree than do poor students. Engagement in both formal self-evaluation and peer-evaluation is probable to serve up as a catalyst for such self-explanations.

Researchers proposed that self-evaluation has the potential to develop student performance (Davies, 2002; Fallows & Chandramohan, 2001). Student self-evaluation is considered to develop critical thinking skills and facilitate students to take a more energetic role in their own learning. Decisively reflecting on their own work may assist students to add an important long-term ability that improves the excellence of their work without relying heavily on others (Gipps, 1994; Wiggins, 1998; Wolf et al. 1991).

2.8.3 Self-evaluation and formative evaluation

Though, studies on the effects of self-evaluation on student learning have capitulated mixed results. Kitsantas and Baylor (2001) established that pre-service teachers by using a self-regulatory tool evaluated themselves scored higher on their posttest than the students who did not use self-evaluation. Likewise, Fontana and Fernandes (1994) described primary school students who were given training in self-
evaluation and assessed themselves in math scored considerably higher on a mathematics test than a control group which did not use self-evaluation. On other side, Andrade and Boulay (2003) clarified no outcome of student self-evaluation on the performance of seventh-grade and eighth-grade students on written essays. Also, in two studies conducted with secondary school students, Olina and Sullivan (2002, 2004) stated that formative self-evaluation of students for their own draft research reports did not result in improvements on the final versions of their research reports.

Peer-evaluation is considered to be valuable to both the assessor and the assessed. Peer-evaluation can enhance the amount of time students spend on task, the amount of practice they receive and their level of engagement. It can also result in a better sense of liability and accountability for their work (Liu, Lin & Yuan, 2002; Topping, Smith, Swanson & Elliot, 2000). In addition, a greater amount of feedback on their own work and by comparison with the work of their peers may result in greater metacognitive awareness and add to development of self-evaluation skills (Topping et al., 2000).

Much of literary evidence on the upshots of peer-evaluation reveals from student reports of its effects, relatively than from measurement of student learning in a peer-evaluation perspective. On the other hand, several researchers have quoted positive effects based on student reports. Brindley and Scoffield (1998) found increased their personal motivation in business undergraduate students who used peer-evaluation in instructional modules and also it increased their understanding of the instructional content. Davies (2000) reported that secondary school students receiving momentous
benefits from marking the work of their peers. A number of students indicated that they had better understanding of the learning material after the peer-evaluation process. Smith, Cooper and Lancaster (2002) noticed that peer-evaluation improved university undergraduate students’ alertness of the evaluation process and their utilization of the scoring criteria in their own work.

If students attitude to teacher, self and peer evaluation are compared, students appear to favor teacher-evaluation over the other two approaches. Zhang (1995) explored school students’ penchants amongst teacher, self and peer evaluators in a formative-evaluation practice. He found a majority of students preferred teacher response over the two other choices, and a big number chose peer-evaluation over self-evaluation while comparing only those two forms. Likewise, Olina and Sullivan (2004) observed that students preferred teacher-evaluation over self-evaluation.

2.8.4 Self and peer evaluation in the classroom

Student’s appraisal about their work is another important issue in the use of self and peer assessment in the classroom. It is important to find out the extent, their evaluations are consistent with that of the teacher. Many researchers have made comparison of student-generated ratings with those generated by the teacher. Falchikov and Boud (1989), reported by their analysis of quantitative self-assessment studies in higher education, results showed no irrefutable evidence concerning consistency of student and teacher ratings, but suggested a tendency among the students in most studies to overestimate or undervalue their own performance. Liu, Lin and Yuan (2002) examined the connection of ratings across different amalgamations of evaluators in the
formative evaluation process: peer-teacher, self-teacher and self-peer. Their results revealed the scores assigned by self-evaluators were notably higher than those assigned by peer evaluators, and peer-evaluator scores were considerably high than those assigned by teacher evaluators.

Student awareness of peer evaluators and their own responsibility as peer evaluator becomes visible to affect the type of feedback they offer and the significance they assign to peer feedback. Researchers explained that students’ posses more encouraging approaches towards evaluating peers’ work especially when the peer assessor is kept unidentified to the student who is under evaluation. Brindley and Scoffield (1998) established findings that when the student being evaluated is well aware with peer evaluator, the peer evaluator often did not allocate a low grade even if it were suitable. Falchikov (1986) quoted an inclination for identified peer evaluators to grade their peers’ work higher than self-graders. Miller and Ng (1994) reported negative attitudes of students towards peer-evaluation when a known peer provided evaluation and feedback. The researcher ascribed this effect to the peer evaluator not willing to make adverse judgments about the person who evaluates. Similar to these results, Cheng and Warren (1997) reported that undergraduate school students did not complain of biasness and inequality when both were unidentified to each other. Likewise, Davies (2000) reported that attitude of classmates who showed a better respect for judging the quality of work when they were unspecified in the peer-feedback process.

Teacher-evaluation is the most widespread form of classroom evaluation, considering at general, and formative evaluation process have normally yielded
encouraging results with regard to student learning. Researchers have quoted quite a few probable benefits of both self- and peer-evaluation, but the relationship of these two evaluation techniques and procedures on student knowledge are visibly less clear-cut than the effects of formative evaluation by teachers. Self-evaluation has diversified learning results across studies, and many of studies of peer-evaluation have used student and researcher attitudes or reports, in spite of direct measures of student learning, as outcome measures.

The classroom is considered as the conventional setting of school education. This situation is improbable to change despite of recent reforms happened in residency training (Ludmerer and Johns 2005; Walter 2006). Until now experimental studies have created a limited understanding of the base and philosophy of residents’ learning in the workplace (Kennedy et al. 2005; Schuwirth and van der Vleuten 2006). The focal point of the leading expert community is on the development of secluded individuals and self-governing of their socio-cultural context (Swanwick 2005; Bleakley 2006). Cognitivist literature on experiential learning supports this point of view (Kolb 1984), reflective practice (Schön 1987), and adult learning theories (Knowles 1973). Even though the Cognitivist dissertation provides helpful insights into the mental development of individuals, research shows learning processes in workplaces as a more intricate picture. Residents learning occurs from interactions within the complex setting of clinical practice (Bleakley 2002; Pope et al. 2003). Socio-cultural perspective on learning shows that work-based experiences promote the socialization of residents as members of a variety of ‘communities of practice’ (Wenger 1998; Fuller et al. 2005). Billett (2004) narrated it: “workplaces provide interactions with human partners and non-human
artifacts that contribute to individuals’ capacity to perform and to the learning that arises from their performance”. Teunissen and colleagues (2007) sketched out a theoretical scaffold combining Cognitivist and socio-cultural features for clarification of what happens when residents learn by contributing in work-based activities. Residents dispense importance and connotation to their work-related activities and this makes ‘personal experiences’ connecting to different aspects of activities, such as treatment, communication, or teamwork. Personal knowledge grows as a result and residents progress through the training program. This unremitting development of personal understanding will be revealed in future interpretations of activities and events and conceptualization.

2.8.5 Formative evaluation and social psychology

Elstein (1988) has found that since the 1970s social psychologists have traced out when, how, and in what direction ‘activated mental concepts’ affect people’s impressions and judgments. Tversky and Kahneman (1974) and (Kahneman and Tversky 1973) have supported this area of research by their publication on “the rules that determine intuitive predictions and judgments”. Tversky and Kahneman depicted that people’s choices can be persuaded by problem presentation, even though the problem itself is the same in nature, result may be different. This effect is known as ‘framing’ (Tversky and Kahneman 1981). McNeil et al. (1982) quoted framing effects by taking experienced physicians, students trained in statistics, and patients to choose between surgery or radiation therapy in a hypothetical case of having lung cancer as sample. The results were presented in provisos of chances of survival or death. Results
show that 90% chance of short-term survival seems less intimidating than a 10% possibility of immediate death; surgery was chosen considerably more often in the survival frame (85.5%) than it was in the mortality frame (65.5%). This large framing effect was experimented in each of the three groups. The general effect of framing is enhanced ease of understanding of particular mental concepts. ‘Accessibility’ refers to the ease with which particular knowledge structures or mental concepts come to mind (Bruner 1957). It became an important concept in judgment analysis.

Higgins (1977) and Srull and Wyer (1979) investigate the concept of ‘priming’. A priming effect is considered to happen when a mental notion is activated in one circumstances and consequently used in another, dissimilar situation, because its ease of access has been enhanced through activation. Srull and Wyer (1979) have experienced by giving psychology students two unrelated tasks in an experiment for investigating priming effects. The first assignment was a test of word comprehension, requiring participants to form complete sentences by underlining three words in sets of four. Two priming conditions were present there, one in which neutral sentences were mixed with sentences priming for antagonism and one in which neutral sentences were mixed with sentences priming for kindness. The second, apparently dissimilar, assignment required the students to judge a squat vignette presenting a man, named Donald, whose behavior looked ambiguous on the hostile/kind dimension. It resulted out that participants’ finding of Donald’s behavior was in harmony with the concept that was made more nearby to them by the priming task. In other words, an ‘assimilation’ effect was observed.
After the studies of Srull and Wyer (1979), plentiful studies have shown similar effects, not just of traits but also of accessible stereotypes, attitudes, moods, and emotions (Stapel and Koomen 2001). Research also demonstrated ‘contrast’ effects due to available knowledge structures (Herr 1986; DeCoster and Claypool 2004). Dijksterhuis (1998) narrated that after priming with a stereotype linked with high intelligence, contributors in a general knowledge test scored higher than participants primed with a low intelligence stereotype; it seems that stereotype priming leads to assimilation. Recently many advancement occurs in the cognitive and learning sciences that they started working on the issues of learning, performance and evaluation in clinical domains.

2.9 Overview of cognitive and learning sciences theories

Learning deals with the ability to transfer already acquired knowledge to new and strange problems and situations (Eva et al. 1998; Holyoak 2005; Patel et al. 1993). Previous research, which worked in behaviorist perspective, did not considered the cognitive and individual differences in learners, but stressed on repetition and practice. Existing perspectives highlight the understanding and application of knowledge as crucial to effective learning. Contemporary learning theories deal with different domains of learning.

Anderson’s ACT-R theory emphasizes on complex cognitive processes, such as clinical reasoning and diagnostic decision-making, and these are a result of the interaction between two types of knowledge: procedural (how to do something) and...
declarative (facts) (Anderson 1983, 1993, 1996; Anderson and Schunn 2000). As per this theory, sufficient amount of both types of knowledge is requirement for learner to have a concept (working knowledge as well as knowledge of action) and skill for solving the problems associated with this concept. The ACT-R theory has positive implications for education and instruction. Anderson and Schunn (2000) support wide-ranging practice for developing a high level of competence, with the argument that the time spent practicing particular skills is the most significant factor for developing life span competencies. It is also a concept which is in line with the importance given to purposeful practice in the development of expertise and the influence of carrying out practice using manifold problems or cases to aid in transfer to new situations (e.g., Gentner et al. 2003; Holyoak 2005). The practice may not extend competence if the incorrect knowledge and competency is being emphasized and learned. Therefore, ongoing response on students’ learning is required. As a resultant factor, computer-based cognitive tutors have been developed that by monitoring students’ learning and providing immediate response on students’ weaknesses, in this way efficiency of the learning process has increased. This theory proved helpful in teaching-learning process and it has numeral implications for learning from errors, effective feedback and development of unrelenting proficiency and competency for education and training.

2.9.1 Cognitive Load Theory (CLT)

Cognitive load theory (CLT) (Sweller et al. 1998) addresses and endeavors to characterize and account for the role of memory and its associated figures and the complication and complexity of learning content and materials in the learning process.
The theory takes into account a number of hypotheses about the organization of human memory. Firstly, it works on assumption that working memory (WM) is limited regarding to the amount of information it can grasp and hold. Secondly, and in contrary to working memory, it supposes no limits for long-term memory (LTM). It has third assumption that LTM is organized in the form of schemata. Schemata are mental structures that serve up for organizing information in usual ways;

- These are easily retrievable from memory;
- These are often automatic,
- These require no effort to use;
- And also are used to interpret new and unfamiliar information.

Following these assumptions, CLT has been used for designing the instructional interventions which help out the learning process by averting or limiting the learner’s high memory load, these can result from any of two sources: Another term ‘extraneous’ cognitive load is also used, it is he kind and amount of information accessible to the learner as part of the instructional intrusion. Intrinsic cognitive load and second well famous term which deals with the complexity of the information itself such as the number of idea units inherent in the information and the interaction among those units.

This theory deals with the mechanism of information organization and storage and has implications for how information is organized in working and long-term memory for storage space and retrieval of pertinent information in a timely way. It also deals with the concerns of information overload during learning and during multitasking in clinical practice and procedure. Particularly, this theory is important and helpful for
the designing of e-learning programs, as many irrelevant components may be introduced that boost the cognitive load on the learner (Clark and Mayer 2007; Mayer and Moreno 2003). Clark, Mayer, Moreno and others have carried out many studies on the application of cognitive load theory to e-learning (Mayer et al. 2001).

2.9.2 Situative Learning Theory

Situative learning theory deals mainly with the performance of groups, rather than individuals (Greeno 2006; Lave 1988; Lave and Wenger 1991; Suchman 1985). One major principle of the situative approach takes learning as context-dependent (Anderson et al. 1996), which means that what content or material is learned depends on the specific teaching environment, and meaning is energetically constructed within the particular learning environment. It can also be explained as all interaction is vigorously constructed and negotiated by the learners using the accessible information and materials (termed artifacts) within the context of the learning activity or learning context. Additionally, opportunities for learning in any social organized activity are present there, even though this may not be formal and prearranged learning. A research study carried out from the situative perspective, analysis focuses on the performance of the whole movement and action system, which may be a group of people and other systems, and personality cognition is only measured in relation to the group’s patterns of interface. Greeno (2006) has quoted “The goal is to understand cognition as the interaction among subjects and tools in the context of an activity”, (p. 84). It is also known as “distributed cognition.” In the situative study, figures and material are
considered and analyzed as records of interfaces rather than verbal reports of one’s thought and reasoning processes.

Principles of situative theory emphasize that learning environments may be planned to encourage learning of the desired knowledge and achievement of particular educational goals (Greeno 1998). In the biomedical domain, situative learning theory considered to be most practical in the classification of learning in real-world practice settings, where the nature of these environments and resources available are continuously varying and shifting. Furthermore, by considering the emphasis on social components of learning, this theory presents a practical scaffold for understanding association and teamwork in educational practice, especially with regards to how academicians construct their representations of mutual educational practice.

2.9.3 Cognitive Flexibility Theory (CFT)

Cognitive flexibility theory (CFT) (Spiro 1991) focuses the learning and its nature in complex and ill-structured domains. Cognitive flexibility addresses of the learner’s aptitude to adapt to the learning material and content along with the conditions of learning in domains where problems can be analyzed diversified approaches acquiring the learning to be elastic in response to different assignment demands (Spiro 1991). CFT is based on the conception of “constructivism.” while there are different adaptations of this concept, the concept refers to the situation and status in which learners build up their understanding of the world by considering and constructing models of reality in their minds. When learner is given a text or a problem, he
constructs its meaning by utilizing the given data in combination with one’s previous knowledge to come to an sufficient understanding, or representation, of the text or problem.

Both the cognitive and situative programs of research proved beneficial in creating important knowledge about human learning. This knowledge domain can help in designing effective learning environments and instructional methods (Konecni, 1974). The perception of learning in this context has very clear message for educationists and experts of instructional training.

2.10 Fostering and Assessing Competence

Cognitive theories of complex learning provide guidance about assessing competence by suggesting techniques of testing that accentuate the elasticity inherent in conceptual understanding. It needs to be recognized that conventional methods of assessing achievement and competence are not adequate for testing for flexible understanding with more difficult and multipart material. Therefore, instruction and evaluation needs to be reformed to efficiently test such deep understanding and bendable problem-solving. For instance, medical instruction should comprise a analytical component to recognize problems, where student’s presumption are identified and elucidated, and a authoritarian component, where direct challenges to areas of knowledge that may present blockades to understanding are provided (Feltovich 1993).

An additional way to review capability is to evaluate the analysis strategies used when solving clinical problems. As per cognitive theory, the expansion of data-driven
interpretation as exemplified by professionals is allied with the development of automaticity of clinical skills (Patel, 2000b). Students need to obtain both biomedical and clinical knowledge for supporting this development. Clinical knowledge is usually adequate to resolve most usual problems in clinical practice. Nevertheless, to resolve difficult problems, biomedical knowledge is essential, which is connected with the use of rearward reckoning, consisting of chains of fundamental clarifications that serve up to clarify the biomedical justification for the existing difficult problem. The biomedical knowledge is fine-tuned to account for inconsistent findings in the multifaceted problem. It ought to be noticed that biomedical knowledge can subsist without any clinical background and it serves a particular reason when provided in a clinical context. Therefore, the category of reasoning and knowledge used when solving usual and difficult clinical problems can be an pointer of the level of proficiency (Patel et al. 2000a).

Appraisal of ability in medicine has been characteristically based on the conceptions resulting from the Bloom’s taxonomy of educational objectives (Bloom; 1956). By using this taxonomy, assessment of educational programs generally found objectives to be in the Knowledge sort, thus emphasizing mere recognition or recall of information. Conversely, the most significant educational objectives are normally considered to be interrelated to the understanding and use of knowledge. Anderson and Krathwohl (2001) has revised the taxonomy which provides a better outset of how educational goals and objectives can be associated with competencies and assessment of performance, by unscrambling types of knowledge from the cognitive processes used(Krathwohl; 2002). By using this improved classification in the form of a table
with Knowledge on the vertical axis and Cognitive Process on the horizontal axis, educational objectives can be categorized according to the type of knowledge used as well as the cognitive process used to comprehend and use the knowledge.

2.11 Teachers and teaching

It is usually recognized that the eminence of teaching influences student learning and concert, so estimation of teachers and teaching is significant to evaluation of an educational curriculum (Griffith et al. 1997, 1998; Jolly et al. 1996). For instance, Patel and colleagues (1991) established the verity that after the first year of medical school, student’s knowledge of pulmonary physiology had difficulties in emergent and precise patho-physiological model of a given problem and were not capable to associate clinical symptoms with the core pathophysiological methods. These difficulties were the outcome of the type of teaching and training, which did not formulate relationships between clinical symptoms and underlying mechanisms adequately unambiguous. Teachers need to communicate robust fundamental science models to students so they can make use of these models for clinical problem solving.

2.11.1 Student learning and performance

Evaluation of learner’s performance is usually used to determine competence, forecast performance as student learning is closely attached to student performance, determine performance improvement after response from assessor, and give a summative grade (Metheny et al. 2005). Combination of manifold assessment techniques is essential to evaluate learning outcomes, and different methods are more
helpful for different learning outcomes. For instance, Miller (1990) proposed a pyramid of learning and van der Vleuten (1996) planned assessment systems according to their suitable place in the pyramid. Similarly, Shumway and Harden (2003) positioned their 12 acknowledged learning outcomes within the pyramid. By combining both pyramids, the best possible assessment techniques for each learning outcome are explicated. Moving one step further, the learning achievements and assessment techniques could be planned along the cognitive procedures and knowledge dimensions of Bloom’s modified taxonomy. Moving in same way all systems like, type of knowledge learned, cognitive processes used, clinical performance and associated techniques of assessment can be connected, providing a strong foundation for the educational experiences and instructional means used. This would turn the focus of assessment to acquired knowledge along with its use in clinical performance. At the same time it will be suggesting that knowledge-based assessments are not inferior to performance-based assessments, both seem necessary for the process of evaluating student learning. Norman (2005a) has recommended use of acquired knowledge for the predicting of actual physician performance, but at the same time, should not be synchronized to the bottom of Miller’s pyramid.

Student learning has strong relationship with the method used to assess clinical competency (Howley; 2004). Consequently, assessment methods are closely linked to the learning process and student performance, and eventually patient outcomes. There is lack of assessment techniques which may provide the perfect picture of an individual’s knowledge, skills, and performance, but the assessment requires being valid and reliable enough to provide helpful data for evaluating clinical competence. Norman and
Feightner (1981, p. 26) define clinical competence as “the ability to gather data from the patient by history and physical examination, integrate this information into a diagnostic formulation, select appropriate investigations to confirm the diagnosis, and institute efficacious management.”

Howley (2004) depicted three areas for performance assessments, together with creating evidence-based locally developed assessments, an understanding of educational results, and assessment factors considered as non-cognitive, transfer of information and skills from classroom to bedside. These are valid areas for conducting effective performance assessment. In recent times, many studies focused on investigating the effects of different factors on students’ Objective Structured Clinical Examination (OSCE) performance. For exemplar, Junger (2005) noticed that a skills training course had a strong impact on OSCE performance. Additionally, Blaskiewicz (2004) reported that the testing context has strong relationship with performance and it influences OSCE performance, in spite of the fact that the OSCE station was matching in both contexts. This judgment supports the need of using case scenarios that envelop a broader range of diagnoses, not limited to that of the urgent context.

Many researchers concentrate highly on state tests, it is vital to deem that over the course of a year, teachers can review students’ learning styles and then use this information to make valuable changes in instruction. This investigative use of assessment for providing feedback to teachers and students over the course of instruction is also considered as formative assessment. It is contrary to summative
assessment, which usually takes place after a phase of instruction and acquires making a judgment about the learning that has happened.

Assessment addresses the real curriculum in school education and dominates the lives of students. Assessment of students' work is a gravely vital part of any subject principally for students as well as teachers.

Rowntree (1977) reported that, “If we wish to discover the truth about an educational system we must look into its assessment procedures” (p.1). This citation from Rowntree strengthens the significant function of assessment in higher education. Assessment tasks deliver students a message about what academicians think is important in relation to studying of a subject, the work standard and the amount of work required. Students study keeping in view the assessment, thus, for them the content of a unit is considered important not the objectives of the curriculum.

2.11.2 Nature of Evaluation Tasks

Students' approaches to learning are influenced by the nature of the assessment tasks. The assessment task may simply require a shallow engagement with the subject, for instance to reproduce information, or the task may entail greater logical effort to resolve problems, or to critically analyze an issue (Winters; 1992). If the assessment tasks in a unit are unsuccessful for test understanding, then misunderstanding of a subject matter may never be revealed.

Assessment is also used for encouraging students' interest and dedication to the study of a subject. It is used for helping students in developing a self-critical approach to their work. Assessment provides intellectual challenges and enhances independence and responsibility for learning. Assessment addresses a number of other reasons as well:
• For tracing out the students' background knowledge at the beginning of the unit so that teaching
  • is efficiently focused on what students still do not know;
  • it attempts to find out the achievement extent of the learning objectives
  • to find out what students have actually learned and failed to learn;
  • it identifies and compares the knowledge of the weakest with the brightest in the class;
  • it diagnoses students' strengths and weaknesses, misunderstandings, interests and needs;
  • it screens or selects students for some predefined function, e.g. for entrance to university, to a specific course, or for a scholarship;
  • it helps in maintaining desirable standards of achievement in the Faculty standards.
  • It provides staff with the data information about the efficiency of their teaching;
  • it motivates students for learning; and provides useful feedback to students.

During the past, assessment tasks were planned to determine and demonstrate students learning of what the Faculty member decreed they should know in a unit. It has great concern with ranking and comparing students according to their achievements in quantitative terms. The judgments of evaluators were considered to be definitive statements of a student's capability. The modern focus of assessment in secondary education attempts to link assessment tasks with expected learning outcomes and criteria of competence.
Assessment in educational phenomena is crucial for its implication in teaching-learning process. It guides student and teacher at the same time for improving educational system and institutional arrangements.

According to Boud (1995), good assessment:

- is designed to assess a broader range of student abilities, e.g. problem solving, critical thinking, effective communication, working in groups
- along with feedback shapes learning in positive and negative ways, e.g. promotes rote learning or learning in depth
- involves students in the assessment process and develops students' capacity to monitor their own performance, e.g. by understanding the criteria used in assessment
- pays more attention to self-assessment as a result of concern for reflective and self-sustained learning
- assesses not only what students know but also what they can do
- should reflect desired learning outcomes and have a beneficial effect on the learning process
- should promote search for meaning and understanding, and self-directed and independent learning
- should provide contextualized, complex challenges, not fragmented and static bits or tasks
- should expand learning opportunities to include active collaboration with others leading to assessment of projects produced by groups of students (pp 79-80),

Learning objectives are focus point in educational phenomena. Assessment ensures the attainment of predefined educational objectives. It is a tool which can be used for redesigning the system. It provides feedback and input for making necessary changes in curriculum. Educational values coerce what we decide to evaluate as well as our assessment techniques.

As Ramsden (1992) explains, “Good teaching implies a considered selection among the content of the subject area of which aspects will be formally and informally assessed, together with explanations of their relative importance”.(p-34)
Educational policy manipulates the base for aims and objectives of educational process. Aims are the long-term, aspirational declarations about enviable achievements in a unit it guides student to become and independent learner in a specific discipline, field or subject area. They serve to afford general trend for unit development but they are too broad to be quantifiable and functional for assessment.

Learning objectives are derived from aims, these are more precise, measurable knowledge, skills, attitudes and values which students are expected to develop and demonstrate. Learning objectives guides assessors for creating desired attitudes and changes in learner. These may portray the changes, teacher want to see happen in students in terms of apparent behaviors or in terms of thinking skills to be contingent from their performance on a required task. Clearly stated educational objectives are considered as the key to the design of good assessment tasks.

Curriculum development is process which establishes the standards of learning outcomes. Recent importance in curriculum development and assessment is upon “outcome-based learning” (Spady, 1993). The focus is more on desired results and less on educational inputs, content and time allocation. The plan is to work back from desired results, establish learning experiences, to plan the experiences which will help the students in achievement of these outcomes. The focus of assessment is on learning outcomes. The outcomes may be set out in a series of increasing intricacy along with patterns for what attributes need to be demonstrated, when and what standard should be followed.

Determining what outcomes educationist would like to see in students is reliant upon diverse factors. A few outcomes may be related with good performances in a
profession, some may be associated with a discipline of knowledge, some with the competencies, some with the study of a certain culture, and others with meeting specific student needs and demands.

2.11.3 Learning Objectives and Formative Evaluation

Learning objectives play a vital role in educational process. These provide guidance towards desired destination and especially expected behavioristic changes in student. Learning objectives formulation is a complex and time consuming task. There are a number of taxonomy schemes which prove helpful for teachers. Teachers can evaluate knowledge, skills or attitudes or all three, depending upon subject matter and following these objectives. Bloom's (1956) hierarchical taxonomy of educational objectives is considered the prototype taxonomy among systems developed over the last 50 years. It deals cognitive process. It is divided into the cognitive and the affective domains and would be helpful in helping to recognize the cognitive and affective processes; teacher can develop educational objectives and assessment tasks by going through bloom taxonomy.

The cognitive domain lies at the top of Bloom's classification scheme. Other categories of objectives are set in descending order. Each category includes a number of sub-categories in the original taxonomy which have not been included here:

- Evaluation of knowledge - Judging value for purposes
- Synthesis of knowledge - Putting elements into a whole
- Analysis of knowledge - Breaking down into parts to clarify organization
- Application of knowledge - Using abstractions in concrete situations
• Comprehension of knowledge - Understanding without relating to other situations
• Recall of knowledge - Remembering something previously encountered.

Educational systems try to develop skills in learners. The skills enhance learning desire among students. The performance of much academic work and of assessment tasks has strong relation with well developed writing skills. Teachers need to consider the extent the assessment is reliant on the possession of written communication skills. It is also mandatory for them to know the extent of the weighting of the assessment be based on the content rather than on writing skills? It will develop their approach for understanding the process.

Nightingale (1996) describe the difficulty in the following terms: “The problem with trying to separate assessment of content from assessment of communication skills is that it fails to recognized just how deeply intertwined are the effectiveness of communication and a student's understanding of the content” (p. 207).

It is worthwhile to take in account a low weighted assessment task of a analytical nature, early in the study of a subject, in order to judge students' level of literacy skills and judge the need for including specific learning activities for the further development of literacy skills. Assessor will also require deeming how he will weight content, organization, expression, grammar, spelling etc. in every one assessment task and make sure that these weightings are made recognized to students.

Reading and writing both skills are evaluated from the assessment process. At this point evaluator may desire to take a brake from reading by examining the educational objectives in a unit with which is linked. Proviso these objectives have not
yet been expressed, try to list some of the key characteristics, desired to be in the students to possess upon completing the unit. Students can also be questioned about the important learning outcomes. Bloom's (1956) tried to summarize these into about five or six statements.

- Review objectives from the following perspectives:
  - Are they written in a way that will enable to assess whether or not they have been achieved?
  - Do they clearly articulate the knowledge, skills and or attitudes which are wanted the student to achieve?
  - Are they written in a way which elaborates the different levels of cognitive, affective and psychomotor abilities and skills will be wanted to assess?
  - Consider what form of assessment tasks e.g. essay, objective test, practical test will best fit the purpose or require students to demonstrate the specified outcomes and allow teachers to discriminate between their performances. (p.45)

2.11.4 Evaluation Task Setting

Good teaching requires well settled tasks so setting suitable assessment tasks is a principle of good teaching. There are a never-ending number of ways and techniques that assessment may be undertaken. All possess advantages and disadvantages, costs and benefits. Numerous key concepts related to assessment and some main assessment means are in practice also there are many available useful reviews of methods. Formative or diagnostic assessment is the teaching characteristic of assessment. Students may appear to understand imperative concepts without essentially understanding them (Brinko; 1990). During formative assessment teacher can monitor students’ level of understanding and how efficiently teacher are teaching. Small-scale assessment strategies are utilized to help decide what students have learned in a lecture or a class and to get better teaching.
Criterion-referenced assessment engages establishing criteria or aspects for judging the extent how well a student has taken out an assessment task. For illustration, if the assessment mission is the writing of an essay, the established criteria may include the quality of the argument, the suitability of the literature quoted, the quality of expression. The evaluator will decide how poorly or how well, or to what standard, the essay was written, following each of these criteria. By giving students an assessment task for completion, educator should also give them the criteria teacher will apply for judgment of their attempts and the relative importance or weighting of each criterion. In the article, the criterion of the worth of the argument may be decided equally important with the criterion of apposite literature referred. Students should be well familiar with each of the criteria to be used.

Brown (1994) has defined this process as;

Making the criteria by which a piece of work is assessed explicit and out in the open is advantageous for a number of reasons
- it avoids trying to guess what is it the mind of the marker,
- it avoids students wasting time and energy on things they are not required to do, and
- students tend to achieve the required outcomes more effectively. (p-54)

Criteria development or criteria selection is also important in evaluation process. The problem is to establish the criteria for a high standard of achievement in relation to a task.

Lysne, (1984) has reported some standard for criteria;

In assessment policy, for example, the following criteria may be associated with a grade of High Distinction, 85% and above, in an assessment task:
- excellent performance indicating complete and comprehensive understanding and application of the subject matter;
achieves all basic and higher-order intended goals for the assessment tasks;
minimal or no errors of fact, omission and application present;
clear and unambiguous evidence of possession of a very high level of required skills;
demonstrated very high level of interpretive and/or analytic ability and intellectual initiative;
very high level of competence. (p. 78)

Norm-referenced assessment referrers to arriving at a grade for a student's performance, by overtly comparing it with the norm established by the review of other’s performance. Standards for specific grades or for passing or failing are settled from the level of performance of the group. The usual distribution is by standardization of marks in a population or by using the “bell curve” which is considered as a valid example of norm referencing. Institutions apply norm referencing if they set anticipations that a assured percentage of students will acquire passes, credits, distinctions and high distinctions.

2.11.5 Nature of Knowledge and Skills for Evaluation

Knowledge and skills are topic of discussion and importance in educational debate. The important question to inquire when considering which assessment method will be used for assessing knowledge and skills, which are expected as successful student after having acquired knowledge and skills as a result of studying presented content (Lysne, 1984). It should be the part of unit objectives. As a general principle, the type of assessment tasks selected “should connect clearly with the goals and objectives of the unit.”

Seldom one assessment method satisfies all educational objectives. If teacher believes assessment as a means of providing opportunities for students to demonstrate
how much they understand, this will guide to employ a greater range of methods and afford more opportunities for students to demonstrate their knowledge.

As Ramsden, (1992) states:

Uniformity of methods makes comparisons superficially easy but forces students into a situation where they may not be able to display what they have learned and where there are often hidden rewards for conformity rather than originality. On the other hand, the use of a variety of assessment methods makes it more difficult to combine the results than to do so using one method, e.g. essay writing. Students perform inconsistently on different tasks. It is not unusual that marks on practical assignments, project reports, and examination results correlate poorly. Three of the more common assessment strategies are the traditional exam, the essay and the multiple-choice test. (p. 191)

Examinations are means to determine the worth of educational institution and learning outcomes of students as well. Exams are continuously used because of their usefulness and cost effectiveness, on the other hand, like any other assessment strategy, exams are highly criticised for not testing conceptual understanding adequately, or for accentuating quantity rather than quality and rote learning.

Knight (1994) narrated the same fact;

Assessment by means of an essay assumes that students know how to write a good essay. If so, it can be used to assess a range of abilities such as:
• explaining cause and effect relationship,
• describing application of principles,
• presenting arguments and hypotheses,
• explaining methods and procedures, and
• evaluating ideas. (p. 78)

Multiple choice questions have received attention of educators during last three decades. MCQ have not fully replaced essay but mostly are used in testing process. Essays of two to three thousand words are time consuming to mark. A difference on the essay which is comparatively easier to mark is the “part-essay” or the “constrained
format” essay, suggested by Brown and Knight (1994). This variety of essay initiate in leader articles in a good newspaper or in a book review. It is consisted of a 300 word introduction; an sketch out of the content of the essay in heading form; minimum two linking paragraphs amounting to 100 words each in the previous section; and a 500 word conclusion or summary.

Marking of Multiple-choice tests is rather easy, these can be machine marked and can reduce marking load. The construction of a test that measures more than recall of information requires a good deal of skill (Isaacs 1994).

Nightingale (1996) illustrated assessment tasks for a diversity of learning outcomes. For the assessment of critical and reflective thinking he suggests the following:

- use of an interview in relation to a critical incident (which has an impact on the student), drawing on student's experience in the workplace;
- use of written description, analysis and reflection on critical incident;
- critical evaluation of research literature;
- critique of a topic of current debate and/or community interest; and
- reflective journal writing.

For solving problems and planning subsequent actions, they suggest:

- use of essay question,
- simulated interview with a client,
- social history report, and
- research poster exercise. (p.107)

2.11.6 Peer Assessment

Peer assessment is getting positive response from practitioners with the passage of time. Research confirmations about subjective marking advocate that peer assessment by students is about as steadfast as that carried out by lecturers (Armstrong
This involves that it may be rational to use this approach for some of the time, for a number of reasons. Peer assessment facilitates students understand the subject and the criteria for high achievement. It also provides them access to a diversity of approaches to tackling assignments, to differences in material and conceptualization, presentation and overall standard of work.

Exemplification of student produced work, and to the standard of quality required both are available to the student for having a better understanding of what is expected. Conversely, when the academic unit coordinator is ultimately responsible for marking, it is difficult to decide whether peer marks should be included in a final grade or not and what should be the percentage?

Armstrong and Conrad (1995) explained the concept of portfolio,

“A portfolio is a collection of a student's work from designated sources over a designated time for a designated purpose” It is a form of assessment now being widely used for a number of reasons

• it encourages students to engage in self assessment and critically reflect on their own learning and performance, e.g. in selecting their best work for inclusion in the portfolio;
• demonstrates development over time, e.g. by cumulative track record rather than the ‘sudden death play-off’ of an examination;
• negotiation between teacher and student on the form and content of the portfolio encourages integration of assessment and learning; and
• multiple skills can be assessed, e.g. writing, critical thinking, and research skills. (p. 85)

Small group work is assessed by this technique. It involves assessing the work of a small group of students in a mode that all students in the group obtain the similar grade regardless of differences in contribution. Possible inequities exist within the group assessment and assessor will require considering ways of overcoming them.
Black (1998) defined assessment

It broadly include all activities that teachers and students undertake to get information that can be used diagnostically to alter teaching and learning. Under this definition, assessment encompasses teacher observation, classroom discussion, and analysis of student work, including homework and tests. Assessments become formative when the information is used to adapt teaching and learning to meet student needs. (P.43)

Teacher have direct interaction with the students. When they are fully aware how students are progressing and where they are facing trouble, they can make necessary instructional adjustments by using this information, for example re-teaching, trying substitute instructional approaches, or offering more opportunities for practice. These efforts prove helpful in improving student success.

William (1998) conducted a widespread research analysis of 250 journal articles and book chapters winnowed from a much larger collection to conclude whether formative assessment enhances academic standards in the classroom. He accomplished that efforts to reinforce formative assessment fabricates considerable learning gains as measured by comparing the common improvements in the test scores of the students involved in the improvement with the variety of scores found for typical groups of students on the identical tests.

Feedback helps learners if given as part of formative assessment to become familiar of any fissures that subsist between their desired objective and their existing knowledge, understanding, or skill and guides them through actions obligatory to attain the goal (Ramaprasad, 1983; Sadler, 1989). Feedback has important role in educational process. It enhances the learning opportunities as well. The most accommodating type of response on tests and homework provides particular comments about errors and
precise suggestions for development and encourages students to spotlight their attention attentively on the chore rather than on simply getting the right answer (Bangert-Drowns, Kulick, & Morgan, 1991; Elawar & Corno, 1985).

Ames, (1992) explained his experience;

This type of feedback may be particularly helpful to lower achieving students because it emphasizes that students can improve as a result of effort rather than be doomed to low achievement due to some presumed lack of innate ability. Formative assessment helps support the expectation that all children can learn to high levels and counteracts the cycle in which students attribute poor performance to lack of ability and therefore become discouraged and unwilling to invest in further learning. (p-119)

Teacher can improve teacher-learning process by using feedback as a tool. Whereas feedbacks normally originate from a teacher, learners can use formative assessment through self-evaluation for learning improvement. Experimental research studies show that students who comprehend the learning objectives and assessment criteria and also have the opportunities to imitate on their own work show greater improvement than those who do not have this opportunity. (Fontana & Fernandes, 1994; Frederikson & White, 1997).

Learning disabilities also affect learning. Students who are taught to use self-monitoring strategies connected to their understanding of reading and writing tasks also show performance gains (McCurdy & Shapiro, 1992; Sawyer, Graham, & Harris, 1992).

The ultimate goal of formative assessment is to increase an understanding of what student’s knowledge and what he doesn't know. For making approachable changes in teaching and learning process, different techniques such as teacher observation and
classroom discussion have an important place alongside analysis of tests and homework.

2.11.7 Evaluation Techniques

Black (1998) persuades teachers to make use of questioning and classroom discussion as the prospect to increase their students' knowledge and develop understanding. On the other hand, teachers must make it sure to ask thoughtful and reflective questions in spite of simple ones, and then provide students enough time to respond. In order to involve everyone, Black (1998) suggests following strategies:

- Invite students to discuss their thinking about a question or topic in pairs or small groups, then ask a representative to share the thinking with the larger group.
- Present several possible answers to a question, then ask students to vote on them.
- Ask all students to write down an answer, then read a selected few out loud.
- Teachers might also assess students' understanding in the following ways.
- Have students write their understanding of vocabulary or concepts before and after instruction.
- Ask students to summarize the main ideas they have taken away from a lecture, discussion, or assigned reading. (p.73)

Questioning is effective techniques especially for creative learning. Black (1998) also recommends that chance should be given to students for completion of a few problems or questions at the end of instruction. He also recommends interviewing the students individually or in groups about their thinking for solving the problems.

Tests and homework is also recommended as tool for formative evaluation. Additionally if teachers investigate where students are in their learning and give precise,
focused feedback concerning performance and ways for improving. Black (1998) has given the following recommendations:

- Frequent short tests are better than infrequent long ones.
- New learning should be tested within about a week of first exposure.
- Be mindful of the quality of test items and work with other teachers and outside sources to collect good ones.
- Portfolios, or collections of student work, may also be used formatively if students and teachers annotate the entries and observe growth over time and practice. (p-124)

Capacity building of teaching staff regarding modern assessment techniques is another important area which needs immediate action. Training and professional development in the discipline of classroom evaluation are indispensable in order to make available individual teachers with the time and support essential to make changes. Teachers require time for reflecting upon their assessment practices and advantage from observing and conferring with other teachers about effectual practices and desired changes.

Duschl & Gitomer (1997) recommended setting up confined groups of elementary and secondary schools; urban, suburban, and rural- for tackling formative assessment at the school level and at the same time collaborating with other local schools. The assessors will face different challenges in different subject areas and external evaluators may help teachers with their work and in collecting facts about effectiveness. Research also has pointed out towards potential conflicts between state assessments and classroom assessments. In this situation the external tests can outline what is going on in the classroom in a negative way. The importance should be given to practice and test preparation in spite of teachers’ best judgment about learning.
Teachers need information for upgrading their methodology for the purpose of improvements in teaching-learning process. Assessment for learning is considered to provide teachers information for modification of strategy and differentiate teaching and learning activities. Individual students can learn in eccentric ways, but it is also recognized that there are many conventional patterns and alleyways which are followed by many students. Careful design is required for it from teachers so that they use the resulting information for determining not only what students know, but also to increase insights into how, when, and whether students apply what they learned and know. Teachers can better employ this information for streamlining the target instruction and resources, and also for providing feedback to students for helping them in their learning enhancement.

Judith (1989) defines link of assessment with meta-cognitive orices;

"Assessment as learning is a process of developing and supporting meta-cognition for students. It focuses on the role of the student as the critical connector between assessment and learning. When students are active, engaged, and critical assessors, they make sense of information, relate it to prior knowledge, and use it for new learning". (p-24)

Prior knowledge is used for learning new ideas. It works as the regulatory process in meta-cognition. It takes place when students observe their own learning process and use the opinion or feedback from this monitoring to make modification, adaptations, and also some vital changes in what they learned. Teachers are supposed to help students in developing, practicing, and becoming comfortable with reflection, and with a critical analysis of their own learning.

Assessment of learning is generally considered as summative in nature and is used for confirmation what students know and can do. Student also try to express
whether they have achieved the curriculum objectives. Teachers focus on ensuring that they have used assessment techniques for providing precise and sound statements of students’ ability, so that this information can be used for making reasonable and defensible decisions.

Kebdall (1996) explained the assessment for learning;

“Assessment for learning, assessment as learning, and assessment of learning all serve valuable, and different, purposes. It is not always easy, however, getting the balance right. If one wants to enhance learning for all students, the role of assessment for learning and assessment as learning takes on a much higher profile than assessment of learning” (p-45).

Conventionally, assessment of learning remained the focus point of classroom assessment. It uses the data for making decisions about students’ performances and provides input for up-gradation of educational system.

2.11.8 Evaluation as Diagnostic Process

Diagnostic process has its own importance in educational stream. Teachers usually use assessment for learning when they incorporate diagnostic processes, formative assessment, and feedback at different steps in the teaching and learning process, although it was over and over again informal and embedded. When students become critical analysts of their own learning, it is called systematic assessment but it is very rare in educational process. (Norris, Stephen, & Linda Phillips. 1987). While a few teachers have integrated practice of self-assessment into their programs, very few have systematically or explicitly used assessment for developing students’ capability of evaluating and adapting their own learning.

Assessment of learning is usually used only when summative judgments are required although it has an important role in learning process. The mechanism and
context of constructing and using assessment is important. If the purpose of assessment is based on enhancing learning, the assessment should give students an opportunity to make their learning obvious without any anxiety or reprimand. If checking learning for reporting stands as the purpose then teachers need to be particularly fretful about the quality of the assessment, and how it will be used by others. The process of serving three different assessment purposes at the same time is very difficult, and for a while not impossible. Understanding the three assessment purposes is important for educators for recognizing the need and sense of balance among them, know which purpose they are using and why, and especially wise use of all these.

Norris, (1997) narrated;

“Careful planning is required to ensure that there are logical connections among the purpose, methods, and use of the results Classroom assessment is planned in relation to purpose and in alignment with curriculum and instruction. Curriculum, assessment, instruction, and learning are interconnected and interact in an iterative and sometimes cyclical process. All four need to be aligned and coherent for the learning to be effective and meaningful”.(p-119)

Planning process makes possible the attainment of educational objectives. This process of planning provides an outline that centers on the principle, serves as making the connections explicit, and creates a logical organizational structure.

Even though teachers need not to be stick strictly to their plans, without proper planning it is difficult to ensure balance and coherence. Teachers start planning a unit or sequence of learning activities after identification of a topic. Then they design special lessons and activities which optimize the resources they already possess, and next they proceed with teaching the material. On the end corner of this process, teachers review
students learning for discovering the lessons or the assessment instruments did not line up well with curricular objectives.

2.11.9 Mapping as Evaluation Tool

Mapping is used as evaluation tool. It creates the essential association among desired objectives, measurement tools, and teaching techniques by revolving the planning process on its head. It prompts assessors to start at “the end” with the purposes and outcomes, teacher expect to achieve.

Ornstein (1989) has defined mapping process;

“Backward mapping requires us not only to think about the curricular goals teachers want students to meet, but also to deconstruct the complex learning processes involved to identify the stages of learning. It also requires us to consider the misconceptions and confusions teachers might encounter along the way, and decide how teachers will assess whether students are progressing toward the goals”. (p-89)

Reporting has the fundamental purpose of enabling parents and students to understand the student’s performance at a particular point in time, and then make a decision about what is required for future progress (Rogosa, 1999). Further complex structures of classroom assessment entail new techniques of reporting. New developed ways require including a frame of reference and sufficient information as well, so that an interloper can make sense of the information and data. Propositions should be included about the proper use of information for making some reasoned judgments. The reports may obviously inscribe the purpose and learning outcomes, and should provide an precise profile of a student in relation to these.

Wiggins (1998) recommended:

Reporting should be outcome-based, honest yet fair, rich in context and user friendly. This kind of reporting;
• explicitly identifies the purpose of the assessment
• provides sufficient context and points of reference to make interpretation reasonable
• uses a variety of descriptors and symbols (e.g., letter or percent grades) that have clear, agreed-upon, and stable meaning
• provides rich, detailed information and evidence (not just a single grade) (p-37)

The methods normally used for collecting, interpreting, and reporting information about students knowledge and learning are in vast variet, and there are many outstanding resources available for teachers regarding educational assessment process.

Rogosa (1999) narrated this process;

“Although some methods have come to be associated with assessment during instruction and learning, and others with assessment at the end of a unit or term, there are a variety of methods that can be used for all three purposes: assessment for learning, assessment as learning, and assessment of learning”. (p-72)

These different methods have interrelationships among them and it is imperative to note down that some methods feel right in multiple categories. These methods have been ordered by function such as gathering information, interpreting data, keeping records, and communicating. (Sanders,1993). Assessment for learning takes place throughout the learning cycle. It is planned to make each student’s understanding observable, so that teachers can make a decision about guiding and helping the student’s progress. Students become skilled at in individual and idiosyncratic levels and ways. There are many expected patterns of connections and presumptions that some students may experience when they move about along the range from emergent to proficient.
2.11.10 Assessment for Learning

Assessment for learning usually used by teachers as an investigative tool to find out as much as they can find about student’s learning and knowledge and what confusions, preconceptions, or gaps they might posses.

The extensive diversity of information that teachers accumulate about their students’ learning processes provides the foundation for determining the measures for student’s learning enhancement. It also affords the basis for providing evocative response for students and deciding about groupings, instructional techniques, and allied material..

Sanders (1993) explained the phenomena as;

- Assessment for learning occurs throughout the learning process. It is interactive, with teachers
  - aligning instruction with the targeted outcomes
  - identifying particular learning needs of students or groups
  - selecting and adapting materials and resources
  - creating differentiated teaching strategies and learning opportunities for helping individual students move forward in their learning
  - providing immediate feedback and direction to students. (p-112)

Students’ motivation and commitment to learning is also enhanced by teachers who use assessment for this purpose. When teachers are ready to learn by focusing the assessment, they can change the classroom culture and make it near to student’s success. They make noticeable what students suppose to be true, and use that data to facilitate students shift forward in convenient, efficient, and respectful ways.

Shepard (2000) addressed the issue in his words;

When the intent is to enhance student learning, teachers use assessment for learning to uncover what students believe to be true and to learn more about the connections students are making, their prior knowledge,
preconceptions, gaps, and learning styles. Teachers use this information to structure and differentiate instruction and learning opportunities in order to reinforce and build on productive learning, and to challenge beliefs or ideas that are creating problems or inhibiting the next stage of learning. (p.15)

Black and Harrison (1991) conducted a research study. They worked with teachers for changing their attitude towards questioning with the students. They tried to construct a pattern which include a longer wait time and it gives equal chance to every student to be prepared to respond teachers at any time. Researcher found that many students teachers remained involve in the discussion. This exercise increased the sophistication of their contributions, and made them able for creating a healthy educational environment which supports inquiry in which all members of the group, not just the keen respondents, were working collectively to investigate new ideas and alternatives, not simply to find the “right” answer.

By carefully framing questions to challenge students’ thinking and to examine issues that are critical to the development of students’ understanding of complex ideas, teachers could gather detailed information about student beliefs, preconceptions, and alternative perspectives, and more students were engaged (Shepard, 2000). They found that questioning can be a powerful tool in assessment for learning. Some examples of “questions worth asking” are:

- What would a penny tell future generations about our civilization?
- Is gravity a fact or a theory? What evidence supports your answer?
- In what ways are the animals in the story like humans? In what ways are they not like humans?
• If plants need sunlight to make food, do you think the biggest plants would grow in the desert? Why?
• Describe what you think is the temperature of the poem.
• What do you suspect happened to the slain knight? Why?

Students’ understanding can be exposed not only through their responses to the teacher’s questions, but also through the questions they formulate to advance their understanding.

Teachers use the curriculum as the starting point in deciding what to assess, and to focus on why and how students gain their understanding. Assessment for learning requires ongoing assessment of the curriculum outcomes that comprise the intended learning. Teachers create assessments that will expose students’ thinking and skills in relation to the intended learning, and the common preconceptions.

2.11.11 Tools for Evaluation

Teachers use focused observations, questioning, conversations, quizzes, computer-based assessments, learning logs, or whatever other methods are likely to give them information that will be useful for their planning and their teaching. Each time a teacher plans an assessment for learning, he or she needs to think about what information the assessment is designed to expose, and must decide which assessment approaches are most likely to give detailed information about what each student is thinking and learning. The methods need to incorporate a variety of ways for students to demonstrate their learning (Stiggins, 2000). For example, opportunities for students to complete tasks orally or through visual representation are important for those who are struggling with reading, or for those who are new English-language learners.
Assessment for learning is of high quality when a teacher can use it to make decisions about students’ learning with enough specificity to be able to provide descriptive feedback, and to design the next stage of learning.

Reporting in assessment for learning is based on open, frequent, and ongoing communication with students and their parents about progress in learning, methods that the teacher is using to ensure ongoing progress, and ways that students, teachers, and parents might help move learning forward with minimal misunderstanding and confusion for the student (Tyree, 1993). The reports might focus on a single outcome but more often on a series, or cluster, of outcomes. Reporting should take into account what learning is expected, provide good models of what students can achieve, and identify strategies for supporting students. Assessment for Learning

Assessment as learning focuses on students and emphasizes assessment as a process of meta-cognition, knowledge of one’s own thought processes, for students. Assessment as learning emerges from the idea that learning is not just a matter of transferring ideas from someone who is knowledgeable to someone who is not, but is an active process of cognitive restructuring that occurs when individuals interact with new ideas (Tyree, 1993). Within this view of learning, students are the critical connectors between assessment and learning. For students to be actively engaged in creating their own understanding, they must learn to be critical assessors who make sense of information, relate it to prior knowledge, and use it for new learning. This is the regulatory process in meta-cognition; that is, students become adept at personally monitoring what they are learning, and use what they discover from the monitoring to
make adjustments, adaptations, and even major changes in their thinking. Assessment as learning is based in research about how learning happens, and is characterized by students reflecting on their own learning and making adjustments so that they achieve deeper understanding. Too many students have reading assessment done to them, or for them. Only reading assessment, that is done with students and eventually by students, can foster true independence and success in reading.

Accomplished readers are flexible in their routines of meta-cognition and comprehension monitoring, as demanded by the particular act of reading. The ability to self-assess is multifaceted, and good readers apply their self-assessment strategies on demand (Shepard, 2000). Although Aflerbach’s comment is specifically about reading, it is applicable to many other areas of learning as well. Students become productive learners when they see that the results of their work are part of critical and constructive decision-making. If young people are to engage in continuous learning in environments where knowledge is always changing, they need to internalize the needing to know and challenging-of-assumptions as habits of mind.

2.11.12 Evaluation for Skill Assessment

Educational assessment has diversified roles to play. The ultimate goal in assessment as learning is for students to acquire the skills and the habits of mind to be Meta-cognitively aware with increasing independence. Assessment as learning focuses on the explicit fostering of students’ capacity over time to be their own best assessors, but teachers need to start by presenting and modeling external, structured opportunities for students to assess themselves.

Ornstein (1989) explained the process;
A high level of student participation in the assessment process does not diminish teachers’ responsibilities. Rather, assessment as learning extends the role of teachers to include designing instruction and assessment that allows all students to think about, and monitor, their own learning. (p-39)

Learning and assessment are interrelated parts of the educational phenomena. Learning is based on the conviction that students are capable of becoming adaptable, flexible, and independent in their learning and decision-making. When teachers involve students and promote their independence, they are giving them the tools to undertake their own learning wisely and well.

Students need to develop refined combinations of skills, attitudes, and dispositions for ensuring the status of independent learners. Self-monitoring and evaluation are considered as multifaceted and difficult skills that may not expand quickly or spontaneously. By comparing with many other complex set of skills, becoming meta-cognitively aware needs modeling and teaching on the side of the teacher, and practice on the part of the student (Ornstein, 1989).

Success and efforts have very close link. Students require experiencing continuous and genuine success. This does not mean that students may never experience stoppage but, relatively, that they need to become easy with identifying different perspectives and confront these perspectives; they need to study to look for misapprehension and inaccuracies and work with them toward a more absolute and rational understanding.

Ornstein (1989) quotes this process;
Students are often unwilling to confront challenges and take the risks associated with making their thinking visible. Teachers have the responsibility of creating environments in which students can become confident, competent self-assessors by providing emotional security and genuine opportunities for involvement, independence, and responsibility. (p-73)

2.11.13 Evaluation as Support for Students

For analyzing the steps for supporting students’ independence in learning, teachers make use of assessment as learning to attain rich and detailed data about ways and styles of student’s progress in developing the habits of brain and ability to monitor, challenge, and regulate their own learning (Quellmalz, 1984). Students study to monitor and challenge their understanding, forecast the results of their existing level of understanding, make rational decisions about their development and difficulties, make a decision about their needs, organize and reorganize concepts, check for constancy between different pieces of data and information, illustrate correlations that may prove beneficial for them in enhancement of their understanding, and place down personal goals. Teachers have much interest in how students understand concepts in the process of assessment as learning, and in how they are using meta-cognitive analysis for making adjustments to their understanding (Quellmalz, 1984). Students’ goal-setting process requires teachers monitoring and observing their learning, and the strategies students employ to maintain or challenge, regulate, and advance their learning.

A variety of methods in assessment as learning is used by teachers as long as the methods are designed to illustrate out detailed information both about students’ learning and about their meta-cognitive processes. Teachers can easily monitor their learning by teaching students how to use the methods, they can also think where they feel comfortable in their own learning and how they become confused or uncertain, and
come to a decision about a learning plan. A variety of assessment methods have the probability to persuade reflection and review, what is important in assessment as learning is that the techniques facilitate students to judge their own learning in relation to models, examples, criterion, rubrics, frameworks, and checklists that afford descriptions of flourishing learning.

Quality is also very important factor in assessment as learning and it totally depends on how the assessment employs students in considering and challenging their thinking, and enables them to make judgments about their views and understanding (Quellmalz, 1984). Teachers set up a high quality of assessment by ensuring that students are using accurate tools and are accumulating the facts needed to make rational decisions about what they supposed to understand or find confusing, and what moreover they require to do to get deeper their understanding.

Quellmalz (1984) described the nature of reliability in assessment;

Reliability in assessment as learning is related to consistency and confidence in students’ self-reflection, self-monitoring, and self-adjustment. As students practice monitoring their own learning and analyzing it in relation to what is expected, they eventually develop the skills to make consistent and reliable interpretations of their learning. In the short term, however, teachers have the responsibility of engaging students in the meta-cognitive processes. (p-39)

Students are capable of assessing themselves only when they have a very clear understanding of skillful learning and the range of steps that necessitate going through for attaining the preferred expertise. A very clear criterion and diverse pool of examples of what good work also needed for this process. They require imitating on their own and others’ work in the framework of teacher feedback and recommendation about future planning.
2.11.14 Feedback

Feedback works as back bone in assessment for learning. It is predominantly vital in assessment as learning. Learning is improved when students observe the effects of their past efforts, and can foresee substitute strategies for understanding the material. Feedback provides models for independent learning by enhancing understanding of students and they inclined to be assiduous and more engaged.

Phillips (1987) states his observation;

Although assessment as learning is designed to develop independent learning, students cannot accomplish it without the guidance and direction that comes from detailed and relevant feedback. Students need feedback to help them develop autonomy and competence. (p-67)

Skills like monitoring and self-regulation become as practice if constant feedback and follow up is involved in using the skills. Valuable feedback challenges ideas, facilitates in introducing additional information, offers substitutive interpretations, and generates conditions for self-reflection and evaluation of new concepts. It facilitates students with data about their concert on a task, and how they can draw conclusions on their own.

Feedback provides direction for what changes students need to bring about and if the feedback is unable to refer to students’ own roles in moving ahead to the next stage of learning they will be continuously asking questions about its accuracy. Feedback in assessment as learning supports students to spotlight their concentration on the target, in spite of getting the answer right.

Linda (1987) narrates it;

It provides them with ideas for adjusting, rethinking, and articulating their understanding, which will lead to another round of feedback and
another extension of learning. Although teachers are the main providers of feedback, they are not the only ones. Peers, family, and community members also are important players. Students learn a great deal within their families and their communities. (p-81)

Students come across new information; they use their accessible beliefs and thoughts and also use community and cultural norms for its filtration. They make a comparison of the new data with the beliefs and ideas of the people around them.

Assessment occurs by students as well as teachers; students are active with their own meta-cognitive ability of self-reflection, self-analysis, construal approach, and restructuring of knowledge. While mentioned skills become well developed, students will start to direct their own learning. They will have learned to request for support, look for new information, and reinforce or challenge their conclusion by peer reviewing and discussing (Phillips: 1987). Assessment as learning enables the students and teachers to discuss what the students are learning, what can be the substitutes enhancement of learning, what is the level of personal goals achievement, and what is possible addition for future goal setting.

Reporting is another very important area in assessment as learning. It is the responsibility of students to report, who must learn to eloquent and protect the nature and worth of their learning. Learning can be observed by observing students. Students reflect on their own learning and they must exchange with others, they are escalating their appreciativeness about a topic, their own learning potency, and the areas for further development.

Quellmalz (1984) elucidates;

Student-led parent-teacher conferences have become a popular reporting forum that fits with assessment as learning. However, the success of these
conferences depends on how well they are structured and how well the students prepare. The students need to have been deeply involved in assessment as learning throughout the instructional process, and be able to provide their parents with evidence of their learning. The evidence needs to include an analysis of their learning progress and what they need to do to move it forward. (p-13)

2.11.15 Evaluation Strategies

Assessment of learning encompasses many areas. Quellmalz (1984) has given its detail description.

Assessment of learning refers to strategies designed to confirm what students know, demonstrate whether or not they have met curriculum outcomes or the goals of their individualized programs, or to certify proficiency and make decisions about students’ future programs or placements. It is designed to provide evidence of achievement to parents, other educators, the students themselves, and sometimes to outside groups (p-54).

Assessment it considered as tool which later on becomes public by announcing its outcomes. Mostly it contributes to key judgments that will influence students’ futures. Underlying judgment and measurement of assessment of learning is plausible and justifiable. The consequences of assessment of learning are mostly extensive and influence students seriously, teachers have the liability of reporting student learning accurately and reasonably based on verification obtained from a range of contexts and applications.

Quellmalz (1984) has given description of assessment of learning;

Effective assessment of learning requires that teachers provide

- a rationale for undertaking a particular assessment of learning at a particular point in time
- clear descriptions of the intended learning
- processes that make it possible for students to demonstrate their competence and skill
Students can make an effort with the guidance of their teacher for assessment of their own learning for demonstration of their competence, as well as the intensity and span of their learning.

Rogosa (1999) has given the purpose of assessment of learning as to measure, verify, and document the level of students’ learning, for making logical decisions about the students. He extracted;

There are many potential users of the information:
- teachers (who can use the information to communicate with parents about their children’s proficiency and progress)
- parents and students (who can use the results for making educational and vocational decisions)
- potential employers and post-secondary institutions (who can use the information to make decisions about hiring or acceptance)
- principals, district or divisional administrators, and teachers (who can use the information to review and revise programming) (p-129)

The prerequisite of assessment of learning is the collection and interpretation of information about students’ achievements in significant curricular areas; it also represents the nature and intricacy of the proposed learning. Indisputable learning for understanding is considered more than recognition or evoke of facts, assessment of learning tries to enable students to demonstrate the complexity of their perception and learning. Application of key concepts and knowledge is also needed by students, in addition, skills and attitudes which are genuine and consistent with existing thinking in the knowledge domain are also needed. Assessment of learning requires selection of appropriates methods to address the proposed curriculum outcomes and the continuum of knowledge that is mandatory to achieve the outcomes. The methods must facilitate
students to demonstrate their understanding and create sufficient information for supporting realistic and justifiable statements about the nature and quality of their learning for appropriate use of this information by others.

2.11.16 Assessment of learning methods

Tests and examinations are main tools for assessment of learning, at the same time a diversity of products and demonstrations of learning assortment are included in methods. Some other ways like exhibitions, performances, presentations, simulations, multimedia projects, and a variety of other written, oral, and visual methods are also used in assessment of learning.

Assessment of learning requires very cautiously designing and developing approach to ensure quality the information upon which decisions are made. Assessment of learning is planned to be summative in nature, and expected to produce justifiable and accurate descriptions of student competence in relation to pre-defined objectives and, rarely, in relation to other students’ assessment results.

Ornstein (1989) explains the process of certification of students’ proficiency; it should be based on an exact, dependable, suitable, and reasonable process of assessment and evaluation;

Reliability in assessment of learning depends on how accurate, consistent, fair, and free from bias and distortion the assessment is. Teachers might ask themselves:
- Do I have enough information about the learning of this particular student to make a definitive statement?
- Was the information collected in a way that gives all students an equal chance to show their learning?
- Would another teacher arrive at the same conclusion?
- Would I make the same decision if I considered this information at another time or in another way? (p-13)
Learning outcomes serve as the reference points for assessment of learning and are identified in the curriculum which makes up the course of study. Assessment targets comprise measures of these learning upshots, and a student’s recital is interpreted and reported with relation to these predefined learning objectives and outcomes.

Ornstein (1989) has quoted different situations;

In some situations where selection decisions need to be made for limited positions (e.g., university entrance, scholarships, employment opportunities), assessment of learning results are used to rank students. In such norm-referenced situations, what is being measured needs to be clear, and the way it is being measured needs to be transparent to anyone who might use the assessment results. (p-23)

Student’s performance and proficiency is reflected by results. Assessment of learning results in reports about students’ ability in extensive areas of study, assessment of learning tasks may focus on the key knowledge, ideas, skills, and temperaments already set out in the curriculum document. Teachers use different approaches for assessment of learning, their records offer the details about the worth of the measurement (Rogosa, 1999). Comprehensive detail of the different components of the assessment of learning are necessary, with a portrayal of each and every component’s capacity to measures, with the accuracy and following the criteria and reference points. It needs to take in account the supporting proof related to the outcomes as validation.

Teachers maintain records in detail and in descriptive form for providing significant reports to parents and community. Simply a figurative illustration of a student’s achievements is inadequate. Reports to parents and others should recognize
the planned learning, the assessment techniques being used for gathering the supporting information, and developing the criteria used to make the decision.

Rogosa (1999) depicts the process;

Assessment of learning operates at the end of a learning cycle, response to students has a less obvious effect on student learning than assessment for learning and assessment as learning. Nevertheless, students do rely on their marks and on teachers’ comments as indicators of their level of success, and to make decisions about their future learning endeavors. (p-99)

2.11.17 Assessment of Learning

Assessment of learning is self-explanatory process. Kebdall (1996) has quoted example of asking a near-sighted person for demonstration of driving proficiency without glasses can better convey basic mechanism of assessment of learning. The examiner can get better and precise picture of the driver’s capability, and to verify him or her as skillful. Following the same way, demarcation in assessment of learning needs the essential space that may allow students to formulate the exacting learning noticeable. Rogosa (1999) has given detailed context;

Multiple forms of assessment offer multiple pathways for making student learning transparent to the teacher A particular curriculum outcome requirement, such as an understanding of the social studies notion of conflict, for example, might be demonstrated through visual, oral, dramatic, or written representations. (p-76)

While assessment of learning never always helped teachers in discriminating instruction or resources, it possess a reflective effect on the position and promotion of students and on the character and delineation of the future education and training that
students will be given. Assessment results need an amount of accuracy for making some
cwise recommendations.

Student proficiency can be reported by many possible approaches. Reporting
assessment of learning acquires to be suitable for the audiences for whom it is proposed,
and must offer all of the information required for them to make rational decisions.
Despite the consequences of the form of the reporting, it needs to be honest, fair, and
present adequate detail and background information for clear understanding.

Conventional reporting, depends only on a student’s normal score, gives very
petite information about learner’s skill development or knowledge. The alternating
mechanism, which identifies many kinds of achievement and makes a sketch of a
student’s level of performance on an emergent-proficient continuum, is the parent
student-teacher conference. This forum encourages parents with sufficient information,
and reinforces learners’ responsibility for their learning.

Classroom has very complicated process in it. Teaching-learning process has
multidimensional approaches for achievements of predefined learning outcomes.
Many innovations have been initiated to educational systems; only some had a primary
effect on classroom environment. Change requires learning and nothing can actually
transform in schools except teachers and administrators have educated new knowledge
and skills, and can transfer learning to the classroom objects.

Norris (1987) provides detail on this process;

Thinking about assessment as a major facilitator of learning is likely to be
one of the most significant changes in classroom practice. This change will
challenge many educators’ fundamental beliefs about their work and about
education, and it will require of them new knowledge and skills.(p-97)
Norris (1987) offers a frame of action for teachers, administrators, and professional developers because they all work jointly to formulate primary changes in classroom assessment techniques.

Teacher holds central place in classroom mechanism. Change can only be launched if teacher is fully practical in teaching-learning process. The accomplishment of embedding and sustaining any sober modification to classroom performance fully relies on changes in the personality of individual teachers, facilitators, administrators, and district or division leaders.

Changing practices needs deeper understanding by the educators. Instant action and the motivations for the change and the delicate dissimilarities between the mature and the new ideas look essential for classroom change. Mostly people have conservative approach to processing new information because human beings are tending to protect existing beliefs and habits relatively transforming them or construct new ones.

Phillips (1987) depicted the change process;

Teachers tend to assimilate new information into our current knowledge structures, rather than create new structures to fit the new information. Teachers may integrate information into our comfort zones, and feel teachers are practicing the innovation. But teachers may not have fully understood the innovation, and so have not made the intended changes. (p-86)

### 2.11.18 Alternative assessment techniques

Educational landscape has been using alternative assessment techniques from several decades and some of them give the impression of adaptation. Considerable changes in classroom assessment function have not been observed. Teachers necessitate
suitable time to plan about existing practices, decide about differences, and support conscious adaptations and innovations.

Stephen (1987) addressed the changing assessment practices;

It is not just intellectual work; indeed, assessment has an inherent emotional component that impacts on motivation. Consider, for example: Assessment for learning is premised on a belief that all students are capable of learning the intended curriculum, and that teachers have the requisite content knowledge and the pedagogical skills to find ways to facilitate students’ learning. If a teacher does not hold this view, he or she may feel conflicted and may focus negatively on why it can’t work. (p-116)

Learning is a complex phenomena and assessment as learning entails re-conceptualizing not only assessment including teaching and learning also. It also means giving up the more conventional erects of transferring knowledge, managing and maintaining the classrooms, and upholding the control, and as an alternative process redistributing liabilities in classrooms. This main transfer in approach can result in a sense of disequilibrium and dissonance.

Every teacher will perceive changes in classroom assessment practices following by his own conception and history. These perceptions require to be tired up, clarified, and investigated as part of any new learning and innovation. The proper conception of change is also very important factor in educational stream. It is also imperative to understand that dissonance is an essential fraction of change. Teachers who are working as agent of change, especially in their own understanding of assessment, at the same time they are learning new ways of assessing and revising their views about children learning process and also about teachers role in supporting learning for every student. They are reviewing, monitoring, adapting, and reflecting on their own effectiveness in the classroom. Definitely, these teachers are utilizing the identical processes as their students are
practicing to become their own best evaluators. They are also following their own learning path.

Mayo (1967) has described same;

Although many teachers have very large repertoires of assessment methods, they may need to revisit and enhance their knowledge and skills in identifying purpose, deciding what to assess, choosing methods, ensuring quality, interpreting evidence, and using the assessment for the intended purpose. Changing classroom assessment depends on teachers building repertoires of knowledge of learning theories, content knowledge, and pedagogical knowledge.(p-34)

2.11.19 Changing Focus of Assessment

Teachers’ knowledge about curriculum and its basic domains helps them to improve classroom practices. A significant facet of changing assessment to spotlight on learning is teachers being well-informed about the curriculum particular to the subject areas that they teach and the learning styles of students and learning material. Teachers are also considered to be well aware with the different misunderstanding which can happen by students to a subject. Teachers can offer valuable expressive feedback by using their own experience. They can also generate targets which can help for learning enhancement.

Marzano (1996) cited;

Attending to the purposes of assessment, and putting the emphasis on assessment for learning and assessment as learning, directs differentiating instruction for all students. When teachers have considerable expertise in tailoring pedagogical practice, they are in a good position to address the needs of groups and individuals. They can plan some learning contexts that are the same for all students, some for groups of students, and some for individuals.(p-156)
Teachers can sketch out an extensive variety of resources, activities, and approaches to fit into place students in their own learning and offer familiarity with the concepts of practice and support. Their planning will make available the outline which can be used for identifying the intents, make the associations overt, reinforce the affiliations, and point out the misunderstandings that can influence the process. Assessment facilitates on-the-spot modifications, or, guiding towards new destination.

Kebdall (1996) narrates it;

Even when high-quality professional development and communities of practice are in place, changes will not occur unless there is also strong instructional leadership and creative management on the part of school administrators. Administrators have the responsibility for creating the conditions necessary for growth in teachers’ professional knowledge. (p-73)

Policy maker have need of a systematic understanding of the theories and the application procedures of classroom assessment therefore that they can efficiently examine and adjust school policies, allot funding, observe changing practices, and fashion a safe culture within the school environment which allows teachers to experience safety while challenging their own beliefs, and change their practices. Schooling is assumed as a powerful and vital support of every community, when a few practices or policies change in schools, public should be informed about the reasons for the change, practice, and the consequences of the change.

Marzano (1996) advocates the public awareness about change;

Assessment is the “public face” of education, and changes need to be shared with all who are affected. Students comprise the first and most important group that needs to understand the changes being made. When students come to understand that the primary goal in the assessment process is learning, through assessment for learning and assessment as learning, and that the teacher is there to help them, they begin to trust that learning is not a competition. (p-39)
Parents have significant importance in educational process. They encompass an additional group that desires to understand the change process within educational setup. Assessment and evaluation have usually been considered as private, inexplicable activities, repeatedly escorted by a sense of apprehension. A concerted effort can be made by shifting parents’ perceptions, and captivating their support. General public is third very important and influential group that wishes to understand the change process and also wants to take active part in it, some time they create problems for change as well. It is imperative that public come to understand the rationale of assessment, and also observe the clarifying and separating purposes which may contribute for improved decisions for all concerned communities.

McMillan (2001) gives his views;

If teachers are to support students’ efforts at becoming lifelong learners who are capable of meeting the challenges of a complex and constantly changing society, it is important that they, as well as their students, • think and work with a mindset of being in charge of their own destinies, always hungry to know more
  • value deep understanding
  • reserve judgment and maintain a tolerance for ambiguity
  • think about a range of perspectives while systematically posing increasingly focused questions. (p-129)

2.11.20 Evaluation for Professional Development

Continuous professional development is unavoidable component of any professional organization. Cavernous learning and its application in practice involves a deeper conception and perception. It is much more than attending workshops and courses. Efficient professional development is more than a uniform delivery of information to teachers. It considers teachers’ varied backgrounds and the assorted
working contexts. Teachers have a liability of acquiring pedagogical knowledge and publicizing it to others by expanded networking. Professional development can occur in formal way, it can be in-service, pre-service and in form of life long learning. It can also be informal closely related to classroom assessment practices. It can occur in initial teacher training or throughout teachers’ careers, at individual level or collectively.

Lysne (1984) has described the process;

For some educators, it is not easy to have direct personal contact with colleagues. An electronic bulletin board or conference related to assessment allows them to ask one another questions, post examples, participate in discussions, and share ideas. There are many excellent books and articles about classroom assessment, some of which are noted in the margins and included in the resource list at the end of this document. (p-18)

Reading educational resources and going through professional journals provides a chance for educators to analyse their own assessment experiences, imitate on their students’ assessment experiences, look at their assessment beliefs and practices, encourage new ideas and apply all learned experiences in practical stream.

By going through the semester time frame or any time frame in any educational setting, teachers can maintain detailed logs of their routine assessment practices, all together with a portrayal of the assessment and its objectives. They also deal with issues of quality, and techniques of using the assessment information. Teachers can appraise their logs and can locate the proportion of their assessment as per each category and objective, this practice normally occurs at the end of a term. They can also decide changes for the improvement of educational process.

Infante (1995) gives his opinion about teachers involvement in assessment process;
Teachers generally need to undertake or participate in some summative assessment as a basis for reporting grades or meeting accountability standards. However, the task of summative assessment for external purposes remains quite different from the task of formative assessment to monitor and improve progress. While state tests provide a snapshot of a student's performance on a given day under test conditions. (p-67)

2.11.21 Role of Teacher in Evaluation

Teaching-learning process is initiated, facilitated and completed with the active participation of an efficient teacher. The teacher in class environment is understood to have the ability of evaluating all of the students by using a variety of skills throughout the school year. That teacher also has the liability to mediate when he or she recognizes an insufficiency in a learner. Usually, teachers assess their classes only some times, by using tests that shrivel all of the related skills into a single number or letter grade. Therefore, they are unable to understand their students' performance in a comprehensive way. Preferably, the system has to work like the self-repair method of biological organisms. Students and teachers give the impression to look at courses in very different ways.

Infante (1995) stated;

As teacher view his courses in terms of learning goals and material coverage, with examinations and assignments as a necessary chore. Students, on the other hand, seem to view them as a certain number of assignments and a certain number of examinations or papers. (p-29)

Teachers have the responsibility to bridge up the gap between conventional assessment techniques and modern techniques. They have to make unambiguous associations between the skills they anticipate students to bring with them and customary modes of assessment. When a specific skill or way of thinking tries to make
the proper solution to a problem, teacher needs to endeavor to find ways for letting the student to assess this skill in advance, rather than letting him or her stagger up against especially while trying to solve a larger problem.

Langer (1989) gives description;

When choosing assessment strategies, one also needs to remember a few simple points about how students learn. First, the single most important factor influencing learning is what the learner already knows. This observation has two important ramifications. One is that which is expected by the students to be able to call on knowledge and skills from prerequisite courses. The other is that which is expected from them to come prepared to class each day. (p-57)

Generally students are supposed to learn all desired skills and content and also it is vital to reveal what was their learning task. The second aspect to keep in mind is that responsiveness of learning motivates learning. Teacher possibly will attempt to find ways for students to understand the payoff when they try to improve their learning and thinking skills. Results can highly motivate students for learning.

Black (1998) has tried to distinguish between measurement and evaluation by indicating the emphasized points to both domains. The stress is upon single aspects of subject-matter achievement or precise skills and abilities in the measurement, while evaluation emphasizes upon extensive personality changes and key objectives of the educational program. Black (1998) states;

In function, evaluation involves;
- The identification and formulation of a comprehensive range of major objectives for a curriculum,
- Their definition in terms of pupil behavior to be realized, and,
- The selection or construction of valid, reliable, and practical instruments for appraising major objectives of the educative process or characteristics of personal growth and development. (p-91)
Test is a vital tool for evaluation of learning outcomes of students. Achievement test are used as a device in measurement. Educational measurement, acquiesces measures of learner achievement in subject-matter areas by the using of achievement tests, especially in acquisition of skills and information. Although earlier test technicians recognized both the ability to apply informative data and skills in problem-solving and interrelated factors, such as aptitude, interests, attitudes, and personal-social adjustment, as achievements and learning outcomes, these technicians revealed slight unease about the assessment of a lesser amount of tangible aspects of educative growth.

2.12 Scope of Modern Evaluation

Modern evaluation has extensive role to play in educational process. It is very comprehensive in nature and by design. Rogosa (1999) has worked out on its scope. His report on the appraisal gives a meticulous account of the development and application of present-day evaluation devices. Basically these instruments were intended to measure the major objectives of high-school instruction. He has given following objectives:

- For aspects of thinking tests of interpretation of data, application of principles, logical reasoning, and nature of proof.
- For social sensitivity; tests of application to social problems of social values, social facts, and generalizations.
- For civic and social beliefs, scales of social, political and economic beliefs.
- For aspects of appreciation in literature and art: a variety of techniques.
- For interests: an inventory of personal, social, and school interests.
- For personal and social development: various self-reporting scales and anecdotal records. (p-72)

Additionally, a range of pupil-record forms for taking down reading and listening are being used for evaluation. These techniques include tests, scales,
inventories, questionnaires, check-lists, pupil logs, and other records and applied for assembling data about the accomplishment of objectives in the curriculum.

2.13 Recent Trends in Evaluation

Modern evaluation has focused all three domains of learning process, if referred to bloom taxonomy. The modern teachers have major concern about vital functional learning outcomes, several of which are less concrete and comparatively easy to measure as compared to the concepts, skills, and abilities included in subject-matter tests of the past.

Impara (1996) has given some detail;

The concern for the total development of the child physical, emotional, social, and intellectual has resulted in an emphasis upon a sound understanding of child growth and development and of individual and group differences, as well as upon the personal and social adjustment of the pupils. (p-97)

General educational development test typically reflect information in different forms like verbal, graphic or written. Normally test exercises are designed to measure the capability of the character to understand and interpret the content and material presented. This is in contrast with the inaccessible test item which mostly emphasizes the recall or identification of items of information. These generally cover areas such as, science, language arts, social studies, including mathematics and literature.

Another significant development in current evaluation studies is the shift towards the role of the individual in measurement process. Concentration is also on the role of small groups and group dynamics.

Sander (1993) and Schafer (2000) are of the opinion;
To enable the right pupils to receive the right education from the right teachers may be considered the aim of the good educational system. Although at first glance this definition seems merely a truism, it does provide the basis for an approach to the question of why pupils should be evaluated. The right pupil is the pupil whose personal attributes and opportunities enable him to profit fully from the education offered him. (p-25)

Society and educational institutions are inter-related. School provides input for society and society is assumed as feedback provider for the school. The accurate education best suits to the requirements of both society and the learner. Apparently, the mission of determining this "appropriateness" of learners and educations is a assignment for the process of evaluation. Learner’s characteristics and opportunities must be ascertained. Student behavior may be evaluated preferably at all stages of the interface between learner and education to decide the fitness of each other.

In the same way, pupils vary, mutually within themselves and among themselves, in about all facets of their personalities and particularly in all those facets which settle on their fitness for accomplishment in the different subject-matter areas and social roles expected by contemporary society. Learners have different personalities and different attitudes towards learning process. They differ, like all living organisms, in the behavior and reactions of which they are gifted, in intellect, concerns, and opportunities. Individual differences result in dissimilarities in the different types of activities like educational and vocational.

Society has the responsibility to provide a favorable environment for continuity of educational activities. Society can take measures for preventing the intellectually weaknesses from filling positions requiring intellectual potency. Schools are the centers where society provides the chance for educating the mental strengths; it needs to keep
itself going on. The schools are the proper places for fitting of pupils to curricula and citizens to social roles and responsibilities. This whole process is carried out with the best results obtained from individual adjustment and social welfare.

Schafer (2000) has narrated;

In the schools this fitting process is guidance, i.e., the guidance of pupils differing in all ways within and among themselves into activities, curricula, and vocations which differ within and among themselves in the capacities they require, so as to make the best fit between pupil and education.(p-49)

2.14 Guiding as a Trends in Evaluation

Guidance is also used as a device to gear up educational process. It involves the evaluation of learners for determining of their specific capacities including strengths and weaknesses. Markle (1969) has explained the process of guidance in educational phenomena;

As pupils proceed up the educational ladder in secondary school, with its general curriculum providing the common core of skills like reading, writing, arithmetic etc, that everyone in civilization may acquire, the process of evaluation may operate upon them so as to reveal differences in their aptitudes, abilities, achievements, interests, environmental backgrounds, and all other relevant attributes. (p-91)

Students follow their own interest, attitude and aptitude for selection of future areas of study. Mostly during secondary schooling pupils set out future paths, some of them follow one set of interests and objectives and some toward another.

The knowledge which the teacher has gained about the pupil for the duration of the secondary schooling can directly influence the choice of the learner and his parents
about their desired type of education. Normally this information is collected by continuous evaluation all the way through the ten years of the educational process. Usually the evaluation is considered comprehensive during these ten years by involving some imperative aspects of the pupil's personality as it is probable for a teacher to recognize it. Likewise, guidance and evaluation go together during the secondary schooling from start to end.

McMillan (2001) has discussed the continuity of the evaluation process. He has the opinion that this process may serve during all the time that the teacher can easily monitor the pupil during different occasions. Teacher’s observational data includes different recitation, assignment, conversation, behavioral features and these will be used as material for the evaluation process. It will provide a baseline for a record whereby he may mount up knowledge of the pupil and handover this accumulated record of evaluations and facts to the next teacher.

Whole personality of the learner is affected by the comprehensiveness of the evaluation process, not only his intellectual achievement. Extremely imperative in guiding process is the evaluation of a pupil's knowledge. At the same time his aptitudes, interests, attitudes, temperament, social adaptability, habits of work and play, physical characteristics have significance fitting him into the world of work and as well as for his happiness. A learner has diversified aspects of personality and each of these aspects must be evaluated with respect to its position regarding other different aspects or facet of the learner and also in relation to similar aspects of the population at large.

Impara, (1996) describes the relation between guidance and evaluation,
Given the instruments, i.e., the achievement tests, with which to make such rankings the teacher or administrator may then use them for the following purposes:

- To maintain standards
- To select students
- To motivate learning
- To guide teaching
- To furnish instruction
- To appraise teachers, teaching methods, books, curricular content, etc. (p-18)

Guidance helps out educational institutions in the process of selection of pupils. The function of educational evaluation is a form of guidance in reverse. (Marzano; 1996). Generally selection is carried out with the advantages of the institution; as the main contemplation, and advantages to students are rejected.

Educational measurement is used in selection purposes following the assumption that it can facilitate by forecast of future success. Such selection purpose may be illustrious from the maintenance-of-standards purpose.

### 2.15 Motivation of learning by examinations

Examinations also serve as a motivating factor. The motivation of learning by examinations has quite a few aspects. Such motivation may be a issue of crystallizing the complete mass of the social and self-criticism into one test score which is requirements of the learning process. Thorndike's Law of Effect is used as an application for this use of evaluations especially when the student comes to understand the different types of behaviors which will be rewarded in positive ways and equally for other types of behavior.
When trying to furnish such motivation, the variety and nature of the test can have critical effects on the learning process covering all the aspects. Expectation about the measurement of learning will guide the learner about areas of study. Rote memory only becomes method of study when past experience of student leads him to expectation that his learning will be assessed on the basis of his rote memorization.

Nash (2000) has elaborated this process,

If on the other band, he expects the test to be one of ability to apply principles, or of ability to interpret data, or of broad "acculturation" in the lore of a subject, or of acquisition of practical technical skills, then he will direct his study habits and learning processes toward these ends. (p-114)

Educational evaluation motivates learning in multiple directions. Evaluation of irrelevant and provisional outcomes of education directs towards detrimental study habits whereas valuation of meaningful, long lasting, and congenial to a pupil's interests and capacities outcomes leads to attractive and pleasant study habits. In view of the fact that measurement directly affects learning, the purpose of measurement has a vital part in the educational process.

Evaluation guides teaching at the time of furnishing of identification of precise strengths and weaknesses in the learner's achievements or capacities. The teacher can plan for eliminating the weaknesses by utilizing special teaching methods and techniques, or to avoid them by turning learning in the direction of areas where the pupil's efforts will be more productive. The causes of flaws in a specific theme can be traced by tests to a range of potential learner’s shortfalls underlying it. Weakness in a math or physics course can be traced because of lack of either conception power or speed in reading or it can happen due to lack of general aptitude, or to lack of the
presupposed locale substance, a lack which possibly be exposed by tests planned to assess attainment at a lower level of the subject.

Learner's difficulties in any subject area can easily be traced out by utilizing tests to a explicit insufficiency, such as incapability to handle certain combinations of facts, or be short of some correction technique, or the use of a burdensome method of short division. Diagnostic evaluation is also helpful in teaching-learning process. Diagnostic testing can disclose the accurate sources of a learner’s shortcomings and can better guide the teacher towards most favorable way of surmounting them.

Students who are not well-equipped may be observed and must provided special attention intended at the basic roots of the "failure." They may be guided about their weak areas.

Cizek (1997) has explained it in detail;

Teachers may discover what parts of a topic or unit need to be re-taught, or taught differently. Those pupils who are capable of doing exceptional work may be discovered by evaluation procedures and the teacher's efforts may be guided toward special assignments and references for them for the fuller realization of their potentialities. (p-189)

### 2.16 Evaluation as Appraisal

Appraisal works as reinforcement device and the use of evaluations for the appraisal of different educational agents such as teachers, teaching methods, and text books is straightforwardly recognized but not very simple to be practiced in school environment. Chapman (1989) has given brief description on use of evaluation as appraisal;

The principle underlying all such uses of evaluations is that whatever yields the greatest realization of the educational objectives desired is the best
teacher, teaching method, textbook or other instrumentality of the educative process. And since most educational objectives may be stated in terms of desired changes in pupils, educational evaluations of those pupils will provide a measure of the degree to which the objectives are attained and of the effectiveness of the instrumentality concerned. (p-138)

Educational evaluation can better serve the purpose of appraisal if reliability of results is not questioned. Because the complexity lies in the status of results, when educational evaluation is used for appraisal purposes. Literature shows that no worthwhile, trustworthy results can be attained except the appraisal becomes a restricted scientific experiment with important variables detained constant and with statistical tests of implication applied to the results.

Well skilled researchers can help out in the valid use of educational evaluation. The details of experimental design that may be involved in the applicable use of educational evaluations for appraisal research (Ebel, 1962).

The authorities who make appraisals like principals and superintendents must get your hands on a mastery of these scientific methods for themselves or turn over the problems to experts. It also seems suitable that authorities who are not well expert in using these may be stopped to make use of evaluations for the appraisal of educational instrumentalities.

Evaluation is required for determining the rightness of pupils and educations for each another. The reason behind this is the heterogeneity of pupils in all phases distressing their adjustment to themselves and society as well. The heterogeneity of activities in context of educational and economic domains is also important for students in a democracy regarding social life. The major role of educational evaluation is enabling the guidance of pupils, so for serving the purpose it must be continuous and
comprehensive. The one important function of evaluation is standards-maintenance; if it is appropriately conceived then it becomes extra standards-revelation function by serving in a more active and beneficial manner. (Ebel, 1962).

Quality of learning needs keen interest of teacher as well as institutional management. The motivational purpose must be observed to the quality as well as the quantity of the learning endeavor. Evaluation attempts to trace out the weaknesses and strong points of learner and facilitators. The diagnostic task of revealing pupil strengths and weaknesses entails special practices other than general practices such as single-score tests. The instructional role may be realized only if assessments are permissible to be self-obligatory and self-criticisms. The controlled and statistically analyzed experimentation can ensure appraisal function of evaluation by undertaking its significant function of guidance.

Ebel (1962) described guidance,

Guidance is not to be considered as operating only at the students' entrance into or exit from a particular level of schooling or vocation, although this is its most striking occurrence. Rather guidance should operate within the schooling period. For example, evaluations should assist in the solution of such guidance problems as choice of elective subjects. (p-58)
2.17 Simulation-based learning

Interactive learning is emerging trend now-a-days in teaching learning process. A rising tendency in secondary education has been observed towards interactive learning during last decade. Simulation techniques can be presented as example of this new trend (Magee: 2006). The key benefits associated with this approach include having a scientific setting without risk of causing learner impairment and having a learner-centered environment. Students learning by using simulation have articulated a feeling of amplified quantifiable competency as well as interest for the learning material.

Gaba (2004) advocates the exercise of simulation for the development of learner’s learning. The motivation for the exercise of simulation techniques for education and training in teaching-learning process has promoted its use in other relevant industries as well.

Gaba (2004) outlines quite a few various uses of simulation in education phenomena, including

(1) education and training of clinicians, (2) assessment of clinicians’ performance, (3) research and evaluation of organizational practices and exploration of human factors (e.g., fatigue), (4) the usability of clinical equipment, and (5) as a tool for helping create the desired “culture of safety” in the workplace. (p-84)

Simulation is also used to facilitate students in acquiring new knowledge and develop conceptual frameworks between detached pieces of information, also try to
integrate learning fundamental concepts with performance and development of the appropriate experimental skills. Stimulation is useful in educational process because it provides a risk free chance of learning. One key benefit associated with the use of simulations is risk-free environment, which lets students to make errors and learn from them without any fear.

Friedrich (2002) has given its importance in the context of training and development;

This is especially important for crisis training and training for clinicians in emergency medicine (discussed in a later section). However, there needs to be evaluation of simulation-based learning and training that shows increases and efficiency in learning in addition to the nature of learning that takes place with the use of technology.(p-48)

Simulation-based learning has been given keen concentration during last few years. Recently it has also been compared in experiential studies with different learning approaches, for example, conformist and problem-based learning approaches. Steadman (2006) assessed the performance of learners by using simulation (SIM) vs. students using problem based learning (PBL) techniques. Results showed that the SIM group performed considerably better on this assessment than the PBL group. One possible justification for this result is that the simulator experience needs the learner to be more betrothed in learning, activating manifold learning corridors e.g., auditory, visual, and tactile.
Simulations are being in practice in formal educational stream as well as are progressively more used in medical education, teaching and assessment, though their exercise is not very prevalent. Morgan and Cleave-Hogg (2002) has quoted;

An international survey was conducted of the use of simulation in education, evaluation, and research in anesthesia found that most simulation centers responding to the survey (38%) used simulation in undergraduate and postgraduate teaching, but very few centers used simulation for evaluation or practice assessment. The biggest challenge to incorporating simulation into curricula is the high cost of equipment and their maintenance. (p-172)

Stimulations facilitate learning and these have very positive impact on learning process. Regardless of their slow consumption, simulations have been revealed a positive impact on learning. Recently reviewed medical simulations (Issenberg et al. 2005), showed that high-fidelity medical simulations facilitate learning under specific conditions. A pilot study by Gordon (2006) investigated that ability to learn basic concepts can be increased by simulation. Study showed that the simulation group showed higher scores both immediately and 1 year after intervention. These all evidences reinforce the idea of utilization of stimulation based learning in educational process.

2.18 Evaluation of Performance

Evaluation also plays an effective role in appraisal of performance. Devitt (1998) established fact that a performance rating scale can be used for discriminating between resident and faculty anesthesiologists. The existing defy in utilization of simulators is the capability to generate scenarios that are suitable and consistent in assessing performance on an assortment of targets.
Kapur and Steadman (1998) described the situation;

Furthermore, it is not quite clear what the nature of the learning is that is acquired through simulation. However, although much research needs to be conducted on the uses and effects of simulations on medical learning, it is possible to speculate about the reasons simulations may be more effective than traditional methods for fostering learning and improving performance. (p-28)

Normally simulations are considered as good ways to get hold of skills and acquire knowledge. Pellegrini (2006) found that simulation-based training provides an massive opening for learners to extend and purify their skills in a diversity of scenarios, without any risk. Stimulation based learning proved much effective in medical education. Simulators that have been designed and developed for exercise in surgical training consisted of mechanical simulators, virtual reality environments, computer-based simulators, and hybrid simulators. These mentioned simulators proved beneficial for the acquisition of different types of skills.

In the early hours constructivists, such as Piaget and Vygotsky, reinforced the knowledge and skill attainment through engaging vigorously in learning and through management which opens up the opportunity for reflection and better understanding.

Spiro (1991) has given detail discussion on nature of simulations;

Given the dynamic nature of simulations, and therefore the opportunities for active engagement, they appear to afford the possibility of developing the needed cognitive flexibility to adapt performance to the changing tasks demands, as suggested by CFT as critical for complex learning. (p-93)
Simulations may not be adequate to promote learning merely on their part, cognitive learning principles are required to guide the design of instruction. It will be most practical and efficient way which needs deeper forms of assessment such as conceptual understanding of modern techniques e.g., rating scales, multiple choice items.

2.19 Simulators

Team work performance is also significant area which is under focus now-a-days. Team performance can be improved by using stimulators in group process. Simulators are frequently used in group situation especially in different programs offering team-based simulation training.

Baker (2005) narrated team work in these ways;

Research has shown that such training may significantly increase teamwork performance; and that didactic teamwork training can be enhanced by the use of high-fidelity medical simulations, possibly due to the task representing actual clinical care involving several patients (Shapiro et al. 2004). However, further research needs to be conducted to see how much simulation training is needed to produce a significant increase in team performance. (p-15)

The application of computer games to learning has been topic of discussion among educationists for more than two decades (e.g., Gee 2003; Gredler 1996; Malone 1981; Prensky 2001; Rieber 1996). Conversely, a huge body of literature indicates the educational prospective of computer programmes; this zest is tempered by the acknowledgment that the experiential confirmation does not clearly ascertain enhanced learning outcomes for such activities. Mostly important reviews on instructional
computer games (Dempsey et al. 1996; Mitchell and Savill-Smith 2000; Randel et al. 1992) showed that the pragmatic findings on the learning efficiency of computer programs are repeatedly conflicting. Experiential studies on the use of computer programs are very limited in number, particularly well-specified and employed in-situ research that scrutinizes what occurs in real contexts over time (Fletcher and Tobias 2006).

Computer programs can easily be used in classroom environment but a widespread uncertainty articulated about using computer programs for learning exists in the shortage of empirically-grounded techniques of incorporating computer programs into classroom and lesson time. The massive number of programs and learning research has been apprehensive with learning abstractly that does not stipulate knowledge of subject area but requires connection to curriculum in school (Egenfeldt-Nielsen 2005, Unpublished doctoral dissertation).

Miller et al. (1999) and Papert (1998) has investigated the problem and suggested;

The investigation into computer games for learning should focus on how games can be carefully aligned with sound pedagogical strategies to be beneficial. Kaptelinin and Cole (2002) further argued that learning outcomes achieved through educational games depend largely on the external instructional activities context that structures the way students use and interact with computer games. (p-138)

The interaction between student and teacher is also important in teaching-learning process. Johnson (1985) has stated that classroom goal structure is the
specification of “the ways in which students will interact with each other and the
teacher to achieve the goal” (p. 669).

Student learns in individual capacity as well as in group dynamics. Learner
works independently, in an individualistic goal arrangement, for ensuring that learning
meets a predetermined criterion autonomously from the efforts of other students. For
this reason students are not anxious about what anybody else is accomplishing. The
viable objective arrangement focuses a learner’s attempt on performing quicker and
more accurately than colleagues. Therefore learners recognize they will be awarded a
reward by comparison with other learners.

Johnson and Johnson (1996) has given their findings;

A cooperative goal structure is the instructional use of small groups in
which students work together to maximize their own and each other’s
learning. Hence learners perceive that they are working together with
other students to gain rewards. According to social interdependence
theory, the way in which the goals in a situation are structured determines
the interaction patterns among participants, which, in turn, determines the
situational outcomes. (p-118)

Research reveals that cooperation is significantly efficient than antagonism and
characteristic efforts in endorsing achievement and retention, hitherto the styles in
which educational technology interrelates with substitute classroom goal structures are
comparatively unknown (Johnson and Johnson 1996, p. 804).

2.20 Cognitive Load Theory

Cognitive load theory deals with working memory load which may be affected
by the intrinsic nature of the learning assignment or targets and the way in which these
tasks are presented. Intrinsic cognitive load can be condensed by developing cognitive schemata. Cognitive schemata can be defined as a large number of interacting elements for a learner; these can be considered as a single element for a specialist who has a schema that accommodates the elements. Extraneous cognitive load has close relation with rate of delivery of information and possibly can be reduced through effective task information representation. Extraneous cognitive load and intrinsic cognitive load are generally considered as additive. Consequently, it was surmised that cognitive learning achievement can be facilitated by a combination of cooperative classroom goal structure and computer game application. Palincsar (1996) has described that a cooperative goal structure would help to reduce the intrinsic load of a learning task by enabling peer discussion and cognitive schema development for both expert students and novice ones.(p-19)

Tuckman (1998) has discussed the academic achievements of secondary school students which depend on factors like previous knowledge, study practices, motivation and confidence. He further explains;

Faculty uses a number of strategies to affect these factors, hoping to achieve positive results in both studying and learning. Perhaps the most entrenched strategy is that of tests and grades, which operate in a punishment reward fashion. While the appropriate use of standardized admission tests is debated and controversial, the appropriate form and use of tests within the school classroom have received far less attention and research. (p-57)

An exam is used in multi-dimensions; it serves as feedback device and at the same time tool for measuring the extent of academic achievements. It reinforces the learning and also provides input to teacher for improvement of teaching methodologies
and techniques. The use of midterm or hourly exams as a testing tactic is the common norm in educational institutions, particularly in lecture-style courses. This approach has many shortcomings, for example, the leaning for students to force for preparation for term exams, which is inefficient for knowledge and its retention.

(Donovan & Radosevich, 1999) and Willingham (2002) have worked over it. They state;

A small body of literature suggests that alternative testing strategies may better promote learning and achievement. To this end, the purpose of this study was to test empirically the effects of two assessment methods on academic achievement among undergraduate students enrolled in two sociology statistics courses a midterm exam structure, i.e., two exams given throughout the semester control group and a bi-weekly exam structure experimental group. (p-45)

Change is a bit complicated process. It is also time consuming activity. The intricacy of putting into practice the educational change has been a current hub of numerous educational leaders concerned in understanding the nature of impact in educational system, enforced by curricular and pedagogical reform efforts. Complex systems theory was introduced by Fullan (2003) as an organizing framework. This framework uses to make known core concepts such as non-linearity, impulsiveness and multi-level activity that are significant issues to challenge the actual educational systems prevailed over the world.

Elmore (1996) described the problems occur while shifting from local to global contexts;

He states that failures, historically, in generating successful large-scale reforms can be attributed to an “absence of practical theory that takes
account of the institutional complexities that operate on changes in practice” (p. 21).

Coburn (2003) supports the thought that educational restructuring and upgrading are matters of complication addressing that problems of extent curtail from the incapability of research to tackle the intrinsic multi-dimensionality between and within diversified educational institutions. She recommends utilization of better research designs for building a more refined and complex vision.

2.21 Learning Layers of Complexity

Complexity of learning has many layers. A huge body of research addressing the learning sciences has painted extra layers of complexity especially when the use of educational technologies has reinforced the educational reforms in society. Fishman et al. (2004) discuss difficulties in implementation that arise when computer access is viewed from the perspective of sustainability of innovations. (p-29). It was recommended that a few cognitively slanting learning mechanisms have traced out their means into regular mainstream practice because of the fissures that subsist between culture, competence and strategy norms operating at different system levels.

Cuban (2002) recommends;

Likewise remark that changes of deeply entrenched systemic organizational and operational factors such as how teaching time is allocated, how technical support is provided and how reliable technical tools are, must take place in order to move beyond simple fleeting modifications to practice. (p-78)

Complex systems approach is considered as effective device for the design, development and execution of an educational upgrading. Many research studies indicate
that complex systems processes are helpful in organizing and evaluating professional development activities along with educational technology tools that educate and guide the complex systems modeling.

Educational systems are using complex learning approaches as effective tool for enhancing the effectiveness of teaching learning process. Complex systems are being used almost in every part and in all aspects of our world.

Benveniste (1985) depicted;

They span the relatively micro scale of structures and behaviors such as molecules of water organizing to form vortices in fast running rivers or the single fertilized egg developing through embryogenesis to create differentiated cells that eventually become the human form. On more macro scales, schools, businesses, cities, animal populations and ecosystems are also thought of as complex systems. (p-67)

A diversity of disciplines arraying in compass and miscellany from physics and chemistry, Math, zoology, botany, biology, psychology and economics educate us about complex systems, how they function and operate. The concept is receiving reinforcement from many research organizations which are working to strengthen it. The investigation for universal principals which can prove helpful in describing and unifying these seemingly disparate domains is in process. Numerous significant pathways have been added in the discipline; on the other hand, confederacy of our knowledge is not on the horizon (Bar-Yam, 1997).
Regardless of disparity in physical components or means, complex systems can be usually explained as accessible when any given numbers of consistent elements correspond in non-linear ways.

Gell-Mann (1994) explained the process;

The patterns of interactions form a collective network of relationships that exhibit emergent properties that are not observable at subsystem levels. When perturbations occur, the network self-organizes in, often unpredictable ways where new properties can emerge.(p-87)

Simple observation is unable to accurately determine the conduct of the system. The style of complex systems communication, acts in response to perturbations and self-organize which is probably understood by going through the dynamical processes by which they develop over a spam of time. Complex systems spot out reliabilities exists in certain information and transform behavior in the real world by using those, it all happens after acquiring information from their environment through feedback.

2.22 Evaluation for Conflict Resolution in Curriculum

Conflict is referred to a situation which arises due to imbalances of interests over certain domain. Learning to keep away from conflicts and making an effective attempt to resolve conflicts is an imperative needed characteristic for any active and productive member of society in usual and the classroom in specific. The skills and information which help people in resolving conflict may take broad view crosswise circumstances, conflicts and their resolutions are always surrounded in diversified contexts like personal, interpersonal, and cultural.
Barab (1999) gives his expert opinion about it;

This inherent embeddedness suggests that fostering conflict resolution is as much about social awareness and critical thinking as it is about learning a pre-defined set of conflict resolution strategies. From this perspective, an important part of conflict resolution curricula is to place learners in rich situations, to scaffold their generalization of conflict resolution strategies and skills, to provide opportunities for them to apply skills to diverse cases in a variety of settings, and to enable ongoing reflection. (p-92)

Learning is directly associated with the situation in which learner is being treated. Situation can enhance learning as well as can reduce its impact by influencing all the learning process. Learning about conflict resolution with reference to different situations can help in developing a situated appreciation for the value of conflict resolution skills among students.

It is necessary to examine the use of technology for establishing these conflict situations and offer measures to students for working collaboratively for resolving conflict. Video-based technologies can be used to place the content followed by thorough discussion to support the collaborative grilling of the videos. Mostly students observe the situations presented in the videos, they become familiarized with a range of conflict resolution skills. This innovative pedagogical approach is unswerving with the sited cognition standpoint considering that observed information and skills are expected to be functional in future when learned specially with the background of exercise.

Greeno (1998). Supports these efforts:

(1) What might a multimedia-enhanced curriculum intended to support conflict avoidance and resolution look like? (2) How do teachers and students use a multimedia conflict resolution curriculum in the classroom?
and (3) What impact does the use of a multimedia conflict resolution curriculum have on learners’ conflict resolving skills and self-efficacy toward conflict resolution? (p-36)

The experience of United States can be quoted as an example where the classroom is considered as a major arena for socialization (Dewey, 1980). Furthermore to the conventional approach “three R’s,” a growing credit about school’s additional responsibility of facilitates students to master the fourth “R” of Relationships. Programs focusing the conflict resolution and violence prevention are progressively being incorporated into classroom, school, and district-wide efforts (Lupton-Smith, Carruthers, Flythe, Goettee, & Modest, 1996; Department of Justice, 1995; Johnson, Johnson, Dudley, Mitchell, & Fredrickson, 1997; Thompson, 1996; Lovett & Walzak, 1997; McCarthy, 1992;). Consequently, educationists, practitioners, researchers and evaluators started giving more concentration toward developing more efficient curriculum, apparatus, and evaluation methods for addressing the diminution of aggression, violence, and conflict situation among teens, adolescents and children.

Lowry (1999) denoted;

For some populations, over 30% of students report having been involved in a physical fight sometime during the previous year. Although the most recently available data reveal a modest decline in the overall percentages of students who report being involved in violent situations, overall rates remain unacceptable (CDC, 2004). (p-94)

2.23 Evaluation and Emotional Intelligence

Learning has manifold associations, and these associations are interlinked with each other to a extended extent. Conflict evasion and resolution skills are covered under
multiplicity of name tags, however the core fundamental competences are alike. The sphere of influence under which conflict prevention and declaration skills are being found, can be listed like intrapersonal and interpersonal intelligence, emotional intelligence, conflict resolution, and social problem solving skills. Goleman discussed in detail the concept of emotional intelligence in his book titled Emotional Intelligence. His popularization reinforced the work of Salovey and Mayer (1990) who introduced the term. Gardner (1993) portrayed an individual’s capability to evaluate emotions in learners and in others and to tie together those emotions for interact effectively with the world as intrapersonal and interpersonal intelligence. Goldsworthy (2002) proposed an Emotional Intelligence Framework (EIF) for guiding interference and anticipation efforts in the socio-emotional domain.

Modern education has developed by using the notion of competence. Different educational institutes of professional education especially in Europe, Australia and New Zealand adopted the conception of competence for guidance and improvement of educational programs. Netherlands has implemented the competence-based professional education in the government institutions at very large scale.

(Poell, 2003) has illustrated;

Competence-based education may be regarded as a response to societal changes. Working situations have become more dynamic and complex, thereby posing new and specific demands on employees. (p-82)

Competency based curriculum also helps teachers to impart instructions effectively. This curriculum is designed to enhance desired competencies among learners. The term competence endows with a way to reflect about changes and
curricular requirements. Uniqueness of competence-based education comprises: a professional approach centered situations, targets, and roles and responsibilities, that may prove helpful for development of learning content; valid assessment techniques and mechanism; integrated curriculum for learning; perception about teachers as learning facilitators and about students as center of educational process.

(De Bie, 2003; Schlusmans, Slotman, Nagtegaal, & Kinkhorst, 1999).

Competence map is another imperative document which is used in developing competence-based education. Program of study and the final attainment levels can be described by this document. Instructional development requires documentation describing the curriculum content, but terminology and focus make competence map as a different document.

Bruner (1960) portrayed the phenomena;

In a competence map, curriculum content is described in terms of interrelated competencies rather than in terms of fragmented or disassociated knowledge, skills and attitudes. Competence maps typically consist of three parts. The first part contains competence descriptions, which provide detailed information about each competency that is distinguished in a certain domain or profession. (p-33)

All competencies are interrelated. Every competency derives from domain or base of other competency. Competence description usually contains information regarding its output or results, its relationships with other competencies and different example of the competency which is in practice. Mostly competence depictions are used for the purpose of designing the instruction, learning assignments, and evaluation procedures. Competence framework or figure falls in the second part of a competence
map which is usually comprised of a visual summary of the competence descriptions. A competence figure is basically used for communicating quickly what a competence map is asking for or what are its contents. General information about the relevant domain comes in the third part of a competence map which describes the goal and definitions used.

Characteristically, a diverse team consisting of skilled and knowledgeable experts like curriculum designers, field experts, practitioners, teachers, educational managers, branch representatives and many others develop a competence map. It is a complicated process and have more and more challenges during developmental phase. (Nieveen, 2002).

2.24 Evaluation and Empirical Research

Competency map is under consideration of many researchers now-a-days. Empirical research has revealed that the designing and development of a competence map faces numerous challenges (Stoof, 2006). Definition of competence and the existing difference between competence and interrelated terms such as knowledge, skills, ability, and expertise comes as a major challenge during this developmental process.

Hendriks (2002) has given his findings;

People may not know what competence means or how it should be defined. This problem has been reported in theoretical explorations of this topic and is characteristic of the terminology problems often encountered in instructional development (e.g., Stoof, Martens, van Merriënboer, & Bastiaens, 2002; Van Merriënboer, van der Klink, &). Another major
An instructional design device can prove helpful for designers for defining the concept of competence and it also can guides them during the development of a competence map, it is a possible resolution to challenges faced by designers of competence maps. This device or tool is expected to escort for improving task performance, augmented task-related knowledge, amplified contentment and increased inner constancy of the productivity of instructional plan and development teams (Gery, 1991; McKenney et al., 2002; Stevens & Stevens, 1995). Presently in-practice instructional design tools principally spotlight the development or production in spite of only on front-end planning and analysis (Van Merriënboer & Martens, 2002).

Research has indicated that standard scientific approaches to assessment and evaluation have very partial and restricted application and worth for education promotion practice, it may be because of the fact that they fail to produce anything valuable or they don’t work with concentration on target (Wright, 1999). Significant features extorted from education promotion may facilitate psychological research and performance and especially psychologists’ awareness about these concepts, it may help to protect and develop the future of psychology and its application in education. Association between education promotion and psychology may be conceptualized by going through the evaluation, which is the common context for the both domains.

(Carr, 1992) discussed the criteria about evaluation;
Evaluative criteria can deal with formative, process, or outcome evaluations. Formative evaluation is an ongoing process of collecting information to be used for further program development and improvement. Formative evaluation is often referenced in contrast to summative evaluation. Summative evaluation aims to make a determination about the overall impact of a program once it is complete. (p-293)

Formative evaluation is recommended as best practice throughout the program planning, development and implementation. In this phase feedback can be incorporated for the program modification. Preferred activities that can be involved in formative evaluation contain redefining objectives, targets and tasks, piloting content and materials, testing interferences and by reviewing the literature (Isreal et al., 1995).

2.25 Process evaluation

During the implementation phase of the program process evaluation is mostly used for tracking the activities. In general, process evaluation establishes the degree to which the program was implemented according to pre-determined procedure. Process evaluation identifies the strengths and weaknesses of the program or intervention by assessing the way how it achieves its planned end results (Dehar, Casswell, & Duignan, 1993). Procedures that can prove helpful in process evaluation consist of: program activities, program quality, and program’s desired goal or destination (Witherspoon & Wilson, 1990).

Coombes, (2000) states;

Finally, outcome evaluation aims to determine the relationship between a program and an outcome and measure results that are attributable to program activities. This is a much more global measure and is difficult
to capture because the effect of any program has resulted from a whole chain of previous events.

Changes in attitudes, knowledge and behaviors can be measured by outcome evaluations and also the costs after implementation of a specific program. (Higginbotham, 1992). The majority of education promotion researchers agree on the two primary tasks in program evaluation, which are; to evaluate results for determining the level to which the intervention achieved and to appreciate the process for identification of basic conditions for flourishing execution of an intervention. (Nutbeam, 1998). Untimely education promotion programs were assessed by using the education criteria and focusing on achieving desired educational objectives. This evaluation was descriptive in nature and also was based on some comparison standards. It seems problematic due to its non-possibility to differentiate between problems with execution and lack of program impact (Thompson, 1992). Eventually, evaluation of education promotion programs has developed to propose both process and outcome evaluation.

Process evaluation along with outcome evaluation are being used for assessing the impact of initiatives which are supposed to increase the possibility that the goals, condensed risk, improved defensive factors, and compact inequities, are assumed as confined and sustainable. Outcome actions may point out a program’s effectiveness but without process evaluation it may not be possible to discover that only a segment of the proposed addressees has been impacted by the initiative.

Nutbeam (1998) described;
Process evaluation allows the research-practitioner flexibility and the opportunity to respond to this discovery and fully understand the impact of the program (WHO-European, 1998). Wellness evaluations should not only include designs that combine different measures, they can include both qualitative and quantitative components in the evaluation. (219)

In 1993 an assessment system showing the capability of school graduates to demonstrate certain skills was developed by the National Education Goals Panel, established as part of the “Goals 2000: Educate America Act,” (Newburger, 1996, p. 70).

Students’ capacity to commune efficiently was among those prioritized skills (Newburger, 1996). Simultaneously, the United States Department of Labor developed a commission which recognized the competencies students have to build up while preparation for the global work environment. The ability to listen and speak well for carrying out certain work-related assignments and tasks, was one of those listed competencies. The objective of teaching necessary skills, for example oral communication, in secondary education is to get ready students to be more efficient employees and accountable citizens of state or country.

2.26 Evaluation and Communication

Communication is most important and desired skills in teaching learning process. Communication education is positively associated with academic and professional achievement for students (Rubin & Morreale, 1996). Students require speaking and listening skills for success in future courses and in the place of work.
Students can have the knowledge of effective communication techniques by a basic communication course which can provide a secure chance for developing and practicing these skills, which can generate positive feelings among students about future communication. The evaluation of these oral communication skills can assist and support students and their academic programs in multi-dimensional ways, also skill insufficiencies can be focused by using solutions at the classroom environment, department, or institutional levels.

Walvoord, Bardes, and Denton (1998) argued that;

“closing the feedback loop” when assessing skills is important at all three levels because it means that teachers get data on their students’ progress early, their initiatives for change can be supported, problems that require institution-wide solutions can be addressed, and assessment activities can be documented for accreditors and other external audiences. (p-61)

Establishing communication-based competencies has become gradually more imperative for a diversity of disciplines such as Business, Health, Engineering, and Biology. Technical disciplines have started to distinguish and discover the function of oral performance in their curricula (Dannels, 2002). Oral communication is overloaded with background inspirations, rationales, audiences, and policies precise to each field of inquiry. Each discipline has to be capable to evaluate communication for ensuring appropriate skills are being urbanized in the classroom. Evaluation practices that assess the degree to which students attain the communication outcomes predefined by certain disciplines to be appreciated, salient, and significant, must be developed. Dannels (2001) assessed suitable skills for communication as educated in a general education setting. The National Communication Association in United States is
dynamically concerned for assessment procedures within the discipline of Communication. In 1970, the Committee on Assessment and Testing was constituted in United States to spotlight on the testing of speech communication skills. A dire need for assessment strategies and tools was observed by different research studies, CAT produced numerous publications on the procedures of oral communication skills assessment for multi-grade students of all ages. A conference on communication assessment was held in July 1990, and produced recommended criteria for the assessment of oral communication skills (Rubin, Welch, & Buerkel, 1995; Speech Communication Association, 1993).

The NCA has recognized numerous communication skills through different academic meetings which are very important for students for learning at basic and advanced levels (Morreale, Rubin, & Jones, 1998). These skills contain the capability to be familiar with appropriate place and chance to speak, speak clearly with good expressions, present thoughts in an organizational way which facilitate others to comprehend them, listen conscientiously, selecting the most suitable and efficient medium for communication, structuring a appropriate message, recognize others’ level of receptiveness for their message, to give information with the support of illustration and exemplification.

Allen (2002) narrated the conference actions;

One result of the summer 1990 conference on communication assessment was the development of an evaluation instrument called The Competent Speaker, an evaluation form that identifies standards for evaluating students’ eight basic speaking competencies: being able to choose an appropriate topic and restrict it according to the purpose and the audience;
communicating the purpose of the speech in a manner appropriate for the audience and the occasion. (p-219)

The basic speaking competencies also include using suitable supporting substance (Dannels, 2001) to complete the rationale of the oral conversation; (Dannels, 2002) using an organizational pattern proper to the theme, spectators, and event; (Dougan, 1996) employing language suitable to the selected addressees; (Erwin, 1991) employing verbal diversity in rate, pitch, and intensity; (Erwin & Sebrell, 2003) articulating plainly, and using accurate grammar and pronunciation; representing nonverbal behavior which provides support for the verbal message (Morreale, Moore, Taylor, Surges-Tatum, & Hulbert-Johnson, 1993; Morreale et al., 1998).

2.27 Evaluation and Speaking Competencies

Advanced skills are considered as intermingle of knowledge, skill, and attitude; these entail larger altitudes of behavioral elasticity and adaptableness. A fundamental skill like “recognize communication goals” changes into “manage multiple communication goals” at an advanced level. This advanced level skill needs recognition of the goals and the behavioral factors of managing the goals both need adaptability (Morreale et al., 1998). Furthermore to speaking competencies at the school level for serving the general education courses, the NCA has also acknowledged competencies for basic and advanced levels, for example skills in speaking, persuading, informing, listening, and connecting interpersonally with different areas of communication.

Reasoning and audience analysis is another major requirement for advanced skills. For instance, skills comprise being capable of understanding people from other
different cultures, associations, or clusters; being capable to classify significant issues or problems; the capacity to illustrate conclusions; and being competent enough to appreciate others to handle conflict better. These skills engage adapting messages to the difficulty of the circumstances or framework. (Jones, 1994) These also call for a greater prominence on creating suitable and efficient messages, describes as two key components of competence (Spitzberg, 1983).

A basic communication course can be designed if expected student outcomes are conveyed to instructors and administrators well in-time so that they can use some or all of them for developing the desired course. Educational institutions can make use of these outcomes as recognized by NCA for explaining the campus anticipations for students in regard to the broad education curriculum (Morreale et al., 1998;). Mostly scholars have defined significant communication skills and also suggested measurements that can be used as helping aid to evaluate communication competence. Research on communication competency is tremendously diverse, and no one definition prevails (Jones, 1994). Consequently, in measuring student learning in a basic speaking course, by utilizing guidelines set forth by the NCA.

Rosenbaum (1994) has given description;

In assessing communication competencies among students, it is especially important to gauge student ability in general education courses where large numbers of students need to learn fundamental skills. The assessment of student learning has historically focused on standardized practices such as exams and the use of survey data. (p-73)

Rosenbaum (1994) estranged customary assessment tools such as tests and surveys from non-conventional tools such as portfolios, capstone courses, oral
assessments, and exit interviews. Several approaches are desirable in the measurement of learning (Erwin, 1991; Rosenbaum, 1994), several of which need the make use of a rubric. Non-conventional performance-based assessment is on preferred grounds these days and assessment techniques are getting fame at all educational levels such as the analysis of portfolios or performance events. Educational stream is adopting new assessment techniques and making use of them at much extended scale. Numerous disciplines are adopting new means for assessment which are authentic and discipline-specific (Borko, 1997; Dougan, 1996; Goldberg, Roswell, & Michaels, 1996; Hambleton & Plake, 1995). Prioritization about the effectual design, understanding, and proficient use of rubrics for these different types of assessment methods is dominant.

2.28 Criterion-Referenced Evaluation

Rubrics are also importance devices and tools in educational assessment and evaluation. Peculiarities can be used for clarifying the ways of using rubric effectively. (Mager, 1997). Criterion-referenced evaluation is used for measuring student performance and is rated according to preset standards by experts, educationists, the discipline or department whereas norm-referenced evaluation are used when students are evaluated on comparisons basis with respect to the performance of other students (Mager, 1997).

All components of the rubric should quantify precisely the objective of that specific component. Whenever communication targets are given as assignment, such as a speech or presentation, rubrics are mostly used in grading student oral presentations; therefore they are considered suitable for the assessment of basic oral communication
skill level. Furthermore, for the purpose of evaluation of fundamental oral communication skill level crossways a whole program, criterion-referenced evaluation is most suitable if it is based on preset standards by the discipline or department.

While using a rubric, assessors use a logical rating system in which each module is scored individually and performance is rated as a whole on by considering an overall impression (Pomplun, Capps, & Sundbye, 1998). Student presentations assessors normally use a logical rating system to facilitate students in understanding the communication competency area, which requires to work on for future public speaking opportunities.

Mager (1997) has explained the process;

For each competency on a rubric, evaluators must always understand the conditions and terms for each competency, and they must make inferences and approximations of those standards. To ensure evaluators infer similar ratings for each competency on a rubric, adequate training should be conducted. When training evaluators to use. (p-84)

Student evaluations show performance of faculty. These reflect the exact position of teaching learning process. These also serve as the annoyance for faculty member’s survival and a recognized source of genuine faculty concern (Ameen, Guffey, & Jackson, 2002; Gardner & Leak, 1994).

Instructors invest time, power and energy, and personal identity in teaching. They obviously anticipate students to be grateful for teaching them and recompense their “performance” with good quality marks. Whenever these anticipations fail to become visible, instructors become disappointed, perplexed, and irritated. Their
response leads student to markdown different sources of information and even throws in to disenchantment with the worth of work.

Several researchers have written about the different features that affect student evaluations by numerical ratings (Theall, Abrami, & Mets, 2001). Mostly written comments by student are hard to interpret (Lewis, 2001) and materialize distinctive and subjective, missing the heuristic power of numerical statistical facts.

2.29 Evaluation and Students’ Intellectual Development

Perry (1999) was the first researcher who worked out on students’ intellectual development; his research work was originally published in 1968. During his study he planned in-depth interviews with Harvard men and sorted out nine steps of intellectual development that can be submerged into four general areas. In the step of dualism students reveals a right/wrong approach to knowledge. Typically students analyze instructors as all-knowing authorities, at this step, and recognize their function as delivering this knowledge to students and students will repeat it back at suitable times. Multiplism is next stage in this process. By the time students get developed, they enter in this stage. At this stage they start recognizing some important questions which ambiguous answers. By end results they may start considering that since some knowledge is unsure, all views or opinions are evenly applicable. Instructors’ criticism of their work can confuse them and they can develop assumption that it is based on personal notion.
Effective problem solving is recognized as a complex cognitive activity by research, it also needs disciplinary knowledge, a range of strategies, recognition use of known knowledge, and productive beliefs about the process.

Schoenfeld (1985) has explained it in detail;

Good problem-solvers typically think through the processes they are using, feeling free to explore, adapt, and reject various paths. Novice problem solvers often have simplistic beliefs about the nature of problem solving, e.g., that there is an equation for every problem and that all problems can be solved quickly. (p-119)

Apprentice learners learn to solve the problems based on very specific learning domain. A paradigm can be considered that apprentice learners in applied sciences mostly categorize problems according to the kinds of objects involved, e.g., levers or pulleys, in spite of focusing underlying principles (Chi, Feltovich, & Glaser, 1981). Secondary School students normally rush to an answer, giving very little time for thinking about different choices of procedure (Schoenfeld, 1985). They focus the significance of problem solving on the answer rather than the procedure.

advised an approach which is capable of meeting high theoretical demands, and can be effortlessly implemented. It can also be practiced at some stage in regular classroom teaching as well as through homework assignments. Zimmerman et al. (1996) have given five modules such as time management, test preparation skills, comprehension and summarization skills, writing skills and note taking. The social-cognitive approach is behind this model. (Bandura 1977, 1986; Zimmerman 1989, 2000a). In this approach self-regulation is assumed as an interaction among personal, behavioral, and environmental processes. Daily achievement measurements and
systematic feedback serve at the core of these training modules. It also includes some actions which contribute to amplify self-regulatory skills. The training arbitrates practical, procedural and provisional knowledge (Renkl et al. 1996; Winne and Butler 1994).

Bandura and Schunk (1981) explained these modules;

Each of the five training modules lasts for five weeks. During these weeks, the students work through all four steps of the self-regulated learning cycle a total of four times. The students’ self-evaluation and monitoring of their own learning which is conducted by the pupils at the beginning of the training, as well as the ensuing systematic observation and discussion of their learning and performance behavior with the assistance of standardized forms, helps them to become actively aware of their strengths and weaknesses. (p-95)

Learning and performance behavior can be improved by appropriate goal setting and strategic planning and this all can occur if reasonable and workable information is available. Teachers facilitate learners for setting transitional goals, based on their requirement, difficulty level, and immediacy in time. A standardized handout for the second training week is used for documentation of goals.

The students are instructed on how to use strategic planning for finding out the appropriate method of attaining their goals. Learners accomplish decisional processes for the selection or modification of self-regulatory plans and techniques. Zimmerman (1989) recommends that the chosen methods should be appropriate for the task and setting. The instructors persuade this process by facilitating the students with suitable strategies for each new assignment, or discuss with them about suitability of the already decided strategies for the tasks at hand. Given that every approach is not equally
suitable for each learner, and the personal, behavioral and environmental factors continuously change, recurring adjustments are obligatory over the track of the training.

Strategy implementation and monitoring comes next in the cycle of self-regulated learning. Zimmerman (2000a) has described;

Performance or volitional control, which is ensured through, among others, self-control processes (e.g. self-instruction, imagery, attention focusing) and self-observation processes (e.g. self-recording, self-experimentation) (for details see plays an important role here. In the modules, self-control is primarily assured through attention focusing. This increases concentration among the pupils and eliminates potential disruptions, and should lead to improvements in strategy application and learning behavior. (p-96).

Implementation of strategies happens mainly at the same time as a student completes homework tasks, topics such as the appropriate place of work; institute and evasion of distraction by television etc. are discussed or pre-decided.

Self-reflection or strategic outcome monitoring comes at the end of the week. The students establish a association between their learning end product and the planned processes they use for the purpose by guidance of the teachers. A systematic evaluation is carried out for reaching the desired goal a using the self-compiled records of learning behavior and performance results produced each day. It facilitates the students to distinguish the efficiency of the planned processes they decide and incessantly adapted over the course of the preceding week.

Mastery and previous performance are used as evaluation criteria in the Zimmerman developed modules. Covington and Roberts (1994) showed that these
variables are predominantly beneficial because they permit persons to examine their personal learning development. It facilitates students in making adaptive or self-protective inferences, for coming to results about alteration of their self-regulatory approaches. Adaptive inferences facilitate students for adopting new and potentially better forms of performance self-regulation, such as shifting the goals hierarchically and adapting or choosing a more effective strategy (Zimmerman and Martinez-Pons 1992). Protective inferences, in dissimilarity, simply safeguard individuals from future disappointment and aversive influences and destabilize suitable adaptations. Following this rationale, by feedback and support in self-evaluation in the training modules, efforts are made to encourage first and foremost adaptive inferences, in this manner heartening advantageous forms of self-reactions.

2.30 Self-reactions and Adaptations in Learning

Adaptations in learning behavior comes through self-reactions and therefore influence learning in the subsequent week because the learners are now have reasonable capacity for using their self-reflections to acclimatize their goals and strategies. The key benefit of the Zimmerman et al. (1996) developed training modules is the sequence of self-regulated learning which repeated numerous times over the track of the 5-week training phase and the learners pass through their self-regulation stages to steady monitoring, development and strong performance.

Curran (2001) summarized the whole process;

To summarize, emphasis should be placed on the fact that the modules developed by Zimmerman et al. (1996) are based on a cyclic model of self-
On-line learning is very effective mechanism, now-a-days being used for imparting instructions and developing conceptual framework in many educational institutions. It also facilitates for the opportunity of fast and bendable communication among educators and their students. It assists learners in accessing information quite easily. Le Grange (2004) states However, the mere exchanging or retrieving of information does not in itself imply that learning takes place; rather there needs to be “a critical engagement with information” (p. 89).

A number of advantage are associated with e-learning, but at the same time, it looks significant to concentrate on the issues and standpoints of students for improving educational and pedagogical practices (Al-Mahmood and McLoughlin 2004; Cook-Sather 2002). Researchers are desired to reflect on the political, cultural, social and economic aspects of e-learning (Selwyn 2000), also the dissimilarities that can influence the comprehensive use of online learning (Huntington and Sudbery 2005).

Research studies has focused the issue and findings point out that students have very positive feeling about e-learning that their learning was very effective due to mutual and insightful interaction with other students (Littlefield and Roberson 2005; Quinney 2005; Roberts-DeGennaro et al. 2005). The distribution and mutually sharing of information by using online discussion forums has proved as a foundation of support for students concerned with practical social work placements (Quinney 2005; Roberts-DeGennaro et al. 2005).
Sweeney et al. (2004) has driven out from a research work that some students went through web-based learning which enabled them for a confident liberty of speech and offering a deeper learning approach. Second group learnt by face-to-face work as comparatively easy and computer-based work which required insightful thinking and a considerable time assurance. Online tutorials have some challenges associated with them regarding the soft flow of interface and communication, because students view most of the communication as lacking coherence.

Literature shows the impact of online learning on education and especially on science education. Statistics from the 2005 national assessment of educational progress (NAEP), point towards science education improvement efforts, which have been made by Benchmarks for Scientific the 1990s and American Association for the Advancement of Science in 1993 and the national science education standards (NSES) National Research Council (1996) have had very little impact on pressing forward student accomplishment in science.

Grigg (2006) has elaborated;

Specifically, a comparison of 1996 and 2005 NAEP data indicate the percentage of eighth grade students achieving at or above the Proficient level in science has remained relatively unchanged at a disappointing 32%. Likewise, the percentage of 12th grade students achieving at or above the same level in science declined from 24% in 1996 to 20% in 2005. (p-114)

2.31 Evaluation and Teachers Qualification

Teacher’s qualification is directly associated with students learning. Highly qualified teachers can play effective role in improving student achievement. Teachers’
subject matter knowledge has positive relationship with instructional practice. Particularly, subject matter knowledge manipulates teachers teaching, content organization, and teaching methodologies and the pedagogical strategies they use for teaching to students (Grossman et al. 1989; Grossman 1995; Leinhardt and Smith 1985; Thompson 1984; Wilson et al. 1987). Furthermore, teachers’ deepness of subject matter knowledge influences the intensity of questions teachers pretend to students (Carlsen 1991), their aptitude for developing additional explanations regarding students’ queries (Smith and Neale 1989), and the stage of activities teachers devise for supporting student’s learning (Smith and Neale 1989). Furthermore, the shortage of content knowledge can negatively influence what content is being delivered or taught. Carlsen (1991) and Smith and Neale (1991) established teachers tendency to teach familiar content and avoid teaching non-familiar content. Gess-Newsome and Lederman (1995) reported analogous findings. Sanders and Rivers (1996) reported a collective and negative end product on achievement levels, when students were positioned with numerous unproductive teachers in consecutive years.

School teachers, working at almost all levels, have the responsibility of helping students in developing an in-depth perception about foundational science concepts and associated concerns in learning science, and especially preparing students for highly developed scientific study, developing teachers’ science content knowledge is very indispensable in nature. Grossman (1995) emphasized the significance for teachers to practice lifelong learning goals for developing a cavernous knowledge pedestal
obligatory to efficiently teach science. It is considered very indispensable for school
teachers who are involved in delivering content from the key science disciplines for
example life science, earth and space science, physical science. This type of venture
needs more than an preliminary echelon of knowledge in every related discipline.

Collins (1999) stated;

However, attracting and retaining highly qualified mathematics and
science teachers in rural regions is a significant challenge (McClure and
Reeves 2004). Nearly half (49%) of all school districts in the United
States are rural and enroll one-fifth of the nation’s student population
(Provasnik et al. 2007). Being rural can be associated with many benefits,
such as smaller class sizes, fewer discipline problems, increased parental
involvement, and the personal attention inherent to small locales (Dunn

Continuous professional development programs for teacher community are
unavoidable initiative for improvement of quality of education and also educational
institutions. The professional development programs can be used effectively improving
in-service teachers’ appreciativeness of science conceptions they are anticipated to
teach (Darling-Hammond 1998; Loucks-Horsley and Matsumoto 1999; Garet et al.
2001). Additionally, programs that have an unambiguous concentration on content
knowledge and employ teachers by using scientific experiments are predominantly
effectual (Borko 2004; Smith and Neale 1991). Literature also points out professional
development programs as identifier of teachers’ previous conceptions, keep busy
teachers in interpretive discussions regarding the concepts being focused, and provide a
support to teachers’ meta-cognitive awareness for developing their understanding
(Beeth 1998; Driver et al. 1994; Hennessey 2003; Vosniadou 2007). Limited admittance
to content, resources and experts is a major problem in rural areas in implementation of teachers capacity building programs.

Mostly rural communities experience very little change in population, close-knit associations among community members are widespread, as teacher turnover is very low there. (Reeves 2003). On the other hand, inadequate industry, high number of joblessness man power, and negligible economic growth results in making few resources accessible to rural school districts for be a magnet for retaining and outstanding extremely qualified teachers.(Collins 1999; Inverness Research Associates 2001). Likewise, geographic location and especially isolation, also diminishes chances for professional growth and development.

Formative assessment advocates the need of an insight into successfully tasks completion mechanism as an important element. While summative assessment is conducted following a learning phase, covering a single course to an entire curriculum and serves accountability or certification purposes. Formative assessment has the goal of promoting learning and is conducted during a learning phase.

Birenbaum (1996) has given description;

Both formative and summative assessment can be directed at the learners’ level of acquired knowledge, however, the preference has started to shift in the past decades towards alternative assessment forms that allow measurement of acquired knowledge as well as cognitive and/or physical skills through performance on relevant whole tasks (p-98).

Formative assessment promotes learning. Learning is cultivated by formative assessment when it becomes successful in helping students recognize their weaker and
stronger points, and facilitating them for overcoming the weaker points during the learning process (Boud 1990; Dierick and Dochy 2001; Topping 2003). Learners need to develop a perception of the performance criterion and standards.

2.32 Instruction and Assessment

While discussing the formative assessment, instructional responsibilities and formative assessment tasks can be indistinguishable. Learners can work on an assignment, evaluate their performance on that chore, make a decision about performance improvement aspects, and choose a next target for work. It will also assist them to develop these performance aspects. They have the option to select someone else for their performance assessment and diagnosis.(Sluijsmans et al. 1998, 2006). It can be logically derived that for gaining success, instruction and assessment should be associated with each other. (; Biggs 1996; Dochy and McDowell 1997) Learners may have the capacity to practice the different tasks that may summatively assessed on, the reason for this initiative is that understanding performance criterion will not facilitate if the predefined target are not aligned with those during instruction.

Arter (1996) has explained;

In complex domains, however, the questions of how to perform a certain task or what constitutes “weak” and “strong” performance, are not answered easily, because these are also partly influenced by the changing situational demands. In other words, in complex domains, defining assessment criteria and standards is difficult, and so is learning to understand and apply them. (p-49)
2.33 Behavioral Model of Reinforcement

Learning can be defined as the process leading to relatively permanent behavioral change or potential behavioral change. Watson (1958) studied the process of learning and its affects on behavior very first time. Behaviorism is a school of thought that focuses only observable behaviors and considers them worthy for research since other constructs such as a person’s frame of mind or thoughts are too subjective. Skinner (1938) was of the opinion that internal states could persuade behavior just as external stimuli. Behavioral Psychology deals with behavior and its associated mechanisms. It is mainly interested in behavior modification from the stimuli both in the environment and within organism. Behavioral learning can be divided in classical and operant conditioning.

Classical conditioning discusses the learning process by linking it with stimulus-response stipulation. Unconditioned stimulus (UCS) may produce unconditioned response (UCR) and conditioned stimulus (CS) may produce the conditioned response (CR) respectively. Operant conditioning is another type of learning which deals with operating mechanism of individual. The term "Operant" takes in account how an organism operates within the environment, and therefore, operant conditioning studies the process of responding of an organism to what is presented in environment. Learning occurs due to the natural consequences of actions can be classified under this domain.

2.34 Reinforcement

The term reinforces means to strengthen. It refers to anything stimulus which strengthens or increases the probability of a specific response. Skinner (1938) defined
reinforcer as the treat, which increases the response. Reinforcement has strong association with learning and playing effective role in learning mechanism. Four types of reinforcement are important in teaching-learning process: positive, negative, punishment, and extinction.

Reinforcer means any stimulus which, when contingent on a response, serves to increase the rate of responding. The probability of occurrence of desired behavior will have direct relation with stimulus which follows it and increases its chances. The operational scope of reinforcement has direct relation with the situation upon whether an event is presented or removed after a response is made, and whether the subject's responding increases or decreases. Reinforcement can be defined as an event which increases responding and if it is presented event, it is called positive and any event that is removed is called negative.

2.34.1 Positive Reinforcement

The effort for increasing response is called positive reinforcement. It is adding something in environment for increasing a response. For instance, adding a treat will raise the response of sitting; adding eulogize will amplify the chances of child cleaning his or her room. The most common types of positive reinforcement are praise and rewards.
2.34.2 Negative Reinforcement

Negative reinforcement is referred as an effort for taking something away for increasing a response. A teenager can be quoted as an example, who is nagged by his mother to take out the garbage on weekly basis. He made complain to his friends about the nagging and one day he performs the task and to his amazement, the nagging stops. The removal of this negative stimulus is reinforcing and will probably amplify the chances that he will take out the garbage next week.

2.34.3 Punishment

Punishment is an attempt to decrease the probability of some behavior; it refers to adding something aversive in order to decrease a behavior. Disciplining a child for misbehaving can be quoted as an example for it. The child begins to associate being punished with the negative behavior. The punishment is not a pleasure act and avoiding it, he or she will stop behaving in that manner.

2.34.4 Extinction

Extinction can be defined as an act when something is removed in order to decrease a behavior. Positive reinforcement is proved as most powerful by research. Adding up a positive reinforcement to amplify a response not only works better, but facilitates both parties to spotlight on the positive aspects of the situation. Punishment can be effective if applied immediately following the negative behavior, but results in
extinction when it is not applied consistently. Punishment can also invoke other negative responses such as anger and resentment.

2.35 Schedules of Reinforcement

Schedules of reinforcement correlate with time or frequency of responses. These are the specific rules which are used for removing or presenting reinforcers or punishers following a specified operant behavior. These rules can be explained in terms of the time or the number of responses essential for presenting or to removing a reinforcer or a punisher. Diversified schedules of reinforcement generate distinctive effects on operant behavior.

Continuous schedule is a situation when any one of the four types of reinforcement is applied every time the behavior happens, getting a elevation after every flourishing project or getting smacked after every negative behavior. It’s continuous because the application occurs continuously after every project or behavior. It can be taken as a practical approach especially when using punishment. Inconsistencies during the process of punishment of children often results in perplexity and antipathy. Mainly two types of continuous schedules are taken in account:

2.35.1 Interval Schedule

Interval schedules are related to time span. These involve a lowest amount amount of time between consecutive reinforced responses. In interval schedules time is a key element which strengthens the responses. Responses which are made before this
time has elapsed are not reinforced. Interval schedules may identify a predetermined time period between reinforcers, fixed interval schedule, or a variable time period between reinforcers. Time can affect this reinforcement process in two forms; one is fixed and second is variable. Fixed Interval schedules generate an accelerated rate of response as the time of reinforcement approaches.

Variable Interval schedules produce a stable rate of response. In this process intervals have different time span. Presses of the "redial" button on the telephone are constant at a stable rate when some is trying to reach target and find a "busy" signal on the other end of the line.

When ever a desired behavior is reinforced after a phase of time, it is defined as an interval reinforcement schedule. This schedule is functional for behaviors which needs to be calculated in stipulations of their period, for example, in-seat behavior, on-task behavior, etc.

Normally there are two types of interval reinforcement. The first one is fixed interval reinforcement. It has a predetermined time span between two responses and it delegates a particular interval of time, after which reinforcement is conveyed contingent on appropriate behavior. Fixed interval schedule reinforcement is considered very important in educational stream especially in learning process. (Masuda, 1997). Examples of fixed interval reinforcement can the process of delivering reinforcement after every five minutes of on-task behavior. This is a very methodical and reliable schedule which is admirable for intensification a behavior. On the other hand, if reinforcement is just stopped, research indicates that the gains completed in behavior will also depreciate. In variable interval reinforcement, normally reinforcement is
delivered after an average length of time. Already established behaviors can be dealt effectively with this approach and also can be used when fading out a fixed interval schedule.

Every student has its own credentials in reinforcement process. It is very important that the reinforcement delivered is apposite for each individual student. The interest of students has its own value in this reinforcement process. If delivered substance to the student as reinforcement is not liked by student, none of these schedules will prove effective.

Reinforcement process is time consuming. When the desired behavior is gained as a result of reinforcement, reinforcing again and again becomes too time-consuming (Cizek, 1997). The ultimate stoppage of the reinforcement may affect the performance of the desired behavior to weaken. When desired behavior is achieved, gradual thinning of reinforcement is recommended. This gradual process will ensure perfection of behavior and also the stepping back of stimuli.

Basically schedules of reinforcement guide about presentation of reinforcer following a behavior. Time intervals and numbers of responses or ratio are important figures in reinforcement process. Amount of reinforcers and responses have main importance in whole process. One reinforcer may be presented if at least one correct response has come out during the time interval.

Socher (2002) has found during his research reinforcement can be helpful for managing classroom discipline, especially for managing behavior in children who have attention deficit disorder and other behavioral disorders. He gives detail;

Creating an orderly and stable classroom environment has helped provide the essential foundation for improving classroom behaviors, study habits,
and organizational skills. The key is to be consistent in applying the positive and negative consequences. When students are learning new behaviors such as positive social skills, a combination of the following strategies has demonstrated the most success. (p-93)

2.35.1.1 Fixed Interval Schedule Reinforcement

When a predetermined time period is involved in reinforcement and reinforcer is applied after a specific amount of time, it is called fixed interval schedule reinforcement. A main problem associated with this schedule is that target group or individual are inclined to develop their performance right earlier than the time period expires, just only to secure their performance. This type of reinforcement may also facilitate habitual groups because if they are well aware with schedule, they may win the game by very small labor or work. This problem can be solved by applying reinforcement on uneven schedules. When reinforcement is applied on an irregular basis, they are called variable schedules.

2.35.1.2 Variable Interval schedules

The fourth schedule of reinforcement deals with variable time intervals. Reinforcing somebody after an inconsistent amount of time is called variable interval schedule reinforcement. When teacher checks the note books of students without following any time frame, it is variable interval schedule. The variable schedules are more influential and result in more reliable behaviors. This principle may not work for punishment because constancy in the application is so significant, but for all further types of reinforcement they have a propensity to result in stronger responses.
2.35.2 Ratio Schedule

Ratio reinforcement schedule is a situation when an intended behavior is reinforced after a number of events or happenings. This schedule can be proving practical after establishing a contingency between the reinforcement and apposite behavior with the continuous reinforcement schedule.

Gullickson (1985) quotes;

It is also a good option when continuous reinforcement would be too cumbersome. There are two types of ratio reinforcement, each with its own benefits. Fixed ratio reinforcement is delivered after a given number of occurrences. Examples of fixed ratio reinforcement are reinforcing a child after every fifth math sheet is completed or after every third time a child exhibits sharing behavior.(p-28)

Ratio schedule entail a specific number of operant responses for generating the next reinforcer. These desired responses number may be preset from one reinforcer to the other, fixed ratio schedule or possibly will diverge from one reinforcer to the next Variable Ratio schedule. Usually fixed ratio schedules maintain a high velocity of response in anticipation of a reinforcer, when it is received, till the time a perceptible pause in responding can be observed, particularly with large ratios. Variable ratio schedules shore up a high and sturdy pace of response.

Fixed ratio reinforcement is a systematic process and, if used consistently, it is useful in establishing a contingency between behavior and reinforcement. Teachers prefer it for continuous reinforcement for establishing the presentation of a suitable behavior; student is already familiar that it is desirable, as it can be much more easily managed.
Fedrik (2001) has publicized that just the once a fixed ratio reinforcement schedule is completed, the gains in positive behavior will also depreciate if the behavior failed to continue reinforcing with some other schedule. Moreover, a student can study to maneuver this schedule if he or she point out the frequency of reinforcement received. Fedrik (2001) recommended that;

A fixed ratio schedule not be used for very long. A variable ratio reinforcement schedule involves delivering reinforcement after an approximate number of times the target behavior is exhibited. Reinforcement might be delivered on average after every fifth math sheet is completed but could range in delivery from every third to every eighth sheet. (p-45)

This scheduled is practical in classroom if teacher is intended to impart learning skill relevant to previous behavior or teachers desires to strengthen an established behavior. This schedule is functional for beginning continuation of a rational well-established behavior and can be proving effective when evaporation out a fixed ratio schedule. Mostly it is considered less systematic or consistent by comparing with continuous or fixed ratio reinforcement, it is not good for teaching a new behavior.

2.35.2.1 Fixed Ratio Schedule Reinforcement

In ratio schedule reinforcement, time factor is very important. If reinforcement is applied after a specific number of behaviors, it is called fixed ratio schedule. This reinforcement mechanism has a problem that the object will instigate to comprehend that he has one to two chances before taking any action. This complexity of this
schedule can be source of wastage of time and energies as well. Consequently, the behavior does not be liable to change in anticipation of right before the preset number.

2.35.2.2 Variable Ratio schedules.

Next schedule of reinforcement changes the time schedule. Here a reinforcer is applied after a variable number of responses. Variable ratio schedules have been established to function effectively under different situations. These schedules are applicable in educational environment and especially in teaching-learning process. Catania, (1992) defined reinforcement schedules as prescriptions for arranging reinforcers in time and in relation to responses, as the rules used to present reinforcing stimuli, or as specifications of the criteria by which responses become eligible to produce reinforcers (p. 394).

Many definitions are functional in indicating the form or structure of the schedule, but at the same time they fall short in other domains. These all schedules are interrelated with each other. Teaching-learning process involves active application of all schedules subject to situation and nature of desired learning or behavior. The dynamic nature reinforcement schedules influences the consequent arrangement between time, responses, and reinforcers which results in performance changes. The previous history or the operation of other related factors for example, in the case of humans, instructions that may have contradiction with other elements or complement the rules specified by the schedule.

Morse & Kelleher, (1977) give their analysis;
The ubiquity of schedules has made them a focal point of behavior analysis. They sometimes have been labeled the “amino acids of behavior” and often have been discussed as fundamental determinants of behavior. Reinforcement schedules are central in the experimental analysis of behavior because of what can be learned about the reinforcement process from them and because they serve as useful baselines for the study of other behavioral processes. (p-129)

Zeiler (1984) has the point of view that reinforcement schedules provide analogous functions in applied behavior analysis and are embedded in most treatment programs, either directly or indirectly.

2.35.2.3 Current Schedule Performance and Schedule History

Behavior analysis assumes that operant behavior is controlled by current reinforcement schedules. Recent research has established relationship between past and present experiences. Current explorations, by involving both animal and human subjects, have analytically investigated how earlier experiences also persuade current schedule-controlled behavior. New techniques and methodologies have been devised for it.

The research illustrates techniques for ascertaining functional relations between overtly arranged precedent experiences. It offers comparatively a broader framework for discussing reinforcement schedule performance and have imperative questions for applied behavior analysts with reference to mechanisms of studying historical variables are how they are conceptualized in relation to applied problems. It is same as situational analysis and then problem analysis, after both these findings, facts and figures, strategy can easily be developed for improvement. Behavior modification involves same steps, if teacher is well aware with the prior knowledge and skills of students, he can plan
effectively for learning enhancement of student.

Matthews(1982) explained the historical variables;

Historical variables have been examined in basic research by establishing different baseline histories of responding under separate schedules and evaluating subsequent performance under a common third schedule as a function of prior schedule experience. For example, Freeman and Lattal (1992) examined behavioral history effects in three experiments with pigeons. In the first experiment, subjects initially were exposed daily to a fixed-ratio (FR) schedule during one session and to a differential-reinforcement-of-low-rate (DRL) schedule during the other session, each under different stimulus conditions.(p-187)

Experiment was comprised of 60 days duration. During above cited experiment each schedule was presented for 50 or more sessions, and developed a historical framework of responding at a high rate that was about 4.5 times higher under the FR than under the DRL schedule. Afterward, indistinguishable fixed-interval (FI) schedules were executed under the relevant stimulus conditions throughout both of the sessions for 60 days. FR condition response rates remained higher than response rates in the DRL condition, and they were likely to congregate only after predetermined prolonged exposure to the FI schedule. The findings of a second experiment were same as first. Third experiment was conducted to very and strengthens the findings of previous experiments. In third experiment a multiple schedule was used to generate high and low response rates within individual baseline sessions.

The findings of these studies established that established behavior under stimulus control in the past, can affect the past schedule performance by current responding in the presence of the stimuli. It is also suggested that assured histories of reinforcement can be comparatively constant.
2.36 Research and Reinforcement

Research has established significance of reinforcement and its effective utilization measures. Human has different nature from animals in molding of responses and control by schedules of reinforcement. Freeman & Lattal (1992) has quoted an example,

“whereas the scalloped or ‘break-and-run’ pattern predicted by FI schedules readily occurs with rats and pigeons i.e., post-reinforcement pause followed by positively accelerated response rates, human performance, particularly human performance that is instructed, is resistant to control by the temporal variables that are implicit in those schedules (p-281).

Mooney (1989) has hypothesized that, in case of humans, variable-ratio (VR) schedules which produce high-rate responding could impede with succeeding FI performance. Conversely Baron and Leinenweber (1995) probed that variable-ratio history may not report itself for FI performance differences between humans and other animals. The sample of study was consisted of 18 rats and study evaluated the performance of rats under an FI 30-s schedule, almost half of experiment subjects had a conditioning history of high-rate responding under single or compound VR schedules. Many previous studies has showed results concerning high-rate responding. The high response rates primarily recognized in the rats showing to VR schedules gradually diminished with constant revelation to the FI schedule.

Variable ratio schedules serve as an important figure in reinforcement process. Even though history effects for rats with previous experience to VR schedules were manifest in high in general response rates under the FI schedule, findings were identical in case of experiment on within-interval response rates and these can be compared to
those of rats without a VR history. Despite the history, the response of rates cannot be assumed similar to human because characteristic of adult human performance are much different in many aspects. It is not possible to apply all rules derived by rates on humans.

Mauro (1994) reported that

Changes in the concurrent VI schedules failed to generate patterns of time allocation that matched the relative rates of reinforcement. Instead subjects generally tended to persist in their allocation patterns of the previous schedule condition. (p. 593).

History effects remained target of researcher for a long time. A prospect for the dearth of history effects in explored as the changes in independent variables are mostly associated with exceptional discriminative stimuli, which may moderate the effects of behavioral history. This finding points out the control of history effects and it is reinforced by the results of experimentation conducted by Freeman and Lattal (1992) which indicated towards stimulus control of history effects. Hanna, Blackman, and Todorov (1992) have invested time over it by conducting studies, they also supported previous results.

Baron and Leinenweber’s (1995) experiment point up that the level to which history effects endure with humans is not possible to be predicted with confidence, especially by the results obtained from involving animals. Perone, Galizio, and Baron (1988) disclosed a variety of dissimilarities between human and animal by different studies. This effort may serve as guideline to the differences in schedule performance as a function of schedule history.

Lovaas (1993) has explained in detail the differences in behavioral history
effects;

Specifically, in the case of differences in behavioral history effects, unlike the animal subjects’ relatively simple and brief e.g., 2 months or less, “simulated” histories, the behaviors of individuals treated by applied behavior analysts often have a history of many years’ duration. The duration of a history might affect its persistence and may, for example, partially account for Lovaas’ finding that the success of even prolonged intensive treatment for autism was related to the age of the child. (p-171)

2.36.1 Behavioral Histories

Behavioral history is regarded as important operational element in behavior modification.Ironically, applied behavior analysts consider its role as paramount and irrelevant whereas it is accepted as fact that history profoundly affects human behavior.

In this domain behavior modification becomes a complex issue. Iwata, Dorsey, Slifer, Bauman, & Richman (1982/1994) has scrutinized the phenomena and argued;

In fact, it could be argued that for applied behavior analysts, arranging conditions to alter subsequent behavior is itself a matter, and goal, of generating a different history that will produce durable changes in the targeted behavior. On the other hand, until the development of functional analysis methods, behavior analysts generally disregarded the historical conditions under which behavior developed. (p-193)

Behaviorist and psychoanalysis experts have different opinion about role of history. Baer (1977) narrated the aspect;

It is recognized that “any difference in behavior will always be subject to interpretation as a product of some currently unknowable fluctuation in those unknown variables” (p. 168).

History has diversified phases. Human nature is complex and its memory and its operational status is another debate for experts. One prominent rationale for paying less attention to behavioral history is a conviction that history is irrelevant; reason may be
the problem behavior, whenever comes across, may be exaggerated by different situations rather than those which add to its improvement. Iwata et al. (1982/1994) depicted his findings, “behavioral researchers and clinicians generally have dismissed the importance of etiology, since the conditions that are necessary to develop or maintain a response may be totally unrelated to the conditions that are sufficient to alter or eliminate it” (p.198).

2.36.2 Human Verbal Behavior and Schedule Performance

Human behavior is considered as a complex mechanism. Reinforcement can be defined as an added value but it cannot be declared as final source for behavior modification. Schedule performance is not indomitable entirely by contemporary requirements for reinforcement. There are chances that human schedule performance may prove affective by characteristically long-standing histories of verbal behavior. Hackenberg and Joker (1994) scrutinized alternatives available for adult humans. As soon the communication between instructions and contingencies started progressively, gradual shifts occurred in the schedule of reinforcement. The purpose behind their research was to observe effects of a history of erroneous instructions on compliance, instructional control under contradictory schedule requirements and the changeover from instructional to schedule control. These three areas were targeted in the study.

Joker’s (1994) has depicted the system of the study;

The procedure involved presenting two different stimuli on a computer screen, one of which was associated with a fixed-time (FT) schedule and the other with a progressive-time (PT) schedule. Neither schedule required a response; the reinforcers were simply delivered at the end of the scheduled time period. After each successive reinforcer, subjects selected one or the other schedule, which then remained in effect throughout that trial. (p-118)
During the study process, subjects were given the identical set of instructions all the way through the experiment. First experimental condition were framed by ensuring that the instructions precisely characterized the series of PT and FT choices for producing the desired reinforcement. Successive experimental conditions guided for the size of the increments in the PT schedule in a style that the similar instructions increasingly became less perfect in describing the most favorable choice sequence. This exceptional practice allowed inspection of the conversion from instructional to schedule control on a range of changing stimulus conditions, and examination of schedule control as a function of history.

First condition quickly ensured the establishment of instructional control, where instructions precisely narrated the schedules. Instructional control and schedule of responding both need accuracy in planning and careful monitoring throughout the process. Nevertheless, orthodoxy with instructions essentially guarded the range of behavior, by this means mutually precluding contact with changes in the schedule and reducing consequent control of responding.

Joker (1994) recited choices as;

Choices continued to be controlled by instructions as the PT step size increased across several experimental conditions. As the instructions became progressively more inaccurate, however, choice patterns became more variable and produced more reinforcement, resulting in an abrupt transition from instructional to schedule control.(p-96)

Different situations in natural environment have different operational demands. The schedules in Hackenberg and Joker’s (1994) experiment are comparable to numerous situations in the natural environment in which conditions steadily change in a way that instructions from others schedule become progressively less accurate.
Hackenberg and Joker’s outcomes show that the history and scale of communication between instructions and consequences of behavior possibly will influence the level to which choices are resolute by those changing conditions or persist to be controlled by instructions.

Hackenberg and Joker’s (1994) findings has effective application for applied behavior analysts which involves interventions for reducing rebelliousness, a major behavior problem of per-childhood (Egel, 1986).

Parrish et al. (1986) stated that reinforcing acquiescent behavior created security reductions in unsuitable behavior for which there were no scheduled consequences. In the same way, needed al. (1983) established that reinforcing observance to a subset of instructions amplified compliance with instructions of analogous types of requests that were not reinforced. The expression of widespread rejoinder classes in both of these studies essentially also reveals that the reinforcement schedule remained unable to control the unreinforced members. While these particular findings had understandable improvements, but there is also a chance that such histories might also have long-term disadvantages provided if they generate rigid rule which is not facilitated and regulated by environmental contingencies.

Wulfert, Greenway, Farkas, Hayes, and Dougher (1994) have traced out some rules during their study on the effects of instructional histories and measures of behavioral rigidity on sensitivity to schedule contingencies. First experiment showed the responding of subjects who were facilitated by precise and explicit instructions under a multiple DRL FR schedule of reinforcement. This group demonstrated more determination when the situation was switched over to extermination than did the
responding of subjects who were not facilitated by such specific instructions.

### 2.37 Experiments in Reinforcement

Hackenberg and Joker’s (1994) conducted an experiment to examine the effects of fixed ration schedule. In this experiment, subjects were given accurate and specific instructions with an FR schedule in effect. When a little change was introduced and schedule was changed to DRL, half of the target group was imparted accurately and the rest half was instructed inaccurately. The group who received accurate instructions responded accordingly and group fed with inaccurate instructions initially responded according to the instructions but the subjects who had formerly been characterized as rigid persisted during this prototype of behavior; mostly responses of the nonrigid subjects ultimately conformed to the schedule.

Rosenfarb, Newland, Brannon, & Howey, (1992) have discussed self generated rules;

Some recent research suggests that self generated rules may control nonverbal behavior similarly to instructions from others. For example, in some preliminary work on matching by Neef and colleagues, a student whose completion of different sets of math problems was reinforced on a concurrent VI 30-s VI 60-s schedule devoted her time exclusively to one set of problems, saying, “I should finish what I start before doing something else.”  (p-12)

Perceptibly in mostly circumstances self generated rules can control the behavior and it can prove adaptive in obtaining associated reinforcers with successful task completion. Following this rule, desired behavior may not be achieved into contact with the programmed contingencies, reason behind this fact is that mostly rule competed with the optimal response strategy. Exclusive responding to a single substitute which yielded
a comparatively low rate of reinforcement continues over numerous sessions in anticipation of adjunct procedures which were included as addition that nonverbally explained the contingencies.

In the same way, the degree to which defective or inaccurate rules have been generated by different subjects in research studies on matching by Neef, Mace, and Shade (1993), Neef, Mace, Shea, and Shade (1992), and Mace et al. (1994) could have added to the preliminary lack of control of their responding by concurrent VI schedules in anticipation of adjunct procedures were used for establishing matching.

Mace (1994) worked over it and stated;

These data suggest that adult humans’ own verbal behavior may influence their behavior in the presence of reinforcement schedules, although in other studies the relation between subjects descriptions and accompanying verbal responses has been less clear (Hackenberg & Axtell, 1993; Jacobs & Hackenberg, 1996). The relation between verbal description and schedule performance may be represented as a continuum. On the one hand, it is likely that schedule performance influences the verbal description. (p-314)

The relationship between schedule and performance control have gained intention of research during last few years. Research has supported the facts that schedules might control performance only by indirect way. It is also assumed that these influence individuals’ verbal behavior with the support of rules which govern behavior. The contemporary observations have interesting propositions for treatment. For instance, logically accessible reinforcement schedules are time and again indistinct, and in treatment patterns it is difficult to position them with the constancy and accuracy that are attributing of laboratory research.

Konarski (1987) explained the phenomena;

To the extent that individuals formulate and follow faulty rules, nonverbal
behavior that is controlled by those rules may be maladaptive or restricted such that it does not come into contact with schedules that might lead to other rules. In counseling situations, for example, the therapist can then shape more appropriate performance rules or prompt clients to follow testable rules that will contact contingencies that support alternative behaviors. (p-104)

Verbal behavior has its own style of corresponding and effects as well. In the early hours applied research focused the role of verbal behavior between verbal and nonverbal correspondence in direct phase (Israel, 1978) or indirect phase during self-instruction training (Meichenbaum & Goodman, 1971). Conversely, some studies in JABA point towards the role of verbal behavior during past few years. Behavior analysts have treated this area as verbal behavior and pretreatment history.

Correspondingly because of the reason that history or covert verbal behavior, no one can be observed directly, and such findings or report cannot be assumed as reliable or valid source of information, mostly these variables are assumed to be incidental to forecast and management of behavior.

A research article which was written by Baer, Wolf, and Risley (1968) explained this process as;

Although the focus of behavior analysis appropriately remains on what an individual does, the research by Horne and Lowe (1993) and others (Catania et al., 1982; Hayes, Brownstein, Zettle, Rosenfarb, & Korn, 1986) suggests that behavior analysts need to consider that “what subjects can be brought to do” may, in many situations, be a function of “what they can be brought to say.” It seems that the analysis of verbal behavior in relation to reinforcement schedules may warrant a more central role in applied behavioral research. (p-213)

2.38 Dynamic Reinforcement Schedules

Research has provided evidence about role of distal experience or relationship of
rules in reinforcement process. Concerns about the role of distal experiences have been showed during research studies, these experiences may occur due to previous reinforcement schedules or features of a history of rule following, on existing schedule performance. The performance is an area of interest of experts and the control of schedule performance by existing, or proximal, events is more characteristic of behavior and analytic research.

Repetition is considered as important functional element in reinforcement process. Reinforcement schedules normally engage cyclical, basically inert, arrangements whereby the identical obligations for reinforcement are in consequence on successive cycles. In fixed response schedule repetition is important element. For instance, following every reinforcer the identical fixed response condition is repeated.

An interaction can be manipulated between schedule performance at a desired step in a sitting and succeeding performance in same position or process. For illustration, during fixed ratio schedule, reinforcement rate can be indomitable by response rate with a condition that more quick responding acquiesces a higher rate of reinforcement than the lower rate responding. Reinforcement schedules are correlated with each other. Each schedule has close relation or implication with other schedule. Same findings have been derived by different research studies.(Baum,1973, 1989; Nevin, 1984).

Humans are different in nature and operations from animals. It may not prove workable to apply rules derives from animals on humans but to some extents it can work where domain may come closer to each other. The relations between humans’ rules and schedules also exemplify dynamic schedule effects. For illustration, if the requirements
of a schedule and a rule are in conflict, after certain time the subject’s behavior may be conventional to the schedule even though it was primarily controlled by the rule (Galizio, 1979). Self-generated rules have the possibility to might interact in a comparable way with schedule-controlled behavior for changing the behavior from time to time. Some time the person’s rules change during the sessions because of behavior changes and it results in further changes in the rules and this process goes on.

Self-motivated reinforcement schedules offer a system for unequivocally studying the different types of interactions. Following these dynamic schedules, the obligations for reinforcement transform after each reinforcer, or a few series of reinforcers. In theory, dynamic schedules attract because of the reason that they may disclose the procedure of adaptation of behavior in swiftly changing circumstances. Applied behavior analysts may have interest in these dynamic schedules because they deal or represent with primary researchers for addressing complex situations without scrutinizing them into more elemental schedules. (Lattal, 1991).

Dougherty, Cherek, and Roache (1994) has worked on PI schedules and stated;

PI schedule performance was investigated on human subjects. Subjects were seated at a console and earned points by pushing a button according to a chained PI t-s FI 20-s schedule. In the presence of the letter A on the computer screen, the first response after t s on a button changed the stimulus on the screen to B, indicating that an FI 20-s schedule was in effect. The first response after 20 s was reinforced in the presence of B. The FI schedule remained in effect until five reinforcers were earned. At that time, the letter a reappeared on the screen and the PI t-s schedule was incremented. (p-87)

Interval gap and requirements of intervals need intensive care during conducting
research or performing some activity related reinforcement. Human PI performance can be illustrated by withdrawing response rates and escalating post reinforcement hiatus durations, it may be considered as requirements of a function of progressively increasing interval. Post reinforcement gaps are inclined towards increasing arithmetically under the arithmetic progressions and geometrically under the geometric ones.

Harzem (1969) has given same results, with rats as subjects. During the experiment, Dougherty et al (1984) made use of PI schedules for assessing sequential control of behavior in humans. He used different doses of marijuana for observing its functions. The findings revealed a schedule with unidentified possessions on human behavior and in next phase they used the schedule for assessing another behavioral process.

Different schedules have different requirements. These requirements may be in shape of time intervals or nature of stimuli or nature and condition of response. In PR schedules, it is usual to maintain increase in the ratio condition for reinforcement till the time responding comes to an end for a pre-established time period, described as the break point (Thompson, 1972, and Keesey & Goldstein, 1968).

PI schedules proved helpful for studying break points. These studies address a question about how PR and PI schedules can be evaluated and also compared by considering engendering response persistence.
2.38.1 Progressive Reinforcement Schedules

Teaching-learning is a complicated process which involves many factors for its successful and effective completion. Difficulty level of learning material is one of target area for reinforcement. A progressive reinforcement schedule can easily be compared with applied scenery in which difficulty level of learning material amplifies autonomously of the subject’s behavior. This process has a condition that changes in material are carried out as instructor-paced criteria. Apart from of the individual student’s performance, aims, objectives or plans and the requirements for reinforcement boost scientifically, as prescribed in a specified set of rules.

Reinforcement proved helpful in launching many developments in numerous educational systems and it is very common for a student’s performance to depreciate as the difficulty level or work requirements boosts up with higher frequency. There are chances that the difficulty level may keep constant, at the same time the subject’s performance improves. The conventional negatively hastened learning curve articulates the latter relation.

Stewart (1975) has described Progressive ratio schedules as source of assessing response persistence. For instance, Hodos (1961) found a correlation between food deficiency and the break point on PR schedules. Mace et al. (1988, 1990) have given a helpful series of experiments with attentive theoretical analyses which may explicate the issues adjacent the confrontation of beleaguered behavior for change in applied settings. PR schedules Performance could be a practical accompaniment to different measures of conflict for making a change in applied settings.
2.38.2 Interactive Dynamic Schedules

Reinforcement schedules deals with organism’s present and past behaviors. Every organism goes under different experiences in life and every experience provides base for another experience. An important dynamic schedule arrangement is considered in which change occurs in the requirements for upcoming reinforcers as a response of the organism’s current or past behavior. This schedule deals with organism’s diversified experiences as pre-established criterion within personality. The last described schedule by Ferster and Skinner (1957) was like an adjusting schedule which explains the response prerequisite of a ratio schedule. If it is increased or decreased after each reinforcer, it will be direct proportional with the time interval that how long the organism paused before responding after each reinforcement. Some experts described these adjusting procedures as titration procedures (Lea, 1976; Weiss & Laties, 1959).

Hackenberg and Axtell (1993) studied the control of human behavior by long and short-term consequences by using an interactive dynamic reinforcement schedule. They state;

Human subjects were provided with choices between a PI schedule and an FI schedule that also, in some conditions, reset the PI schedule to 0 s. On each of a series of trials, subjects chose one of two schedules, each correlated with a distinct stimulus on a computer monitor. All operant responses were made on a computer keyboard space bar and, according to the schedules described below, yielded points that could be exchanged for money. Each trial started immediately after the preceding reinforcer with the simultaneous presentation of a red and a blue stimulus on the monitor.

Dynamic interactive schedule can be exemplified in many ways. A very simple example of a dynamic interactive schedule can be quoted as the shaping mechanism of
some desired response through the differential reinforcement of successive approximations. Shaping engrosses methodical and progressive changes in the reinforcement requirements since the subject meets consecutive behavioral objectives in relation to the desired behavior. Behavior and target response both are interconnected and the necessary conditions for reinforcement are accustomed as the behavior strongly approximates the desired response. Therefore, the subject’s behavior amends the arrangement of reinforcers.

The relationship between behavior and reinforcement requirements has gained meticulous concentration during last couple of decades. The topic of significant speculation with restricted experimentation was to find out the optimal relationship between behavior and changing reinforcement requirements for the period of shaping. A common understanding about the changing position and condition of contingencies is that it goes gradually but the behavior modification is considered as an art.

Eckerman, Hienz, Stern, and Kowlowitz (1980), has added an exemption to this issue by proposing the measures of quantification of the shaping process and recommended that outsized, swiftly changing requirements for reinforcement escort to the greatest shaping of behavior. Alleman & Platt, 1973; Kuch & Platt, (1976) have given same findings about the dynamic changes in shaping. Galbicka (1994) has given his opinion in recent times about the applied implications of Platt’s shaping procedures.

Long and short-term consequences have intensive relation with the control of human behavior and this can be considered as an example of an interactive dynamic schedule. In this process, individuals past experiences or learning can affect positively
or negatively to future reinforcers.

Hackenberg and Axtell’s (1993) has conducted an experiment. They give it detail;

The value of the PI schedule, and the subsequent overall rate of reinforcement, depended on the choices of either of two schedules over trials. Applied behavior analysis historically has emphasized the immediate consequences of behavior, but further development of the type of analysis and theory derived from Hackenberg and Axtell’s analysis of a dynamic interactive schedule illustrates how one might conceptualize and assess the effects of longer term consequences on both individual and group behavior. (p-218)

The use of different techniques for controlling self-injurious behavior in the case of individual behavior can be quoted as an example. This generated some controversial discussion among applied behavior analysts (Iwata, 1988). Some of analysts have argument about the short-term advantages of condensed maladaptive behavior are possible to be compensated by subject’s long-term negative effects of shock. Hackenberg and Axtell’s (1993) studied the dynamic reinforcement schedules and proposed a framework for adjustment of the problem of challenging results of individual behavior or individual action programs on a gamut with different problems that engage the short and long-term consequences analysis.

2.38.3 Continuous Reinforcement

Continuous reinforcement is a phenomenon when some desired behavior is reinforced each and every time and it is exhibited. CR is practical in teaching learning process especially while teaching new concepts. This schedule is workable especially when teaching the student a new behavior which is not part of the student's repertoire previously and this schedule particularly functional with younger students due to its
systematic nature. When it is preferred to establish an effective association between the 
desired behavior and the reinforcement, its application is recommended.

Continuous reinforcement is expected to establish an association between 
behavior and reinforcement but long-term changes in behavior may not develop with 
this effort or strategy. In case only continuous reinforcement is used for behavior 
modification, there are chances that the desired behavior will be cease after the 
withdrawal of reinforcement. Another important point in its application will be its too 
long use; a student may start manipulating the intervention by appropriate behavior only 
when reinforcement is desired. After establishment of an association schedule change is 
suggested by experts.

2.39 Principles of Reinforcement

Reinforcement theory is encircled by three basic principles. There are the rules 
of consequences and these rules illustrate the logical outcomes which normally 
expected to occur after consequences.

- Consequences which provide rewards, they amplify a behavior.
- Consequences which bestow punishment, they lessen a behavior.
- Consequences which support neither rewards nor punishments, they snuff 
out a behavior.

These are very basic and simple rules and offer an outstanding blueprint for 
influence. If increase in a behavior is desired, then a reward will be required or a 
consequence of reward. If some body looks for decrease in a behavior, then a 
consequence of punishment is required. And when a behavior is desired to extinguish, 
then ignorance will be better than providing any consequence.
2.40 The Process of Reinforcement

The process of reinforcement is completed in three steps. The rules of consequence are also used in a three step. Below are given steps;

Step 1: When some one is involved in some situation,

Step 2: When some one demonstrates some behavior,

Step 3: When some one gets in some consequence.

Reinforcement theory explains that people learn numerous things for the duration of reinforcement process. The process can be exemplified as humans when involve in some situation, they learn how to behave in that situation and it is ensured by certain consequences. This third phase is the most noticeable application of the rules of consequence. If this process in observed in teaching-learning process, when a student realizes to well on a task, then she/he will get a rewarding consequence. Another student practices speaking out inappropriately, then he/she will have chances to get the punishing consequence.

These rules are very practical in nature, especially when these are applied in teaching-learning process. A student whenever encounters a new situation, he tries to be familiar with it and during this familiarity process he learns more and more. There are only three consequences, rewarding, punishing, and ignoring for any situation and for any type of learning.

2.40.1 The Limitations of Reinforcement

Reinforcement Theory is considered as a practical and influence tool but at the same time it does have a number of severe limitations. For effective and efficient use of this theory, awareness with the difficulties in application is necessary.
It seems difficult to categorize, in advance, rewards and punishments for different situation because reinforcers are known as per their function and every function has its own pros and cons. Therefore, no hard and fast rule is available for rewards and punishments. Candy can boosts up student collaboration, but may not work as financial incentive for a factory worker. So researcher has to watch students very cautiously to find out the things they assume most rewarding or punishing. And when these things and their functions are identified effectively by student, he/she can be critically frustrated to realize that these rewards lose their worth over time. As the students develop as familiar to receiving some reward, they may feel bored with the passage of time. This point looks very critical and also is the utmost challenge for teachers. To trace out good rewards and punishments needs a great deal of knowledge, experience and insight.

All sources of reinforcement can be controlled and monitored. Teaching environment needs intensive care in monitoring these resources. Teachers need to compete with the student's peer group because these peers present an tremendously imperative source of reinforcement, occasionally it proves greater than any reward or punishment which a teacher can offer or can get hold. Reinforcement can be given by the child's parents and family and they are very effective source for it. Sometimes teachers assume that their reinforcement applications are fading, and selection of "right" reward or punishment may be the reason for it. The major cause for it can be the peer group who offers more than teacher. So this reason often remains untraced and teacher remains worried about his strategy.
Another main area for consideration is internal changes which can be difficult to create. Children learn to perform behaviors and it can be one side effect of reinforcement theory. The nature of child is different monkey, it becomes diminutive additional than a well-trained monkey who can perform a swindle, and after performance he holds out a hand waiting for the banana. The child needs the complete process for internalizing the behavior. This situation demands a full time devoted teacher who has nothing to do except this responsibility. It means that the teacher possibly will always be running just about providing the accurate consequences for the desired behaviors at the right time. This situation points out the training and capacity building of both student and teacher.

Punishment is complex phenomena to operate. It looks difficult to do well. Punishment is a tremendously influential consequence for all living creatures whether it is a pigeon, monkey, child, or adult. Punishing consequences can create awfully swift, tough, and unforgettable changes. Efficient punishment needs assured requirements. The research studies have given some measures for effective punishment:

1) Immediate,
2) Intense,
3) Unavoidable, and
4) Consistent

These are obligatory conditions for punishment. If these conditions are not followed then the punishment is expected to fail. Therefore, the best form of punishment would be something similar to this. When a child does the undesirable thing, then he is instantaneously given some concept of things which are considered
dangerous by him or he is placed on some uneasy place but to follow all conditions will not be an easy task.

Punishment can damage the conceptual framework of teacher in student mind. Students may start hating teachers who use punishment during teaching process. Punishment is painful and aversive consequence by definition. Students have very negative emotional feeling when they were punished by teacher. Consequently, when a teacher gives punishment, his students will possibly experience anger or fear or hopeless and they will then attach or relate these negative feelings with the cause of the punishment, and that is the teacher.

Reinforcing a pigeon is comparatively easy, but to reinforce a whole flock is not an easy task. Reinforcement theory has mostly focused animals for experimentation, particularly pigeons. This research effort with pigeons has capitulated exceptional results. The difference between work of these researchers and teacher’s work is the situation of work. The researchers work out on one pigeon at a time whereas teachers have to teach a whole group at one time. Here quantity of subjects matters. An experiment conducted on one subject cannot be generalized on whole group. The absolute size of a classroom brings a very hard aspect into the appropriate application of reinforcement theory.

Reinforcement learning is a type of learning which deals with what to do, how to plan situations for actions and how to make best use of a statistical reward signal. In this process learner is guided about proper stepping. The learner is not informed which steps to take like the machine learning but as a substitute, learner must determine which actions acquiesce the good number reward by making some attempts for them. Actions
and rewards have direct relationship. Mostly, in the appealing and exigent cases, actions may influence not only the instantaneous recompense but also the subsequently circumstances and all consequent rewards. The two main distinguishing features of reinforcement learning are, trial and error and search and delayed reward.

Fleming & Chambers (1983) has worked out on characterization of a learning problem. They describe;

Reinforcement learning is defined not by characterizing learning methods, but by characterizing a learning problem. Any method that is well suited to solving that problem, it is considered to be a reinforcement learning method. A full specification of the reinforcement learning problem in terms of optimal control, but the basic idea is simply to capture the most important aspects of the real problem facing a learning agent interacting with its environment to achieve a goal. (p-62)

Supervised learning has its own principles and rules. Reinforcement learning is dissimilar from supervised learning, the variety of learning reports in most existing literature and research are; in machine learning, statistical pattern recognition, and artificial neural networks. External supervisor has nothing to perform in reinforcement learning. Supervised learning is a type of learning which is supported by examples provided by a knowledgeable external supervisor. An effective interaction is required for this learning. This is a significant sort of learning, but unaccompanied it is not sufficient for learning from interaction. For solving the interactive problems, mostly it is unreasonable to attain examples of preferred behavior that are both accurate and agent of different situations in which the mediator has to take action. In unexplored region where one would anticipate learning to be most valuable, an negotiator must be competent enough to gain knowledge from its own experience.
2.41 Challenges in Reinforcement Learning

Exploration and exploitation are two main functions which are associated with challenges that influence the reinforcement learning. Reward is an incentive at the same time and for obtaining a lot of reward, a reinforcement learning manager gives priority to those actions which he has tried in the earlier period and found effective for getting reward. Discovery of such actions is not so easy; he has to endeavor for actions which it has not selected previously. The manager needs to exploit his learning and previous knowledge only for obtaining reward at the same time he also has to explore new actions for making a better action selections in the future.

Furham (1988) puts down the challenges;

The dilemma is that neither exploration nor exploitation can be pursued exclusively without failing at the task. The agent must try a variety of actions and progressively favor those that appear to be best. On a stochastic task, each action must be tried many times to gain a reliable estimate its expected reward. The exploration-exploitation dilemma has been intensively studied by mathematicians for many decades.(p-215)

An additional prominent feature of reinforcement learning is that it unambiguously regards as the entire problem of an objective intended for interacting with an undecided environment. It looks contradictory with several other approaches that reflect on sub-problems devoid of addressing their adjustment mechanism into a larger picture. For instance, a large amount of machine learning research is apprehensive with supervised learning lacking overtly identifying how such ability would finally be practical and beneficial. Theories of planning have been developed by many researchers aiming at common goals. This approach is missing the component of
consideration of role of planning in immediate decision making or the question of origination of predictive models which are essential for planning.

Reinforcement learning adopts the contradictory tack by operating with an inclusive, interactive and goal-oriented agent. All reinforcement learning managers have clear goals. They can observe different aspects of their learning environments and also can select actions for influencing their environments.

Furham (1988) has given detailed discussion;

Moreover, it is usually assumed from the beginning that the agent has to operate despite significant uncertainty about the environment it faces. When reinforcement learning involves planning, it has to address the interplay between planning and real-time action selection, as well as the question of how environmental models are acquired and improved. When reinforcement learning involves supervised learning, it does so for specific reasons that determine which capabilities are critical and which are not. For learning research to make progress, important sub-problems have to be isolated. (p-97)

Emerging trends of reinforcement learning include a contact between artificial intelligence and other engineering disciplines. Artificial learning has gained attention of researcher during last two decades. In past, artificial intelligence was considered as separate subject from control theory and statistics. It deals with logic and symbols. Griffin (1973) has worked on artificial intelligence. He explains;

Artificial intelligence was large LISP programs, not linear algebra, differential equations, or statistics. Over the last decades this view has gradually eroded. Modern artificial intelligence researchers accept statistical and control algorithms, for example, as relevant competing methods or simply as tools of their trade. The previously ignored areas lying between artificial intelligence and conventional engineering are now among the most active, including new fields such as neural networks, intelligent control, and our topic, reinforcement learning. (p-117)

In reinforcement learning ideas are derived from optimal control theory and stochastic approximation. These ideas address the aims and goals of artificial
intelligence. A policy defines the learning agent's way of behaving at a given time. Tentatively speaking, a policy is a mapping from perceived states of the environment to actions to be taken when in those states. It corresponds to what in psychology would be called a set of stimulus-response rules or associations. In some cases the policy may be a simple function or lookup table, whereas in others it may involve extensive computation such as a search process (Griffin, 1973). The policy is the core of a reinforcement learning agent in the sense that it alone is sufficient to determine behavior. In general, policies may be stochastic.

Reinforcement learning problems and rewards have very close relation with each other. A reward function defines the goal in a reinforcement learning problem. It maps every one supposed state of the environment to a single figure by indicating the intrinsic prestige of that state. The major objective of reinforcement learning is to maximize the total reward it takes as a delivery in the long run. The reward function can easily be defined by explaining the good and bad events for the agent. While discussing a biological system, it will work as an appropriate approach to spot out rewards with pleasure and pain. These are the instantaneous and defining features of the problem faced by a manager. Reward function may work as constant not as variable and the reward function have to essentially be irreversible by the manager. It may, however, dole out as a foundation for changing the policy. For instance, if an accomplishment chosen by the policy is followed by comparatively squat reward, after that the policy may be altered to pick some other act in that circumstances in the future. Usually, reward functions may be stochastic.
Reward function and value function both go together as a reward function designates the good which exists in an instantaneous sense, at the same time a value function specifies what is good in the long run. Gullickson (1985) as depicted the situations as;

To make a human analogy, rewards are like pleasure and pain, whereas values correspond to a more refined and farsighted judgment of how pleased or displeased one is that our environment is in a particular state. Expressed this way, one hope it is clear that value functions formalize a basic and familiar idea. (p-37)

Rewards are placed in primary place where as predictions of rewards stood at secondary. Values have no value or existence without rewards. The only rationale of estimating values is to accomplish additional reward. Decision making is key element in this reinforcement process and all decision depends on values. Nevertheless, values are mostly given importance when making and evaluating decisions. Value judgments provide a baseline for action choices. The actions which facilitate the states of highest value are not places at highest reward because these actions acquire the greatest amount of reward for targets over the long run. Value is the most important one which is mostly concerned in the process of decision-making and planning.

Environment provides baseline for rewards directly, but values are anticipated and re-estimated following the sequences of observations which a manager carries over its whole life span. The most significant constituent of approximately all reinforcement learning algorithms is a process for competently estimating values. The innermost function of value inference is arguably the most imperative thing have learned about reinforcement learning over the last few decades.
2.42 Reinforcement Learning Methods

Mostly all the reinforcement learning methods established during last three decades are prepared around estimating value functions. These methods include such as genetic programming, genetic algorithms, simulated annealing, and other purpose optimization methods are being used for solving the reinforcement learning problems (Gullickson, 1985). These methods directly concern with value functions and serve for policy matters as well. These are also tagged as evolutionary methods, the reason behind this is their operation which is analogous to the biological evolution, which produces organisms with accomplished behavior but they are unable to learn any thing behavior during their individual lifetimes.

Gullickson (1985) describes the policies;

If the space of policies is sufficiently small, or can be structured so that good policies are common or easy to find, then evolutionary methods can be effective. In addition, evolutionary methods have advantages on problems in which the learning agent cannot accurately sense the state of its environment. (p-191)

Reinforcement learning has direct interaction with the environment. Nevertheless, reinforcement learning involves different approaches for learning whilst interacting with the environment. It is common understanding that methods capable to take benefit of the mode of individual behavioral interactions can be more competent than evolutionary methods in many cases. Normally evolutionary methods pay no attention to the functional structure of the reinforcement learning problem. These methods normally not employ the reality that the policy is formulated to convert states
to actions. Usually they do not become aware of which phases an individual covers during its life span.

This information can be used more effectively for policy making and implementation although some time it can prove misleading. Both evolution and learning go long way together by sharing many features. Evolutionary methods can’t be considered as well suited to reinforcement learning problems due to many other reasons.

Infante (1995) has discussed model of the environment as the fourth element of reinforcement learning systems. It may act as something that mimics the behavior of the environment. For instance, provided that a state and action, the model might forecast the resulting next state and next reward. Planning is next very important stage and models are used for planning purpose. These models also take into account potential future situations before they are experienced in reality. The amalgamation of models and planning into reinforcement learning systems is a comparatively new development.

In Early days, reinforcement learning systems were overtly based on trial and error theories; mostly findings were opposite of the planning. Gradually reinforcement learning methods become clear in approach and in nature and proved closely related to dynamic programming methods. These methods use models and also are closely related to state-space planning methods. Modern reinforcement learning helps in deliberate planning and spans the spectrum from low level and moves with trial and error learning towards high level.

Educators the mostly use relevant homology in learning while following operant conditioning. The simple logic behind it lies that reinforcing a behavior augments the
probability of that behavior recurring. Skinnerian reinforcement is also known as operant conditioning is experienced crossways many species but extensively studied in rats, pigeons, and humans. In educational practices these operant techniques are operational including token economies, behavior modification, contingency contracting and various forms of programmed instruction.

Herbert Walberg (1984) has tried to find out effectiveness of are these operant techniques as compared to other teaching methods in teaching learning process. He made an attempt to respond this question through an inclusive review of the research literature. He discussed in detail whole theory and concluded. His conclusion was unambiguous: “Skinnerian reinforcement or reward for correct performance has the largest overall average effect 1.17 deviations” (p. 23).

Teaching needs competency and to provide each and every person a well skilled teacher, is not a simple task. It looks easy to realize the reason of shortage of individual tutoring for teaching purpose because society simply cannot manage to provide every student with a competent individual teacher. In 1989, Skinner felt that his thoughts on teaching were being disregarded. Same was reported by Axelrod, Moyer, and Berry (1990).

Application of reinforcement in education faced two main arguments. First the instructional techniques were drawn from animal models are intrinsically inappropriate and secondly the principle about the use of rewards and incentives which in counterproductive and harmful to students.

Stephen Ray Flora’s (2004) discussed the use of reinforcement especially in educational environment. He explained different issues, the relevance of rat psychology
to human learning was first issue, Flora gives persuasive examples of the educational application of approaches first developed with animal models. He further discussed;

Pediatricians and child psychologists now routinely recommend the use of timeout instead of spanking to modify children’s behavior. Flora notes that the term timeout is derived from the phrase “timeout from positive reinforcement” and that the technique was first developed in experiments with rats and pigeons. (p. 9)

It is general perception that animal psychology is bit different from human psychology and many studies of animal behavior have nonentity to add to for perception of human psychology. On the other hand instructional and therapeutic intercessions based on operant conditioning have faced huge amount of criticism as tyrannical and appalling.

Packard (1977) has given its characterization;

In education this characterization of operant approaches became conventional wisdom with the publication of Alfie Kohn’s Punished by rewards, The trouble with gold stars, incentive plans, A's, praise, and other bribes in 1993. Kohn’s position is extreme, arguing against the use of all forms of rewards, including verbal praise. (p-75)

Effects of reinforcement are central point for research from last many years. The question regarding these effects cannot be established by an application of common sense. It needs evidence of some examination as necessary action. Desirable behavior is target of all educationist and psychologists; they are struggle for its effective achievement. Reward may not provide surety for achieving some desirable behavior. Kohn criticizes the programs that pay reward to students for getting some desirable behaviors. He also wants to disclose the negative consequences, associated with it.
2.42.1 Conditioned to Specific Learned Response

Conditioning is another contributing factor in reinforcement learning. Often animals learn behaviors that are totally different from those which were conditioned. Furthermore, it is observable that these specific behaviors, which the animals adapt quickly, are very clear examples of instinctive behaviors. It seems noticeable that these animals are fascinated by strapping instinctive behaviors. These behaviors are clear demonstration of the prepotency of such conditioned behavior patterns.

Similar prepotencies in humans are a question which needs proper response. The situation of phobias can be exemplified for this question. The humans can be conditioned to fear through an extensive assortment of objects and situations. It looks practical to disposed to certain types of phobias, such as fear of snakes.

Mineka (2002) worked over this question and have proposed the continuation of a fear module, entrenched in particular brain structures and formed by evolution that disposed to acquire phobias to stimuli that endangered mammalian ancestors. Such proposed model may prove effective for classroom learning but not applicable at levels.

Behavior modification is central point of this theory. Reinforcement theory is an effort of determining behavior by controlling the consequences of the behavior. In reinforcement theory, normally an amalgamation of rewards or punishments is used for reinforcing some desired behavior or extinguishes unnecessary behavior. Operant behavior is a behavior which elicits a consequence because the human being operates on own environment. Reinforcement theory tries to focus on finding out the relationship between the operant behavior and the associated consequences. It is also called operant conditioning. Mineka (2002) portrayed the situation;
Behavioral theories of learning and motivation focus on the effect that the consequences of past behavior have on future behavior. This is in contrast to classical conditioning, which focuses on responses that are triggered by stimuli in an almost automatic fashion. Reinforcement theory suggests that individuals can choose from several responses to a given stimulus, and that individuals will generally select the response that has been associated with positive outcomes in the past. E.L. Thorndike articulated this idea in 1911, in what has come to be known as the law of effect. (p-211)

The law of effect has its effective role in reinforcement process. Principally this law deals with stimuli and its response. It states that if all other things are being equal, responses to stimuli will be strengthened if followed by satisfaction, but responses that are followed by discomfort will be weakened.

B.F. Skinner was a psychologist who made a great contribution for reinforcement theory and in the development of modern ideas concerning reinforcement theory. Skinner argued that;

The internal needs and drives of individuals can be ignored because people learn to exhibit certain behaviors based on what happens to them as a result of their behavior. This school of thought has been termed the behaviorist, or radical behaviorist, school. (p-39)

Reinforcement is the most vital principle of reinforcement theory. Normally, two types of reinforcement are more important, positive reinforcement and negative reinforcement. Some behavior can be positively reinforced if the incidence of an appreciated behavioral consequence has the outcome of strengthening the probability of the behavior being repeated. The reinforcer is referred as the specific behavioral consequence. A positive reinforcement can be exemplified by a salesperson that tries additional effort for selling out the quota (behavior) and is then rewarded with a bonus.
(positive reinforcer). The continuation of the positive reinforcer will ensure that the salesperson will continue to exert the necessary effort in the future.

Positive reinforcement is also applicable in classroom environment. When a student receives appreciation for his work, his interest increases and he puts more efforts to get more scores during assessment. Negative reinforcement results when an unwanted behavioral consequence is suspended, with the outcome of intensification the prospect of the behavior being repeated.

Negative reinforcement is often mixed up with punishment whereas both have different approaches and operational mechanisms. Punishment endeavors for diminishing the probability of a particular behaviors; negative reinforcement tries to augment preferred behavior. As a consequence, both positive and negative reinforcement try for increasing the chance that a particular behavior will be learned and repeated.

Extinction also makes efforts for reducing unwanted behavior. The extinction procedure instigates when an appreciated behavioral consequence is suspended in order to lessen the prospect that a educated behavior will prolong. With the passage of time, it will result in the ceasing of that behavior. Extinction may work as alternate source for reducing a required behavior. For exemplar, if an employee is constantly admired for the swiftness in which he does his work for more than a few months, but remains unable to receive any eulogize in succeeding months for such behavior, his gorgeous behaviors may weaken. Therefore, for avoiding unnecessary extinction, positive behavioral consequences are needed to be offered by managers.
Reinforcement schedule are the timing of the behavioral consequences that pursue a given behavior. Basically, two types of reinforcement schedules are used in educational phenomena, these are continuous and intermittent. Continuous reinforcement is situation when a behavior is reinforced each time it occurs. Behavior modification is common point of interest for educationist and psychologists. Research proposed that continuous reinforcement is the best way to set up new behaviors or to get rid of undesired behaviors. Such practice is not recommended for an organizational setting. Consequently, intermittent schedules are generally practices. Intermittent reinforcement refers to a practice when each instance of a desired behavior is not reinforced. Four types of intermittent reinforcement schedules are discussed in literature: fixed interval, fixed ratio, variable interval, and variable ratio.

Fixed interval schedules of reinforcement deals with set periods of time and it occurs when desired behaviors are reinforced after the predetermined time intervals. A fixed interval schedule of reinforcement may not work as strong device for establishing a desired behavior, and behavior educated in such approach needs a rapid extinction. The fixed ratio schedule of reinforcement applies the reinforcer after a set number of occurrences of the desired behaviors. One organizational example of this schedule is a sales commission based on number of units sold. Like the fixed interval schedule, the fixed ratio schedule may not produce consistent, long-lasting, behavioral change.

Variable interval reinforcement schedules are employed when desired behaviors are reinforced after varying periods of time. Examples of variable interval schedules would be special recognition for successful performance and promotions to higher-level positions. This reinforcement schedule appears to elicit desired behavioral change that
is resistant to extinction. Fixed interval schedule of reinforcement is selected for this study. This study will also try to find out impact of formative evaluation and fixed interval schedule reinforcement on academic achievement of secondary school students.
CHAPTER 3

METHOD AND PROCEDURE

3.1 Research design

The research design for exploring out the impact of formative evaluation and fixed interval schedule reinforcement on academic achievements of secondary school students was a one treatment (teacher made test- formative evaluation) × 2 phases (formative evaluation and fixed interval schedule reinforcement) design. This design permitted analysis and comparison of student performance on pretests and posttests both prior to formative evaluation and after formative evaluation and fixed interval schedule reinforcement. Additional quantitative data were collected using teacher made tests during the treatment.

Phase I

Pre-test ➔ Control Group ➔ No evaluation ➔ Post-test

Pre-test ➔ Experimental Group ➔ Evaluation by Teacher made tests with unequal intervals ➔ Post-test
Experimental research designs are usually used for the controlled testing of causal processes. One or more independent variables are manipulated for determining their effect on a dependent variable. These independent and dependent variables operate on their extents. These designs are practical in following situations;

1. If There is time priority in a causal relationship (cause precedes effect),
2. If There is consistency in a causal relationship, and
3. The magnitude of the correlation is great.

One of the most significant necessities of experimental research designs is the requirement of eliminating the effects of imitation, overruling, and precursor variables. A true experimental design entails an simulated environment so as to control for all spurious, intervening, and antecedent variables.

### 3.2 Population

The target population in this study were secondary school students of Attock district, who were admitted to the secondary classes and taking classes during 2003-05. For the present experimental study one government high school was chosen from 160 government high schools of the district Attock, as it was convenient.
3.3 Sample

A group of 60 students out of 200 students studying in 9th class in government high school Fateh Jang was selected randomly for the first phase and a group of 60 students out of 220 students studying in 10th class in government high school Fateh Jang was selected randomly for the second phase. The selected sample for each phase was divided into two equal groups randomly, of which, one served as experimental group and other as control group.

Criteria for selecting the participants included: (1) being secondary class student; (2) time period of 2003-05; (3) must have passed subject of English and Maths in last exam; A total of 200 students met the criteria for first phase and 220 students met the criteria for second phase.

Table 3.1 Strength of selected sample

<table>
<thead>
<tr>
<th>1 Group</th>
<th>Phase I</th>
<th>Phase II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Experimental</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Total</td>
<td>60</td>
<td>60</td>
</tr>
</tbody>
</table>
#### 3.4 Instrument development

Separate achievement tests were developed for pre-test and post-test by consultation of experts. Pre-test of Mathematics for class 9th consisted of 50 marks with the duration of 90 mints.

**Pre-test Math**

**Per-test Math**

<table>
<thead>
<tr>
<th>Class 9th</th>
<th>Time: 90 mint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marks: 50</td>
<td></td>
</tr>
</tbody>
</table>

If \( A = \{1, 2, 3, 4, 5\} \) \( B = \{3, 5, 7\} \) and \( C = \{2, 4, 6, 8, 10\} \) then find out \( (B \cup A) \) and \( (A \cup C) \)

1. If \( N = \{11, 12, 13, 14, 15\} \) \( M = \{13, 15, 17, 19\} \) and \( L = \{2, 4, 5, 6, 7, 8, 10\} \) then find out \( (N \cap M) \) and \( (N \cap L) \)

2. If \( (a + b) = 5 \) and \( ab = 12 \) then find out value of \( a^2 + b^2 \)

3. Solve \( \sqrt{64} a^2 b^2 \)

4. Find fraction \( x^2 - y^2 \)

During the experiment teacher made test were used for formative evaluation of experimental group. Each test consisted of 30 marks with the duration of 45 mints.
1. Choose the correct answer below each part:

i- \{1,3,5,\ldots\} is the set of:
   a) Even no  b) odd no  c) prime no  d) integers

ii- The union of two sets is denoted by:
   a) A \cap B  b) A \cup B  c) A-B  d) B-A

iii- 2/3 is:
   a) a rational no  b) a whole no  c) an odd integer  d) a natural no

iv- The set \{2\} is:
   a) Infinite  b) singleton  c) power set  d) subset

v- If A has n elements then number of elements in P(A) is:
   a) n^2  b) 2^n  c) 3^n  d) 2n

Reinforcement was provided with the strategy mentioned by hilgurd in his book
1. Write a story on any of given topics;
   Honesty is the best policy, Truth is ever green, Haste makes waste

2. Write essay on any one topic
   A rainy day, My favorite hero, A football match

3. Use given words in your own sentences;
   Fate, currency, opportunity, pleasure, hurdle

4. Give plural:
   Factory, storey, machine, class, thief

5. Change the narration from active voice to passive voice;
   (a) Most of his colleagues appreciated his work.
   (b) Will the old man forgive all his sons?
   (c) Why are you submitting a new plan?
   (d) He and his friends always start the quarrel.
   (e) Block all the roads.

6. Convert the following into Future Perfect Tense:
   a) The man was searching for his lost relatives.
   b) He had left his family during the war.
   c) He did not get any news about them after that.
   d) Had the enemy imprisoned them?
   e) Will they be going through torture?
Test 1 English

Marks: 30
Time: 45 mint

1-Answer the questions;
1. Who is a reformer?
2. What was the result of Hazrat Muhammad’s (SAW) teaching on the people of Arabia?
3. What condition was imposed on the prisoners of the battle of Badr?
4. What was the condition of women before the advent of Islam?
5. What was the system introduced by Hazrat Muhammad’s (SAW) to check justice?

2- Fill in the blanks.
1. Hazrat Muhammad’s (SAW) was the greatest ___________ of the world.
2. He (SAW) reformed the people of Arabia.
3. He (SAW) advised the Muslims to get themselves ___________.
4. Hazrat Muhammad’s (SAW) disliked ________ and _________.
5. Hazrat Zainab (R. A) was married to a slave named ___________.

3- Write “T” for true statement and “F” for false statement.
1. Hazrat Muhammad (SAW) started preaching God’s message to the people of China.
2. Hazrat Abdullah (R. A) was allowed to show off his wealth.
3. Hazrat Muhammad (SAW) gave women the right of inheritance in property and wealth.
4. Before the advent of Islam, women were treated inhumanely in Arabia
5. The reforms of Hazrat Muhammad (SAW) laid the foundation of a new Islamic society.

3.5 Instrument validation

The achievement tests were administered to 20 students for each phase, members of same population but excluded from the sample. After data collection, by item analysis, its validity was found.

Technical or instrument validity was found by reviewing literature (Eisenhart and Howe 1992, pp. 657–662; Hitchcock and Hughes 1995, pp. 105–106) and by following the guidelines about evaluation instruments (LeCompte et al. 1993, pp. 322–349; Heron 1996, pp. 159–170; Denscombe 1998, pp. 213–214; Reeves and Hedberg 2003, p. 34), which occasionally needs resorting to simpler English language or to translating statements into vernacular and negotiations. The possibility of different meanings and related difficulties in analysing them also exists there.

During item development, experts were especially concerned with content and face validity (DeVellis 2003). For higher content validity (DeVellis 2003) scale items were developed with input from experts in teaching English and Maths and its evaluation. Although face validity is not necessarily required for a valid scale, experts were interested in student evaluation and wanted this interest to be transparent to the students in the study.

Item reliability for the scale was assessed for all 4 tests, 2 pretest and 2 posttest and was good to excellent, Cronbach’s $\alpha = .70$ and above.
3.6 Experimentation

3.6.1 Procedures

The instructors were trained in the procedures for the study by the researcher. During the experiment the instructors were provided with a detailed syllabus for the lessons and an instructor guide, which they were directed to follow closely. The researcher then randomly assigned the six sections to treatment conditions, teacher made test-formative evaluation, and students were evaluated formatively. All groups completed the same instruction for the 60 days. During this time, students were taught subjects of English and Maths by subject specialists. After completion of 60 days a posttest was conducted. The experiment was completed by administering the achievement tests for pre-test and post-test to control and experimental group separately. During the experiment, teacher made test were used to measure the learning outcome as formative evaluation.

In first phase 60 students were randomly selected from 9th class of government high school Fateh Jang. This sample was divided into two equal groups. One served as control group and other as experimental group.

A pretest was conducted to both the groups in the subjects of Math and English. Groups were taught by subject specialist of Math and English. Control group was only taught and was not given any treatment whereas experimental group was taught and evaluated by teacher made test during session. After administering the teacher made tests, relevant teachers marked the tests and given back to students to observe what were their mistakes and how to remove those. After completion of 60 days teaching, a post test was conducted in Math and English.
In second phase 60 students were randomly selected from 10th class of government high school Fateh Jang. This sample was divided into two equal groups. One served as control group and other as experimental group.

A pretest was conducted to both the groups in the subjects of Math and English. Both groups were taught by subject specialist of Math and English. Control group was only taught and was not given any treatment whereas experimental group was taught and evaluated by teacher made test and reinforcement was given with a fixed interval schedule, after every 5th day, during session. After administering the achievement test, relevant teacher marked the test and gave remarks on test page as, Excellent, Very Good, Good, Satisfactory, Needs Improvement, Struggle Hard and also mentioned their weak areas. Students were asked to show their test page to parents and have parent’s signature over it. This process was repeated after every 5 days. After completion of 60 days teaching, a post-test was conducted in Math and English.

3.7 Data Analysis

Data was analyzed by using SPSS, statistical package for social sciences. There were 8 null hypotheses. Each hypothesis was analyzed at 0.05α or 0.01α Pearson correlation was applied on data and with the help of table value, correlation between variables was found. For finding out the impact of formative evaluation and fixed interval schedule reinforcement on academic achievements, t test was applied. By using this strategy data analysis was completed.
CHAPTER 4

ANALYSIS OF DATA

The study designed to explore out the impact of formative evaluation and fixed interval schedule reinforcement on academic achievements of secondary school students. Population of the study was comprised of all secondary schools of district Attock. Two groups of students having strength of 60 each were selected from Govt. High School Fateh Jang as a sample of the study. First group was selected from 9th class and second group was selected from 10th class randomly. The selected sample of each group was divided into two equal groups, of which, one served as experimental group and other served as control group.

The study comprised 8 objectives, 4 null hypotheses and 4 sub-hypotheses. Experimental design study was conducted to find out the impact of formative evaluation and fixed interval schedule reinforcement on academic achievement of secondary school students. Study was completed in two phases. Each phase consisted of two groups, control group and experimental group. A pretest was conducted for both the groups in the subjects of Math and English to assess their knowledge baseline.
Phase I

Pre-test Math

Per-test Math

Marks: 50

class 9th

Time: 90 mint

1. If \( A = \{1, 2, 3, 4, 5\} \) \( B = \{3, 5, 7\} \) and \( C = \{2, 4, 6, 8, 10\} \) then find out
   \( (B \cup A) \) and \( (A \cup C) \)

2. If \( N = \{11, 12, 13, 14, 15\} \) \( M = \{13, 15, 17, 19\} \) and \( L = \{2, 4, 5, 6, 7, 8, 10\} \)
   then find out \( (N \cap M) \) and \( (N \cap L) \)

3. If \( a + b = 5 \) and \( ab = 12 \) then find out value of \( a^2 + b^2 \)

4. Solve \( \sqrt{64} a^2 b^2 \)

5. Find fraction \( x^2 - y^2 \)

After conducting the pre-test, the control group was only taught and was not given any treatment whereas experimental group was taught and evaluated by 10 teacher made tests during session. After administering the teacher made tests, relevant teachers marked the tests which were given back to students to observe what their mistakes and how they could remove those
**Post-test Mathematics**

Post-test Math  
Marks: 50  
class 9th  
Time: 90 mint

1. If \( N = \{8, 10, 12, 14, 16\} \) and \( M = \{7, 8, 9, 10, 11, 13, 15\} \)
then find out \( N - M \) and \( M - N \)

2. If \( A = \{0, 1\} \) \( B = \{0, 1, 2, 3\} \) \( C = \{1, 3, 5, 7, 9\} \)
then prove that \( A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \)

3. If \( X = \{a, b, c\} \) and \( Y = \{d\} \) then find out \( X \times Y \) and \( Y \times Y \)

4. If \( X = \sqrt{7} + 3 \) then find value of \( X + \frac{1}{X} \) , \( X^2 + \frac{1}{X^2} \)

5. Find the values with the help of logarithm \( 47.5192 \times 5.1678 \)

**Pre-test English**

Pre-test English  
Marks: 50  
class 9th  
Time: 90 mint

1. Write a story on any of given topics;
   
   Honesty is the best policy, Truth is ever green, Haste makes waste

2. Write essay on any one topic
   
   A rainy day, My favorite hero, A football match

3. Use given words in your own sentences;
   
   Fate, currency, opportunity, pleasure, hurdle

4. Give plural:
   
   Factory, storey, machine, class, thief
5. Change the narration from active voice to passive voice;

( a ) Most of his colleagues appreciated his work.

( b ) Will the old man forgive all his sons?

( c ) Why are you submitting a new plan?

( d ) He and his friends always start the quarrel.

( e ) Block all the roads.

6. Convert the following into Future Perfect Tense:

a) The man was searching for his lost relatives.

b) He had left his family during the war.

c) He did not get any news about them after that.

d) Had the enemy imprisoned them?

e) Will they be going through torture?

Post-test English

Post-test English Marks: 50

class 9th Time: 90 mint

1- Recall the lesson “Hazrat Muhammad (SAW), The Reformer” and answer the questions;

1. What happened to idol worshippers when the Holy Prophet (SAW) preached Allah’s message?

2. What happened to people of Arabia after embracing Islam?

3. What kinds of learning were promoted?

4. In which fields did the Muslims make great progress?

5. How did the Arabs settle their disputes in the earlier days?

2. Change these into male gender;
lady, hen, hostess, lioness, tigress

3. Write the plural of these;
Mouse, foot, city, chief, calf

4. Use these words in sentences of your own.
Little, mighty, moment, humble, errors,

5. Write ten sentences about the rural life.

6. Match with the opposites from columns given below

<table>
<thead>
<tr>
<th>Column A</th>
<th>Column B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rich</td>
<td>Lose</td>
</tr>
<tr>
<td>Left</td>
<td>Then</td>
</tr>
<tr>
<td>Dark</td>
<td>Dry</td>
</tr>
<tr>
<td>Wet</td>
<td>Light</td>
</tr>
<tr>
<td>Now</td>
<td>Right</td>
</tr>
<tr>
<td>Find</td>
<td>Poor</td>
</tr>
</tbody>
</table>

2. Translate into Urdu;

There are a large number of government and private hospitals and clinics. They have the latest equipments, and highly qualified staff of doctors, nurses and their assistants.

A large number of primary and secondary schools, colleges and universities have been opened in the cities.
Phase II

Pre-test Math

Pre-test Math
Marks: 50
class 10th
Time: 90 mint

1. If \( N = \{2, 3, 4, 5, 6\} \) and \( M = \{1, 3, 5, 7, 9\} \)
   \[ (N \cup M) = (N \cap M) \]
2. Solve \( (x \square y \square) \square \)
3. Find the value of \( x \) when \( \log x = 0.9009 \)
4. Solve by using formula \( (2a + 2b)(2a - 2b)(4a^2 + 4b^2) \)
5. Find L.C.M by factorization;
   \[ 18 a^2 b c d , \quad 42 a b c d \quad 60 a b c d \]

Post-test Mathematics

Post-test Math
Marks: 50
class 10th
Time: 90 mint

1. Find the value of \( x \) and \( y \) when \( 4x + 2y = 6, x + y = 8 \)
2. Solve \( 4x + 1 \leq 7, \quad (x, y \in \mathbb{R}) \)
3. Solve by quadratic formula \( 8(x + 2) = x(x - 6) \)
4. Eliminate \( x \) by formula; \( y - 1/y = m, \quad y^2 + 1/y = n \)
5. Find value of \( x \) when \( 4:6 = 6 + x : 4 + 8x \)
Pre-test English

<table>
<thead>
<tr>
<th>Pre-test English</th>
<th>Marks: 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>class 10th</td>
<td>Time:</td>
</tr>
<tr>
<td></td>
<td>90 mint</td>
</tr>
</tbody>
</table>

1. Write a letter to a friend telling him how you and your class went on a study tour and what happened there.

2. Write a paragraph on any one topic;
   Uses of internet, Importance of newspaper, Education for women

3. Use following words in your own sentences;
   Market, journey, industry, leader, interesting, study, respect, solution, normal, literate

4. Convert the following into Passive Voice;
   a) The President delivered a beautiful speech.
   b) They were not working out the solutions to all the problems.
   c) Have you broken up with all your friends?
   d) Will you supply me the whole lot of the medicine?
   e) Isn’t the company looking for new executives?

5. Read the passage and answer the questions:

   Because my father was super fit. When we went for a swim he would walk down the beach on his hands then flip into the water. My puny physique was a torment to him. He applied himself ruthlessly to putting condition on my bones. Several times I was actually loaned out to family friends for fattening, like a weaner pig.
They would squeeze my arms or prod me with their fingers and threaten: 'You're a picky little eater, aren't you? Give me six weeks and I'll soon put flesh on you.' But my body always resisted. I once came home from an intensive course of porridge, scrambled eggs, cottage pie and rice pudding thinner than when I went away. Food went through me and never stuck. My father regarded the problem as a contest of willpower: my stubborn stomach versus his furious and uncompromising good intentions. But no matter what he promised, bribed or threatened, there was some food I simply would not swallow. After Sunday dinners were over I would often be made to stay at the table, sometimes for hours on end, trying to force down a plateful of roast mutton and soggy cabbage. I can still remember how slimy and cold and evil the muck on my plate became; its greyness and sickening stench.

**Questions**

1. Why was the author loaned out to family friends?
2. Why was he forced to stay at table after dinner?
3. What did his father regard the problem as?
4. What happened after he was back from a comprehensive eating course?
5. What was the physical difference between the father and the son?
1. Recall the lesson, “A Father’s Advice” and answer the following questions.
   I. What is a respectable way of earning a livelihood?
   II. Who is master of death and life?
   III. What should we like for others?
   IV. What should we do to command greater respect?
   V. How should we treat others?

2. Use given words in your own sentences;
   Battle, refresh, care, across, survive, relative, worst, injured, brave, humble

3. Match the words from list “A” with the words having opposite meanings in list “B”

<table>
<thead>
<tr>
<th>List A</th>
<th>List B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worst</td>
<td>Low</td>
</tr>
<tr>
<td>High</td>
<td>Best</td>
</tr>
<tr>
<td>Humble</td>
<td>Cruel</td>
</tr>
<tr>
<td>Kind</td>
<td>Enemy</td>
</tr>
<tr>
<td>Friend</td>
<td>Proud</td>
</tr>
</tbody>
</table>

4. Fill in the blanks with proper word given against each sentence.
   a. She----a glance at me. (caste, cast)
b. He is Syed by---------. (caste, cast)

c. He put the--------on fire. (cattle, kettle)

d. The ------ are grazing. (cattle, kettle)

e. I bought some------from the bazaar. (cloth, clothe)

f. He is wearing beautiful------------.(cloth, clothe)

g. Pakistan exports-------cloth. (course, coarse)

h. This question is out of--------------. (course, coarse)

i. A------breeze was blowing. (cool, cold)

j. -------wind blows in winter. (cool, cold)

5. Change the narration of the following sentences.

I. I said to him, “will you lend me your book?”

II. He said to me, “Were you absent yesterday?”

III. The man said, “May I come in?”

IV. He said to me, “Can you open the big door?”

V. He said, “May you live long”

6. Translate into Urdu;

Islam teaches us to treat our enemies without discrimination. It advises us to treat our enemies humanly and politely as far as it is possible. Islam is great protector of human rights. As Islam stresses that all are equal in the eyes of Allah, so the Jews got equal rights.
The following statistics were applied for the analysis of data:
Mean, standard deviation, variance, correlation and t-test

**Tables Phase I**

**Table 4.1 Descriptive Statistics for Pre-test Control Group in the subject of Math**

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>21.33</td>
<td>4.51</td>
<td>20.36</td>
</tr>
</tbody>
</table>

**Table 4.2 Descriptive Statistics for Pre-test Experimental Group in the subject of Math**

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>20.76</td>
<td>4.51</td>
<td>20.39</td>
</tr>
</tbody>
</table>

Table 4.1 and 4.2 show that both the groups, control and experimental, have almost equal score in pre-test in the subject of Math. Mean of control group is 21.33 and mean of experimental group is 20.76, standard deviation of control and experimental group is same, variance of control and experimental group is 20.36 and 20.39 respectively. Data shows that learning outcome of both groups in the subject of Math is almost equal.
Table 4.3 Comparative Statistics for Pre-test Control Group and Pre-test Experimental Group in the subject of Math

<table>
<thead>
<tr>
<th></th>
<th>Pre-test</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control Group</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>21.33</td>
<td></td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>4.51</td>
<td>4.51</td>
</tr>
<tr>
<td>Variance</td>
<td>20.36</td>
<td>20.39</td>
</tr>
<tr>
<td>Range</td>
<td>17.00</td>
<td>16.00</td>
</tr>
<tr>
<td>Percentiles 25</td>
<td>16.75</td>
<td>16.75</td>
</tr>
<tr>
<td>Percentiles 50</td>
<td>21.50</td>
<td>21.00</td>
</tr>
<tr>
<td>Percentiles 75</td>
<td>25.00</td>
<td>23.00</td>
</tr>
</tbody>
</table>

Table 4.3 shows that both the groups, control and experimental, have almost equal quantitative values in pre-test in the subject of Math. Mean of control group is 21.33 and mean of experimental group is 20.76, standard deviation of both groups is 4.51 and variance of control and experimental group is 20.36 and 20.39 respectively. Range of control group is 17 whereas range of experimental group is 16. Percentiles show that at 25th percentile, value of both groups is same, at 50th percentile, control group has 21.50
and experimental group has value of 21, at 75\textsuperscript{th} percentile, value of control group is 25 and value of experimental group is 23. Data shows that students of both groups have almost equal learning outcome in the subject of Math.

Graph 4.1 Showing comparative statistics for pre-test control group and pre-test experimental group in the subject of math

Graph 4.2 Showing percentile values for pre-test control group and pre-test experimental group in the subject of math
Table 4.4 Descriptive Statistics for Pre-test Control Group in the subject of English

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>24.70</td>
<td>3.18</td>
<td>10.14</td>
</tr>
</tbody>
</table>

Table 4.5 Descriptive Statistics for Pre-test Experimental Group in the subject of English

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>24.36</td>
<td>3.02</td>
<td>9.13</td>
</tr>
</tbody>
</table>
Table 4.4 and 4.5 show that both the groups, control and experimental, have almost equal score in pre-test in the subject of English. Mean of control group is 24.70 and mean of experimental group is 24.36, standard deviation of control and experimental group is 3.18 and 3.02, variance of control and experimental group is 10.14 and 9.13 respectively. Data shows that learning outcome of both groups in the subject of English is almost equal.

Table 4.6 Comparative Statistics for Pre-test Control Group and Pre-test Experimental Group in the subject of English

<table>
<thead>
<tr>
<th>Pre-test</th>
<th>Control Group</th>
<th>Experimental Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>24.70</td>
<td>24.36</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>3.18</td>
<td>3.02</td>
</tr>
<tr>
<td>Variance</td>
<td>10.14</td>
<td>9.13</td>
</tr>
<tr>
<td>Range</td>
<td>12.00</td>
<td>12.00</td>
</tr>
<tr>
<td>Percentiles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>22.75</td>
<td>22.00</td>
</tr>
<tr>
<td>50</td>
<td>25.00</td>
<td>24.50</td>
</tr>
<tr>
<td>75</td>
<td>27.25</td>
<td>27.00</td>
</tr>
</tbody>
</table>
Table 4.6 shows that both the groups, control and experimental, have almost equal score in pre-test in the subject of English. Mean of control group is 24.70 and mean of experimental group is 24.36, standard deviation of control and experimental group is 3.18 and 3.02, variance of control and experimental group is 10.14 and 9.13 respectively. Range of both groups is 12. Percentiles show that at 25th percentile, value of control group is 22.75 and experimental group is 22, at 50th percentile control group has 25 and experimental group has value of 24.50, at 75th percentile value of control group is 27.25 and value of experimental group is 17. Data shows that learning outcome of both groups in the subject of English is almost equal.

Graph 4.3 Showing comparative statistics for pre-test control group and pre-test experimental group in the subject of English
Graph 4.4 showing percentile values of pre-test control group and pre-test experimental group in the subject of English.
Table 4.7 Descriptive Statistics for Post-test Control Group in the subject of Math

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>15.96</td>
<td>2.22</td>
<td>4.93</td>
</tr>
</tbody>
</table>

Table 4.8 Descriptive Statistics for Post-test Experimental Group in the subject of Math

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>25.93</td>
<td>2.74</td>
<td>7.51</td>
</tr>
</tbody>
</table>

Table 4.7 and 4.8 show that both the groups, control and experimental, have prominent difference in post-test in the subject of Math. Mean of control group is 15.96 and mean of experimental group is 25.93, standard deviation of control group is 2.22 and experimental group is 2.74, variance of control and experimental group is 4.92 and 7.51 respectively. Data shows a prominent difference in learning outcome of both groups.
Table 4.9 Comparative Statistics for Post-test Control Group and Post-test Experimental Group in the subject of Math

<table>
<thead>
<tr>
<th></th>
<th>Post-test Control Group</th>
<th>Post-test Experimental Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>15.96</td>
<td>25.93</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>2.22</td>
<td>2.74</td>
</tr>
<tr>
<td>Variance</td>
<td>4.92</td>
<td>7.51</td>
</tr>
<tr>
<td>Range</td>
<td>9.00</td>
<td>9.00</td>
</tr>
<tr>
<td>Percentiles 25</td>
<td>14.75</td>
<td>23.00</td>
</tr>
<tr>
<td>Percentiles 50</td>
<td>16.00</td>
<td>26.00</td>
</tr>
<tr>
<td>Percentiles 75</td>
<td>17.00</td>
<td>28.00</td>
</tr>
</tbody>
</table>

Table 4.9 shows that both the groups, control and experimental, have prominent difference in post-test in the subject of Math. Mean of control group is 15.96 and mean of experimental group is 25.93, standard deviation of control group is 2.22 and experimental group is 2.74 , variance of control and experimental group is 4.92 and 7.51 respectively. Range of both groups is 9 . Percentiles show that at 25th percentile, value of control group is 14.75 which is lesser than 23, the value of experimental group, at 50th percentile control group has 16 and experimental group has value of 26, at 75th percentile value of
control group is 17 which is lesser than 28, value of experimental group. Data shows a prominent difference in learning outcome of both groups.

Graph 4.5 Showing Comparative Statistics for Post-test Control Group and Post-test Experimental Group in the subject of Math

<table>
<thead>
<tr>
<th></th>
<th>Control Group</th>
<th>Experimental Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>25.93</td>
<td>15.96</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>2.22</td>
<td>2.74</td>
</tr>
<tr>
<td>Variance</td>
<td>4.92</td>
<td>7.51</td>
</tr>
<tr>
<td>Range</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

Graph 4.6 Showing percentile values for Post-test Control Group and Post-test Experimental Group in the subject of Math
Table 4.10 Comparative Statistics for Pre-test and Post-test of Control Group in the subject of Math

<table>
<thead>
<tr>
<th></th>
<th>Control Group</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
<td>Post-test</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>21.33</td>
<td>15.96</td>
<td></td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>4.51</td>
<td>2.22</td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td>20.36</td>
<td>4.92</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>17.00</td>
<td>9.00</td>
<td></td>
</tr>
<tr>
<td>Percentiles</td>
<td>25</td>
<td>16.75</td>
<td>14.75</td>
</tr>
</tbody>
</table>

![Bar chart showing percentiles for Control Group and Experimental Group](chart.png)
Table 4.10 shows comparative statistics of Pre-test and Post-test of Control Group in the subject of Math. The data shows prominent difference between pre-test and post-test. Mean value of pre-test is 21.33 and mean of post-test is 15.56, standard deviation of pre-test is 4.51 and post-test is 2.22, variance of pre-test is 20.36 and of post-test is 4.92. Range of pre-test is 17 and of post-test is 9. Percentiles show that at 25th percentile, value of pre-test is 16.75 and post-test is 14.75, at 50th percentile pre-test has 21.50 and post-test has value of 16, at 75th percentile value of pre-test is 25 which is greater than 17, value of post-test.

Graph 4.7 Showing Comparative Statistics for Pre-test and Post-test of Control Group in the subject of Math
<table>
<thead>
<tr>
<th></th>
<th>Pre-test</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>21.33</td>
<td>15.96</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>4.51</td>
<td>2.22</td>
</tr>
<tr>
<td>Variance</td>
<td>20.36</td>
<td>4.92</td>
</tr>
<tr>
<td>Range</td>
<td>17</td>
<td>9</td>
</tr>
</tbody>
</table>

Graph 4.8 showing percentile values for Pre-test and Post-test of Control Group in the subject of Math.
Table 4.11 Comparative Statistics for Pre-test and Post-test of Experimental Group in the subject of Math

<table>
<thead>
<tr>
<th>Experimental Group</th>
<th>Pre-test</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>20.76</td>
<td>25.93</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>4.51</td>
<td>2.74</td>
</tr>
<tr>
<td>Variance</td>
<td>20.39</td>
<td>7.51</td>
</tr>
<tr>
<td>Range</td>
<td>16.00</td>
<td>9.00</td>
</tr>
<tr>
<td>Percentiles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>16.75</td>
<td>23.00</td>
</tr>
<tr>
<td>50</td>
<td>21.00</td>
<td>26.00</td>
</tr>
<tr>
<td>75</td>
<td>23.00</td>
<td>28.00</td>
</tr>
</tbody>
</table>

Table 4.11 shows the comparative statistics of Pre-test and Post-test of Experimental Group in the subject of Math. Mean value of pre-test is 20.76 and mean of post-test is 25.93, standard deviation of pre-test is 4.51 and post-test is 2.74, variance of pre-test is 20.39 and of post-test is 7.51. Range of pre-test is 16 and of post-test is 9. Percentiles show that at 25th percentile value of pre-test is 16.75 and post-test is 23, at 50th percentile pre-test has 21 and post-test has value of 26, at 75th percentile value of pre-test is 23 which is lesser than 28, value of post-test. The data shows that a prominent difference exists between pre-test and post-test.
Graph 4.9 showing Comparative Statistics for Pre-test and Post-test of Experimental Group in the subject of Math

![Comparative Statistics Graph](image)

Graph 4.10 Showing percentile values for Pre-test and Post-test of Experimental Group in the subject of Math
Table 4.12 Descriptive Statistics for Post-test Control Group in the subject of English

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>14.8333</td>
<td>3.2279</td>
<td>10.420</td>
</tr>
</tbody>
</table>

Table 4.13 Descriptive Statistics for Post-test Experimental Group in the subject of English

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>26.8667</td>
<td>2.2702</td>
<td>5.154</td>
</tr>
</tbody>
</table>
Table 4.12 and 4.13 show that both the groups, control and experimental, have significant difference in the post-test in the subject of English. Mean of control group is 14.83 and mean of experimental group is 26.86, standard deviation of control group is 3.22 and experimental group is 2.27, variance of control group is 10.420 and experimental group is 5.154.

**Table 4.14 Comparative Statistics for Post-test Control Group and Post-test Experimental Group in the subject of English**

<table>
<thead>
<tr>
<th>Post-test</th>
<th>Control Group</th>
<th>Experimental Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>14.83</td>
<td>26.86</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>3.22</td>
<td>2.27</td>
</tr>
<tr>
<td>Variance</td>
<td>10.42</td>
<td>5.15</td>
</tr>
<tr>
<td>Range</td>
<td>17.00</td>
<td>7.00</td>
</tr>
<tr>
<td>Percentiles</td>
<td>25</td>
<td>13.00</td>
</tr>
</tbody>
</table>
Table 4.14 shows that both the groups, control and experimental, have significant difference in the post-test in the subject of English. Mean of control group is 14.83 and mean of experimental group is 26.86, standard deviation of control group is 3.22 and experimental group is 2.27, variance of control group is 10.42 and experimental group is 5.15. Range of control groups is 17 and experimental group is 7. Percentiles show that at 25th percentile, value of control group is 13 and experimental group is 25, at 50th percentile, control group has 14 and experimental group has value of 27, at 75th percentile, value of control group is 15.25 which is lesser than 29, value of experimental group.

**Graph 4.11 Showing Comparative Statistics for Post-test Control Group and Post-test Experimental Group in the subject of English**
Graph 4.12 Showing percentile values for Post-test Control Group and Post-test Experimental Group in the subject of English.
Table 4.15 Comparative Statistics for Pre-test and Post-test of Control Group in the subject of English

<table>
<thead>
<tr>
<th>Control Group</th>
<th>Pre-test</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>24.70</td>
<td>14.83</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>3.18</td>
<td>3.22</td>
</tr>
<tr>
<td>Variance</td>
<td>10.14</td>
<td>10.41</td>
</tr>
<tr>
<td>Range</td>
<td>12.00</td>
<td>17.00</td>
</tr>
<tr>
<td>Percentiles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>22.75</td>
<td>13.00</td>
</tr>
<tr>
<td>50</td>
<td>25.00</td>
<td>14.00</td>
</tr>
<tr>
<td>75</td>
<td>27.25</td>
<td>15.25</td>
</tr>
</tbody>
</table>

Table 4.15 shows the comparative statistics of pre-test and post-test of control group in the subject of English. The data shows prominent difference between pre-test and post-test. Mean value of pre-test is 24.70 and mean of post-test is 14.83, standard deviation of pre-test is 3.18 and post-test is 3.22, variance of pre-test is 10.14 and of post-test is 10.41. Range of pre-test is 12 and of post-test is 17. Percentiles show that at 25\textsuperscript{th} percentile, value of pre-test is 22.75 and post-test is 13, at 50\textsuperscript{th} percentile pre-test has 25 and post-test has value of 14, at 75\textsuperscript{th} percentile, value of pre-test is 27.25 which is greater than 15.25, value of post-test.
Graph 4.13 Showing Comparative Statistics for Pre-test and Post-test of Control Group in the subject of English

Graph 4.14 Showing percentile values for Pre-test and Post-test of Control Group in the subject of English
Table 4.16 Comparative Statistics for Pre-test and Post-test of Experimental Group in the subject of English

<table>
<thead>
<tr>
<th>Experimental Group</th>
<th>Pre-test</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>24.36</td>
<td>26.86</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>3.02</td>
<td>2.27</td>
</tr>
<tr>
<td>Variance</td>
<td>9.13</td>
<td>5.15</td>
</tr>
<tr>
<td>Range</td>
<td>12.00</td>
<td>7.00</td>
</tr>
<tr>
<td>Percentiles</td>
<td>25</td>
<td>22.00</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>24.50</td>
</tr>
</tbody>
</table>

![Graph showing comparative statistics](image)
Table 4.16 shows the comparative statistics of pre-test and post-test of experimental group in the subject of English. Mean value of pre-test is 24.36 and mean of post-test is 26.86, standard deviation of pre-test is 3.02 and post-test is 2.27, variance of pre-test is 9.13 and of post-test is 5.15. Range of pre-test is 12 and of post-test is 7. Percentiles show that at 25th percentile value of pre-test is 22 and post-test is 25, at 50th percentile pre-test has 24.50 and post-test has value of 27, at 75th percentile value of pre-test is 27 which is lesser than 29, value of post-test.

Graph 4.15 Showing Comparative Statistics for Pre-test and Post-test of Experimental Group in the subject of English
Graph 4.16 Showing percentile values for Pre-test and Post-test for Experimental Group in the subject of English
Table 4.17 Correlation coefficient between pre-test and post-test of control group in the subject of Math

<table>
<thead>
<tr>
<th>Control Group</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>pre-test</td>
<td>0.156</td>
</tr>
<tr>
<td>post-test</td>
<td></td>
</tr>
</tbody>
</table>

The findings of table 4.17 show that correlation between pre-test and post-test of control group in the subject of Math is 0.156, which is not significant at $\alpha=0.05$ or $\alpha=0.01$. It shows that score of pre-test have not significant relationship with the score of post-test in the subject of Math. It also shows that learning outcome did not remain sustainable.

Table 4.18 Correlation coefficient between pre-test and post-test of experimental group in the subject of Math

<table>
<thead>
<tr>
<th>Experimental Group</th>
<th>Correlation coefficient*</th>
</tr>
</thead>
<tbody>
<tr>
<td>pre-test</td>
<td>0.364</td>
</tr>
<tr>
<td>post-test</td>
<td></td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.05 level

The findings of table 4.18 show that correlation between pre-test and post-test of experimental group in the subject of Math is 0.364, which is significant at $\alpha=0.05$. It shows that score of pre-test have significant relationship with the score of post-test in the subject of Math.
Table 4.19 Correlation coefficient between pre-test control group and post-test control group in the subject of English

<table>
<thead>
<tr>
<th>Groups</th>
<th>Correlation coefficient*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>0.159</td>
</tr>
<tr>
<td>Experimental</td>
<td></td>
</tr>
</tbody>
</table>

The findings of table 4.19 show that correlation between pre-test and post-test of control group in the subject of English is 0.159, which is not significant at $\alpha=0.05$ or $\alpha=0.01$. It shows that score of pre-test have not significant relationship with the score of post-test in the subject of English. It also shows that learning outcome didn’t remain sustainable.

Table 4.20 Correlation coefficient between pre-test experimental group and post-test experimental group in the subject of English

<table>
<thead>
<tr>
<th>Groups</th>
<th>Correlation coefficient*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>0.445</td>
</tr>
<tr>
<td>Experimental</td>
<td></td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.05 level

The findings of table 4.20 show that correlation between pre-test and post-test of experimental group in the subject of English is 0.445, which is significant at $\alpha=0.05$. It shows that score of pre-test have significant relationship with the score of post-test in the subject of English.
Table 4.21 Mean difference between pre-test control group and pre-test

1.1.1.1.1 Experimental group in the subject of Math

<table>
<thead>
<tr>
<th>Groups</th>
<th>Mean</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>21.33</td>
<td>1.172</td>
</tr>
<tr>
<td></td>
<td>20.76</td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table 4.21 shows that the mean differences between pre-test control group and pre-test experimental group in the subject of Math is significant, $t=1.172$.

The data reveals that both the groups possess almost equal learning outcome in the subject of Math.

Table 4.22 Mean difference between pre-test control group and pre-test

1.1.1.1.2 Experimental group in the subject of English

<table>
<thead>
<tr>
<th>Groups</th>
<th>Mean</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>24.70</td>
<td>0.541</td>
</tr>
<tr>
<td></td>
<td>24.36</td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table 4.22 shows that the mean differences between pre-test control group and pre-test experimental group in the subject of English is significant, $t=1.541$.

The data reveals that both the groups possess almost equal learning outcome in the subject of English.
Table 4.23 Mean difference between post-test control group and post-test

1.1.1.1.3 Experimental group in the subject of Math

<table>
<thead>
<tr>
<th>Groups</th>
<th>Mean</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>15.96</td>
<td>-15.06</td>
</tr>
<tr>
<td></td>
<td>25.93</td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The 4.23 shows that the mean differences between post-test control group and post-test experimental group in the subject of Math is not significant, t=-15.06.

The data shows that learning outcome of experimental group is better than learning outcome of control group.

Table 4.24 Mean difference between post-test control group and post-test

1.1.1.1.4 Experimental group in the subject of English

<table>
<thead>
<tr>
<th>Groups</th>
<th>Mean</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>14.83</td>
<td>-16.60</td>
</tr>
<tr>
<td></td>
<td>26.86</td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table 4.24 shows that the mean differences between post-test control group and post-test experimental group in the subject of English is not significant, t=-16.60.

The data shows that learning outcome of experimental group is better than learning outcome of control group.
Tables Phase II

Table 4.25 Descriptive Statistics for Pre-test Control Group in the subject of Math

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>21.16</td>
<td>4.16</td>
<td>17.38</td>
</tr>
</tbody>
</table>

Table 4.26 Descriptive Statistics for Pre-test Experimental Group in the subject of Math

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>21.10</td>
<td>4.80</td>
<td>23.05</td>
</tr>
</tbody>
</table>

Table 4.25 and 4.26 show that both the groups, control and experimental, have almost equal learning outcome in pre-test in the subject of Math. Mean of control group is 21.16 and mean of experimental group is 21.10, standard deviation of control groups is 4.16 and experimental group is 4.80, variance of control and experimental group is 17.38 and 23.05 respectively.
Table 4.27 Comparative Statistics for Pre-test Control Group and Pre-test

<table>
<thead>
<tr>
<th>Experimental Group in the subject of Math</th>
<th>Pre-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control Group</td>
</tr>
<tr>
<td>Mean</td>
<td>21.16</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>4.16</td>
</tr>
<tr>
<td>Variance</td>
<td>17.38</td>
</tr>
<tr>
<td>Range</td>
<td>18.00</td>
</tr>
<tr>
<td>Percentiles</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>75</td>
</tr>
</tbody>
</table>

Table 4.27 shows that both the groups, control and experimental, have almost equal learning outcome in pre-test in the subject of Math. Mean of control group is 21.16 and mean of experimental group is 21.10, standard deviation of control groups is 4.16 and experimental group is 4.80, variance of control and experimental group is 17.38 and 23.05 respectively. Range of control group is 18 whereas range of experimental group is 16. Percentiles show that at 25th percentile, value of control groups is 18.75 and
experimental group is 16.75, at 50\textsuperscript{th} percentile control group has 21.50 and experimental group has value of 21, at 75\textsuperscript{th} percentile value of control group is 24 and value of experimental group is 23.75.

**Graph 4.17 Showing Comparative Statistics for Pre-test Control Group and Pre-test Experimental Group in the subject of Math**

<table>
<thead>
<tr>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Group</td>
<td>21.1621.1</td>
<td>4.16</td>
<td>17.38</td>
</tr>
<tr>
<td>Experimental Group</td>
<td>23.05</td>
<td>4.8</td>
<td>18</td>
</tr>
</tbody>
</table>

**Graph 4.18 Showing percentile values for Pre-test Control Group and Pre-test Experimental Group in the subject of Math**
**Table 4.28 Descriptive Statistics for Pre-test Control Group in the subject of English**

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>22.46</td>
<td>3.29</td>
<td>10.87</td>
</tr>
</tbody>
</table>

**Table 4.29 Descriptive Statistics for Pre-test Experimental Group in the subject of English**

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>23.26</td>
<td>3.51</td>
<td>12.34</td>
</tr>
</tbody>
</table>
Table 4.28 and 4.29 show that both the groups, control and experimental, have almost equal quantitative values in pre-test in the subject of English. Mean of control group is 22.46 and mean of experimental group is 23.36, standard deviation of control and experimental group is 3.29 and 3.51, variance of control and experimental group is 10.87 and 12.34 respectively. Range of both groups is 13. Percentiles show that at 25th percentile, value of control group is 20 and experimental group is 21, at 50th percentile, control group has 22 and experimental group has value of 23, at 75th percentile, value of control group is 25 and value of experimental group is 26.25.

<table>
<thead>
<tr>
<th></th>
<th>Pre-test Control Group</th>
<th>Pre-test Experimental Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>22.46</td>
<td>23.26</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>3.29</td>
<td>3.51</td>
</tr>
<tr>
<td>Variance</td>
<td>10.87</td>
<td>12.34</td>
</tr>
<tr>
<td>Range</td>
<td>13.00</td>
<td>13.00</td>
</tr>
</tbody>
</table>
Table 4.30 shows that both the groups, control and experimental, have almost equal quantitative values in pre-test in the subject of English. Mean of control group is 22.46 and mean of experimental group is 23.36, standard deviation of control and experimental group is 3.29 and 3.51, variance of control and experimental group is 10.87 and 12.34 respectively. Range of both groups is 13. Percentiles show that at 25th percentile, value of control group is 20 and experimental group is 21, at 50th percentile, control group has 22 and experimental group has value of 23, at 75th percentile, value of control group is 25 and value of experimental group is 26.25.

Graph 4.19 Showing Comparative Statistics for Pre-test Control Group and Pre-test Experimental Group in the subject of English
Graph 4.20 Showing percentile values for Pre-test Control Group and Pre-test Experimental Group in the subject of English
Table 4.31 Descriptive Statistics for Post-test Control Group in the subject of Math

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>15.96</td>
<td>3.31</td>
<td>10.99</td>
</tr>
</tbody>
</table>

Table 4.32 Descriptive Statistics for Post-test Experimental Group in the subject of Math

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>26.03</td>
<td>2.57</td>
<td>6.65</td>
</tr>
</tbody>
</table>

Table 4.31 and 4.32 show that both the groups, control and experimental, have prominent difference in post-test in the subject of Math. Mean of control group is 15.96 and mean of experimental group is 26.03, standard deviation of control group is 3.31 and experimental group is 2.57, variance of control and experimental group is 10.99 and 6.65 respectively.
Table 4.33 Comparative Statistics for Post-test Control Group and Post-test Experimental Group in the subject of Math

<table>
<thead>
<tr>
<th></th>
<th>Control Group</th>
<th>Experimental Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>15.96</td>
<td>26.03</td>
</tr>
<tr>
<td><strong>Std. Deviation</strong></td>
<td>3.31</td>
<td>2.57</td>
</tr>
<tr>
<td><strong>Variance</strong></td>
<td>10.99</td>
<td>6.65</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>13.00</td>
<td>9.00</td>
</tr>
<tr>
<td><strong>Percentiles</strong></td>
<td>25</td>
<td>14.00</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>15.00</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>18.25</td>
</tr>
</tbody>
</table>

Table 4.33 shows that both the groups, control and experimental, have prominent difference in post-test in the subject of Math. Mean of control group is 15.96 and mean of experimental group is 26.03, standard deviation of control group is 3.31 and experimental group is 2.57, variance of control and experimental group is 10.99 and 6.65 respectively. Range of control groups is 13 and experimental group is 9. Percentiles show that at 25\textsuperscript{th} percentile, value of control group is 14 and experimental group is 24, at 50\textsuperscript{th} percentile,
control group has 15 and experimental group has value of 26.50, at 75th percentile, value of control group is 18.25 and value of experimental group is 28.

Graph 4.21 Showing Comparative Statistics for Post-test Control Group and Post-test Experimental Group in the subject of Math

Graph 4.22 Showing percentile values for Post-test Control Group and Post-test Experimental Group in the subject of Math
Table 4.34 Comparative Statistics for Pre-test and Post-test of Control Group in the subject of Math

<table>
<thead>
<tr>
<th>Control Group</th>
<th>Pre-test</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>21.16</td>
<td>15.96</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>4.16</td>
<td>3.31</td>
</tr>
<tr>
<td>Variance</td>
<td>17.38</td>
<td>10.99</td>
</tr>
<tr>
<td>Range</td>
<td>18.00</td>
<td>13.00</td>
</tr>
<tr>
<td>Percentiles</td>
<td>25</td>
<td>18.75</td>
</tr>
</tbody>
</table>
Table 4.34 shows the comparative statistics of pre-test and post-test of control group in the subject of Math. The data shows prominent difference between pre-test and post-test. Mean value of pre-test is 21.16 and mean of post-test is 15.96, standard deviation of pre-test is 4.16 and post-test is 3.31, variance of pre-test is 17.38 and of post-test is 10.99. Range of pre-test is 18 and of post-test is 13. Percentiles show that at 25th percentile, value of pre-test is 18.75 and post-test is 14, at 50th percentile pre-test has 21.50 and post-test has value of 15, at 75th percentile value of pre-test is 24 which is greater than 18.25, value of post-test.

Graph 4.23 Showing Comparative Statistics for Pre-test and Post-test of Control Group in the subject of Math
Graph 4.24 Showing percentile values for Pre-test and Post-test of Control Group in the subject of Math
<table>
<thead>
<tr>
<th></th>
<th>Experimental Group</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
<td>Post-test</td>
</tr>
<tr>
<td>Mean</td>
<td>21.10</td>
<td>26.03</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>4.80</td>
<td>2.57</td>
</tr>
<tr>
<td>Variance</td>
<td>23.05</td>
<td>6.65</td>
</tr>
<tr>
<td>Range</td>
<td>16.00</td>
<td>9.00</td>
</tr>
<tr>
<td>Percentiles 25</td>
<td>16.75</td>
<td>24.00</td>
</tr>
<tr>
<td>Percentiles 50</td>
<td>21.00</td>
<td>26.50</td>
</tr>
<tr>
<td>Percentiles 75</td>
<td>23.75</td>
<td>28.00</td>
</tr>
</tbody>
</table>

Table 4.35 shows the comparative statistics of pre-test and post-test of experimental group in the subject of Math. The data shows prominent difference between pre-test and post-test. Mean value of pre-test is 21.10 and mean of post-test is 26.03, standard deviation of pre-test is 4.80 and post-test is 2.57, variance of pre-test is 23.05 and of post-test is 6.65. Range of pre-test is 16 and of post-test is 9. Percentiles show that at 25th percentile, value of pre-test is 16.75 and post-test is 24, at 50th percentile, pre-test has value of 21 and post-test has value of 26.50, at 75th percentile, value of pre-test is 23.75 which is lesser than 28, value of post-test.
Graph 4.25 Showing Comparative Statistics for Pre-test and Post-test of Experimental Group in the subject of Math

Graph 4.26 Showing percentile values for Pre-test and Post-test of Experimental Group in the subject of Math
Table 4.36 Descriptive Statistics for Post-test Control Group in the subject of English

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>14.83</td>
<td>3.10</td>
<td>9.66</td>
</tr>
</tbody>
</table>

Table 4.37 Descriptive Statistics for Post-test Experimental Group in the subject of English

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>25.66</td>
<td>2.83</td>
<td>8.02</td>
</tr>
</tbody>
</table>
Table 4.36 and 4.37 show that both the groups, control and experimental, have significant difference in the post-test in the subject of English. Mean of control group is 14.83 and mean of experimental group is 25.66, standard deviation of control group is 3.10 and experimental group is 2.83, variance of control group is 9.66 and experimental group is 8.02.

<table>
<thead>
<tr>
<th></th>
<th>Post-test Control Group</th>
<th>Post-test Experimental Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>14.83</td>
<td>25.66</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>3.10</td>
<td>2.83</td>
</tr>
<tr>
<td>Variance</td>
<td>9.66</td>
<td>8.02</td>
</tr>
<tr>
<td>Range</td>
<td>17.00</td>
<td>10.00</td>
</tr>
<tr>
<td>Percentiles</td>
<td>25</td>
<td>13.00</td>
</tr>
</tbody>
</table>
Table 4.38 shows that both the groups, control and experimental, have significant difference in the post-test in the subject of English. Mean of control group is 14.83 and mean of experimental group is 25.66, standard deviation of control group is 3.10 and experimental group is 2.83, variance of control group is 9.66 and experimental group is 8.02. Range of control groups is 17 and experimental group is 10. Percentiles show that at 25th percentile, value of control group is 13 and experimental group is 23, at 50th percentile, control group has 14.50 and experimental group has value of 25.50, at 75th percentile, value of control group is 16 which is lesser than 28.25, value of experimental group.

Graph 4.27 Showing Comparative Statistics for Post-test Control Group and Post-test Experimental Group in the subject of English
Graph 4.28 Showing percentile values Post-test Control Group and Post-test Experimental Group in the subject of English
Table 4.39 Comparative Statistics for Pre-test and Post-test of Control Group in the subject of English

<table>
<thead>
<tr>
<th></th>
<th>Pre-test</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>22.46</td>
<td>14.83</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>3.29</td>
<td>3.10</td>
</tr>
<tr>
<td>Variance</td>
<td>10.87</td>
<td>9.66</td>
</tr>
<tr>
<td>Range</td>
<td>13.00</td>
<td>17.00</td>
</tr>
<tr>
<td>Percentiles 25</td>
<td>20.00</td>
<td>13.00</td>
</tr>
<tr>
<td>Percentiles 50</td>
<td>22.00</td>
<td>14.50</td>
</tr>
<tr>
<td>Percentiles 75</td>
<td>25.00</td>
<td>16.00</td>
</tr>
</tbody>
</table>

Table 4.39 shows the comparative statistics of pre-test and post-test of control group in the subject of English. The data shows prominent difference between pre-test and post-test. Mean value of pre-test is 22.46 and mean of post-test is 14.83, standard deviation of pre-test is 3.29 and post-test is 3.10, variance of pre-test is 10.87 and of post-test is 9.66. Range of pre-test is 13 and of post-test is 17. Percentiles show that at 25th percentile, value of pre-test is 20 and post-test is 13, at 50th percentile, pre-test has 22 and post-test has value of 14.50, at 75th percentile, value of pre-test is 25 which is greater than 16, value of post-test.
Graph 4.29 Showing Comparative Statistics for Pre-test and Post-test of Control Group in the subject of English

Graph 4.30 Showing percentile values for Pre-test and Post-test of Control Group in the subject of English
Table 4.40 Comparative Statistics for Pre-test and Post-test of Experimental Group in the subject of English

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>Pre-test</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>21.00</td>
<td>23.00</td>
</tr>
<tr>
<td>50</td>
<td>23.00</td>
<td>25.50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experimental Group</th>
<th>Pre-test</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>23.26</td>
<td>25.66</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>3.51</td>
<td>2.83</td>
</tr>
<tr>
<td>Variance</td>
<td>12.34</td>
<td>8.02</td>
</tr>
<tr>
<td>Range</td>
<td>13.00</td>
<td>10.00</td>
</tr>
<tr>
<td>Percentiles</td>
<td>25</td>
<td>21.00</td>
</tr>
</tbody>
</table>

275
Table 4.40 shows the comparative statistics of pre-test and post-test of experimental group in the subject of English. The data shows prominent difference between pre-test and post-test. Mean value of pre-test is 23.26 and mean of post-test is 25.66, standard deviation of pre-test is 3.51 and post-test is 2.83, variance of pre-test is 12.34 and of post-test is 8.02. Range of pre-test is 13 and of post-test is 10. Percentiles show that at 25th percentile, value of pre-test is 21 and post-test is 23, at 50th percentile, pre-test has 23 and post-test has value of 25.50, at 75th percentile, value of pre-test is 26.25 which is lesser than 28.25, value of post-test.

Graph 4.31 Showing Comparative Statistics for Pre-test and Post-test of Experimental Group in the subject of English
Graph 4.32 Showing percentile values for Pre-test and Post-test of Experimental Group in the subject of English
Table 4.41 Correlation coefficient between pre-test and post-test of control group in the subject of Math

<table>
<thead>
<tr>
<th>Control Group</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>pre-test</td>
<td>0.050</td>
</tr>
<tr>
<td>post-test</td>
<td></td>
</tr>
</tbody>
</table>

The findings of table 4.41 show that correlation between pre-test and post-test of control group in the subject of Math is 0.050, which is not significant at $\alpha=0.05$ or $\alpha=0.01$. It shows that score of pre-test have not significant relationship with the score of post-test in the subject of Math. It also shows that learning outcome didn’t remain sustainable.

Table 4.42 Correlation coefficient between pre-test and post-test of experimental group in the subject of Math

<table>
<thead>
<tr>
<th>experimental Groups</th>
<th>Correlation coefficient*</th>
</tr>
</thead>
<tbody>
<tr>
<td>pre-test</td>
<td>0.356</td>
</tr>
<tr>
<td>post-test</td>
<td></td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.05 level

The findings of table 4.42 show that correlation between pre-test and post-test of experimental group in the subject of Math is 0.356, which is significant at $\alpha=0.05$. It shows that score of pre-test have significant relationship with the score of post-test in the subject of Math
Table 4.43 Correlation coefficient between pre-test and post-test of control group in the subject of English

<table>
<thead>
<tr>
<th>control group</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>pre-test</td>
<td>-0.342</td>
</tr>
<tr>
<td>post-test</td>
<td></td>
</tr>
</tbody>
</table>

The findings of table 4.43 show that correlation between pre-test and post-test of control group in the subject of English is -0.342, which is not significant at $\alpha=0.05$ or $\alpha=0.01$. It shows that score of pre-test have not significant relationship with the score of post-test in the subject of English. It also shows that learning outcome didn’t remain sustainable.

Table 4.44 Correlation coefficient between pre-test and post-test of experimental group in the subject of English

<table>
<thead>
<tr>
<th>experimental Groups</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>pre-test</td>
<td>0.381</td>
</tr>
<tr>
<td>post-test</td>
<td></td>
</tr>
</tbody>
</table>

The findings of table 4.44 show that correlation between pre-test and post-test of experimental group in the subject of English is 0.381, which is significant at $\alpha=0.05$. It shows that score of pre-test have significant relationship with the score of post-test in the subject of English.
Table 4.45 Mean difference between pre-test control group and pre-test

1.1.1.2.2 Experimental group in the subject of Math

<table>
<thead>
<tr>
<th>Groups</th>
<th>Mean</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>21.16</td>
<td>0.164</td>
</tr>
<tr>
<td></td>
<td>21.10</td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table 4.45 shows that the mean differences between pre-test control group and pre-test experimental group in the subject of Math is significant, t=1.164.

The data reveals that both the groups possess almost equal learning outcome in the subject of Math.

Table 4.46 Mean difference between pre-test control group and pre-test

1.1.1.2.2.1.1 Experimental group in the subject of English

<table>
<thead>
<tr>
<th>Groups</th>
<th>Mean</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>22.46</td>
<td>1.167</td>
</tr>
<tr>
<td></td>
<td>23.26</td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table 4.46 shows that the mean differences between pre-test control group and pre-test experimental group in the subject of English is significant, t=1.167.

The data reveals that both the groups possess almost equal learning outcome in the subject of English.
Table 4.47 Mean difference between post-test control group and post-test experimental group in the subject of Math

<table>
<thead>
<tr>
<th>Groups</th>
<th>Mean</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>15.93</td>
<td>-12.46</td>
</tr>
<tr>
<td></td>
<td>26.03</td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table 4.47 shows that the mean differences between post-test control group and post-test experimental group in the subject of Math is not significant, $t=-12.46$.

The data shows that learning outcome of experimental group is better than learning outcome of control group.

Table 4.48 Mean difference between post-test control group and post-test experimental group in the subject of English

<table>
<thead>
<tr>
<th>Groups</th>
<th>Mean</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>14.83</td>
<td>-13.06</td>
</tr>
<tr>
<td></td>
<td>25.66</td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table 4.48 shows that the mean differences between post-test control group and post-test experimental group in the subject of English is not significant, $t=-13.06$.

The data shows that learning outcome of experimental group is better than learning outcome of control group.
CHAPTER 5
SUMMARY, FINDINGS, DISCUSSION CONCLUSIONS AND RECOMMENDATIONS

5.1 SUMMARY

5.1.1 Introduction

The study is designed to find out the impact of formative evaluation and fixed interval schedule reinforcement on academic achievements of secondary schools students. It was experimental design and completed in two phases with the duration of 120 days, 60 days for each phase.

5.1.2 Sample

Two groups of students having strength of 60 each were selected from Govt. High School Fateh Jang. First group was selected from 9th class and second group was selected from 10th class randomly. The selected sample of each group was divided into two equal groups, of which, one served as experimental group and other served as control group.

5.1.3 Hypotheses

Following were hypotheses of the study:

1. There is no significant relationship of formative evaluation with academic achievements of secondary school students.

2. There is no significant impact of formative evaluation on academic achievements of secondary school students.

3. There is no significant relationship of fixed interval schedule reinforcement with academic achievements of secondary school students.
4. There is no significant impact of fixed interval schedule reinforcement on academic achievements of secondary school students.

Following were sub-hypotheses of the study:

1. There is no significant difference between the performance of control group and experimental group in pre-test in the subject of Math.
2. There is no significant difference between the performance of control group and experimental group in pre-test in the subject of English.
3. There is no significant difference between the performance of control group and experimental group in post-test in the subject of Math.
4. There is no significant difference between the performance of control group and experimental group in post-test in the subject of English.

5.1.4 Methodology

The study was completed into two phases. In first phase 60 students were randomly selected from 9th class of Govt. High School Fateh Jang. This sample was divided into two equal groups. One served as control group and other as experimental group. A pretest was conducted to both the groups in the subjects of Math and English. Both groups were taught by subject specialist of Math and English. Control group was only taught and was not given any treatment whereas experimental group was taught and evaluated by teacher made test during session without receiving reinforcement. After completion of 60 days teaching, a post test was conducted in Math and English.

In second phase 60 students were randomly selected from 10th class of Govt. High School
Fateh Jang. This sample was divided into two equal groups. One served as control group and other as experimental group. A pretest was conducted to both the groups in the subjects of Math and English. Both groups were taught by subject specialist of Math and English. Control group was only taught and was not given any treatment whereas experimental group was taught and evaluated by teacher made test and reinforcement was given with a fixed interval schedule during session. After completion of 60 days teaching, a post test was conducted in Math and English.

5.2 FINDINGS

Following results have been obtained from the analysis of data.

The data of phase I reveal that:

1. Data shows that both the groups, control and experimental, have almost equal quantitative values in pre-test in the subject of Math. Mean of control group is 21.33 and mean of experimental group is 20.76, standard deviation of both groups is 4.51 and variance of control and experimental group is 20.36 and 20.39 respectively. Range of control group is 17 whereas range of experimental group is 16. Percentiles show that at 25th percentile, value of both groups is same, at 50th percentile, control group has 21.50 and experimental group has value of 21, at 75th percentile, value of control group is 25 and value of experimental group is 23. Data shows that students of both groups have almost equal learning outcome in the subject of Math.

2. Findings show that both the groups, control and experimental, have almost equal score in pre-test in the subject of English. Mean of control group is 24.70 and mean of experimental group is 24.36, standard deviation of control and
experimental group is 3.18 and 3.02, variance of control and experimental group is 10.14 and 9.13 respectively. Range of both groups is 12. Percentiles show that at 25th percentile, value of control group is 22.75 and experimental group is 22, at 50th percentile control group has 25 and experimental group has value of 24.50, at 75th percentile value of control group is 27.25 and value of experimental group is 17. Data shows that learning outcome of both groups in the subject of English is almost equal.

3. Data shows that both the groups, control and experimental, have prominent difference in post-test in the subject of Math. Mean of control group is 15.96 and mean of experimental group is 25.93, standard deviation of control group is 2.22 and experimental group is 2.74, variance of control and experimental group is 4.92 and 7.51 respectively. Range of both groups is 9. Percentiles show that at 25th percentile, value of control group is 14.75 which is lesser than 23, the value of experimental group, at 50th percentile control group has 16 and experimental group has value of 26, at 75th percentile value of control group is 17 which is lesser than 28, value of experimental group. Data shows a prominent difference in learning outcome of both groups.

4. Findings show the comparative Statistics of Pre-test and Post-test of Control Group in the subject of Math. The data shows prominent difference between pre-test and post-test. Mean value of pre-test is 21.33 and mean of post-test is 15.56, standard deviation of pre-test is 4.51 and post-test is 2.22, variance of pre-test is 20.36 and of post-test is 4.92. Range of pre-test is 17 and of post-test is 9. Percentiles show that at 25th percentile, value of pre-test is 16.75 and post-test is
14.75, at 50th percentile pre-test has 21.50 and post-test has value of 16, at 75th percentile value of pre-test is 25 which is greater than 17, value of post-test.

5. Data shows the comparative statistics of Pre-test and Post-test of Experimental Group in the subject of Math. Mean value of pre-test is 20.76 and mean of post-test is 25.93, standard deviation of pre-test is 4.51 and post-test is 2.74, variance of pre-test is 20.39 and of post-test is 7.51. Range of pre-test is 16 and of post-test is 9. Percentiles show that at 25th percentile value of pre-test is 16.75 and post-test is 23, at 50th percentile pre-test has 21 and post-test has value of 26, at 75th percentile value of pre-test is 23 which is lesser than 28, value of post-test. The data shows that a prominent difference exists between pre-test and post-test.

6. Data shows that both the groups, control and experimental, have significant difference in the post-test in the subject of English. Mean of control group is 14.83 and mean of experimental group is 26.86, standard deviation of control group is 3.22 and experimental group is 2.27, variance of control group is 10.41 and experimental group is 5.15. Range of control groups is 17 and experimental group is 7. Percentiles show that at 25th percentile, value of control group is 13 and experimental group is 25, at 50th percentile, control group has 14 and experimental group has value of 27, at 75th percentile, value of control group is 15.25 which is lesser than 29, value of experimental group.

7. Findings show the comparative statistics of pre-test and post-test of control group in the subject of English. The data shows prominent difference between pre-test and post-test. Mean value of pre-test is 24.70 and mean of post-test is 14.83, standard deviation of pre-test is 3.18 and post-test is 3.22, variance of pre-test is
10.14 and of post-test is 10.41. Range of pre-test is 12 and of post-test is 17.

Percentiles show that at 25<sup>th</sup> percentile, value of pre-test is 22.75 and post-test is 13, at 50<sup>th</sup> percentile pre-test has 25 and post-test has value of 14, at 75<sup>th</sup> percentile, value of pre-test is 27.25 which is greater than 15.25, value of post-test.

8. Data shows the comparative statistics of pre-test and post-test of experimental group in the subject of English. Mean value of pre-test is 24.36 and mean of post-test is 26.86, standard deviation of pre-test is 3.02 and post-test is 2.27, variance of pre-test is 9.13 and of post-test is 5.15. Range of pre-test is 12 and of post-test is 7. Percentiles show that at 25<sup>th</sup> percentile value of pre-test is 22 and post-test is 25, at 50<sup>th</sup> percentile pre-test has 24.50 and post-test has value of 27, at 75<sup>th</sup> percentile value of pre-test is 27 which is lesser than 29, value of post-test.

9. The findings of table 4.17 show that correlation between pre-test and post-test of control group in the subject of Math is 0.156, which is not significant at \( \alpha=0.05 \) or \( \alpha=0.01 \). It shows that score of pre-test have not significant relationship with the score of post-test in the subject of Math.

10. The findings of table 4.18 show that correlation between pre-test and post-test of experimental group in the subject of Math is 0.364, which is significant at \( \alpha=0.05 \). It shows that score of pre-test have significant relationship with the score of post-test in the subject of Math.

11. The findings of table 4.19 show that correlation between pre-test and post-test of control group in the subject of English is 0.159, which is not significant at
α=0.05 or α=0.01. It shows that score of pre-test have not significant relationship with the score of post-test in the subject of English.

12. The findings of table 4.20 show that correlation between pre-test and post-test of experimental group in the subject of English is 0.445, which is significant at α=0.05. It shows that score of pre-test have significant relationship with the score of post-test in the subject of English.

13. The data shows that the mean differences between pre-test control group and pre-test experimental group in the subject of Math is significant, t=1.172.

The data reveals that both the groups posses almost equal learning outcome in the subject of Math.

14. The findings show that the mean differences between pre-test control group and pre-test experimental group in the subject of English is significant, t=1.541.

The data reveals that both the groups posses almost equal learning outcome in the subject of English.

15. The data shows that the mean differences between post-test control group and post-test experimental group in the subject of Math is not significant, t=-1.506.

The data shows that learning outcome of experimental group is better than learning outcome of control group.

16. The findings show that the mean differences between post-test control group and post-test experimental group in the subject of English is not significant, t=-1.60.

The data shows that learning out come of experimental group is better than learning outcome of control group.
The data of phase II reveals that:

17. Table data shows that both the groups, control and experimental, have almost equal learning outcome in pre-test in the subject of Math. Mean of control group is 21.16 and mean of experimental group is 21.10, standard deviation of control groups is 4.16 and experimental group is 4.80, variance of control and experimental group is 17.38 and 23.05 respectively. Range of control group is 18 whereas range of experimental group is 16. Percentiles show that at 25\textsuperscript{th} percentile, value of control groups is 18.75 and experimental group is 16.75, at 50\textsuperscript{th} percentile control group has 21.50 and experimental group has value of 21, at 75\textsuperscript{th} percentile value of control group is 24 and value of experimental group is 23.75.

18. Findings show that both the groups, control and experimental, have almost equal quantitative values in pre-test in the subject of English. Mean of control group is 22.46 and mean of experimental group is 23.36, standard deviation of control and experimental group is 3.29 and 3.51, variance of control and experimental group is 10.87 and 12.34 respectively. Range of both groups is 13. Percentiles show that at 25\textsuperscript{th} percentile, value of control group is 24 and experimental group is 23.75, at 50\textsuperscript{th} percentile, control group has 22 and experimental group has value of 23, at 75\textsuperscript{th} percentile, value of control group is 25 and value of experimental group is 26.25.

19. Table 4.33 shows that both the groups, control and experimental, have prominent difference in post-test in the subject of Math. Mean of control group is 15.96 and mean of experimental group is 26.03, standard deviation of control group is 3.31
and experimental group is 2.57, variance of control and experimental group is 10.99 and 6.65 respectively. Range of control groups is 13 and experimental group is 9. Percentiles show that at 25th percentile, value of control group is 14 and experimental group is 24, at 50th percentile, control group has 15 and experimental group has value of 26.50, at 75th percentile, value of control group is 18.25 and value of experimental group is 28.

20. Table 4.34 shows the comparative statistics of pre-test and post-test of control group in the subject of Math. The data shows prominent difference between pre-test and post-test. Mean value of pre-test is 21.16 and mean of post-test is 15.96, standard deviation of pre-test is 4.16 and post-test is 3.31, variance of pre-test is 17.38 and of post-test is 10.99. Range of pre-test is 18 and of post-test is 13. Percentiles show that at 25th percentile, value of pre-test is 18.75 and post-test is 14, at 50th percentile pre-test has 21.50 and post-test has value of 15, at 75th percentile value of pre-test is 24 which is greater than 18.25, value of post-test.

21. Table 4.35 shows the comparative statistics of pre-test and post-test of experimental group in the subject of Math. The data shows prominent difference between pre-test and post-test. Mean value of pre-test is 21.10 and mean of post-test is 26.03, standard deviation of pre-test is 4.80 and post-test is 2.57, variance of pre-test is 23.05 and of post-test is 6.65. Range of pre-test is 16 and of post-test is 9. Percentiles show that at 25th percentile, value of pre-test is 16.75 and post-test is 24, at 50th percentile, pre-test has 21 and post-test has value of 26.50, at 75th percentile, value of pre-test is 23.75 which is lesser than 28, value of post-test.
22. Table 4.38 shows that both the groups, control and experimental, have significant difference in the post-test in the subject of English. Mean of control group is 14.83 and mean of experimental group is 25.66, standard deviation of control group is 3.10 and experimental group is 2.83, variance of control group is 9.66 and experimental group is 8.02. Range of control groups is 17 and experimental group is 10. Percentiles show that at 25\textsuperscript{th} percentile, value of control group is 13 and experimental group is 23, at 50\textsuperscript{th} percentile, control group has 14.50 and experimental group has value of 25.50, at 75\textsuperscript{th} percentile, value of control group is 16 which is lesser than 28.25, value of experimental group.

23. Table 4.39 shows the comparative statistics of pre-test and post-test of control group in the subject of English. The data shows prominent difference between pre-test and post-test. Mean value of pre-test is 22.46 and mean of post-test is 14.83, standard deviation of pre-test is 3.29 and post-test is 3.10, variance of pre-test is 10.87 and of post-test is 9.66. Range of pre-test is 13 and of post-test is 17. Percentiles show that at 25\textsuperscript{th} percentile, value of pre-test is 20 and post-test is 13, at 50\textsuperscript{th} percentile, pre-test has 22 and post-test has value of 14.50, at 75\textsuperscript{th} percentile, value of pre-test is 25 which is greater than 16, value of post-test.

24. Table 4.40 shows the comparative statistics of pre-test and post-test of experimental group in the subject of English. The data shows prominent difference between pre-test and post-test. Mean value of pre-test is 23.26 and mean of post-test is 25.66, standard deviation of pre-test is 3.51 and post-test is 2.83, variance of pre-test is 12.34 and of post-test is 8.02. Range of pre-test is 13 and of post-test is 10. Percentiles show that at 25\textsuperscript{th} percentile, value of pre-test is
21 and post-test is 23, at 50\textsuperscript{th} percentile, pre-test has 23 and post-test has value of 25.50, at 75\textsuperscript{th} percentile, value of pre-test is 26.25 which is lesser than 28.25, value of post-test.

25. The findings of table 4.41 show that correlation between pre-test and post-test of control group in the subject of Math is 0.050, which is not significant at $\alpha=0.05$ or $\alpha=0.01$. It shows that score of pre-test have not significant relationship with the score of post-test in the subject of Math.

26. The findings of table 4.42 show that correlation between pre-test and post-test of experimental group in the subject of Math is 0.356, which is significant at $\alpha=0.05$.

\textbf{It shows that score of pre-test have significant relationship with the score of post-test in the subject of Math}

27. The findings of table 4.43 show that correlation between pre-test and post-test of control group in the subject of English is -0.342, which is not significant at $\alpha=0.05$ or $\alpha=0.01$. It shows that score of pre-test have not significant relationship with the score of post-test in the subject of English.

28. The findings of table 4.44 show that correlation between pre-test and post-test of experimental group in the subject of English is 0.381, which is significant at $\alpha=0.05$. \textbf{It shows that score of pre-test have significant relationship with the score of post-test in the subject of English.}

29. The table 4.45 shows that the mean differences between pre-test control group and pre-test experimental group in the subject of Math is significant, $t=1.164$. The data reveals that both the groups posses almost equal learning outcome in the subject of Math.
30. The table 4.46 shows that the mean differences between pre-test control group and pre-test experimental group in the subject of English is significant, $t=1.167$. The data reveals that both the groups possess almost equal learning outcome in the subject of English.

31. The table 4.47 shows that the mean differences between post-test control group and post-test experimental group in the subject of Math is not significant, $t=-12.46$. The data shows that learning outcome of experimental group is better than learning outcome of control group.

32. The table 4.48 shows that the mean differences between post-test control group and post-test experimental group in the subject of English is not significant, $t=-13.06$. The data shows that learning outcome of experimental group is better than learning outcome of control group.

5.3 DISCUSSION

This study investigated the impact of formative evaluation and fixed interval schedule reinforcement on academic achievements of secondary school students. Students in experimental groups improved their performance significantly from pre-test to final post-test. On the posttest measure, students in the experimental groups scored significantly higher than students in the control groups.

The results for the formative evaluation are moderately reliable with a number of the potential benefits cited by proponents of these procedures and reported during past. These impending benefits incorporated enhanced student performance, a more active student...
role in their own learning, a greater awareness of the evaluation process and scoring
criteria, and an increased understanding of the instructional content (Brindley &
Scchofield, 1998; Smith et al., 2002; Davies, 2000, 2002; Fallows & Chandramohan,
2001). While students accomplished the greatest development in their academic
attainments under formative evaluation by their teacher, subjects played a more active
role in their own learning and also made important improvements in their plans under
both evaluation conditions. Additionally, all four teachers reported that formative
evaluation enabled students to increase a better understanding of the content and
concepts.

This study explored out impact of formative evaluation and fixed interval schedule
reinforcement on academic achievements of secondary school students. Students in all
two treatment groups improved the quality of their learning significantly as compared to
control groups. The amount of improvement differs from phase I to phase II, which
indicates impact of fixed interval schedule reinforcement. Posttest scores were modestly,
but significantly higher than the pretest scores across the two groups.

An earlier study by Ozogul et al. (2007) explored important developments for the similar
form of evaluation that were used in this study, but considerably better improvements for
student evaluation than for the two types of teacher evaluation. Conversely, the teachers
in the previous study reported that it was so protracted for them to carry out both
formative evaluation and fixed interval schedule reinforcement that they would be very
unlikely to use these forms together on a regular basis in the future. The result was
encouraging in this study in that there was only a fractional difference (16.80–16.61) in
posttest scores between the control and experimental group in phase I and the difference (16.80–15.86) between the control and experimental groups in phase II. This result suggests that the formative evaluation and fixed interval schedule reinforcement was effective in improving learning outcome of secondary school students.

5.4 CONCLUSIONS

This study indicated that, teachers received appropriate training and practice in conducting formative evaluation; they perform at a similar level under formative-evaluation. Not surprisingly, the students still have more positive attitudes toward teacher-based formative evaluation because of the teachers’ greater knowledge of the subject matter and, therefore, greater credibility.

Following conclusions can be drawn from the findings of the study:

1. Formative evaluation has significant relationship with learning and academic achievements in the subjects of Math.
2. Formative evaluation has significant relationship with learning and academic achievements in the subjects of English.
3. Formative evaluation has significant relationship with learning and academic achievements of secondary school students.
4. Formative evaluation has stronger impact on learning and academic achievements in the subjects of Math.
5. Formative evaluation has stronger impact on learning and academic achievements in the subjects of English.
6. Formative evaluation has stronger impact on learning and academic achievements of secondary school students.

7. Fixed interval schedule reinforcement has significant relationship with learning and academic achievements in the subjects of Math.

8. Fixed interval schedule reinforcement has significant relationship with learning and academic achievements in the subjects of English.

9. Fixed interval schedule reinforcement has significant relationship with learning and academic achievements of secondary school students.

10. Fixed interval schedule reinforcement has stronger impact on learning and academic achievements in the subjects of Math.

11. Fixed interval schedule reinforcement has stronger impact on learning and academic achievements in the subjects of English.

12. Fixed interval schedule reinforcement has stronger impact on learning and academic achievements of secondary school students.

5.5 RECOMMENDATIONS

Following recommendations are made on the basis of results obtained from the study:

1. Evaluation is back bone of teaching learning process. In classroom environment evaluation especially formative evaluation is very helpful in strengthening the teaching learning process. Educational institutional management may arrange capacity building programs for teachers to familiarize them with evaluation, its importance and its process.
2. The managers of educational institutions may monitor teaching learning process with a view to implementing formative evaluations policy.

3. Curriculum designers and developers may design and develop formative evaluation tools within curriculum document for each subject.

4. Educational Boards may give weightage to formative evaluation reports, prepared by teachers during the academic year, while assessing the students at their respective level.

5. Reinforcement plays a pivotal role in behavior modification and in pushing some one towards desired destination. Fixed interval schedule reinforcement may enable a student to achieve his target in a better way. Teachers may be given training to reinforce students in teaching learning process by fixed interval schedule reinforcement techniques.

5.6 FUTURE RESEARCH

Future research on these types of formative evaluation could take several possible directions. This study was conducted with a multiple task (formative evaluation and fixed interval schedule reinforcement) and a specific class level (secondary school level). It would be appropriate to extend the research to other types of tasks to investigate the generality of the present findings and to lower grade levels in an effort to identify an approximate level at which students can begin to use such procedures effectively. Also, the teacher training in this study was task specific but not type specific. Training that is more specific to the type of evaluation may yield somewhat better results. An additional possibility would be to reduce the amount of formative teacher feedback by making it
more general to the class, rather than specific to each individual, based on a more cursory
review of the student work. The teacher feedback could then be combined with a form of
student evaluation to determine the effects of formative teacher and student evaluation in
combination. Clearly, formative evaluation has the potential to improve the quality of
student work. Future research in this area should help us to identify the most feasible and
effective procedures for conducting this type of evaluation.
Bibliography


