ignou
THE PEOPLE'S
UNIVERSITY

Indira Gandhi
National Open University
School of Education

BES-127
Assessment for
Learning

Block

2

TECHNIQUES AND TOOLS OF ASSESSMENT AND EVALUATION

EXPERT COMMITTEE

Prof. I. K. Bansal (Chairperson)
Former Head, Department of Elementary
Education, NCERT, New Delhi

Prof. Shridhar Vashistha
Former Vice-Chancellor
Lal Bahadur Shastri Sanskrit
Vidhyapeeth, New Delhi

Prof. Parvin Sinclair
Former Director, NCERT
School of Sciences,
IGNOU, New Delhi

Prof. Aejaz Masih
Faculty of Education,
Jamia Millia Islamia, New Delhi

Prof. Pratyusha Kumar Mandal
DESSH, NCERT, New Delhi

Prof. Anju Sehgal Gupta
School of Humanities,
IGNOU, New Delhi

Prof. M. C. Sharma (Retd.)
B.Ed. Prog. Coordinator till 31.3.17
School of Education
IGNOU, New Delhi

Prof. N. K. Dash
School of Education
IGNOU, New Delhi

Dr. Gaurav Singh
(Programme Co-coordinator-B.Ed.)
School of Education
IGNOU, New Delhi

SPECIAL INVITEES (Faculty of School of Education)

Prof. D. Venkateshwarlu
Prof. Amitav Mishra
Ms. Poonam Bhushan
Dr. Eisha Kannadi
Dr. M. V. Lakshmi Reddy

Dr. Bharti Dogra
Dr. Vandana Singh
Dr. Niradhar Dey
Dr. Elizabeth Kuruvilla

**Programme Coordinators: Prof Saroj Pandey, (From April, 2017) and
Dr. Gaurav Singh, SOE, IGNOU, New Delhi**

Course Coordinator: Dr. Niradhar Dey, SOE, IGNOU, New Delhi

COURSE PREPARATION TEAM

Course Writers

Unit 5   Dr. Elizabeth Kuruvilla
         SOE, IGNOU, New Delhi

Unit 6   Dr. Niradhar Dey
         SOE, IGNOU, New Delhi

Unit 7   Dr. Gaurav Rao
         Dept of Edu., CSJM University,
         Kanpur

Unit 8   Dr. Ajith Ku. C.
         SOE, IGNOU, New Delhi

Content Editing
Prof. SVS Chaudhary
SOE, IGNOU, New Delhi

Language Editing
Dr. Y. Nirmala
SOE, IGNOU, New Delhi

Proof Reading
Dr. Niradhar Dey
SOE, IGNOU, New Delhi

PRINT PRODUCTION

Prof. Saroj Pandey
Director
SOE, IGNOU, New Delhi

Mr. S.S. Venkatachalam
Assistant Registrar (Publication)
SOE, IGNOU, New Delhi

October, 2017

© *Indira Gandhi National Open University, 2017*

ISBN: 978-81-266-

Further information on the Indira Gandhi National Open University courses may be obtained from the University's Office at Maidan Garhi, New Delhi-110068.

Printed and published on behalf of the Indira Gandhi National Open University, New Delhi, by the Director, School of Education, IGNOU, New Delhi.

Laser Typeset : Rajshree Computers, V-166A, Bhagwati Vihar, Uttam Ngr. (Near Sec.-2, Dwarka), N.D.59

Printed at :

BES 127	ASSESSMENT FOR LEARNING
BLOCK 1	**Evaluation in Teaching-Learning Process**
Unit 1	Concept and Purpose of Evaluation
Unit 2	Perspectives of Assessment
Unit 3	Approaches to Evaluation
Unit 4	Issues, Concerns and Trends in Assessment and Evaluation
Block 2	**Techniques and Tools of Assessment and Evaluation**
Unit 5	**Techniques of Assessment and Evaluation**
Unit 6	**Criteria of a Good Tool**
Unit 7	**Tools for Assessment and Evaluation**
Unit 8	**ICT Based Assessment and Evaluation**
Block 3	**Learner's Evaluation**
Unit 9	Teacher Made Achievement Tests
Unit 10	Commonly Used Tests in Schools
Unit 11	Identification of Learning Gaps and Corrective Measures
Unit 12	Continuous and Comprehensive Evaluation
Block 4	**Analysing and Interpreting Learner's Performance**
Unit 13	Tabulation and Graphical Representation of Data
Unit 14	Measures of Central Tendency
Unit 15	Measures of Dispersion
Unit 16	Correlation: Importance and Interpretation
Unit 17	Nature of Distribution and Its Interpretation

BLOCK 2   TECHNIQUES AND TOOLS OF ASSESSMENT AND EVALUATION

Introduction to the Block

This Block consists of four units dealing with various techniques and tools of assessment and evaluation used in the process of teaching-learning. The Block starts with the Unit, '**Techniques of Assessment and Evaluation**' which deals with the conceptual discussions of understanding various techniques used in evaluation. The major techniques of evaluation such as self-report techniques, assignments, observations, peer-assessment, socio-metric techniques, portfolio, seminar, and cooperative learning and social skills are briefly discussed in this Unit.

Unit-6 of this Block, titled, '**Criteria of a Good Tool**' elaborately discusses the essential features of an effective tool. More specifically, it focuses on the criteria of a test namely reliability, validity, usability, objectivity, and norm. The Unit helps the learners to develop conceptual understanding on the above criteria and also equally empower them to determine reliability, validity and other qualities of a tool.

Unit-7 of this Block, entitled, '**Tools for Assessment and Evaluation**' which discusses the tests and tools used in the process of teaching and learning for assessment. The Unit starts with a discussion on various types of tests such as paper-pencil test, performance test, aptitude test, achievement test, diagnostic-remedial test and also intelligence test. The Unit also critically discusses the tools used in assessment such as rating scales, questionnaire, inventories, checklists, interview and observation schedule. The Unit ends with a discussion on alternatives to traditional evaluation such as anecdotal records, learners portfolio and rubrics.

The Eights Unit, '**ICT Based Assessment and Evaluation**' deals mostly with the use of technology in the process of evaluation. It critically discusses the importance of ICT in the process of assessment and it also discusses the role of teachers in technology-based evaluation. The Unit also broadly discusses the online examination system, preparing online tools for examinations, use of e-portfolios and e-rubrics for assessment. The Unit gives a complete out look to use the free online sources for preparing tests, administer them online and also to analyse the results.

UNIT 5   TECHNIQUES OF ASSESSMENT AND EVALUATION

Structure

5.1   INTRODUCTION

In Unit 1, you have studied about the concepts of measurement, assessment and evaluation and by now, you might be able to differentiate these terms from one another. Assessment is an integral part and ongoing process in the educational context. To learn effectively, students need to know as to how they are performing. Similarly, to be an effective teacher, you must be informed about what the student knows, feels, and can do, so that you can help him/her build on her/his skills, knowledge, and attitudes. Therefore, you and your students need continuous feedback on their progress and difficulties in order to plan appropriate learning activities and to make adjustments to those already planned. Assessment means the way by which teachers and other stakeholders involved in students' learning collect information systematically and then use it to know their level of achievement in academic, behavioural and social areas. In this Unit, you will study various techniques of evaluation.

5.2   OBJECTIVES

After going through this Unit you will be able to:

- describe the techniques of assessment and evaluation,

- explain the significance of self-report techniques,

- analyze students' characteristics through observation,

- use interview as a tool of evaluation,

- recognize the use of assignment as an assessment technique,

- identify the importance of peer ratings,

- conduct socio-metric technique in classroom,

- evaluate project work by using a rubric,

- assess portfolios for understanding students' progress, and

- evaluate the techniques used for group processes.

5.3   PURPOSE OF ASSESSMENT AND EVALUATION

Let us start with the discussion on the purpose of evaluation. Generally evaluation is used for the following purposes:

- **To assist in student learning:** A good evaluation provides information to teachers about the progress of students' learning. For students also, it draws a road map to understand their learning progress.

- **To identify students' strengths and weaknesses:** Identification and assessment of students' strengths and weaknesses are necessary for two reasons: i) to structure and restructure the learning activities and ii) to restructure the curriculum.

- **To assess the effectiveness of a particular instructional strategy:** It is important for a teacher to know as to how well a particular teaching strategy helps to achieve the learning objectives.

- **To assess and improve the effectiveness of curriculum programmes:** Components of the curriculum are continually assessed through formative and summative assessments.

- **To assess and improve teaching effectiveness:** To enhance student learning, teachers are continuously evaluated on the basis of a) their commitment to the job, b) their ability to cope with students of a particular age, c) their ability to show mastery of appropriate instructional techniques.

- **To communicate with and involve parents and guardians in their childrens' learning:** Parents, guardians and community need to share with accountability for the effectiveness of learning of the students. Parents or guardians are eager to know the progress of their children in school. Evaluation feedback is necessary to communicate to parents, guardians and the community. For example, distributing progress card of students by conducting a parent-teacher meeting.

Check Your Progress 1

Note : a) Write your answer in the space given below.

         b) Compare your answer with those given at the end of the unit.

1. Why do teachers use assessment?

    ....

    ....

    ....

5.4 TECHNIQUES OF ASSESSMENT AND EVALUATION

Assessing students' performance is an integral part of the teaching-learning process. For this, you need to adopt various techniques of evaluation. Choice of evaluation techniques should be aligned with the learning objectives. Students learn information at different pace. You, as a teacher, have to find out the extent of students' learning and the interventions required for facilitating rather than to master the desired content.

When you assess your students, you collect information about their level of performance where as in evaluation, you compare a student's achievement with other students with a set of standards. Effective assessment is a continuous process and it is not simply that has to be done at the end of a Unit. Evaluation is integrated into all aspects of the curriculum, thus providing both students and teachers useful and relevant data to gauge progress of students. Not only teachers but also students play an important role in assessing their own learning progress. While assessing students, you need to keep the following points in mind:

● Which technique should I use to get the adequate feedback?

● How will I apply this technique in my class?

● Will the assessment technique provides adequate information about what students are learning in classrooms?

● Does the technique focus on various variables that can be changed to promote better learning?

● Are the results derived from the use of a particular evaluation tool easy to analyze?

● How will I know that the technique is useful or not to assess students?

● What effort should I make to arrange better teaching-learning practices based on the information I received?

In the subsequent section, we will discuss about the various techniques of evaluation. The various techniques of evaluation are summarized in the Figure 5.1.
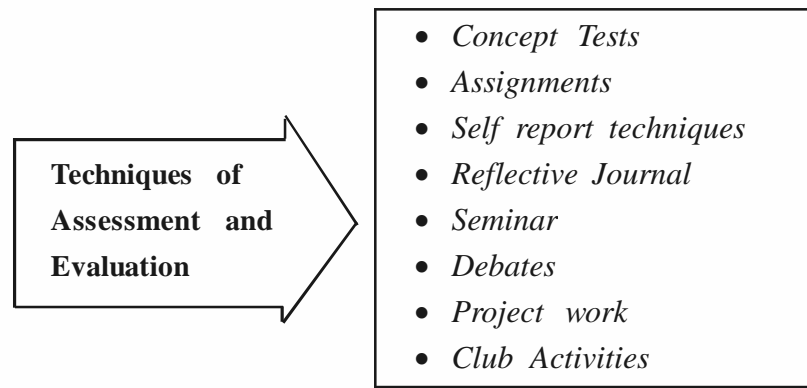
Techniques of
Assessment and
Evaluation →

- *Concept Tests*
- *Assignments*
- *Self report techniques*
- *Reflective Journal*
- *Seminar*
- *Debates*
- *Project work*
- *Club Activities*

Figure 5.1: Techniques of Assessment and Evaluation

5.4.1 Concept Tests

Concept tests are informal and short tests are conducted by teachers to evaluate students' understanding of key concepts. These tests can be used to assess students' previous knowledge about a concept or after the instruction. In this technique, class time is divided between short lectures and conceptual multiple-choice questions. In the concept test, the understanding of students about the basic concepts are evaluated. These questions are framed to assess students' understanding of the principal concepts in the content. Next, let us discuss how to design a concept test.

Frame a question that describes an event, problem or situation. Read the following example:

Example: This problem is designed to help students understand the functioning of the circulatory system in the human body. Suppose you are staying on the second floor of the building and due to low pressure of the water, the water reaching the taps in your floor is not sufficient. Then you are advised to install water pump at your house. Predict:

- *What will happen?*

- *What if there is any blockage in the water pipes supplying water to your house taps?*

- *Which concept is relevant here?*

(Source: Dogra, B., 2010)

In the above example, the teacher presented a situation and then asks some questions related to prediction, problem-solving and clarification of concept. Such kind of test helps you to know how much your students attain understanding of concepts.

5.4.2 Self-report Techniques

Self-report techniques are data gathering tools where respondents provide information about themselves. There are various self-reporting techniques such as interview, diaries, questionnaire, reflective journal, etc. These techniques are widely used to assess various traits or attitudes possessed by the individuals. Though these are subjective techniques, they provide means for extracting the hidden treasure of one's own behaviours and patterns of personality. Let us discuss in detail about the various self-reporting techniques such as interview, focus group discussion, questionnaire and reflective journal.

Interview: An interview is a data collection method through which an interviewer asks questions to an interviewee on a particular topic. It can be semi-structured

or structured and may use different media (e.g., phone, e-mail, and in person). In semi-structured interview, the interviewer has an idea about the types of questions but the order and way of asking questions may vary. A structured interview on the other hand, is a specified set of questions that are asked in the same order during each interview. Interviewing is considered as an interpersonal encounter. Establishing rapport is an important element in interviewing. Keeping impartiality to the responses given by the interviewee is important; otherwise it leads to bias responses.

During interviewing, you may know whether the student understands a particular concept. The student may be asked to give a personal explanation of a specific concept and to use that concept to solve a problem. Through this technique, besides getting information, it is also possible to observe a student's body language and facial expressions. For assessment purpose, interview may serve the following functions:

- to identify the areas of difficulty and gaps in understanding a concept among students,

- to understand how students are able to apply their knowledge in new situations,

- to examine how students' understanding change with teaching-learning process,

- to obtain verbal feedback about teaching techniques, new concepts, etc.

Through the feedback received from students, you may be able to refine your teaching along with assessing the understanding level of students.

Focus group discussion: It is a systematically planned discussion led by the teacher who acts as a moderator to collect detailed information on a particular topic. In focus group discussion, a moderator leads the discussion with a small group of individuals to analyze in detail the thinking and feeling of the group members about a topic. It is called 'focus group' because the moderator keeps the individuals in the group focused on a particular topic taken for discussion. The moderator generates group discussion through the use of open-ended questions, and he/she acts as a facilitator of the group process. It is used for multiple purposes as follows:

- Generating new ideas and concepts,

- Obtaining general background information about a topic of interest,

- Generating research hypotheses that can be submitted to further research,

- Diagnosing the problem in a new way,

- Examining how respondents talk about the selected topic, and

- Interpreting previously obtained results.

You might be thinking how an interview differs from a focus group discussion. In interview, the interviewer has a prominent role in asking specific questions. The depth of information received from an interview is less. In the case of focus group discussion, the number of participants varies from four to eight members and you will get a great deal of information.

Questionnaire: A questionnaire is a self-report data collection technique that each participant fills out in the given form. There are open-ended and closed-ended questionnaires. Through questionnaires, one can obtain information about the thoughts, feelings, attitudes, beliefs, values, perceptions, personality, and

behavioural intentions of participants. Open-ended questions are valuable when the researcher needs to know what people are thinking. Such questions help the interviewer to understand participants' inner worlds in their natural languages and categories.

When do we use questionnaires?

Questionnaires will be used for the following purposes:

*   To collect feedback from a large number of students,

*   To know students' opinions, attitudes, feelings and perception about a particular issue, and

*   To allow each student to give anonymous feedback.

Reflective Journal: Reflective journal is a tool where students can reflect on and write about their progress in learning. They can write their achievements and difficulties in learning. Students can write whatever they found to be innovative or what made them think and ponder about their learning experiences. Journal writing is a non-threatening way to get information about students' thinking and feeling. The following points give you an idea about how to use and what to be included in the journal.

*   **When should I use it?**
    *   Daily or weekly

*   **How could I use it?**
    *   As a self- assessment tool for students
    *   At the beginning or end of a unit, topic or project
    *   As a communication tool for students and teachers

*   **Why should I use it?**
    *   It enhances higher order thinking
    *   It leads to self-assessment
    *   It sets goals for future learning
    *   It gives students a sense of ownership and control on their own learning

Students' reflective thinking can be supported by giving opportunities to reflect on their own learning and their own context of learning. Thereby, it promotes students' responsibility of learning. Reflective journal permits students to practice their writing skill in a free environment that encourages critical thinking. In this way, reflective journals can be used as an assessment tool that provides teachers an insight into how students value their own learning and achievement.

Check Your Progress 2

Note :   a)   Write your answer in the space given below.

b)   Compare your answer with those given at the end of the unit.

2. Explain how self reporting techniques can be used as assessment tools.

....
....
....

5.4.3 Assignments

Assignments are teacher-assigned task that students are expected to complete during or after the school hours and enable to assess students' understanding. It is essential to keep in mind that class-appropriate assignments may be chosen for your students. For first-or second class students, hands-on-activities are more preferable than paper-and-pencil assignments. The time required for assignments should be set based on the age of the children. If you want to give a comprehensive project as an assignment, consider breaking it into manageable sections with due dates. The types of assignments you give should motivate the students. Through this technique, you can find out students' perspective, their interest and learning levels. The students' progress in learning, their understanding and critical thinking can also be assessed through assignments.

5.4.4 Observation Technique

As prospective teachers, you might have observed students while they solve problems, interact with peers in different learning situations or in the playground. It provides insight into student learning and growth. Observation is used as a technique to evaluate various aspects of behaviour in controlled or uncontrolled situations. Through observation, behaviour is captured in a particular situation. It is a means of first hand information as experienced at a specific moment. It is pre-planned and purposeful activity that provides immediate recording of events. With the help of observation checklists, teachers could record information quickly. The reliability of observation can be increased by repeated observation or through observation done by many individuals. While preparing observation checklist, the following points may be kept in mind:

● Write down the criteria to be observed

● Inform students about the criteria to be observed

● Determine the specific outcomes to assess

● Develop a data gathering system such as checklist or rubric or anecdotal points

● Target observation on one or two specific outcomes

● Record the date of observation

● Share observation details with individuals or target groups

● Use the collected information to modify your instruction.

Observation is an important technique of collecting information about people because people do not always do what they say. This statement is equivalent to the maxim in behavioural and social sciences that 'attitudes and behaviour are incongruent'. Generally there are two types of observation: i) Participant observation, and ii) Non-participant observation.

i. **Participant observation:** In participant observation, the observer becomes a member of the group. The participant observer plays a dual role such as becoming a member of the group and observing the participants carefully. This kind of observation provides reliable results. An advantage of this technique is that, for ethical reasons, the researcher can request permission to collect and record data as needed. In addition, the researcher can obtain

feedback about his/her observations and tentative conclusions from the participants. A weakness is that the participants might not behave naturally because they are aware that they are being observed.

ii. **Non-participant observation:** In non-participant observation, the observer observes the group passively from a distance without his/her presence in the group. Non-participant observation helps in recording and studying the behaviour of a particular person or group in detail. It is considered as an unobtrusive method of data collection to study focused aspects of a setting in order to answer specific questions within a study.

5.4.5 Peer Assessment

It refers to one student or a group of students assessing other students. It develops interpersonal skills and may help the student to develop an impartial attitude, improve listening skills, inculcate team spirit, leadership quality and time management. Students internalize the quality of the work by evaluating the work of their peers. Peer assessment has its roots in theories of active learning. In active learning, students involve in doing things and think about the things they are doing. The learning environment in the classroom should be supportive for effective peer evaluation. There must be a comfortable and trust atmosphere to promote constructive feedback. The following table shows the benefits of peer assessment.

Benefits of peer-assessment

● An educational process that installs autonomy in learners.

● Empowerment of the learner in a learning environment.

● Development of learner confidence in assessing/ marking peers (through practice).

● Development of learner ability to self-evaluate and reflect.

● Greater understanding of what is required by tutors for assessment at various levels.

● Interactive classes for marking/feedback.

● Reflection on recently completed assessments with full explanation of the answer (improving information and understanding).

● Clear, open marking systems (seeing what is required and improving work).

● Seeing standards set by peers as well as mistakes of others (and avoiding them in the future).

● Gaining an ability to 'stand back' from own work for assessment purposes (an essential ability of an 'objective', 'unbiased' scientist).

● Rapid way for a tutor to assess a large amount of student work and provide specific feedback.

(Source: Langan, Mark A. & Wheater, P., 2003)

Check Your Progress 3

Note : a) Write your answer in the space given below.

b) Compare your answer with those given at the end of the unit.

3. List out the benefits of peer-assessment.

....

....

....

....

5.4.6 Sociometric Technique

The sociometric technique gives an objective picture of social relationships in a classroom. Jacob Levy Moreno developed this technique to study the relationship between social structures and psychological well-being. The words 'sociometry' is originated from Latin word 'socius' meaning companion and 'metrum' meaning measure. The main purpose of a sociogram is to find out the basic network of friendship patterns. Another type of information that can be obtained from a sociogram is to know the relation of any one child to the group. Through sociogram, you will get greater understanding of group behaviour.

We can construct sociogram in a variety of ways. Now, let us learn the procedure of constructing a sociogram. The following questions may be asked to the group members to collect information:

- Who are your three best friends in your class?

- Which three friends in the class would you like to work with in a small group in your class?

- Which three friends would you like to sit with at lunch time?

- Which three friends would you like to go with for a picnic?

The above questions are examples of a 'fixed positive nomination technique'. The term fixed is used in the sense that in the questions, the choices are restricted to three friends. Through these questions, it is possible to find interpersonal acceptance among students. In order to find interpersonal rejection, the following questions may be asked:

- Which three friends do you like the least?

- Which three friends you don't like to include in your group work?

- Which three friends you don't like to go with for a picnic?

The study of rejection helps a teacher to know the unfavourable emotional reaction of students. The graphical representation of relationships between a specific group is known as sociogram. It is a chart to know the interrelationships within the group. An example of a sociogram is depicted below:
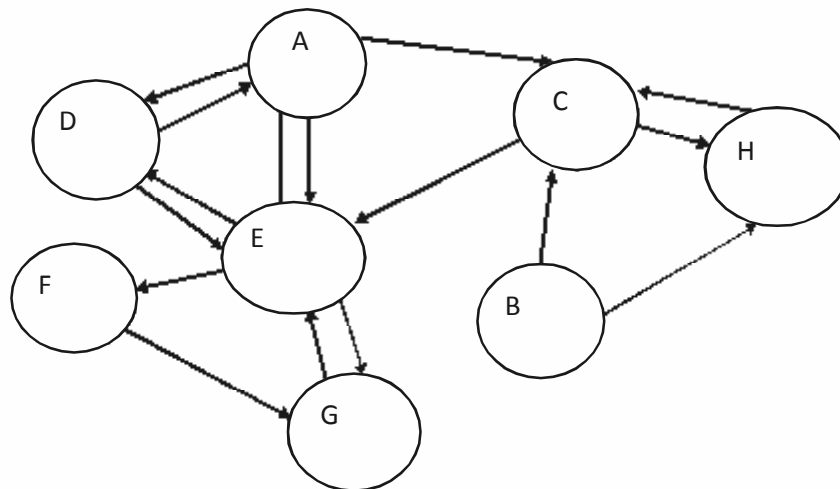
Figure 5.2: Sociogram Showing Pattern of Choices

In the figure 5.2, you notice that student 'B' is not selected by others. Hence, 'B' is considered as an isolate. Students 'A', 'D' and 'E' choose each other and thus form a clique. Student 'E' is liked by most, thus becoming a star. Sociogram also provides details about the relationship among the members of the group. The sociogram, thus helps the teacher to build a supportive learning environment. Between students 'C' and 'H', reciprocity is both sided while for 'E' and 'F' it is one-sided only. Sociometry is thus an important technique for peer ratings.

Sociometric technique as an assessment technique: You might be thinking how sociometric technique can be used for assessment purposes. As an assessment technique, sociometry is useful to assess social behaviour, relatedness and interpersonal trust that exists within the group. You can use the results of sociometry for studying social patterns in the group such as acceptance, status, cohesiveness and different roles. Through sociogram, you can locate sub-groups and mutual attraction patterns of your classroom. It also helps you to know interpersonal acceptance-rejection among students. Through sociogram, you will be aware of the impact of friendships within your classroom, and how it will affect students' academic motivation and achievement. While going through the interpersonal rejection, problems related with academic and socio-emotional adjustment could be found out. Moreover, peer rejection can be treated as a predictor of academic dysfunction.

5.4.7 Portfolios

The details of portfolio have been discussed in Unit-7 of this Block. A portfolio is a systematically organized collection of a student's work that covers a specific period of time. It proved to be an excellent vehicle for gathering together student writing and displaying the changes in the proficiency over a period of time. Portfolio assessment represents a blend of the instructional process and the evaluation process. You might know that portfolio is not something static but it is an ongoing dialogue between teacher and student and, often, between the individual student and others in the class or between the students and parents. Components of portfolios will vary depending on the subject that is being taught and preferences of teachers and students. Portfolios can contain the following items:

- Completed assignments

- Journal writings

- Reflections on discussions that have been held in or out of class

- Photos, sketches, and other visuals

- Summary statements made at different points regarding what has been learned

- Self-assessment statements

Use of portfolios is consistent with reflective teaching. There is an ongoing dialogue between the teacher and student that takes the form of written comments about the materials in the portfolios. Portfolio assessment focuses evaluation of students' progress, processes, and performance over a period of time. Generally there are two basic types of portfolios:

- Process portfolio is used for classroom-level assessment. It reflects both formative and summative assessments.

- A product portfolio is more summative in nature. It is meant for a major evaluation and is accompanied by an oral presentation of its contents.

A variety of tasks that elicit planned performance for a variety of purposes are included in both types of portfolios and demonstrate reflection about learning, including goal setting and assessment. Portfolios can be used to demonstrate processes and growth in relation to the selected learning objectives. They can be used to display 21st century skills such as problem-solving, creativity, and information literacy. It can also reflect strengths and weaknesses, time management skills, and meta-cognitive abilities.

On the point of assessment, a portfolio exhibits the progress of students and provides teachers, students as well as parents the evidence of learning progress. The overall aim of a portfolio is to provide evidence about a set of learning objectives mastered by a student. It is a like a personal album containing various tasks and achievements of a student. It is a symbolic representation of what the student does know. Student portfolios promote and support student learning in three ways:

- Students' is participation in the selection of contents

- Students' written reflection about learning

- Students' communication with teachers regarding their reflections

5.4.8 Project Work

Project work allows students to think beyond the limits of classroom and develop among them various skills, behaviours, inquisitiveness and confidence. Providing learning environments that allows students to question, analyze, evaluate leads them to higher-order thinking. The learning that occurs during the process of project work assesses the skills of collaboration, problem-solving, decision-making, and communication. As a teacher or prospective teacher, you can also assess work quality and understanding of students throughout the project-building process.

For example, for Class IX, you have given a group project work on 'Indian Rivers and Their Uses'. After assigning the task, you must be continuously observing the progress of the project work. Next, you have to assess the project report. How would you assess it? Assessment doesn't happen at the end of the task or unit. It is linked with all three learning aspects namely the content, the process, and the product. It is performance-based and on-going one. We should give multiple opportunities to students to reflect on their learning. For assessing a project report, you may develop a rubric. You should not forget that assessing the project is equally important that of assessing students' achievement. Hence, your assessment rubric should match to the content and format of your instruction. The rubric for project work may contain the following questions:

- Did your rubric contain criteria related with learners' engagement in a real-world task?

- Did your rubric have criteria to assess group work?

- Did your rubric assess students' use of higher-order thinking and problem-solving skills?

- Did your rubric evaluate the students' progress throughout the task?

A format of a rubric related to a project work is given below:

Table 5.1: Rubric for 'Indian Rivers and their Uses

Descriptive Indicators Content	Excellent	Very Good	Good	Poor
Different rivers and its tributaries are covered	The content is of high quality and utilizes many sources.	The content is complete	Content has little coverage	Content is irrelevant
Maps showing potential uses of rivers	Maps are drawn properly and include new and useful information. Symbols are correctly used	Maps are complete	Maps are vague to understand	No maps are prepared
Relevance of the Content	Content is up-to- date and creatively written. There are few or minor gram-matical errors.	Content is complete and have a few grammatical errors.	Content lacks latest information and has numerous grammatical errors.	No up-to-date content
Process				
Use the Internet to search relevant information concerning rivers and its uses	Students express the ability to search for a topic by using several search engines.	Students collect the information through internet	Students search content with minimal assistance	Students do not have the ability to use internet to search for content.

Descriptive Indicators	Excellent	Very Good	Good	Poor
Group participation	All the members in the group take part in the project work effectively	The participation of group members is satisfactory	Some students did not participate up to their level	No group participation
Outcome/ Product				
Presentation of content	Complete coverage of content, clear writing, relevant examples, and have very few errors.	Descriptions are complete, and have few errors.	Incomplete presentation and with several grammatical errors.	Improper presentation of content
Validation of findings	Multiple validation techniques are used that produce summative conclusion	Validation techniques are good and produce conclusive results.	Efforts are made to validate the information but are inconclusive	There is no validation of information
Develop a plan on how to protect rivers	The plan is complete, well organized, and viable.	The plan is nearly complete and minimum viable.	The plan is not currently ready to be implemented.	The plan suggested is not viable

5.4.9 Debate

Debate is a kind of dialogue about a significant topic. It acts as a useful technique to promote independent learning about a topic selected for debating. Sometimes you are familiar with this technique and its procedure. In this technique, you have to play the role of a moderator. The first task is to decide the topic to be debated. You have to give an overview of the topic prior to the task being set. Then you should set the scene and allocate the team. Generally, the ideal team size is 10, out of which 5 present a case in favour of the topic and the rest 5 defend the topic. During the conduct of debate you, as a moderator, can give clear guidelines on the structure of the debate, timings and the roles to be performed.

The next task is how to evaluate the debate. For evaluation, you can prepare a rating scale containing the following indicators:

- Relevance of the content

- Presentation style

- Clarity of the argument

- Logical approach

- Quality of the rebuttal

This technique develops skills such as listening, rapid response, speaking, and debating the issues in a professional manner. It also promotes skills such as team spirit, synthesizing data, negotiation, argumentation, and rebuttal skills.

5.4.10   School Club Activities

You might be familiar with the club activities that have been initiated in your school. Various disciplines have different school clubs such as literary, science, social science, mathematics, sports, tourism, aesthetic, etc. Activities such as quiz competition, seminar, debate, elocution competition, tour, field trip, gardening, awareness campaigns, are undertaken by these various clubs. Through these activities, you can assess students' team spirit, general awareness, self-discipline, individual and group responsibility, initiation, accountability, endurance, time management, dignity of labour, sense of engagement and attachment to their school. By joining clubs, students learn to work together towards a common goal and learn to respect each other. It gives chance to relieve their stress and gain confidence in performing various activities. Students will also understand the process of planning and organizing work.

5.5   ASSESSMENT OF GROUP PROCESSES

Present day classroom instructional strategies emphasize on group work, collaborative and co-operative learning. Through group work, students' analytical and cognitive skills, team work skills, collaborative skills, organizational and time management skills can be assessed. In the next section, we will discuss how co-operative learning and seminar help in assessing the students.

5.5.1   Cooperative Learning and Social Skills

The purpose of a cooperative learning group is that individuals within the group must learn. The achievement of the group then depends on the learning of individuals within the group. Rather than competing for rewards for achievement, members of the group cooperate with each other by helping each other to learn. When small groups of students of mixed backgrounds with their varied, skills and capabilities work together, their likings and respect increase. As a result, there is an increase in each student's self-esteem and academic achievement. Through cooperative learning, students learn several social skills such as respecting others' ideas, sharing materials, taking turns, providing encouragement, paraphrasing, active listening, managing conflict and building consensus. Working in groups can provide your students with valuable learning opportunities. They learn to see other's point of view and learn from others. Working collaboratively is an important life skill. Problem-based learning frequently involves elements of group work and enhanced student learning. During co-operative learning and post collaborative learning activity, assessment takes place either at the individual or group levels. While enacting each role, teacher can assess the behaviour of students.

5.5.2   Seminar and Report as Assessment Techniques

Seminar is an activity where the social skills of students can be enhanced. It can be done individually or in group. The goal of the group may be pre-determined by the facilitator or decide by the individual or the group collaboratively. The use of seminar as an assessment tool can be done at three

phases: pre-seminar stage, seminar stage and post seminar stage. During pre-seminar stage, students can ponder over the topic, drafting a good seminar topic and think how to present. During the seminar stage, students present the topic and in post-seminar stages they reflect on their ideas in the presence of an audience and finally a seminar report will be prepared. Seminar report is used as a self assessment tool or as a group assessment tool. While doing self assessment, a student examines the feedback given by others and how his/her paper could be improved for the next time. The facilitator's assessment indicators may consist of the following:

–   Understanding of concept

–   Critical thinking and reasoning

–   Ability to build new knowledge

–   Self-confidence and self-discipline

–   Presentation style

After the seminar presentation, assessment of the reports can be made based on the following indicators:

–   Clarity of presentation

–   Accuracy of report writing

–   Organization of the report

–   Analysis of team work (if the reports are made in group)

–   Generation of new ideas

Check Your Progress 4

Note :   a)   Write your answer in the space given below.

   b)   Compare your answer with those given at the end of the unit.

4. What points would you keep in mind while assessing a seminar?

....

....

....

5.6   LET US SUM UP

In this Unit, the techniques of evaluation are discussed. You have read about the various techniques of evaluation such as concept tests, self-reporting techniques, observation technique, portfolios, peer assessment, assignments, and project work. Though some of the generally used techniques such as assignments, and concept tests are implemented by teachers, in this Unit you have learnt about the details of how these techniques can be used for proper assessment. While going through self- reporting techniques, you have found that they are of different types such as interview, questionnaire, and reflective journal. By now, you might have developed the ability to assess student learntly by using various

self-reporting techniques. You might have understood how evaluation of your students is possible through observation, socio-metric technique, portfolios, project work, debate and club activities. Also, you might have understood how you can interpret group assessment techniques. In the last section, if was also discussed how seminar and report can be used as an assessment tool.

5.7 REFERENCES AND SUGGESTED READINGS

Angelo, Thomas A., & Cross, K. Patricia. (1993). Classroom Assessment Techniques: A Handbook for College Teachers. San Francisco: Jossey-Bass. Retrieved from https://www.cmu.edu/teaching/assessment/assesslearning/CATs.html on 15/02/2017.

Badders, William. (2000). Methods of Assessment. Retrieved from https://www.eduplace.com/science/profdev/articles/badders.html.

Barbour, A.C. (2010). Learning at Home Prek-3: Homework Activities that Engage Children and Families. Thousand Oaks, CA: Corwin.

British Colombia Institute of Technology (2010). Classroom Assessment Techniques. Retrieved from http://www.northernc.on.ca/leid/docs/ja_assesstech.pdf

Dogra, Bharti. (2010). Constructivist Classroom Activities for Biology Learning. 'Journal of Indian Education' Volume XXXVI, 2, pp. 111-120.

Eberly Centre for Teaching Excellence and Educational Innovation. (2015). Using Classroom Assessment Techniques. Retrieved from https://www.cmu.edu/teaching/assessment/assesslearning/CATs.html 02/03/2017.

Field Tested Learning Assessment Guide (n.d.). *Classroom Assessment Techniques (CATs) - Overview.* Retrieved from http://www.flaguide.org/cat/cat.php on 22/02/2017.

Greenstein, Laura. (2012). Assessing 21st Century Skills. New Delhi: Sage Publications India Pvt. Ltd.

5.8 ANSWERS TO CHECK YOUR PROGRESS

1. To assist in student learning, to identify students' strengths and weaknesses, to assess the effectiveness of a particular instructional strategy, to assess and improve the effectiveness of curricular programmes, to assess and improve teaching effectiveness and to communicate with and involve parents and guardians in their childrens learning.

2. Self reporting techniques such as interview, focus group discussion, questionnaire and reflective journal provide means for extracting the hidden treasure of one's own behaviours and patterns of personality.

3. Autonomy in learning, learner confidence, ability to self-evaluate and reflect, improving information and understanding, improving one's own work, avoiding mistakes in future and able to assess a large amount of student work.

4. Understanding of concept, critical thinking and reasoning, ability to build new knowledge, self-confidence, self-discipline and presentation style.

UNIT 6   CRITERIA OF A GOOD TOOL

Structure

6.1   INTRODUCTION

In Unit-5, you have studied about the techniques of evaluation such as observation, interview, rating scale, socio-metric techniques, seminar, group discussion, etc. From that Unit, you might have understood the techniques of using various assessment tools for assessment and evaluation. In this Unit, we will discuss the types of tools and the essential criteria of a good tool. During the course of discussion you will understand that there are different criteria to be taken into consideration while evaluating a tool or determining the worth of a tool to be used. In this regard, broadly we will discuss the criteria of an effective tool such as reliability, validity, usability, objectivity and norm of a tool and how to determine the index of reliability and validity for making it an effective tool. Again, the effective tools are used for satisfying the purpose of its evaluation. According to this purpose, we can classify the tools into two broad categories, that are self-made tool and standardized tool. This Unit will also make you understand the concept of self-made and standardized tools and also to differentiate between them.

6.2   OBJECTIVES

After going through this Unit, you should be able to:

● define self-made and standardized tools,

● differentiate self-made and standardized tool,

● discuss the different criteria of an effective tool,

● explain the relationship between validity and reliability,

● describe the procedure to find out reliability and validity of a tool,

● calculate the index of reliability and validity of a tool, and

● explain the criteria of an effective tool.

6.3 EVALUATION TOOLS : TYPES AND DIFFERENCES

We use a variety of tools for the purpose of assessment of performances of the students in the school. We also use different types of tools for collecting data to conduct various projects and research studies. It is also equally used for various administrative purposes, for example, recruitment for various jobs. That is why we say that for assessing something, we need to use a tool. Now the question is as to which type of tool can be used for what purpose. On the basis of the purpose, tools are classified into two broad types namely self-made and standardized tools. In this Section, you will learn about both types of tools and also you will be able to differentiate them.

6.3.1 Self-made and Standardized Tools

As discussed, we usually use two types of tools, self-made and standardized, in assessment and evaluation. Let us understand the concept of both the tools.

Self-made tool : Self-made tool is also popularly known as teacher-made tool. This type of tool is prepared for assessing performance of the students in various subjects. These types of tools are usually prepared by those class teachers who are engaged in teaching the subject to the particular class. Self-made tools are meant to limited purpose for the particular groups of students to whom the teacher teaches. You as a teacher, should be engaged in preparing and administering such type of tools in your class. You may be engaged in preparing a unit, quarterly or half-yearly test for assessing the performance of your students. This is an example of teacher-made tool.

Standardized tool : When we prepare a tool by following a set procedure of test construction and develop norms for its use, then it is called standardized tool. This is generally prepared keeping in view the large group. All the criteria for preparing an effective tool are followed in a standardized tool. A self-made tool can also be a standardized tool, provided the procedures and steps are followed in developing the same. Generally, standardized tools are developed by specialists or experts in the field. The norms, such as age, gender, percentile, habitation, etc. are also developed in a standardized tool. A test prepared for classes 10 or 12 Board Examination by Central Board of Secondary Education (CBSE) or by any State Board can be an example of standardized test. Now you might have understood that the standardized tool is used on a large population, say entire country, entire State or a University. The entrance examination conducted by the CBSE for admitting students in medical or engineering programmes is also an example of a standardized test.

6.3.2 Difference between Self-made and Standardized Tool

We hope that you have understood the concept of self-made and standardized tool. Now we shall clarify the difference between the self-made and standardized tool. Read the points given in Table 6.1.

Table 6.1: Difference between self-made and standardized tool

Self or Teacher-made Tool	Standardized Tool
The purpose of teacher-made tool is to know the progress of the students in a particular subject, especially at the school stage.	The purpose of standardized tool is to admit the students in various courses and also to provide them placement or employment.
Teacher-made tool is made by the teacher for the students to whom he/she teaches.	Standardized tools are prepared by the test specialists and the experts in the field.
Teacher-made tool is prepared for the small group students in a school or class.	Standardized tools are prepared for the large group population.
The use of teacher-made tool is very limited.	The use of standardized tool is very vast.
The rigorous procedures of tool construction are not followed in teacher-made tool.	The complete procedure of test standardization is followed in standardized test.
Norms are not developed for developing teacher-made tool.	Developing norms are compulsory in standardized tool.
No statistical techniques are used to determine the reliability and validity of the teacher-made tool.	Rigorous statistical methods are used to determine the reliability and validity of the test.
Item analysis and item discrimination are not calculated in teacher-made tool.	Item-wise analysis is made to know the item difficulty and discrimination before including the item in the final test.
The difficulty level of the item in teacher-made tool depends upon the standard of teaching of the teacher to the group. It may be easy, average or difficult.	The difficulty value of the item in standardized tools is generally average in standard.
Subjectivity in preparing and applying the teacher-made tool is always there.	Standardized tools are always objective in nature.
Teacher-made tool supports the principles and spirit of continuous comprehensive evaluation, specifically the formative assessment.	Standardized tools support the principles and spirit of summative evaluation.
The Unit tool, quarterly and half-yearly examinations conducted in the schools are examples of teacher-made tool.	The tests prepared for the Board Examination may be central or state level and the test prepared for admission, placement and recruitment for various vocations/jobs are also the examples of standardized tool.

The following learning points emerge to summarize the above table :

● Self-made tool is used for limited purpose that is in particular class whereas standardized tool is used for the large group of students that may even cover students of a state or country.

● Self-made tool is prepared by the class teacher whereas standardized tool is prepared by test specialists.

● All steps of tool construction including development of norms is followed in standardized tool whereas teacher-made tool follows only few steps upto getting the content validity.

● The difficulty level of items in teacher-made tool depends upon the standard of teaching by the teacher, that is it may be easy or difficult or average whereas the difficulty level of a standardized tool is always average in standard.

Activity 1

Analyse the Table 6.1 and summarize it on the following aspects :

Purpose of both type of tools :

....

....

....

Process followed to prepare the tools :

....

....

....

Main beneficiaries of the tools and the jurisdiction of its use :

....

....

....

Examples of the tools :

....

....

....

Check Your Progress 1

Note : a) Write your answers in the space given below.

b) Compare your answers with the ones given at the end of the unit.

1. Define self or teacher-made tool in your own words.

....

....

....

....

....

2. Define standardized tool in your own words.

....

....

....

....

....

6.4 ESSENTIAL CRITERIA OF AN EFFECTIVE TOOL OF EVALUATION

Developing an effective tool is a challenging task. You have to follow scientific and systematic procedure to develop or standardize a tool. To get accurate results, you have to prepare or select a proper tool. Before selecting a tool for certain purpose, you have to look into the criteria or qualities of the tool. The essential qualities or criteria of an effective tool may be as follows:

- Reliability

- Validity

- Usability

- Objectivity

- Norm

Let us now discuss each one of these criteria in detail.

6.4.1 Reliability

Reliability is the important criteria of a good test/tool. Reliability refers to consistency. A test which shows a consistent result in its frequent uses in different situations and places is called reliability of the test. The other synonyms that can be used for getting reliability of the test are: dependability, stability, consistency, predictability, accuracy, etc. It implies that the reliable test always provides a stable, dependable, accurate and consistent result in its subsequent uses. Before discussing the methods or techniques of determining reliability, it is worthwhile to observe here some points cited by Gronlund (1981) :

---

i.   Reliability refers to the results obtained with an evaluation instrument and not to the instrument itself. An evaluation tool may have a large number of different reliabilities depending on the groups of subjects and situations of use.

ii. Test scores are not reliable in general. An estimate of reliability always relates to a particular type of consistency – say consistency of scores over a period of time (stability) or consistency of scores over different samples of questions (equivalence), etc.

iii. Reliability is a necessary but not a sufficient condition for validity. A valid test is always reliable, but a reliable test may or may not be valid.

iv. Reliability is primarily statistical in nature in the sense that the scores obtained on two successive occasions are correlated with each other.

v.   The coefficient of correlation is known as self-correlation and its value is called the 'reliability coefficient'.

(**Source :** *ES – 333; Educational Evaluation, IGNOU, 2010*)

Methods or techniques of reliability : As per Singh (2002), there are three common methods of estimating the reliability coefficient of test scores. These methods are :

(i)   Test-retest reliability.

(ii) Parallel-forms reliability.

(iii) Internal consistency reliability.

Let us discuss each method of reliability in detail.

(i) **Test-retest reliability :** Test-retest reliability means the same test is administered twice on the same group of sample within a given time interval and correlation is calculated between the two sets of scores (first and second administration). If the coefficient of correlation is positive and high, it is considered that the test is reliable. Let us discuss the procedures of using test-retest reliability.

Table 6.2 : Precedures of using test-retest reliability

Conditions	Use
Test	Single form
Administer	Twice
Group	Single
Time interval	Ideally 15 days to 6 Months
Determining reliability	By employing correlation statistics
Decision	Reliable if correlation coefficient is positive and high. (r = +0.5).
Threat of error	Error of 'carry over effect', 'memory effect', 'practice effect', etc.
Suggestion	Highly used method as it is easy to determine reliability of the test, but because of the error factor, test–retest method should be the last option for the user.

After analysis of Table 6.2, we can say that test-retest method is one of the easiest method to determine reliability of the test. In this method, the same test is administered twice within a gap of time limit to minimum 15 days to maximum six months. After getting the scores of test and re-test results, any method of correlation either Spearman's Rank Difference or Pearson's Product Moment Correlation method are used to get the coefficient of correlation (please read Unit-16, Block-4, BES-127). As this has already mentioned in the Table, if the coefficient of correlation (r) is = +0.5, we can say that the test is reliable. The following methods can be employed to calculate correlation of the testing and re-testing scores.

Spearman's Rank Difference Method (ρ)

$$\rho = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)}$$

ρ = Is called as 'rho', means correlation in test-retest method.

d = Deviations ($R_1 - R_2$) [Rank of the candidates in first test score (R_1) and Rank of the candidates in second test score (R_2)]

n = Is total number of candidates

Karl Pearson's Product Moment Method (r)

$$r = \frac{\sum xy}{N\sigma x \sigma y}$$

r   = Correlation coefficient

x   = Deviations taken from mean of the distribution (X-M) [X – Individual Score and M – Mean of the first distribution]

y   = Deviations taken from mean of the distribution (Y-M) [Y – Individual Score and M – Mean of the second distribution]

N   = Total number of the candidates

σx = Standard deviation of first distribution

σy = Standard deviation of second distribution

Limitations of using the method : The following are the limitations of the test-retest method :

● As the same test is administered twice on the same group, there will be the threat of carry over effect, it means, during the second administration, the candidates may remember many items from the first administration.

● The scoring of second administration is usually high than the first one.

● Maintaining a gap of time between test and re-test is also again one of the important aspects to determining exact value of reliability. If time gap is very less, then carry over effect will be high and on the other side, if time gap is very high, maturity effects of the candidates may hamper the test results.

● This method is not free from errors. Memory, carry over, practice and maturity effects are high in this technique.

Activity 2

Prepare a small test and find out reliability of the test by using test-retest method of reliability.

....

....

....

....

....

(ii) **Parallel-form Reliability :** Because of the error factors in test-retest method, parallel-form method is one of the alternate methods of the test-retest method and it can minimize many of the errors occurred in the earlier method. In the parallel form method, two parallel tests are prepared keeping in consideration equivalence in all aspects such as similarities in content, objectives, types and number of items, time allowed in both the tests, level of difficulty, discrimination value, conditions of use, etc. The main effort by doing these is to make two equivalent forms of test.

Let us discuss the procedures and conditions of using parallel-form method for determining reliability.

Table 6.3 : Precedures of using parallel-form method for reliability

Conditions	Use
Test	Two equivalent forms
Administer	Twice (First test for first administration and second test for second administration)
Group	Single
Time interval	This is advisable to maintain a gap of minimum 15 days to maximum six months for first and second testing.
Determining reliability	Use of correlation statistics by the Spearman or Pearson formula of correlation.
Decision	Reliable if correlation coefficient is positive and high. (r = +0.5).
Threat of error	Error of 'carry over effect', 'memory effect', 'practice effect', etc. are not totally minimized, but in comparison to test-retest method, the occurrences of error is comparatively less. Here there is no question of splitting the items into two equal halves.
Suggestion	As it is an advanced method of test-retest method, this can be used in many situations, but the main challenge here is to develop exactly two equivalent forms of test.

In Table 6.3 we have stated the processes adopted in the parallel-form method to determine the reliability of a test. In this method, two parallel tests are developed which are equivalent in all respects, such as content, objectives, types of items, time given, difficulty level of the item, etc. When two parallel tests are used to get the reliability of the test, it is quite natural that the carry over, memory, and practice effects are highly minimized. For example, when items are different, there is very less chance to remember the earlier questions. Just like the test-retest method, the first form of the test is administered on the group and after a gap of some time period say, from a minimum of 15 days to a maximum of 6 Months, the second form of the test may be used. Spearman's or Karl Pearson's method can be used to get the coefficient of correlation of the two administrations. If correlation coefficient is, $r \geq +0.5$, then the test is said to be reliable.

Limitation of parallel form method : Parallel form method is also not completely free from errors. There are possibilities of making errors in this method also:

- Practice and carry over effect is not totally minimized, as both the tests are equivalent in nature in many respects except only the items are different and a time interval of 15 days to 6 months is given for testing the second form of the test. During this period, there is a chance that

the students may practice the similar content and items and hence chances for getting better scores in second testing are generally more.

● Preparing two parallel forms of the tests is also a complex task.

● This method is comparatively time taking to get the reliability.

(iii) Internal consistency reliability : Internal consistency reliability indicates the homogeneity of the test. If all the items of the test measure the same function or trait, the test is said to be a homogeneous one and its internal consistency reliability would be pretty high. The most common methods of estimating internal consistency reliability are the (a) Split-half method and (b) Rational equivalence method. Let us discuss split-half method first.

(a) Split-half method : This method is also called as 'odd-even method'. Let us discuss the procedures and conditions of using split-half method for determining reliability.

Table 6.4 : Precedures of using split-half reliability method

Conditions	Use
Test	Single form
Administer	Once
Group	Single
Time interval	Not necessary as the test will administer only once.
Determining reliability	By employing correlation statistics as well as split-half reliability formula.
Decision	Reliable if reliability index is positive and high. ($r = +0.5$).
Threat of error	Error of 'carry over effect', 'memory effect', 'practice effect', etc. are minimized, but still splitting the full test into two halves is also difficult because splitting the full test into two-halves can be made in many different ways such as odd and even, first fifty percent and last fifty percent, and so on.
Suggestion	Is one of the highly used method as most errors are minimized.

Table 6.4 reveals the procedures of using split-half reliability method. Most of the errors that occur in test-retest method are minimized in split-half method as the test is not conducted twice on the same group. In this method, the test is to be administered only once to a single group and after getting the scores of the students, it has to be split into two equal halves like in odd and even items, or any other technique to divide it into two halves. The scores of the candidates for odd items and for even items will be separated and suitable statistical techniques are used to get the reliability of the scores of two halves. The Spearman-Brown Prophecy formula is generally used for calculating reliability of the full test. The formula is as follows:

$$r_{tt} = \frac{n\sigma 2 - M(n-M)}{\sigma 2\ (n-1)}$$

r_{tt} = Reliability of the whole test

$r_{1/2}$ = Correlation of the two halves of the test

Just like, test-retest method, correlation of the two halves of the test ($r_{1/2}$) will be calculated by using the method of Spearman or Pearson's correlation. Splitting of the item can be done as shown in Table 6.4.

Table 6.5: Worksheet for the Odd-Even Reliability

Examinee	Total No. of correct scores	Number of correct scores on odd-numberred items	Number of correct score on even-numbered items
1	50	29	21
2	46	20	26
3	55	28	27
4	39	19	20
5	55	25	30
6	49	27	22
7	60	32	28
8	42	20	22
9	45	24	21
10	53	28	25

Table 6.5 represents the total number of correct scores of the examinees and also the number of correct scores on odd numbered items and on even numbered items. For example, the examinee at serial number 1 scored a total of 50 out of which 29 scored from the odd items and 21 from the even items. Accordingly, all examinee's scores are grouped for odd and even items and made into two groups for getting the reliability of the whole test.

Limitations of this method : The important error in this method is to split the full test into two equal halves. This can be done in many ways like odd-even, first 50 percent items and next 50 percent items, etc. The correlation of two-haves of the test and the reliability of the full test will also be different based on different ways of splitting the test items. Besides this type of errors, the other type of errors such as 'carry over error', 'maturity error', 'memory and practice effect' etc. are mostly minimized in this method.

Activity 3

With the example presented in Table 6.5, calculate reliability coefficient of whole test by using Spearman-Brown Prophecy formula.

....

....

....

(b) Rational equivalence method : This is purely a statistical process of determining reliability of a test. The formula for calculating reliability in rational equivalence method is developed by *Kuder and Richardson.* They developed two formulae popularly known as Kuder and Richardson (KR) formula number 20 and 21 (KR 20 and KR 21). Lee J. Cronbach has called it *Coefficient of Internal Consistency.* This method of reliability is an improvement of all the earlier methods as it minimises errors in calculating reliability. You should also understand that rational equivalence method is the improvement of split-half method. The assumption in *Kuder and Richardson* formula is that all items have the same or equal difficulty value, but not necessarily the same persons solve each item correctly. The difficulty value of an item depends on knowing aspect of a person. If the person knows the item, it is easy for him/her and if he/she does not know the item, it is difficult for him/her. The assumption is that the test is homogeneous in nature.

Table 6.6: Precedures of using rational equivalence method

Conditions	Use
Test	One
Administer	Once
Group	Single
Time interval	No need
Determining reliability	By employing KR-20 and KR-21 formullae.
Decision	Reliable if reliability index is positive and high. (r $\geq$ +0.5).
Threat of error	Errors are mostly minimized
Suggestion	Can be used effectively in research

Table 6.6 reveals the conditions of using rational equivalence method in determining reliability of a test. This method statistically eliminates most of the errors of determining reliability of a test. It requires certain statistical applications to determine the reliability, that's why the skill of applying statistics is required to get the reliability. As indicated, all most all errors are minimized in this method besides the human and calculation errors. This is highly used in tool construction for research. The formula of *KR-20* and *KR-21* are as follows :

KR-20 formula is the more accurate and is given as follows :

$$r = \left(\frac{n}{n-1}\right)\left(\frac{\sigma^2 - \sum pq}{\sigma^2}\right) \qquad\qquad \text{—— } KR\text{-}20$$

r = Reliability Index

σ^2 = Variance of the total test (square of the standard deviation)

n = Number of items in the test

p = Proportion of right responses

q = (1-p) proportion of wrong responses

KR-21 formula is given as follows :

r   =   —— *KR-21*

r   =   Reliability Index

σ^2 =   Variance of the total test (square of the standard deviation)

n   =   Number of items in the test

M   =   Mean of the test scores

Limitations of rational equivalence Method : The following are the main limitations of this method :

● The item variation of item difficulty is not considered in calculating reliability index.

● The reliability coefficient is approximate and lower than the accurate reliability coefficient.

● It can not be used for power test or heterogeneous test.

● The different KR formula yields different reliability coefficient.

Factors influencing Reliability Coefficient : Factors influencing reliability of a test can be classified as extrinsic factors and intrinsic factors.

Extrinsic factors : The important extrinsic factors (i.e. the factors which remain outside the test itself) influencing the reliability are :

i.   **Group variability :** When the group of pupils being tested is homogeneous in ability, the reliability of the test scores is likely to be lowered and vice-versa. It is therefore, the method of randomization is used to get the groups having a representative of all abilities.

ii.   **Guessing and chance errors :** Guessing in test results in to increased error variances and thus reduces reliability. For example, in two-alternative response options there is a 50% chance of answering the items correctly in terms of guessing. It is therefore difficult to determine the true scores of the examinees.

iii. **Environmental conditions :** As far as practicable, testing environment should be uniform like the proper facilities of light, air, seating arrangement, audibility of the instructions, time of testing, qualities of materials used for the test, etc.

iv.   **Momentary fluctuations :** Momentary fluctuations may raise or lower the reliability of the test scores. A broken pencil, momentary distraction by the sudden sound of a train running outside, anxiety regarding non-completion of homework, mistake in giving the answer and knowing no way to change it are the factors which may affect the reliability of the test scores.

Intrinsic factors : The principal intrinsic factors (i.e. those factors which lie within the test itself) which affect the reliability are :

i.   **Length of the test :** Reliability has a definite relation with the length of the test. The more the number of items the test contains, the greater will

be its reliability and vice-versa. Therefore adequate number of items needs to be included in the test.

ii.   **Homogeneity of the items :** Homogeneity of items has two aspects : item reliability and the homogeneity of traits measured from one item to another. If the items measure different functions and the inter-correlation of items is 'zero' or near to it, then the reliability is 'zero' or very low or vice-versa.

iii. **Difficulty value of the items :** Broadly, items having indices of difficulty at 0.5 or close to it yield higher reliability than items of extreme indices of difficulty. It emphasizes that the items should be of the average difficulty values. Too much difficulty and too much easy items are usually avoided in the test.

iv.   **Discriminative value :** When items can discriminate well between superior and inferior, the item total correlation is high, then the reliability is also likely to be high and vice-versa. It implies that the item has the quality to discriminate the lower and the higher group.

v.   **Scorer reliability :** Scorer reliability, otherwise known as reader reliability, also affects the reliability of a test. Scorer reliability speaks of how closely two or more scores agree in scoring the same set of responses. The reliability is likely to be lowered if they do not agree.

(*Source :* *Factors Influencing Reliability, ES-333, IGNOU, 2010*)

Check Your Progress 2

Note : a) Write your answers in the space given below.

b) Compare your answers with those given at the end of the unit.

3. Define reliability in your own words.

    ....

    ....

    ....

4. Which is the best method for determining reliability and why?

    ....

    ....

    ....

5. List any five factors which decrease the reliability of a test.

    ....

    ....

    ....

6.4.2 Validity

In the preceding sub-section you have studied about the reliability. Let us now discuss about validity. Validity tells us about the accuracy and truthfulness of a test. The accuracy of a test can be said if and when the test measures the

the purpose it was constructed. Cronbach (1970) defines validity as 'validity is the extent to which a test measures what it purports to measure'. Further Freeman (1965) defines validity as 'an index of validity shows the degree to which a test measures what it purports to measure when compared with accepted criterion.'

From the above three definitions given we can say that validity talks not only about consistency but also accuracy and truthfulness of the test results. This can also be said that, validity is above reliability and it includes reliability also. See Figure 6.1 given below :
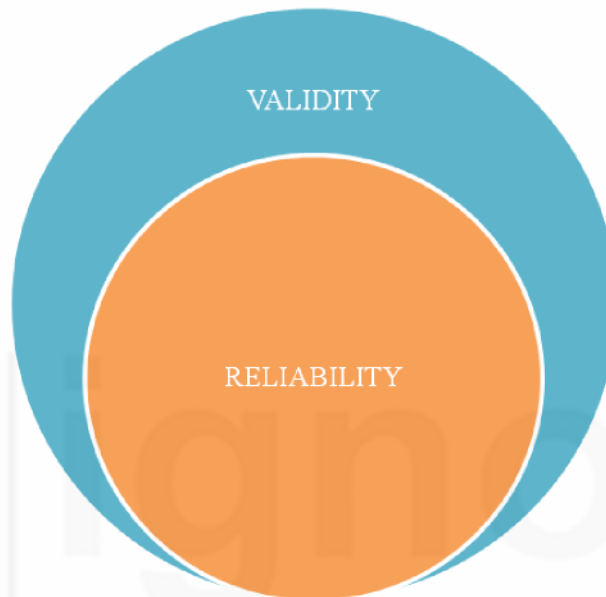
Figure 6.1: Validity and Reliability

From the definitions of validity, let us try to understand the characteristics of validity.

Characteristics of validity : The following are the main characteristics of validity of a test :

i.   Validity is an index of external correlation. The test scores are correlated with external criterion scores such as the test scores will be correlated with an earlier developed valid test prepared for measuring the same aspect.

ii.   The criterion may be a set of operation, purpose or predictor for future course of performance.

iii.   It deals with the psychological construct of a variable which is indirectly measured with the help of behaviours.

iv.   No test in education and psychology is perfectly valid because measurement is indirect.

v.   Validity endorses the reliability of a test. If a test is valid, it must be reliable, but a reliable test may or may not be valid.

vi.   It refers to the truthfulness or purposiveness of test scores.

vii.   It indicates the degree to which the test is capable to achieve the aims for which it is developed.

viii. Validity is best considered in terms of matter of degrees, such as high, moderate and low validity.

Types of validity

Commonly, five types of validity are used in preparing tools. These are :

(i)   Face Validity

(ii)   Content Validity

(iii) Criterion-related Validity

- Predictive validity

- Concurrent validity

(iv) Construct Validity

(v)   Factorial Validity

Let us understand the concept of the above types of validity.

(i)   Face validity : Face validity is the first step to know validity of a test. This is also called validation by face. This method is not widely used because it never analyses the entire test and its items to determine the validity of the test. In the face validity, the appearance of the test, purpose of its construction, objectives it covers, dimensions it measures, language used, etc. are taken into consideration for determining the face validity of the test. This is the lowest level of determining validity of the test. This method can only be used in case of shortage of time for using other methods of validity. Further, before using other methods of determining validity, a judgment is taken, whether the test is validated by face or not. If the test lacks face validity then usually other methods of validity are not determined.

(ii)   Content validity : Content validity is the second level of validity of the test. In this method, the format as well as the content of the test is examined and decision is taken for its validity. Anastasi (1968) defines content validity as, 'it involves essentially the systematic examination of the test content to determine whether it covers a representative sample of the behavior domain to be measured.' Content validity basically matches the test items with the instructional objectives. Content validity is very important for achievement test. To know the content validity of a test, usually the items are examined as per the blue-print/table of specification on which the test is prepared. Let us try to understand, what is a blue-print or table of specification with the help of an example :

Example :

Test : Achievement Test of Social Science for Class-IX

Total Marks – 100

Total Item – 100 (Each item carries 1 mark)

Types of Items – Objective Type Item (multiple choice type with four options)

Weightage to Objective Areas – Knowledge (40%); Understanding (20%); Application (20%); and Skill (20%)

Weightage to Content Areas – History (25%); Political Science (25%); Economics (25%); and Civics (25%).

Blue-print/table of specification : A blue-print is a three dimensional chart, where weightage are given to Content, Objectives and Form of Questions in terms of marks. Here the first dimension is 'content', second dimension is 'objectives' and the third dimension is 'form of questions in terms of marks'.

Table 6.7: Blue-print of an Achievement Test

Content/Objectives	Knowledge	Understanding	Application	Skill	Total
History	10×1 = 10	5×1 = 5	5×1 = 5	5×1 = 5	25
Political Science	10×1 = 10	5×1 = 5	5×1 = 5	5×1 = 5	25
Economics	10×1 = 10	5×1 = 5	5×1 = 5	5×1 = 5	25
Civics	10×1 = 10	5×1 = 5	5×1 = 5	5×1 = 5	25
Total	**40**	**20**	**20**	**20**	**100**

As presented in Table 6.7, the content validity of a test depends upon as to how far the items are written as per the table of specification. The closer the test items correspond to the specified sample, the greater the possibility of satisfactory content validity. Therefore, it is desirable that the items in a test are screened by a team of experts; they should check whether the placement of various items in the cells of the table is appropriate and whether all the cells of the table have an adequate number of items. For determining content validity of the test, the experts require to examine the items with the textbook, syllabi, etc.

(iii) Criterion-related validity : Unlike content validity, criterion-related validity can be objectively measured and declared in terms of numerical indices. The concept of criterion-related validity focuses on a set of 'external' criterion. The external criterion may be data of 'concurrent' information or of a future performance. The criterion related to concurrent information is called as 'Concurrent Validity' and criterion related to future performance is called as 'Predictive Validity'.

● **Predictive validity :** Predictive validity refers to the predictive capacity of a test. It indicates the effectiveness of the test in forecasting or predicting future outcomes in a specific area. In short, predictive validity determines how far the test is able to predict future result. This can be better understood by an example. Suppose we have to prepare a medical entrance examination test to admit students in medical courses. The predictive validity of the test can be determined only when those qualified students who took admission in medical course performed well in the medical final examination. Predictive validity is a time consuming method. To get predictive validity of a test, you have to wait till the completion of the course. Sometimes, prediction can also take much time to correlate it with further criterion such as those students who performed well in the entrance examination, whether they successfully completed the course or not, and further, those students who successfully completed the course did get a job/ placement or not. So, to determine the predictive validity of a test, there is a need to establish correlation of the scores between entrance

examination result and the course completion result. If the correlation coefficient is positive and high, you can say that the test is valid.

This type of validity is sometimes referred to as 'empirical validity' or 'statistical validity' as our evaluation is primarily empirical and statistical. You can test the validity empirically.

● **Concurrent validity :** Concurrent validity refers to the extent to which the test scores correspond to already accepted measures of performance. For example, suppose you have prepared a test of 'intelligence' and you want to know the concurrent validity of the test, you have to correlate the scores of the test administration with the scores of another established standardized test. Let us understand it with the help of another example. The Intelligence test, which you have prepared and the intelligence test prepared by Stanford-Binet can be administered among the same group of students and correlation coefficient of two sets of scores can be determined. If the coefficient of correlation is high, we can say that the test has concurrent validity.

(iv) Construct validity : Construct validity is also called 'psychological or trait validity'. Construct validity means that the test scores are examined in terms of a construct. For example, the construct for achievement of a student may be his/her intelligence, practice, aptitude, interest, attitude, etc. Construct validity can be defined as the extent to which the test may be said to measure a theoretical construct or trait or psychological variable. In construct validity the variables related to the test which contributes to that aspect are correlated and examined.

For example, this is a theoretical fact that intelligence and achievement are positively correlated with each other. Suppose you have to prepare an intelligence test and you want to know the construct validity of the test. For that, you have to correlate the intelligence test scores of the students with their achievement test scores. The assumption here is that those students who have done well in intelligence test will naturally do well in achievement test, because as per the theory, both are positively correlated with each other. In case the correlation is negative, it can be said that the intelligence test is lacking construct validity. This can also be correlated with other theoretical and psychological construct of an intelligence test with the assumptions as follows :

● Intelligence and achievement are positively correlated with each other.

● Intelligence and aptitude are positively correlated with each other.

Construct validity is to the extent test results are interpreted in terms of known psychological concepts and principles. Certain common examples of theoretical constructs of most psychological tests are intelligence, scientific attitude, critical thinking, reading, comprehension, study skills and mathematical aptitude, etc.

(v) Factorial validity : Factorial validity determines the correlation of the different factors/components with the whole test. Factorial validity is determined by a statistical technique known as factor analysis. It uses methods of expansion of inter-correlations to identify factors (which may be verbalized as abilities) constituting the test. The correlation of the test

with each factor is calculated to determine the weight contributed by each such factor to the total performance of the test. This validity tells us about the factor loading. The factors responsible for achievement of students are called factor loading. This relationship of the different factors with the whole test is called the factorial validity. Factorial validity is the clearest description of what a test measures and by all means should be given preference over other types of validity.

Factors affecting validity : A large number of factors influence the validity of the test. Gronlund (1981) has suggested the following factors :

i.   Factors in the test itself :

The following factors that affect validity of a test are included in the test itself. These are also called as intrinsic factors.

- **Unclear direction :** If directions regarding how to respond to the items, whether it is permissible to guess and how to record the answers, are not clear to the pupil, then the validity will tend to reduce. Hence, clear direction should be given in the test.

- **Reading difficult vocabulary and sentence structures :** The complicated vocabulary and sentence structure meant for the student taking the test may fail in measuring the aspects of pupil performances; thus it results in lowering the validity.

- **Inappropriate level of difficulty of the test items :** When the test items have an inappropriate level of difficulty, it will affect the validity of the tool. For example, in criterion referenced test, failure to match the difficulty specified by the learning outcome will lower the validity.

- **Poorly constructed test items :** The test items which provide unintentional clues to the answer will tend to measure the pupils' alertness in detecting clues as well as the aspects of pupil performance which ultimately affect the validity.

- **Ambiguity :** Ambiguity in statements in the test items leads to misinterpretation, multi-interpretations and confusion. Sometimes, it may confuse the good students more than the poor ones resulting in the discrimination of items in a negative direction. As a consequence, the validity of the test is lowered.

- **Test items inappropriate for the outcomes being measured :** Many a times we try to measure certain complex types of achievement, understanding, thinking, skills, etc. with test forms that are appropriate only for measuring factual knowledge. This affects the results and leads to a distortion of the validity.

- **Test too short :** A test usually represents a sample of many questions. If the test is too short to become a representative one, then validity will be affected accordingly.

- **Improper arrangement of items :** Items in the test are usually arranged in terms of difficulty with the easiest items first. If the difficult

items are placed early in the test, it may make the students spend too much of their time on these and fail to reach other items which they could answer easily. Also, such an improper arrangement may influence the validity by having a negative effect on pupil motivation.

- **Identifiable pattern of answer :** When the students identify systematic pattern of correct answer (e.g. T, T, F, F, T, T, F, F or A,B,C,D,A,B,C,D, etc.) they can cleverly guess the answers and this will affect the validity. It is therefore, in objective items, the order of answers should not follow any pattern.

ii. **Functioning content and teaching procedure :** In achievement testing, the functioning content of test items can not be determined only by examining the form and content of the test. The teacher has to teach fully how to solve a particular problem before including it in the test. Tests of complex learning outcomes seem to be valid if the test items function as intended. If the students have previous experience of the solution of the problem included in the test, then such tests are no more a valid measurement for measuring the more complex mental processes and thus affect the validity.

iii. **Factors in test administration and scoring :** The test administration and scoring procedure may also affect the validity of the interpretations from the results. For instance, in teacher-made tests factors like insufficient time to complete the test, unfair help to individual students, cheating during the examination, and the unreliable scoring of essay answers might lead to lower the validity. Similarly, in standardized tests the lack of standard directions and time limits, unauthorized help to students and errors in scoring, would tend to lower the validity. Whether it is a teacher-made test or standardized-test, adverse physical and psychological conditions during testing time may affect the validity.

iv. **Factors in students' response :** There are certain personal factors which influence the students' response to the test situation and invalidate the test interpretation. Students' emotional disturbance, lack of motivation, test anxiety, etc. may affect the validity.

v. **Nature of the group and the criterion :** Factors such as age, sex, ability level, educational and cultural background of the students influence the test measures. Therefore, the nature of the validation group should find a mention in the test manual. The nature of the criterion used is another important consideration while evaluating validity coefficient. For example, scores on a scientific aptitude test are likely to provide a more accurate prediction of achievement in an 'environmental studies' course. Other things being equal, the greater the similarity between the performance measured by the test and the performance presented in the criterion, the larger the validity coefficient.

(**Source :** *Factors affecting Validity, ES-333, IGNOU, 2010*)

Check Your Progress 3

Note : a) Write your answers in the space given below.

b) Compare your answers with those given at the end of the unit.

6. Define validity in your own words.

....

....

....

7. Explain predictive validity with an example.

....

....

....

8. State any four important factors that influence validity of a test.

....

....

....

6.4.3 Usability

Usability refers how successfully you, as a teacher, use the test in a classroom situation. It has been observed that, many highly valid tests lack the quality of usability. The user fails to understand or feels it difficult to use the test. Therefore, a good test should have the quality of usability. While selecting an evaluation tool, you should look for certain practical considerations like easy for administration and scoring, easy for interpretation, availability of comparable forms and cost of testing. All these considerations induce a teacher to use tools of evaluation and such practical considerations are referred to as the 'usability' of a tool of evaluation. In other words, usability means the degree to which the tool of evaluation can be successfully used by the teacher and school administrators. So usability of a test includes comprehensibility, easy for administration and scoring, easy for interpretation, appearance of the test, economy and availability of the test for use.

6.4.4 Objectivity

Objectivity is another important feature for a good test. Objectivity of a test refers to two aspects of the test, viz;

● Item-objectivity, and

● Scoring-objectivity.

Item-objectivity means the item is having only one right answer. Many a times, you might have observed that a single item has two or more related answers. That affects the validity of the test. Apart from these, ambiguous questions, lack of proper direction, double-barreled questions, questions with double negatives, essay type questions, etc. do not have objectivity. So much care has to be exercised while framing the questions.

Scoring-objectivity refers the test paper would fetch the same score. Scoring objectivity can be ensured by carefully maintaining the item-objectivity. In objective-type items, it is easy to ensure scoring-objectivity whereas in subjective item, certain precautions needs to be taken to ensure scoring-objectivity such as carefully phrasing the essay items, making proper directions of scoring, making the items short-answer type instead of essay type item, etc.

6.4.5 Norm

Generally norm is considered as a standard, but technically, there is difference between the concepts of norm and standard. Norm can be defined as the average or standard score on a particular test made by a specified population.' Thorndike and Hegen (1977) defines norm as 'average performance on a particular test made by a standardized sample.' Determining norm is one of the important criteria of a good test. Most standardized tests determine norm. Norm can be characterized as follows :

● It acts as a basis for interpreting test scores and minimize interpretive error of the test.

● It helps to transform the raw scores into standard scores or derived scores and put meaning to it.

● Norm suggests a level and therefore the individual departure from the level can be evaluated in quantitative term.

● Norms are necessary for the purpose of promotion, gradation, selection and classification of examinees.

● It refers to the average performance on a particular test made by standardized sample or specified population.

The procedure of deterring the norm is a challenging task. Without determining the norm of a test we can not say that the test is standardized. It is therefore, in a standardized test, usually the age, level, habitation, etc. are mentioned on which the test can be used and it also equally reflects in the interpretation of the test scores. For interpretation of scores of educational and psychological tests, different norms like age, grade, percentile, standard score, etc. are broadly employed.

Check Your Progress 4

Note : a) Write your answers in the space given below.

b) Compare your answers with those given at the end of the unit.

9. Explain the concept of item-objectivity.

....
....
....

10. What are the different aspects of determining usability of a test?

....
....
....

6.5 LET US SUM UP

In this Unit, we presented the essential criteria of a good tool. In this context, we discussed first the concept of a teacher-made test or self-made test and standardized test. We also differentiated a self-made test and standardized test and also the purpose and context of its use. For clarifying criteria of a good tool, we discussed reliability, validity, usability, objectivity and norm. You were acquainted with the methods of determining reliability and validity of a test and you might have understood how to calculate the reliability and validity of a test. We have also discussed the other criteria of a tool such as objectivity, usability and norm. In this Unit, the concept of usability, objectivity and norm have been presented with examples.

6.6 REFERENCES AND SUGGESTED READINGS

Anastasi, Anne (1976). Psychological Testing, 4th ed., New York : Macmillan Publishing Co. Inc.

Cronback, L.J. (1970). Essentials of Psychological Testing, 3rd ed., New York : Harper and Row.

Ebel, Robert, L. (1996). Measuring Educational Achievement. New Delhi: Prentice-Hall of India.

Ebel, Robert, L. and Fristic, David, A. (1991). Essentials of Educational Achievement. New Delhi: Prentice-Hall of India.

Gronlund, N.E. (1981). Measurement and Evaluation in Teaching. 4th ed. New York : Macmillan Publishing Co. Inc.

Freeman, F.S. (1965). Theory and Practice of Psychological Testing (3rd ed.). Oxford and IBH Publishing Co. Pvt. Ltd., New Delhi.

IGNOU (2010). Criteria of a Good Tool (Unit 6; Block 2), Educational Evaluation (ES-333, B.Ed.), New Delhi: IGNOU.

Kerlinger, F.N. (1973). Foundations of Behavioral Research, 1st ed., New York : Holt, Rinchart and Winston Inc.

Nayak, B.K. and Rath, R.K. (2010). Measurement, Evaluation, Statistics and Guidance Services in Education, New Delhi: Axis Publications.

Nually, J.C. (1972). Educational Measurement and Evaluation. 2nd ed., New York : McGraw Hill Book Company.

Sharma, R.A. (2005). Mental Measurement and Evaluation. 2nd ed., Meerut : R. Lal Book Depot.

Singh, A.K. (2002). Tests, Measurements and Research Methods in Behavioural Sciences. 3rd ed., Patna : Bharti Bhawan.

Thorndike R. L. and E.P. Hagen (1977). *Measurement and Evaluation in Psychology and Education*. 4th ed, New York : John Wiley and Sons.

6.7 ANSWERS TO CHECK YOUR PROGRESS

1. Test prepared by the class-teacher for formative assessment of the students to whom he/she teaches which is not standardize is called as self-made or teacher-made test.

2. Standardized tool is prepared for the large group. The complete procedures for test construction/standardization are followed to standardize a test. The reliability, validity, norm, etc. are usually determined to standardize a tool.

3. Reliability refers to consistency. A test which shows a consistent result in its frequent uses in different situations and places is called reliability of the test.

4. Rational equivalence method is the best method for editermining reliability of a test as statistifically most of the errors are minimised in this method.

5. (a) In case number of items in the test is less; (b) If items are not homogeneous; (c) If memory and carry-over effect works; (d) Error in scoring the test, and (e) In case items fail to discriminate the higher and lower group students.

6. Validity means accuracy and truthfulness of the test and which satisfies the very purpose of the test.

7. Predictive validity indicates the effectiveness of the test in forecasting or predicting future outcomes in a specific area.

8. Self exercise.

9. Item-objectivity means the item is having only one right answer and it does not carry ambiguous questions, lack of proper direction, double-barreled questions, questions with double negatives, broad essay type questions, etc.

10. Comprehensibility, easy for administration, easy for scoring, easy for interpretation, appearance of the test, economy and availability of the test.

UNIT 7 TOOLS FOR ASSESSMENT AND EVALUATION

Structure

7.1 INTRODUCTION

In Unit-6, you have studied about the techniques of evaluation and criteria of a good tool. In Unit-5 of this Block, you have studied various techniques, such as self-report, assignment, projects, socio-metric techniques, peer assessment, debates, school and club activities, cooperative and collaborative learning and assessment, and seminars use for assessing students performance in scholastic and co-scholastic areas. Units 5 and 6 of this Block will help you to understand the discussion in this Unit.

In this Unit, we will come to know about some prevailing tests being used in our schools to comprehend, evaluate and accomplish various aspects of teaching and learning namely achievement, oral, intelligence, and aptitude of the students.

These aspects not only help you to evaluate the learners and diagnose their problems but also assess your own teaching too, so that it can be improved. The most common tools that are used in the schools such as rating scale, questionnaire, inventories, checklists, interview and observation schedules,

---

*(**Note** : Some sections of this Unit have been partially reproduced from Course: ES-333 'Educational Evaluation', B.Ed. Programme, IGNOU, 2010)*

anecdotal reports, portfolio and rubrics are discussed in the Unit. This will help you to do an effective and enjoyable teaching and learning.

7.2 OBJECTIVES

After going through this Unit, you should be able to:

- discuss the concept of tests and tools,

- differentiate between paper – pencil and oral tests,

- explain the concept of aptitude tests, their types, uses and limitations,

- discuss and administer the aptitude, achievement and diagnostic tests,

- explain the concept of intelligence test, its types, uses and limitations,

- explain a rating scale, its types, uses and limitations,

- explain the process of developing and use of questionnaire, inventory, checklist, interview and observation schedule,

- discuss the use of anecdotal records, and

- discuss the use of portfolios and rubrics.

7.3 TESTS

Tests may be of different forms, such as psychological test used to measure mental and behavioural traits, achievement test to assess performances of students, etc. They can be used to assess both the scholastic and co-scholastic abilities of the students. Let us now understand the concept of the test. The tests are those instruments by which you, as a teacher, collect information as data through verbal and non-verbal responses of the students. A concise definition may be : a test is an instrument or systematic procedure for measuring a sample of behaviour (Gronlund, 1990). Further, a psychological test is defined as a standardized, repeatable procedures used to elicit and measure samples of human behaviour' (Kazdin, 2000). From the above two definitions, you can summarize the meaning of a test and the sample behaviour it measures as follows :

- Human abilities, including intelligence, aptitudes, skills, and achievement in various areas.

- Personality characteristics, which include traits, attitudes, interests, and values.

- Adjustment and mental health, which involves detecting signs and symptoms of psychological and neurological disorders and appraising the effectiveness of psychological functioning.

(*Sourse* : *Kazdin, 2000*)

A test can be used for two broad objectives. First, it attempts to compare the same student on two or more than two aspects of a trait, such as attitude and aptitude of the same student; and second, two or more than two students may be compared on the same trait like attitude of two students (Singh, 2002).

45

The tests which you are going to study in this Unit are classified based on the various objectives or purposes to be fulfilled which are not only related to the learner but also useful to the teacher and teaching-learning process. Some tests have been discussed in the subsequent sub-sections.

7.3.1 Paper Pencil Test

The paper pencil tests comprise a standard set of questions which are presented to the student in writing on paper or Optical Mark Recognition (OMR) sheets that requires completion of cognitive tasks in the form of response by the student on those papers by pencil/pen mark. These responses or answers are summarized/scored to obtain a numerical value that represents a characteristic of the student for which the test was administered. The paper pencil tests can focus on what the student knows (achievement), is able to learn (ability or aptitude), chooses or selects (interests, attitudes or values), or is able to do (skill).

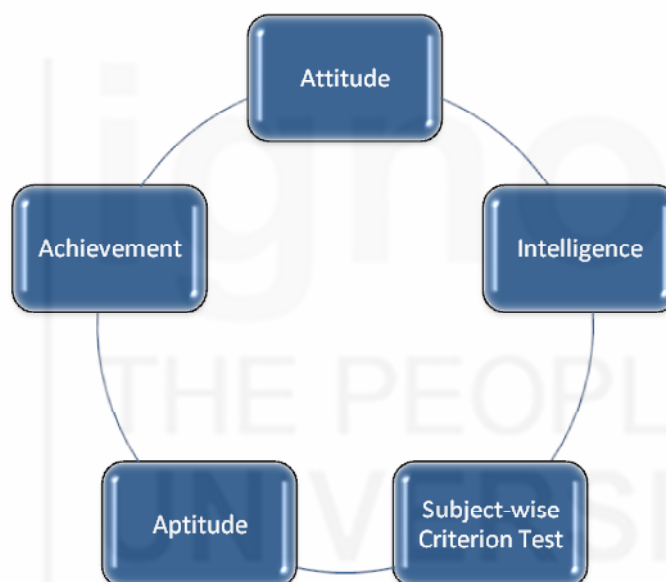Following figure shows different types of paper pencil tests:

Figure 7.1: Paper Pencil Tests

These paper pencil tests are easy to administer and score. They are economic and have standard questions which provide uniform procedure for all the students. These tests provide standard scores to the teachers.

7.3.2 Oral Test

Oral tests are those tests in which the response, solution or the task requires oral response to answer the question. Teachers' conversation with the students for the purpose of assessment, component of viva-voce examination for completion of a course, etc. can be called oral test. You might have noticed that in a written test, very little scope is left for the students to express themselves on any aspect whereas in oral test, students enjoy freedom to express themselves by citing many examples. In oral test, you can also use figures, diagrams, charts, maps, models, signs, etc. for asking students to explain the concepts covered in it. Oral tests have proved most valuable when used with students having language disabilities, the illiterates, shy feeling and the young children.

Actually, the oral tests rely on the students' oral performance to display their cognitive capabilities. This may be in the form of an oral discourse or a question and the answer session between the teacher and the students. These are effective for students with writing difficulties but may not be appropriate for students with hearing and speech difficulties or those who are prone to nervousness. For the students of speech difficulties, sign languages can be used, but in an inclusive classroom setting, sometime it is not possible to apply sign language as it requires sign language training both for the teachers and the students as well.

Oral testing is particularly important in language classes because spoken language plays an important role to express oneself. Oral tests focus on students' ability to communicate with what they know and how they articulate the things. Since spoken language is creative and flexible, students may find that oral testing gives them opportunities to self correct, ask for clarification and use language creatively. Teachers may appreciate the opportunity to hear students who choose not to speak on a daily basis in class.

The oral test is also having its own limitations as hardly it measures the skill of writing of the students. Conducting an oral test is also time consuming method, especially when it involves in-depth questioning. Oral test can be used as one of the components of overall assessment of the performance of the students.

7.3.3 Aptitude Test

An aptitude test is also an instrument used to determine and measure an individual's ability and skills to acquire, through future training. Aptitude tests may be classified into two groups : multiple aptitude test, and special aptitude tests. Multiple aptitude tests are those which intend to measure various areas of aptitude (musical, mechanical, etc.) each by independent sub-tests (Singh, 2002), whereas the special aptitude tests measure only one specific aptitude like teaching, musical, etc. The multiple aptitude tests measure abilities of students in more than one area simultaneously by using different sections of a test while the special aptitude tests measure ability of the students in one area.

Aptitude tests are used to predict the future performance of the students. These predictions are for the performance based on specific criterion (for example you all appeared before an entrance test for admission to the B.Ed. programme. The test included teaching aptitude components required to get into the B.Ed. programme) which are prior to instruction, placement or training.

The aptitude tests are also used for guidance, as well as prediction of success in some occupation. Training or academic courses are possible on the basis of scores on standardized aptitude tests. For example, for pursuing teacher education programmes, one needs to have teaching aptitude; for pursuing medical courses, one needs to have medical aptitude; for pursuing engineering courses, one needs to have engineering aptitude; for musical courses and sports related courses, one requires to have musical and the sports aptitude respectively.

7.3.4 Achievement Test

Achievement tests are administered on students to measure their learning outcomes. These tests are more prevalent in our schools. These tests show as to what has been learned by the students, rather than to predict future performance as in the case of aptitude tests. There are teacher made and standardized achievement tests. The choice of your achievement test will depend upon your purpose. The achievement tests can be used for following purposes:

● to know the learning progress of the students.

● checking any weakness in the instruction.

● in formulation of learning objectives and provide an easy means of critical examination of the content and the methods of teaching.

● adapting the instruction to the need of the individual learner.

● to know the effectiveness of any academic programme.

The achievement tests can be classified as :

▪ General achievement tests (batteries)

▪ Special achievement tests

General achivement tests attempt to measure the general educational achievement of the students at different stages which includes the common subject areas taught at the school in a particular class. Special achievement tests are meant for measuring the achievement of students in selected areas which may be grouped into two distinct groups – the diagnostic tests (discussed later in this Unit separately) and the standardized summative achievement tests. Diagnostic tests are used to know the areas of difficulties of the students and accordingly to provide suitable remedial instructions to them. Standardized summative achievement tests are used to grade the students in a particular standard and also to certify them.

An important precaution before using any achievement test is to assess the content validity with a standardized achievement test because content differs from one school to the other and even in the same subject. For using any standardized achievement test, you need to compare the content included in your curriculum for the level it has prepared and the components of the content included in the standardized test.

Though achievement tests are commonly used in schools to assess the students' achievement, there are some limitations of these tests. The following are the main limitations :

● Scores in the achievement test may not be the exact measure of achievement in the concerned field as there is subjectivity in assessment and the test selected may not completely represent the curriculum.

● Many factors such as standard of teaching, completion of curriculum, testing condition, prior information to the students, practice of the students, etc. may affect the students' achievement, but those factors may not be measured by the achievement test.

● Achievement scores of students differ in similar types of achievement tests. There may be the possibility that a student scores differently in different achievement tests meant for the same class.

You will study in detail the process of developing achievement tests in Unit-9 of Block-3 of this Course.

7.3.5 Diagnostic – Remedial Test

The achievement tests which are used to identify the strengths and weaknesses of the students are called diagnostic tests. The primary purpose of the test is to identify the problems faced by the students in achieving the learning objectives and to suggest remedial measures for them. Such tests are available in various

areas of study, such as in language, mathematics, science, etc. Diagnostic remedial test is specifically required in case the teacher observes certain learning difficulties in the students. For example, if a student scores 90 marks in Mathematics in the Class IX annual examination, but if you observe that the same student scores 70 marks in Mathematics in Class X quarterly examination, thereafter 50 marks in half-yearly examination, and 30 marks in pre-board examination. In this case, there is a need for diagnosing the causes of securing lower marks in subsequent examinations in Mathematics and accordingly provide remedial teaching to the student.

Those tests which are devised to diagnose learning difficulties are considered as remedial tests. In most of the cases the remedial tests are subjective and individual oriented. If the problem is widespread among the group of students in a class or a region then the group remedial test can be used.

You will study the details of diagnostic testing and remedial instruction in the Unit-12, Block-3 of this Course.

7.3.6 Intelligence Test

The term 'intelligence' is difficult to define in a single sentence acceptable to all. But intelligence can be understood as a general set of mental traits, which is often reported as the mental abilities. Test of mental ability or intelligence tests measure convergent thinking (Singh, 2002). Convergent thinking is a process of finding out the solution of a problem. Guilford defines convergent thinking is the ability to give correct answer to standard questions that do not require creativity. For instance, most of the school based tasks can be done through different tests. Many psychologists and educationists opined that intelligence is a product of heredity and environment, but it is a matter of debate to get the real contribution of heredity and environment for intelligence of an individual. Many researches you will find in this regard. In this section, our intension is to understand the measurement of intelligence and the tests used for measuring intelligence.

Intelligence tests are used to provide a very general measure, usually reporting a global test score. As they are general, intelligence tests are useful in predicting a wide variety of tasks. Intelligence tests can be classified on various backgrounds.

On the basis of administration it can be classified into two categories :

a) Individual tests

b) Group tests

Let us talk about each category.

Individual intelligence test is one which can be administered on one person at a time. The first individual intelligence test was the Binet-Simon scale (1904). Individual test requires a highly skilled and experienced test administrator because he/she has to give exact instructions to the students in different sections and sub-sections of the test.

The group intelligence test is one which can be administered to more than one person at a time, that is, it can be administered on a group. The Army Alpha Test, 1917 (Verbal) and Army Beta Test (non-language) were the first group intelligence tests. They are generally used in educational setting for adolescents

and adults. At present you will find many such standardized intelligence tests available in different departments such as psychology and education of various institutions for use in schools, for recruiting personnel in different jobs and also for using in masters and doctoral researches.

Intelligence tests can also be classified on the basis of the nature of items included in the tests such as verbal and non-verbal tests. In verbal tests, the instructions and items are presented to the student in written language, whereas, in non-verbal tests, items are given in the form of figures, charts, diagrams, etc. Let us now discuss the measurement of intelligence, that is intelligence quotient.

Intelligence Quotient (IQ) : Intelligence of an individual can be measured by using the following intelligence quotient (IQ) formula.

$$I.Q. = \frac{Mental\ Age}{Chronological\ Age} \times 100$$

Chronological age : The exact age of the individual whose intelligence will be measured. This can be calculated counting the age from the date of birth of an individual. Chronological age of a child never varies.

Mental age : Mental age is calculated by comparing a child's obtained score from the intelligence test with the average score at his/her age level. For example, if average score in an intelligence test for six years age group children is 70, and a child of this age group scored 70, then it will be considered that the mental age of that child is 7. In case, the same child scored the average score of another intelligence test meant for the age group of eight years, then the child's mental age will be 8. Therefore, mental age of a child may vary from the chronological age. In different cases it may be same as chronological age or above or below of it.

Let us discuss it with an example :

Let the chronological age of a child is nine (CA – 9) and mental age is 10 (MA-10).

$$I.Q. = \frac{MA}{CA} \times 100$$

$$= \frac{10}{9} \times 100$$

$$= 111.11$$

To interpret the IQ score, you can refer the intelligence table as follows :

Table 7.1: IQ Reference Table

IQ score range	Category
Under 70	Feeble mindedness
70-79	Border line
80-89	Dullness
90-109	Average or normal
110-119	Superior
120-140	Very superior
Over 140	Genius or almost genius

(**Source** : https://www.iqtestforfree.net/iq-scale.html, retrieved on 10.8.2017)

As per the above example, you can say that the child having the IQ of 111.11 is in superior category.

Check Your Progress 1

Note : a) Write your answers in the space given below.

         b) Compare your answers with those given at the end of the unit.

1. What do you mean by a test?

    ....

    ....

    ....

2. List different types of paper-pencil tests?

    ....

    ....

    ....

3. What do you mean by oral tests?

    ....

    ....

    ....

4. For whom are the oral tests more effective?

    ....

    ....

    ....

5. What is the purpose of diagnostic tests?

    ....

    ....

    ....

6. Define the concept of mental age.

    ....

    ....

    ....

7.4 TOOLS

In the preceding section, you have studied about the various tests. In this section, you will study about various tools including the scales, schedules and inventories. To measure the scholastic abilities of the students, we usually use the tests such as achievement, intelligence, aptitude, etc., but for measuring most of the co-scholastic abilities, we use the tools such as rating scale, inventories, checklists, schedules, questionnaires, etc. In simple language, you can understand a tool as an instrument to measure something. In education, the word tool is used to measure various traits of students. Now, let us discuss each tool with details.

7.4.1   Rating Scales

Rating scale is one of the important tools widely used in psychology and education. It is used for assessing attitudes of students on any situation, idea, object, character, person or an attribute. In a rating scale, the opinions are given in various degrees such as strongly agree, agree, and disagree; highly satisfied, satisfied, and dissatisfied, etc. A rating scale is prepared always in odd number points like 3-point scale, 5-point scale, 7-point scale, or 9-point scale. It is in odd number points because, a definite middle measuring point will be possible only when the scale is odd points. You might have observed that many persons have the attitude to opine their opinions at the middle rating. Rating scales can be presented in different categories. The most commonly used categories of rating scales used in the schools are as follows :

(i)   Numerical scale

(ii) Graphic scale

Let us discuss the above mentioned scales briefly.

i)   Numerical scale : In this type of rating scale, numbers are assigned in different degrees. It is therefore called numerical rating scale. In a numerical scale a sequence of definite numbers is supplied to an observer. The assigned numbers against the item carry certain meaning/interpretation. See the example given below :

Question : How do you consider Sudhir's behaviour in your class? (in this example, Sudhir is a student)

9   Most pleasant

8   Extremely pleasant

7   Moderately pleasant

6   Mildly pleasant

5   Indifferent

4   Mildly unpleasant

3   Moderately unpleasant

2   Extremely unpleasant

1   Most unpleasant

Instead of nine-point scale you can have 3-point, 5-point or 7-point scale. In a 5-point scale you can have the following characteristics:

5   Most pleasant

4   Moderately pleasant

3   Indifferent

2   Moderately unpleasant

1   Most unpleasant

Numerical rating scales are the easiest to construct and to apply. They are also the simplest in terms of handling the results. However, numerical scales are often rejected in favor of other types of scales because it is believed that they suffer from various biases and errors.

ii)   Graphic Scale : The graphic scale is a popular and most widely used rating scale. In this scale, a straight line is shown vertically and horizontally

with various clues to help the rater. The line is either segmented into units or continuous. If the line is segmented then the number of units can vary from case to case. Let us try to understand from the example given below:

Question : How effective was the teacher in the class?

1	2	3	4	5
Very Effective	Moderately Effective	Average	Moderately Ineffective	Very Ineffective

Question : Semester system of education encourages learners to engage in their studies regularly.

1	2	3	4	5
Strongly agree	Agree	Undecided	Disagree	Strongly disagree

The above two examples are presented in 5-point rating scale. Scoring to this type of items are done as per the instructions given in the test. In case it is a positive item, the highest positive degrees of the response are given highest points and for the negative items lower points. For example, in a positive item, 5 points can be given for 'strongly agree', 4 to 'agree', 3 to 'undecided', 2 to 'disagree', and 1 to 'strongly disagree' in a 5-point scale. This is vice-versa in case of negative items.

Limitations of rating scales : The rating scales have certain limitations. Some of them are discussed as follows :

- **Error of leniency:** There is a constant tendency among the raters to rate those whom they know well or in whom they are closely involved, higher than they should. Such raters are called 'easy rater'. Some raters become aware of the feeling of easy rating and consequently rate individuals lower than they should. Such raters are called 'hard raters'. The leniency error refers to a general and consistent tendency for a rater to rate too high or too low for whatever reasons.

- **Error of central tendency:** Most of the raters hesitate to rate the individuals on the extremes of the scale, instead they follow the middle of the scale. Hence the result gets distorted.

- **Halo effect:** It is an error which obscures the cluster of traits within an individual. The rater develops a general opinion about the person's merit and his ratings on specific traits and is greatly influenced by the impression. It results in a false positive co-relation among the traits which are rated.

- **The logical error:** It is due to the fact that judges are likely to give similar ratings for traits which they feel logically related to each other.

- **The contrast error:** It is due to a tendency of a rater to rate others in the opposite direction (contrasting) from himself in a trait.

- **The proximity error:** It has been seen that adjacent traits on a rating tend to inter correlate higher than remote ones, their degree of actual similarity being approximately equal. This error may be counteracted to some extent by placing similar traits further apart and the different ones close together.

(**Source** : ES-333, B.Ed., IGNOU, 2010)

7.4.2 Questionnaire

A questionnaire is a device comprising a series of questions dealing with some psychological, social, and educational topic(s) given to an individual or a group of individuals, with the object of obtaining data with regard to some problems under investigation. (Koul,1997) Questionnaire is a common tool that we, the teachers usually apply to collect data from a situation, condition or prevailing practice.

The questionnaire can be classified in terms of the nature of the questions constructed to gather the data from the stakeholders. It may be closed or open ended. *Closed questionnaires* are those where the respondents answer in limited way, like responding in 'yes' or 'no'; underlining the replied among the pre-defined responses, putting the sign 'correct' or 'incorrect'. Whereas in *open questionnaires* the respondents are free to share, clarify and put their view. Let us discuss certain examples of closed and open ended questions.

Examples of open-ended questions :

- Describe any one of the wars fought for independence in India.

- Write an essay on Quit India Movement.

Examples of closed-ended questions :

- In which year Quit India Movement took place?

  a) 1941   b) 1942   c) 1943   d) 1944

According to other classification, questionnaire may be Structured or Non-Structured. A structured questionnaire contains definite, concrete and directed questions, whereas non-structured questionnaires are often used as the interview guide which is non-directive. It is an important instrument being used to gather information from widely scattered sources and also used when factual information is desired. (IGNOU, 2007)

Advantages of questionnaires : Some of the advantages of questionnaires are enumerated as follows :

- The responses are gathered in a standardized way.

- Questionnaires are more objective, certainly more so than interviews.

- Relatively quick to collect information.

- Potential information can be collected from a large portion of a group.

Limitations of questionnaires : Although the questionnaire is one of the widely used tools but it has also certain limitations which can be enumerated as follows :

- Not easy to be used with children or illiterates.

- Respondents may not agree to respond in writing.

- Sometimes it is difficult to construct questions on complex and crucial topics.

- Respondent interprets the questions from his/her angle of understanding.

- External factors may affect the response.

7.4.3 Inventories

An inventory is constructed in the form of a questionnaire. But the inventories are more exhaustive than questionnaire. Inventories have been mostly used for measuring personality traits, interests, values and adjustments i.e. for assessing self-reporting affective domain of behaviour. It consists of a series of questions or statements to which the subject responds to by answering 'Yes' or 'No',

or 'Agree' or 'Disagree'. This can also be answered in some similar ways to indicate preferences or to make those items that describes the subject's typical behaviour.

In the inventory, the statements are put in first person, For example, "I think I am comparatively more tense than others". In the questionnaire, there is a question in a second person, for example "Do you think you are more tense than other persons around you?"

The term inventory is also used in another sense. That is, an inventory of various items is available in an office room or inventory of clothes to be given to the washer man, etc, but in the educational evaluation, in a school and classroom situation, inventories are used for the aforementioned concepts.

Guidelines for the administration of the inventory are as follows

- The teacher should explain printed instructions very clearly to the learners.

- The teachers should make it clear to the learners' that the data will be kept confidential.

- The teacher should remove doubts, if any, regarding the manner of filling the inventory or the questionnaire.

- The teacher should take all other timely precautions for preparing a state of mind conducive to response.

7.4.4   Checklist

According to Koul (1997), a checklist is a simple device consisting of a prepared list of items which are prepared by the teacher to be relevant to the problem being studied. After each item a space is provided for the observer to indicate the presence or absence of the item by checking 'yes' or 'no', or a type of number of items may be indicated by inserting the appropriate word or number.

The checklist is a systematic and quick way to gather data of the relevant factors and take actions accordingly. A very simple example is being given below :

For example :

Is the school building fire proof?	Yes/No
Does the school follow rain harvesting policy?	Yes/No
Is the school building earthquake proof?	Yes/No

It is noteworthy that the responses collected on the checklist are as a matter of fact not any judgment. It is a good tool in gathering facts for educational survey, checking your school library, laboratory, game facilities, school buildings, textbooks, etc. It may also be used to check the availability of other facilities in your school.

Checklists are also applied to classroom instructional activities such as studying the working habits of students, supervising classroom instructions, teacher –pupil relationships, etc.

7.4.5   Interview Schedule

Interview is a communication or conversation by which a person asks interviewer and interviewee responds verbally in the face-to-face situation. An interview can also be conducted through skype electronic media. It can be conducted through telephonic conversation or through by using internet. It is a common technique for collecting required information about an individual. You have learnt about interview technique in Unit-5 of this Block. The interview schedule is a tool

with the help of which the interview is conducted.

Interview schedule can be classified according to the purpose for which it is structured and used. If it is to resolve the research hypothesis, it is a research interview schedule and if it is for clinical purpose, it a clinical interview schedule.

On the bases of the structure, the interview schedules are categorized as structured or unstructured.

● **A structured** interview schedule is one in which the procedure to be followed is standardized and it is determined in advance of the interview. The same types of questions are presented to the interviewee and the wordings of the instructions to the interviewee are specified.

● In the **unstructured interview** schedule**,** the series of questions are decided in advance but the interviewer is largely free to reorganize the questions and timing to attain the objective of the interview.

7.4.6   Observation Schedule

Observation has been the first practice of assessment that we do in our classroom. Each observed incident, expression and reaction has useful data for teachers, hence the observation is an effective tool for us. Observation is the process in which one or more persons observe what is happening in a real-life situation, and he/she classifies and records pertinent happening according to pre-planned scheme. It is used to evaluate the overt behaviour use uniformly either American or British English i.e., spellings of words, etc. of individuals in both the controlled and uncontrolled situations. (Koul, 1997, p.170)

The observation schedules are the enumerations, listing of the facts or other data that are observed under observation process. Like questionnaires, the observation schedules are also classified as structured or unstructured. This can also be classified as participant and non-participant observation. You have already studied various types of observation techniques in Unit-5 of this Block. In case the things to be observed are properly defined, the style of recording the observed information, standardized conditions of observation and the selection of pertinent data of observation then it is a structured observations (Kothari & Garg, 2014, p.91) and the schedule is a structured observation schedule. On the other hand when these are not pre-determined aspects of observation, such observation schedules are unstructured schedules.

Merits of observation : The following are some of the merits of observation

● Through observation the data are gathered from a natural setup.

● The data observed are direct or first hand.

● As the data are first hand or direct, while doing observation we can correlate what is being said and what is being shown.

● Does not rely upon people's willingness or ability to provide information.

Demerits of observation : The observation has also the following demerits :

● The observation if not done with planning then it will not bring out authentic information.

● Observer's biases may affect the result.

● There may come '*Hawthorne effect*', that is when the person gets to know that he/she is being observed then the real problems may not be shared or shown.

● Does not develop the understanding why people behave in a particular manner.

Table 7.2: Strengths and weaknesses of various tools

Tool	Strengths	Weaknesses
*Tests**	• Economical • Standard Questions • Commercial tests Strong in Technical Qualities • Objective Tests Easy to Score • Standardized tests provide Uniform Procedure for all subjects and standards	• Norms may be inappropriate • Standardized tests may be too broad and general • Standard Score may distort differences • Standardized tests may give false sense of validity • Locally developed tasks often technically weak test anxiety • Restricted to subjects who can read and write
Scales	• Consumes less time to administer • Interesting to rater • Can be used by little training • Wide range of Application	• Consumes time to construct • Items if not made valid data can be misleading • More than three point scale are difficult for some raters to answer
*Question-naires**	• Economical • Can be anonymous • Standard Questions & Uniform Procedure • Usually easy to score • Provides time for subject to think about responses	• Response rate of mailed questionnaire • Inability to probe and clarify • Scoring open ended items • Faking and social desirability • Restricted to subjects who can read or write • Biased or ambiguous terms • Response set
Inventories	• Exhaustive • Appropriate for assessing self reporting affective behaviour. • Used when quick overview is required	• Can be only used when required to enlist things. • Lacks explanations. • Ill-preparation may lack essential ingredients of inventories.
Checklists	• Simple Device • Easy to Use • Quick collection of data • Systematic	• Limited in response • Restricted to one aspect • Factual • Useful when used in support of other tools
*Interviews**	• Flexible • Adaptable • Ability to probe and clarify • Ability to include non verbal behaviour • High response rate • Used with non readers	• Costly • Time consuming • Interviewer bias • Not anonymous • Subject effects • Effect of interviewer characteristics • Requires training • Leading questions
*Observation Schedules**	• Captures natural behaviour • Mitigate social desirability, response set and subject effects • Relatively unobtrusive • Reliable for low interface observations	• Costly • Time consuming • Effect of observer on subject • Observer bias • Requires training • Reliability difficult for complex behavior and high inference observations • Inability to probe and clarify • Usually not anonymous • Interpretations of high inference observations

(***Source :** McMillan & Schumacher, 2006)**

Table 7.2 summarizes the strengths and weaknesses of tests and tools. All tools have their strengths and also weaknesses. A single tool can not be suitable to use for all types of measurement. The data gathered through a checklist can not be possible by using a scale and accordingly, data received through interview is also somehow different from observation.

Check Your Progress 2

Note : a) Write your answers in the space given below.

b) Compare your answers with those given at the end of the unit.

7. What is a numerical rating scale?

....

....

....

8. How do the questionnaires differ from inventories?

....

....

....

9. Give two dimensions of a school where checklist can be used to collect information.

....

....

....

10. How is an unstructured interview schedule better than structured one?

....

....

....

11. Write any two demerits of an observation schedule.

....

....

....

7.4.7 Anecdotal Records

Every teacher (also parents, friends, relatives, etc.) observes his/her students on the day-to-day basis and those observations can be formally recorded to his own understanding or to that of others who will later deal with the students. Such reports of a student recorded by informal observations in the form of anecdotes (not only of present, but past too, that is childhood, nursery and pre-school days, etc.) by parents, siblings, friends and peer group are called anecdotal records. (IGNOU, 2010)

The anecdotal records provide a means for the observation of one's behaviour. A record of events is one dependable data that will remain unchanged from

the time it is made until the time we want to refer to it. A set of such records provides stable evidence on which later appraisals can be raised. Making a record of an observation of a child's behaviour, a prompt record, while the behaviour is still fresh in the mind, is dependable as it eliminates the limitations and distortions of memory. Such a record can be used, with practice, to provide a relatively direct and objective report of actions, with reactions of some observers as well.

These records serve two purposes. The first purpose may be to give you practice in studying students with a view to develop their understanding and increasing their sympathetic insight. Second purpose is that it provides an informal and usually qualitative picture of certain aspects of an individual's behaviour. It is a primary aspect of social functioning or adjustment to personal problems that one hopes to illuminate by records of incidents of school behaviour. The interaction of a child with the other children in a room, aggression or withdrawal, events that throw light on the child's role in the group and his reaction to it are good materials for records. Indications of personnel tensions and adaptation to them, habitual mood and temper or special crises and adjustment are worth recording.

Anecdotal records should be factual report of an event in a child's life reported with adequate details so that it is a meaningful evidence of his/her behaviour.

A good Anecdotal record has the following features:

- It provides an accurate description of any event,

- It describes a setting sufficiently to give the events meaning.

- If it includes interpretations or evaluation by the recorder, this interpretation is separated from the description and its different status is clearly identified.

- The event relates to the child's personal development or social interactions.

(**Source :** *Anecdotal Record, Unit 7, ES-333, IGNOU, 2010*)

7.4.8 Learners Portfolios and Rubrics

You have already studied about portfolio in Unit-5 of this Block. We will discuss in this section some more details about portfolio as an evaluation tool. A portfolio is a purposeful, systematic collection and evaluation of student work that documents progress towards meeting learning objectives. Portfolios have been used for years in the various fields such as architecture, art and journalism as the primary method of evaluating performance. In the field of education, portfolios are being used with increasing frequency, especially with assessment of reading and writing skills. (McMillan &Schumacher, 2006, pp.193) A portfolio (a collection of the student's work in an area, showing growth, self-reflection and achievement) is a systematic collection of work, often including work in progress, revisions, student self analysis and reflections on what the student has learned. (Popham, 2008)

Portfolio comprises artistic pieces or written work of the student. The student might also include letters in the portfolios describing each entry and its importance. They may include graphs, diagrams, pictures or digital slideshows, power point presentations, etc. They may also include recordings of the student reading, their work, unedited and final drafts of essays or poem, list of books

and references to a particular theme, peer comments, videotapes, laboratory reports and computer programmes. Anything that demonstrates learning in the area being taught and assessed can also be included. (Popham, 2008)

Few common portfolios and their purposes are as follows :

(i) Growth portfolios

● To show growth or change over time

● To help develop process skills such as self-evaluation and goal-setting

● To identify strengths and weaknesses

● To track the development

(iii) Showcase portfolios

● To showcase end-of-year/semester/trimester accomplishments of the students

● To prepare a sample of best work by the students of the school

● To showcase student perceptions of favorite, best or most important works

● To communicate a student's current aptitudes to future endeavours

(iv) Evaluation portfolios

● To document achievement of students' for grading purposes

● To document progress towards standards set by the school

● To place students appropriately in a particular section or community of learning group

Creating portfolios : Students should be involved in selecting the pieces that will make up the portfolios. The criteria for creating a portfolio is as follows :

● Ask students to include a rationale of their selection.

● Have each student write a 'guide' to his or her portfolio, explaining how strengths and weaknesses are reflected in their work.

● Include self and peer critique, indicating specifically what is good and what needs improvement.

● Model self criticism of your own production.

To evaluate these portfolios, scoring rubrics are used that actually determine the quality of students' performance.

*(**Source** : Woolfolk et.al. 2012)*

Rubric

A rubric is an assessment tool for the teachers to assess performances of the students in different areas with fixing certain criteria of assessment. To measure students' performance based on a pre-determined set of criteria, a rubric or scoring scale is typically created which contains the essential criteria for the task and appropriate levels of performance for each criterion.

A rubric comprises two components: criteria (vertical column) and levels of performance (horizontal column). Any rubric has at least two criteria and at least two levels of performance (which may be added as per the requirement of the teacher). The criteria for the performance on a task, are listed in the left-hand column in the rubric below (concepts, facts and figures, organization/presentation and references) and the performance is marked in three levels-poor, average and good. Weightage to each criteria is also given in the second column (Refer Table 7.3).

Table 7.3: Students' performance rubrics

Criteria	Weightages to the criteria	Performance		
		Poor 1	Average 2	Good 3
Concepts	X2			
Facts and figures	X3			
Organization and Presentation	X3			
References	X1			

Finally, the rubric above needs a mechanism of scoring. The weigthage to the above criteria are : two times to concept (X2); three times to facts and figures (X3); three times to organization and presentation (X3); and one time to references are (X1). Accordingly, points for the criteria 'concept' will be: poor–2; average – 4; and good – 6. Scoring to other criteria can be done accordingly.

Criteria	Weightages to the criteria	Performance		
		Poor 1	Average 2	Good 3
Concepts	2-times	2	4	6
Facts and figures	3-times	3	6	9
Organization and Presentation	3-times	3	6	9
References	1-time	1	2	3
Total		**9**	**18**	**27**

Accordingly, score of each student may be calculated. You can also use many other methods to score the rubrics.

Activity 1

Cite an example of an activity that you would like to conduct among the students to whom you teach and prepare a rubric for its assessment. Also specify the scoring procedure.

....

....

....

7.5   LET US SUM UP

In this Unit, you have studied about the tests and tools which can be used in schools easily. We discussed the tests, their types such as paper-pencil tests, oral tests, etc. We also discussed about commonly used tests in the school system such as aptitude tests, achievement tests, diagnostic-remedial tests and intelligence tests. You have also studied when to use the aptitude test and how the achievement tests have become the most used tests in our schools.

You also studied about the different tools like rating scales, questionnaires, inventories, checklists, interview and observation schedules, anecdotal records, portfolios and rubrics and their uses and limitations.

The contents and discussions covered in this Unit will help you to conceptualise and understand Unit-9, 10 and 11 presented in Block-3 of the same Course. The details about developing achievement test, diagnostic remedial test, and commonly used tests in schools have been elaborately discussed in Block-3 of this Course.

7.6   REFERENCES AND SUGGESTED READINGS

Anastasi, A. (1968). Psychological Testing. London: MacMillan Company.

Binet, A. (1916). The Development of Intelligence in Children (translated by E. S. Kite). Baltimore: Warmick and York.

Brody, N. (1976). Intellegence , New York: Academic Press.INC

English, H.B. & A.C. English (1958). A comprehensive Dictionary of psychological and Psychoanalytical Terms. New York: David McKay Company, Inc.

Freeman, S.F (1962). Theory and Practice of Psychological Testing (3rd ed.), New Delhi, Oxford & IBH Publishing Co. Pvt. Ltd.

Gronlund, Norman E. and Robert L. Linn. (1990). Measurement and Evaluation in Teaching. 6th ed. New York, Macmillan, p. 5.

IGNOU (2010). B.Ed. Course ES-333, 'Educational Evaluation' Unit 7, Types of Tools, IGNOU: New Delhi.

IGNOU (2007). *Tools and Techniques of Data Collection* in MES-016-Educational Research : Research Design**.** New Delhi: School of Education, IGNOU.

Kazdin, Alan E. (2000). Encyclopedia of psychology. Washington, D.C., American Psychological Association, vol. 8, p. 47-48.

Kothari, C.R. & Garg, G. (2014). Research Methodology-Methods and Techniques, New Delhi: New Age International Publishers.

Koul, L. (1997). Methodology of Educational Research, Delhi: Vikas Publications.

Popham,W.J. (2008). Transformative Assessment. USA:ASCD

Singh, A.K. (2002). *Test Construction* in Tests, Measurements and Research Methods in Behavioural Sciences. Patna: Bharti Bhawan Publishers.

Sternberg, R.J. (1985). Beyond IQ- A triarchic Thoery of intelligence , Cambridge: Cambridge University Press

Woolfolk, A., Mishra, G. & Jha, A. K. (2012). *Classroom Assessment, Grading and Standardized Testing* in Fundamentals of Educational Psychology, Delhi: Pearson.

Website Referred

http://busyteacher.org/14344-6-assessment-testing-strategies-esl-classroom.html dated 14/-5/16

www.icbl.hw.ac.uk/ltdi/cookbook/info_**questionnaires**/printable.**pdf**

http://www.elon.edu/students/portfolios/dated 10/07/16

http://jfmueller.faculty.noctrl.edu/toolbox/rubrics.htm dated 14/07/2016

https://www.iqtestforfree.net/iq-scale.html, retrieved on 10.8.2017

7.7 ANSWERS TO CHECK YOUR PROGRESS

1. A test is a standardized instrument designed to measure objectively one or more aspects of a total personality of a person through the sample of verbal and/or non-verbal responses.

2. Achievement, aptitude, subject-wise criterion test, intelligence test, etc.

3. The oral tests are those tests in which the response, solution or the task requires oral response to answer the question.

4. Oral tests are effective for students with language disabilities, illiterates, shy feeling and young children.

5. The main purpose of diagnostic test is to identify the problems faced by the students in achieving the learning objectives and to suggest remedial programmes for them.

6. Mental age is calculated by comparing child's obtained score from the intelligence test with the average score at his/her age group.

7. In numerical rating scale, numbers are assigned to different degrees for responding to an item in the rating scale.

8. In questionnaires, the questions are written in second person whereas the questions in inventories are written in first person.

9. School library, laboratory, game facilities, school buildings, textbooks, etc.

10. Unstructured interviews are focused, in depth and non directive hence more data can be gathered on the problem under investigation, whereas in the case of structured interview options of answers are given and it is also restricted.

11. Observer biases and Hawthorne effect.

UNIT 8 ICT BASED ASSESSMENT AND EVALUATION

Structure

8.1   INTRODUCTION

In Units-5 and 7 of Block-2 of this Course, we have discussed various tools and techniques used for assessing learning of children. Many of them such as achievement, aptitude, and intelligence tests are used in the day-to-day classroom situations. Nevertheless, as you know, the constructivist approach in the teaching–learning processes, facilitate teachers to apply innovative information and communication technologies (ICTs) to assess children. For the last few decades, ICT has become an important component of the classroom activities. Teachers gather information via internet and other ICT resources for utilizing it in the teaching and learning process. Children are also using computer and other electronic devices for self-learning and also acquiring more knowledge with the existing curriculum they study. It is under this context; you need to be oriented and equipped in utilizing ICT for evaluating student performance. Various ICT tools are available that help teachers to assess the performance of students. In this Unit we will discuss various aspects such as importance of ICT in evaluation, role of teacher in ICT based evaluation, few forms of ICT techniques used to asses children's progress and use of ICT in preparing and analyzing evaluation tools.

8.2   OBJECTIVES

After going though this Unit, you should be able to:

*   discuss the role of ICT in assessment and evaluation,

*   describe the use of ICT in various types of evaluation such as online and e-examination,

- discuss the of use of different tools to create online/e-exams,

- discuss e-portfolio and e-rubrics as types of evaluation,

- describe the use of different tools to create online/e-exams, and

- identify ICT tools to prepare tests and analyse results.

8.3 IMPORTANCE OF ICT IN ASSESSMENT AND EVALUATION

The nature of teaching, learning and evaluation has undergone changes with the emergence of ICT. The ICT, as you have studied in various units of this course is a broad term that encompasses various tools, techniques, devices and software applications that are used to communicate and disseminate information. The ICTs include television, computer, iPod's, learning management system(LMS), virtual reality, social networking sites, online teaching, online digital repositories, etc. ICT has number of applications in education which develop skills and competencies in individuals to succeed in the workplace and lead a quality life. As we know, in 21st century being complex in nature, education will promote individuals to develop talents and proficiencies to surmount the crisis that they would come across. Even if we have a balanced curriculum, pitfalls exist. In this technology era most of our classrooms fail to adopt and organize teaching-learning sessions integrating technology. What are the applications of ICT in education? This Unit discusses one of such applications of ICT.

In teaching and learning, you may employ ICT for both teaching and evaluating student's performance. You are familiar with the term 'evaluation'. Before moving further, let us be clear about these two terms 'assessment' and 'evaluation'. As you have already studied in Unit-1 of Block-1 of this Course, 'assessment' is a process of determining the present status/condition of something while evaluation is the process of measuring something to judge and value it.' Assessment is a step involved in evaluation. Both assessment and evaluation are integral of the teaching-learning process. These help you to ensure as to how far you are able to transact curriculum and facilitate learning in the children. Now our concern as a teacher is to know as to how we can utilize ICT for both assessment and evaluation? Definitely, there are possibilities which would enable you to find instances where you can use ICT to evaluate student's achievement. Thus ICT acts as an effective medium for teaching-learning and evaluation. Integrating ICT in evaluating children's performance is important for the reasons as follows:

- ICT-enabled assessments provide children scope for engaging in individualised testing situations. In such contexts, the learning proceeds in tune with the pace of the individual attempting tests.

- ICT integrated assessment provides immediate feedback to the children. Technology enabled assessment enhances the confidence level of children as they receive results of their learning instant.

- The assessment could be arranged in such a way that children would get instant feedback and thus can correct their mistakes and move forward.

- The frequency of assessment can be increased that would benefit children and will continuously engage them in their studies.

- It generates interest and increases motivation among leaners as they try out varied types of technology assisted tests

- Technology assisted testing/assessment techniques are cost effective and can be easily prepared.

- The principles of learning like 'own time and own pace' is emphasised while using ICT for assessment.

- It helps to develop higher order thinking and digital literacy skills among children

- The problems like "access and quality" are minimised with the use of ICT in assessment and evaluation.

- It caters to the tastes of multiple audiences who differ in intelligence, creativity, etc.

Check Your Progress 1

Note : a) Write your answers in the space given below.

b) Compare your answers with those given at the end of the unit.

1. Why do you consider ICT important in evaluating students' performance?

....

....

....

....

....

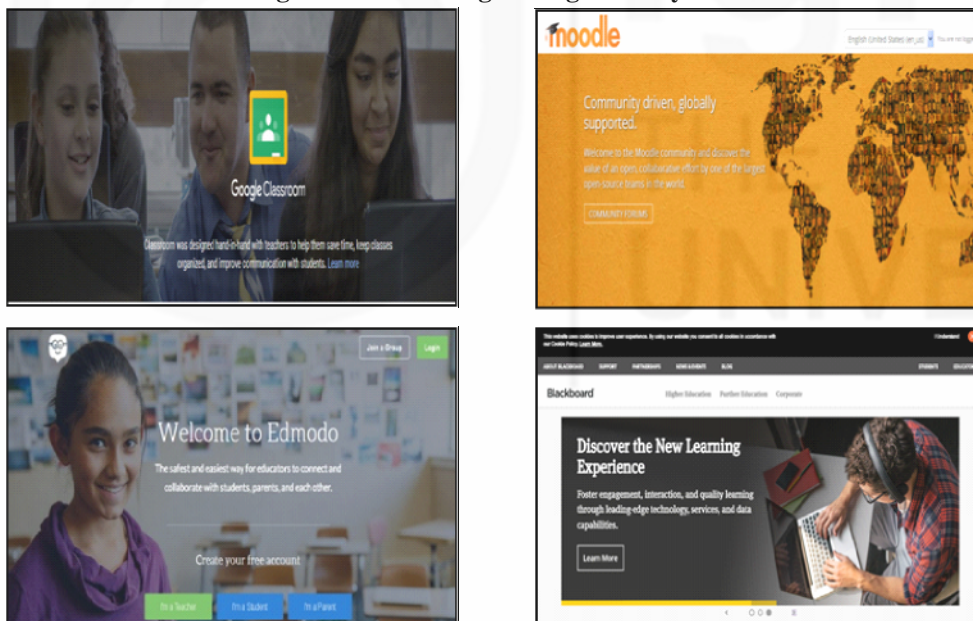## 8.4 USE OF ICT IN VARIOUS TYPES OF ASSESSMENT AND EVALUATION

Evaluation, as you know, is an integral part of teaching-learning process. It enables you to evaluate the instructional activities and learning achievement of children. The paper and pencil tests, unit and term end examinations, oral questioning techniques are some of the traditional evaluation methods. But the emergence of ICT has influenced evaluation methods. A simple example would be recording the marks of children in an excel sheet. Earlier the marks of children were recorded in paper sheets but today application softwares such as Microsoft excel are used to record it. It is to be noted here that the use of ICT in evaluation would be viewed on two dimensions; **technology as a tool and technology as an assisting medium**. Let us discuss two simple examples. A student using a video camera to record the teaching session would be an example of using technology as a tool while analysing the recorded video would be a case of using technology as assisting evaluation. Before moving further, you may go through the graphic, which will help you to differentiate the different types of evaluation.

Figure 8.1: Types of Evaluation

Diagnostic Evaluation

The type of evaluation conducted at the beginning of teaching session to detremine the learning level of children and to group them. As per the learning levels, teacher may organise the learning experiences/activities

Formative Evaluation

This type of evaluation is conducted during the course of teaching to identify the learning performance of children. Help children to mkae out performance in learning and teacher may modify learning actvities accordingly.

Summative Evaluation

This type of evaluation is conducted at the end of an instruction actvity/unit/course. This will help to grade learners and to jugde the effectiveness of teaching. Summative evaluation helps in value judgement.

As a teacher, our concern is about using ICT in various types of evaluation. Let us explore the various ways of it. As mentioned, diagnostic evaluation is performed before the teaching session. At this stage you may utilize ICT in place of verbal questioning. For example, multiple choice questions can be prepared in computers (using any application software) and children may be directed to attempt those questions. While preparing questions, animations and motivating appeal should be provided. Similarly in case of formative evaluation, children may be directed to attempt the questions provided in the learning management systems (platforms like, Google classroom, MOODLE and Edmodo may be used. The home page screen shots of few such platforms are given in Figure 8.2. During summative evaluation, the possibilities of online examination could be tried out.

Figure 8.2: Learning Management System

(*Sources:* *https://moodle.org/.,https://www.edmodo.com/.,http:// anz.blackboard.com/.,https://classroom.google.com/h*)

Activity 1

How will you use technology in formative evaluation? Suggest your plan?

....
....
....
....

67

Check Your Progress 2

Note: (a) Write your answers in the space given below.

(b) Compare your answers with the one given at the end of this Unit.

2. Differentiate between diagnostic, formative and summative evaluation.

....

....

....

3. How will you use technology for summative evaluation? Suggest a plan.

....

....

....

8.5   ROLE OF TEACHER IN TECHNOLOGY ENABLED ASSESSMENT AND EVALUATION

Teachers play a pivotal role in deciding the technology that may be incorporated for testing the students' achievement. Technology can be employed both for evaluation and assessment as assessment is a part of evaluation. In earlier days, paper and pencil tests were the prominent testing devices but today, number of technological tools have been identified and developed that makes testing easier and more joyful for children. So whether it is evaluation or assessment, the creativity and skill of teacher matters in utilizing technology. A framework namely, Technological Pedagogical Content Knowledge (TPACK) could be of great use to teachers at this stage. The framework simply describes the individual and combined knowledge of any teacher in three areas namely Technology, Pedagogy and Content. TPACK is basically defined as a framework of teacher knowledge for technology integration. Teacher knowledge is defined as a complex interaction and intersection among three bodies of knowledge within the framework of TPACK: content, pedagogy and technology (Koehler & Mishra, 2008).
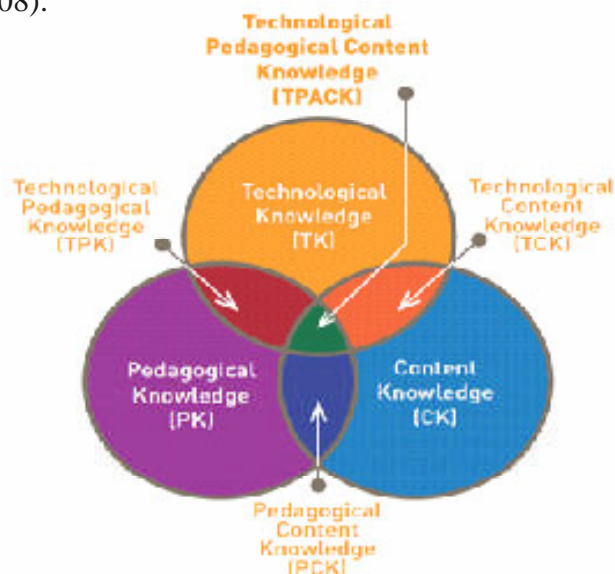

Figure 8.3: Technological Pedagogical Content Knowledge (TPACK)

(*Source*: *https://www.education.ie/en/Publications/Policy-Reports/Digital-Strategy-for-Schools-2015-2020.pdf*)

TPACK could be considered as one of the bases for selection and integration of technology in assessing children's performance. Thus your role in technology-based evaluation is to employ various technologies in testing. This would be easy if you use TPACK framework, which encompasses decisions concerning technology based evaluation. Let us see how the concept of TPACK is employed in assessment. For example, suppose you had taught the concept kinetic energy to eighth class students. Kinetic energy is the energy possessed by an object due to its motion. After teaching the concept, you try to check whether the students have grasped the concept or not? Rather than asking verbal questions, you may employ technology. At this stage, remember that, you should have knowledge of both pedagogy and content which help you to transact the concept in a meaningful way.

Coming to technology based assessment, let us see what you have done. First, you would have listed the technologies and various applications like computer, online testing, rubric, computer-based testing, learning management systems, applications for developing quizzes and multiple choice questions (MCQs), etc. From the list, you may select computer testing and applications used to develop MCQs. In the process, you may display the below given figure and ask children to identify the ball that would make bigger splash. For that, a MCQ was prepared (may use Hot Potatoes, application for creating MCQ's) as given below:

Q.1: Which ball will make bigger splash?

    a)   Small

    b)   Big

    c)   Both

The above example is a simple way of employing technology in assessment. It is up to you what technology is to be adopted for assessing children's learning. So number of factors need to be considered for utilizing technology for assessment. Let us summarize these roles as given below:

- Decide the content/topic to be assessed

- Identification of suitable technologies for assessment

- Preparation of relevant tools and techniques

- Placement of questions to be included in the assessment

- Reviewing the questions

- Tryout of the tool and rectifying mistakes

- Execution of the test

Activity 2

Prepare a teaching module employing TPACK framework. Discuss the factors that you have considered in selecting the technologies while preparing the teaching module.

....

....

....

....

....

Check Your Progress 3

Note: (a) Write your answer in the space given below.

(b) Compare your answer with the one given at the end of this Unit.

4. Discuss the role of a teacher in technology-based assessment and evaluation.

....

....

....

8.6 ONLINE AND E-EXAMINATION

The terms online and e-examination connote the same meaning and are used interchangeably, even though difference exists between them. **In online examination the assessment occurs in an internet environment while e-exams could be through any digital resources**. For example, the LMS platform like MOODLE require an internet connection to work and the assessment occurs in an internet environment. While the LMS platform, eXe software (eXe is a website to develop online content) operate both offline and online. Thus the assessment carried out in eXe is more of electronic assessment in nature. As the name indicates, online exams occur in an internet environment; mostly carried out with the help of computers. While e-exams can use any form of digital devices, the widespread use of intent technologies has exposed more avenues in engaging assessment through digital applications. Similarly, online and e-exams are gaining popularity in the education sector.

How examinees appear in online exams/e-exams? Do examinees have to take any precautions? What is the nature of such exams? Let us try to answer few such questions related to online/e-exams. As you know, in traditional paper and pen exams, the examinees will be provided with a question paper and blank answer sheet to write on, for a fixed interval of time. In online/e–exams, the following procedure is carried out:

● The students will be allotted an **User ID and Password.** The same is entered in the pre-arranged computer ready software. Thus students will get access only if they enter correct user name and password.

● As the student can access the examination software by entering username and password, the examination time starts; time will be displayed at a visible position in the computer screen. So students have to **Be Ready Themselves** to attempt the questions. Unlike in conventional examination, extra time will not be granted.

● Attempt questions and enter the answers by **clicking 'mouse'**. Unless the students click mouse and thereafter '**Save and Next Button** is pressed, the answers will not be recorded. This is one of the important steps students have to remember in online /e-exams; mouse clicking and saving the answers as they proceed.

● Conventional examination has the option of attempting questions at random but in online/e-exams; '**Sequential Move'** is preferred as forth-back movement is time consuming. Thus attempt questions in sequential manner to save time but students can skip questions (if they don't know the answer) and can be attempted at a later stage

● Students can proceed in a sequential manner and the progress of attending exam will be displayed; may be question wise or percentage mode. Thus online exams/e-exams shows the progress of examination, and students would get the option of '**Assessing their Progress'.**

● The final step in online/e-exams is very important; never forget to **Click Submit Button**. Unless submission is done by clicking mouse, the answers will not be recorded in data base.
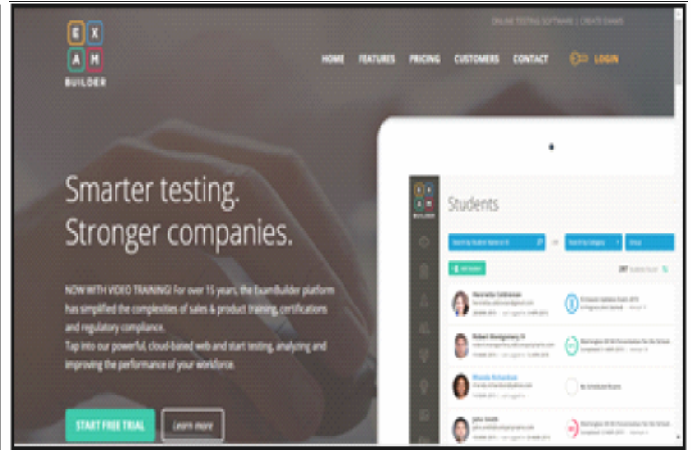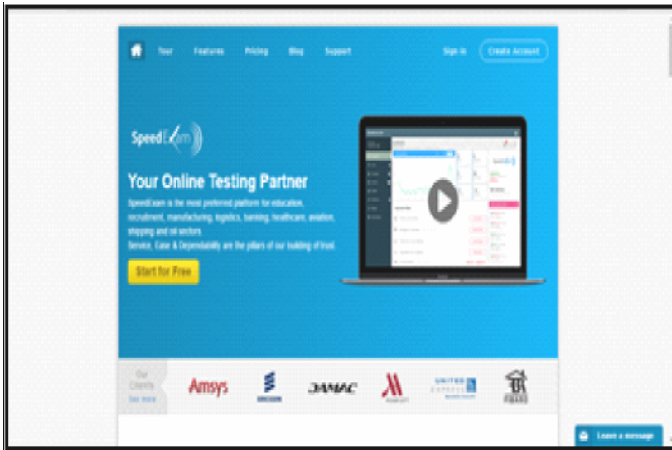
Multiple software's/applications, both paid and free are available. You can use them to develop online/e-exams. You should search the best available options and use some of them in your teaching practice or sessions in your classroom. Home page of some of the online/e-exams creating websites is given as follows:

Figure 8.4: Home pages of online/e-exams creating websites
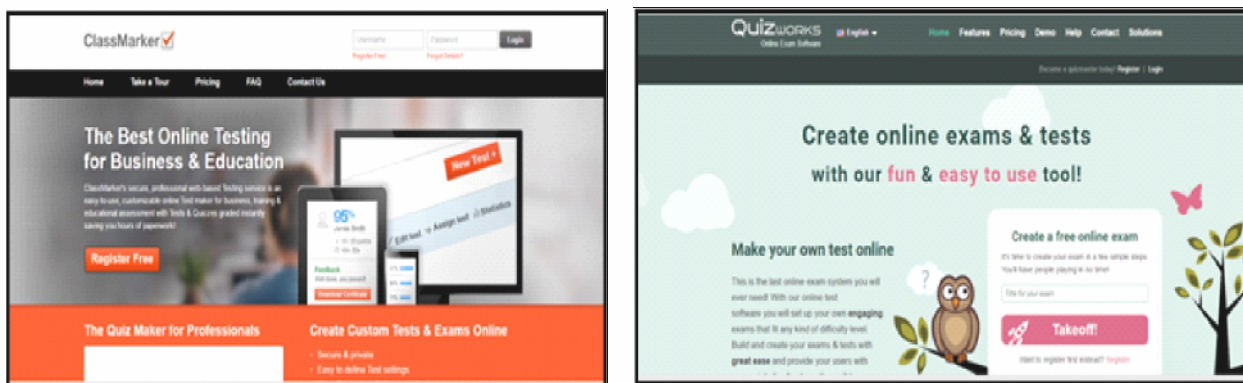
71

Source:http://speedexam.net/Source: https://exambuilder.com/

(*Source: http://www.quiz-creator.com; and http://www.flubaroo.com*)

Activity 3

Select any free online or e-exam creator website/software and prepare an e-exam for the children of your class. Try them in your class and identify the difficulties you have faced in developing and administering it.

....

....

....

....

....

Features of online/e-examination: As you know, pen and paper examinations require well constructed plans for implementation. Similarly, online and e-exams also need to be planned well in advance before they are put into practice. Online and e-exams should have the following features:

● **Examination instructions:** The duration, marks allotted for each section/ question, class, mode of answering, type of questions, etc. must be clearly mentioned.

- **Registration:** Students attempting online and e-exams are required to register or open an account, where in the questions to test their learning will be posted. Either the examination will be available online which can be assessed by entering their login password or else the examiner would set the questions which can be only attempted by the students.

- **Valid time duration of the examination:** The examination will be valid for a fixed time afterwards it is unavailable. This is because the pattern of the exam is set in such a way that the students will not get extra time to attempt question unlike in the traditional examination.

- **Time reminder:** As the students progress in answering, the system itself will generate warnings about the duration of the time left, questions unattended or partially attempted and so on. Such prompts act as reminders for students to complete the examination on time.

- **Submission of answers:** Students can submit answers as a whole or a single question at a time. In both the cases before submitting the answer warning signs would be generated : whether answering is complete, would they like to modify answers, and so on. This helps students in rechecking their answers.

- **Declaration of results:** There are two options as children submit answer. The system may generate the status of answering (whether the answer is right or wrong), or else after submitting the answers of all the questions, the result may appear as a whole.

Advantages of onli]ne/e-examination: It is a fact that lots of preparations are required to conduct paper-pen examinations such as preparing schedule of examinations, development of question papers, arrangement of examination halls, evaluation of answer scripts, declaration of results, arrangement of human power, etc. Even after spending a great deal of time and effort, complaints may arise. In such a situation, online and e-examinations have great relevance. The following are some of the advantages of online/e-examinations:

- Human errors are minimal in e-exams. In conventional exams, the possibility to commit mistakes is high while evaluating answer scripts, entering evaluated grades, identifying malpractices, etc. On the other hand most of the activities in e-exams are digital in nature and hence errors are minimal.

- The habit of indulging in malpractices in multiple ways can be eliminated. As children are provided with computer or other electronic gadgets to attempt examinations, the possibility of malpractices is minimum.

- The conventional exams carry question papers and answer sheets to distant places and within the examination halls but doesn't happen in online/ e-examinations. Thus the physical movement of examination materials is less for online exam.

- The number of the human power required is less in online examinations. Conventional examinations require lots of human power in organising but online/exams require less human power.

- The workload in the conduct of conventional examinations is vast such as in setting question papers, executing of examinations, evaluating answer sheets and so on. While online/exams entail less workload since distribution of question papers is done once, evaluation and declaration of result is ICT based and so on.

Disadvantages of online/e-examination : Let us now discuss the disadvantages of online/e-exams. As we know, in the digital era many of the educational institutions and government agencies utilize technology. But the fact is that technology has some disadvantages. Some of the disadvantages of online/e-exams are as follows:

- The competence is in developing online/e-exams which many of the stakeholders associated with the conduct of e-exams lack.

- Online/e-exams require specialised settings and facilities to conduct, a cumbersome job for institutions and teachers.

- Chances of technical error in electronic machines are expected and as a result examinations may fail.

- The problems of viruses/bugs/internet attacks in software are a threat for online/e-examinations.

- Children need to be trained in attempting online/e-examinations.

- The cost involved in online/e-exams is higher and creates problems on the part of institutions/organisations.

Check Your Progress 4

Note : a) Write your answers in the space given below.

b) Compare your answers with those given at the end of the unit.

5. Discuss the procedures followed in organising online/e-examinations.

....

....

....

8.7 LEARNERS' E-PORTFOLIO AND E-RUBRICS

Generally the schools keep record of marks scored by the children in different subjects. In addition to that, data about their family background, physical records, involvement and achievement in extracurricular activities, etc., are also recorded. Why do we collect such details? No doubt, it is part of evaluation/assessment. The grades/ marks represent children's performance. At present, continuous and comprehensive evaluation (CCE) is practiced in schools in which students are continuously assessed all through the academic sessions on various parameters. Parameters such as learning performance, participation in co-curricular activities, involvement in various activities of the school, etc., are continuously assessed. Thus, the assessment is comprehensive and continuous.

The multiple data relating to a student is recorded systematically. Portfolios document learning performance, activities undertaken, details of participation in co-curricular activities and other relevant data concerning the student throughout the session. Such record is highly useful in evaluating the student's performance during his/her schooling. Since we emphasis the culture of digital teaching-learning, portfolios are replaced with e-portfolios. What are

e-portfolios? **"e-portfolios is a valuable learning and assessment tool. An e-portfolio is a digitized collection of artifacts including demonstrations, resources, and accomplishments that represent an individual, group, or institution. This collection can be comprised of text based, graphic, or multimedia elements archived on a website or on other electronic media such as CD-ROM or DVD. An e-portfolio is more than a simple collection of data-it can also serve as an administrative tool to manage and organise work created with different applications and to control who can see the work. E-portfolios encourage personal reflection and often involve the exchange of ideas and feedback"** (Lorenzo and Ittelson, 2005).

E-portfolio also known as 'digital portfolio' or 'online portfolio' is the record of digital works performed by the student. The digital works include text material prepared in Microsoft word software/PDF or other application software's, video/ audio clips, digital images, digital reports, blog entries, comments posted in discussion for a power point presentations, etc. The digital creations of an individual when stored in digital formats either in online operated portfolio or any digital resources represent an e-portfolio. E-portfolio is the latest trends in teaching –learning as it is an immense tool for evaluating students' performance. By looking at the digital collection, teacher would be able to judge the performance of learning.
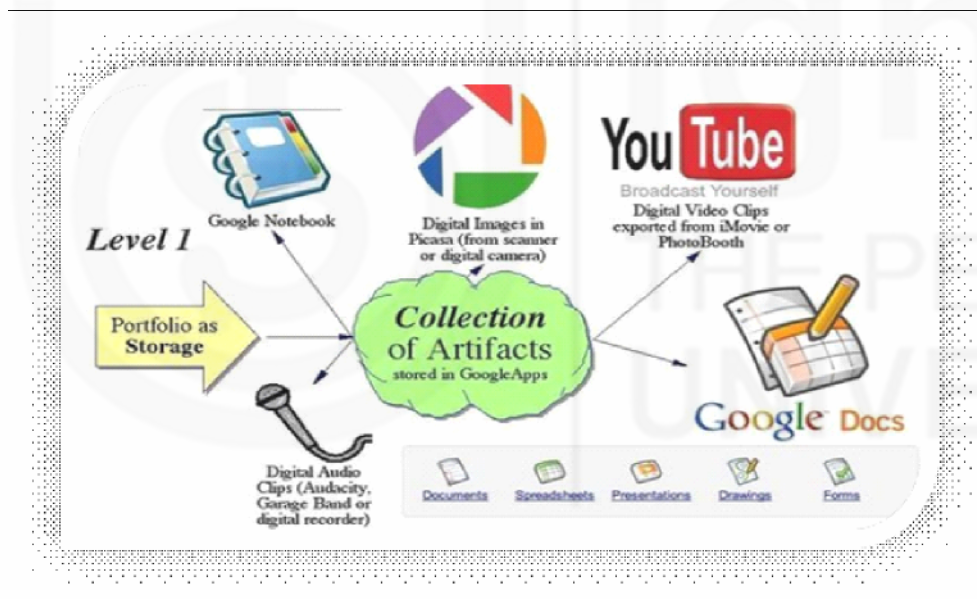

Figure 8.5: Artifacts of e-portfolio

(*Source:* *http://www.informationweek.com/software/7-ways-to-create-e-portfolios/d/d-id/1110673*)

As we know, conventional portfolio is a collection of different non-digital works of a student, such as assignments, reports and so on; while e-portfolios are his/her digital creations like a presentation, project, collection of photography, etc. Why are e-portfolios important in education? First, e-portfolio is a tool for assessment and evaluation. The record of activities performed by the students is accumulated in e-portfolios and at a later stage they are used to judge their performance. Secondly, e-portfolios act as tool for self evaluation. The digital constructions of students will give themselves a status of their performance in various activities. For example, a video clip submitted by a student can be viewed at a later stage, and errors can be identified. Those mistakes can be improved

while developing a video clip second time. Third, it is an evidence of children's work. Many a times, records are being missed in schools but in case of e-portfolios, the digital works are stored in digital formats and the chances of missing them are fairly low.

As a teacher, you would be more interested in gaining knowledge about the ways of developing e-portfolios. The questions might come to your mind such as: How e-portfolios are developed? Is there any software to develop e-portfolio? What digital collections would be incorporated in e-portfolio, and so on. Developing an e-portfolio is not a burdensome activity. As you know, today the teaching-learning process engages children in multiple activities, such as conducting a project, preparing assignments, etc. To prepare the assignments children refer different website, download images, etc. Then they prepare the text in MS word or such application softwares. So here, the assignment prepared by the student is in digital form and the same could be included in the e-portfolio of that particular student. There are number of tools/softwares (free/paid and offline/online), where you can place the digital formations of your students. To create e-portfolio, firstly you have to signup in any particular website. This will generate your account on that particular website. Thereafter, you can start recording the digital works of your children. Some websites are given below (Also the screen shots of home page of some of these websites).

- www.rcampus.com
- www.pebblepad.co.uk
- www.portfoliogen.com
- www.mahara.org
- www.digication.com
- www.taskstream.com
- www.pebblelearning.com
- www.myefolio.com
- www.foliotek.com
- www.epsilen.com
- www.foliospaces.org

Figure 8.6: Commercial/Open Source E-Portfolio Creation Tools/Websites

Source: http://www.rcampus.com/ Source: http://www.pebblepad.co.uk/

Source: https://mahara.org/Source: http://www.portfoliogen.com/
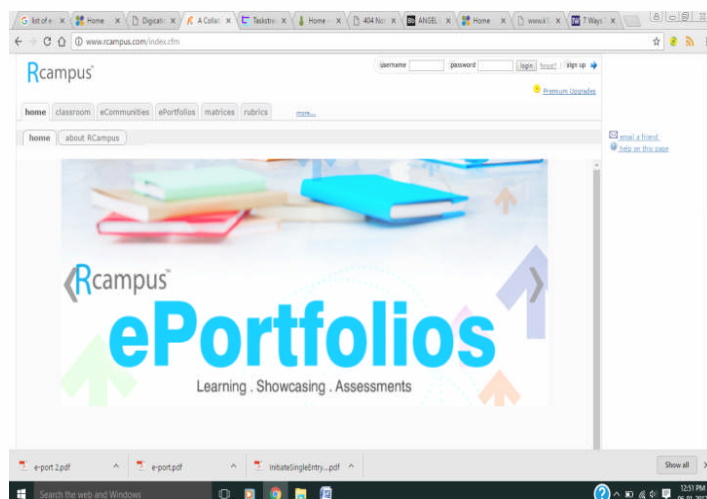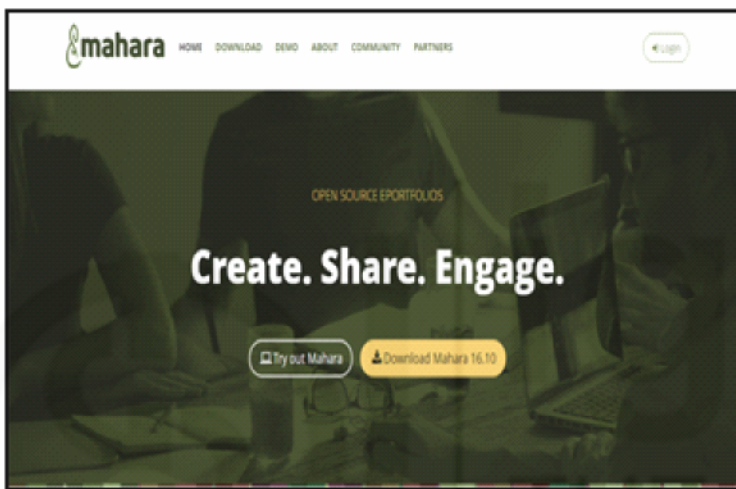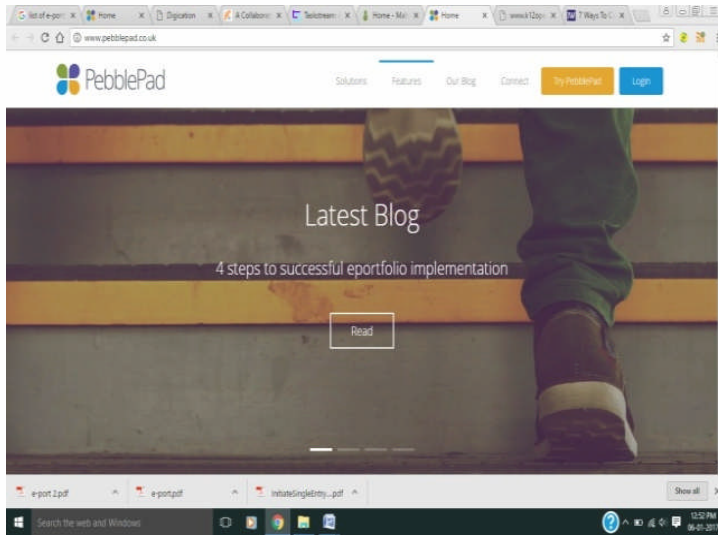
Activity 4

Select a free e- portfolio creator (even google apps) and develop one e-portfolio suitable for children of your class. Describe the process that you have carried out in developing e-portfolio.

....

....

....

Types of e-portfolios: As discussed in preceding sub-sections, e-portfolios are also known as digital portfolios, web portfolios and online portfolios. Basically the nomenclature is based on the format of portfolios. Generally, online portfolio functions with the help of internet while other formats are internet independent. Based on the functions of e-portfolios, they are classified into three; developmental (or working), assessment and showcase.

● **Developmental (working) e-portfolios:** These e-portfolios collect and record students' activities. For example, students are assigned a project of two months as a learning activity. During the period of the project, they have to submit short reports every 25 days in MS Word format. Such a digital report will be stored in the database of e-portfolio. Thus developmental e-portfolios record activities over a period of time. The primary objective of recording the progress is to communicate the progress to the students and the teachers.

● **Assessment e-portfolios:** These e-portfolios collect and record children's activities for the entire duration of the programme/course. The primary objective is to assess and judge the competency of children in specific skills/ courses and programmes. These portfolios are submitted for assessing and evaluating the performance in a learning activity/programme. For example, digital formats of assignments in a single subject submitted during his/her studies in eighth class.

● **Showcase e-portfolios:** These portfolios exhibit exceptional works/skills of an individual. This may include works in preparing reports, inserting graphic/pictures in preparing write-ups, development of quality video/audio clips and so on. These portfolios happen to judge one's performance, selection to a job, etc. Thus, as the name indicates, showcase e-portfolios are images on individual persona.

Check Your Progress 5

Note : a) Write your answers in the space given below.

b) Compare your answers with those given at the end of the unit.

6. Define e-portfolio? Discuss the artefacts 'recorded in e-portfolios.

....

....

....

....

E-Rubric: During learning, children participate in various activities, which are assessed to judge their performance. Generally achievement tests are employed in evaluating the children. What is rubric? Rubric is a set of guidelines on which students' performance is assessed. A rubric is a unified criteria for students' work that includes descriptions of the level of performance. Thus a rubric will have few set criteria and description of levels of performance for those criteria. Rubric is more descriptive in nature such that, the skills to be evaluated are described as various levels and that the level of the students is identified for that particular skill. Even though descriptive, the level of performance is used in judging the performance of individual. Thus the main purpose of rubric is to assess the performance. For example, if you aspire to measure the skill of children in using a thermometer to read temperature; the rubric prepared for

the same would include skills such as holding thermometer, measuring temperature, recording the reading and so on. Thus the parameters included in rubric are holding thermometer, measuring temperature, recording the reading and so on.

To develop rubric, the assessment levels such as very poor, poor, average, good, very good for any particular parameter will be assessed (See Table 8.1). You can also refer Unit-7 of this Block. These parameters are developed as template of rubric. For example, the rubric template would look like the following:

Table 8.1: Rubric Template

Parameter	Level 1 (Very Poor)	Level 2 (Poor)	Level 3 (Average)	Level 4 (Good)	Level 5 (Very Good)	Total Score
Holding Thermometer						
Measuring temperature						
Recording Reading						

Now let us understand e-rubric. In a digital learning atmosphere, the rubric is prepared in digital format. The rubric that we have discussed above (Table 1) is prepared in non-digital format (may be in normal paper or other non-digital resources). While e-rubric (electronic-rubric) is prepared by any digital means. Generally, e-rubrics are prepared with the help of rubric creators/websites/ online tools, etc. The e-rubrics are stored in either computers, mobiles, CD-ROMs, hard drives or any storage devices, which can be retrieved and manipulations are possible at any time. Thus whenever a task is assigned, the teacher would prepare its rubric template and the performance level will be recorded against each criteria. At the completion of task, the rubric will be used to assess the performance of the student.

As a teacher, you may be interested in gauging information on tools used to develop e-rubrics. Similar to e-portfolios, you may find websites either paid/ free (commercial/non-commercial; open source), which may be employed to develop e-rubrics. Few such websites are mentioned below:

- http://rubistar.4teachers.org

- http://www.assessmentfocus.com

- http://www.rcampus.com

- http://www.teach-nology.com

- http://www.learner.org

- http://www.essaytagger.com

- http://www.edudemic.com

- http://edtechteacher.org

\

Figure 8.7: Home page of websites used for creating e-rubric

(*Source*:   *http://rubistar.4teachers.org.,http://www.assessmentfocus.com/.,
http://www.rcampus.com/.,   http://www.teach-nology.com*)

Importance of e-rubric : We have discussed the basic purpose of building of e-rubric. Basically, e-rubric helps us evaluating a students' performance. Let us elaborate its importance in the teaching learning and evaluation process. In a teacher's point of view, e- rubric is a tool to set his learning objectives. While deciding on the topic or activity to be transacted, he/she should think of the learning objectives and skills to be developed among students. Accordingly different criteria will be developed and recorded in the e-rubric template. Later, the students will be evaluated based on the e-rubric and that will help you, as a teacher to know the performance of the students. This will also help you to reflect on your teaching process. In the case of students, e-rubric is an instrument for self reflection. While teacher identifies the criterion for any particular learning activity/task, the same could also be discussed with students. Thus students are aware of the criteria based on which they are going to be assessed. Thus as students proceed with the task, they would pay attention on the criteria leading to development of learning skills as desired by the teacher. Such a process provides scope for students' self-reflection and to become aware of their own learning.

Steps to Create E-Rubric

As we have discussed, e-rubric is an effective tool to assess children's performance. Generally teachers can prepare two types of e-rubric; one general e-rubric and the other specific e-rubric. In general e- rubric, general activities and its criterion will be included. For example, writing an assignment; in writing assignment, the criterion includes, writing a good introduction, body (content), conclusion, supporting examples, figures and so on. Irrespective of the subject, the format followed in writing assignment remains the same. In specific e-rubric, criterion for specific activities is included. For instance, performing experiment in a chemistry laboratory. In the laboratory, each student is required to perform certain activities; may be mixing chemical in a specific proportion, taking measurements and so on. Such tasks are specific tasks and each student is supposed to develop those specific skills. The following steps should be used to develop an e-rubric.

i.   Select the learning activity from your subject of teaching. It can be either general or specific in nature

ii.   List the criterion relating to the learning activity. At this stage, children may also be encouraged to suggest criterion.

iii. Decide levels for each criterion ranging in different modes; can be 1 to 5, low to high, minimum to maximum, poor to high and so on. A five point level would be reasonable for assessing children on varied tasks.

iv.   Develop re-rubric template. The e-rubric template will show the criterion and different levels for specific or general learning activity. You may also leave a column for students to write their comments. The e-rubric can also be shared with students.

v.   Execute the learning activity and record scores against each criterion in the e-rubric template.

vi.   Asses and evaluate children's performance based on the scoring in e-rubric.

vii. If required, revise the e-rubric based on the feedback and performance of children.

Activity 5

Identify a task of your choice and develop its criterion in a rubric template. Carryout the task in your class and record the level of children against each criterion.

....

....

....

....

....

Check Your Progress 6

Note : a) Write your answers in the space given below.

b) Compare your answers with the one given at the end of the unit.

7. What is e-rubric? Why e-rubrics are important for students?

....

....

....

8. Describe the steps for developing an e-rubric.

....

....

....

8.8 USE OF ICT TOOLS FOR PREPARING TESTS AND ANALYZING RESULTS

In the earlier section, you have studied that ICT has immense opportunities in assessment and evaluation. Till now we have discussed, various ICT resources, online/e-examination, e-portfolio and e-rubric. As a result, today, the potential of ICT is being explored in multiple ways. In this Section we will discuss specific tools, websites and resources which can be used to construct questionnaires and tools that have inbuilt options to analyse test results.

ICT may be employed at three stages for assessing children, they are:

- before beginning of the teaching session,

- during the teaching session, and

- after the teaching session.

In the beginning, assessment will help you to identify the learning level of children, while during teaching session, it will give you an idea of how far children have achieved learning objectives and at the end, assessment helps to find out the achievement of the learning objectives. As we have discussed earlier, teachers
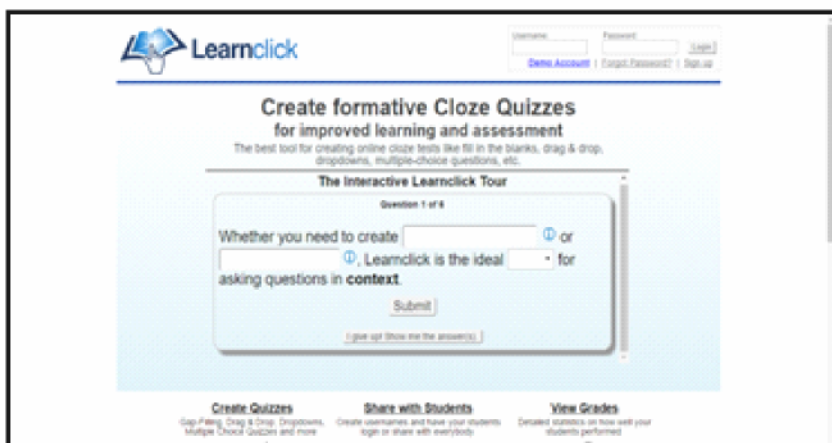
have important role in utilizing ICTs. It is up to the teacher to select appropriate technology or tool to assess children at various levels of teaching. You can make use of mobile technology to assess children's learning in the beginning or during teaching session. Even computer made tests can also be employed. At the end of the teaching session, computer made tests and online or e-exams are preferred. The following websites (home page of some of the websites are also given in Figure 8.8) can be used to create test of your choice.

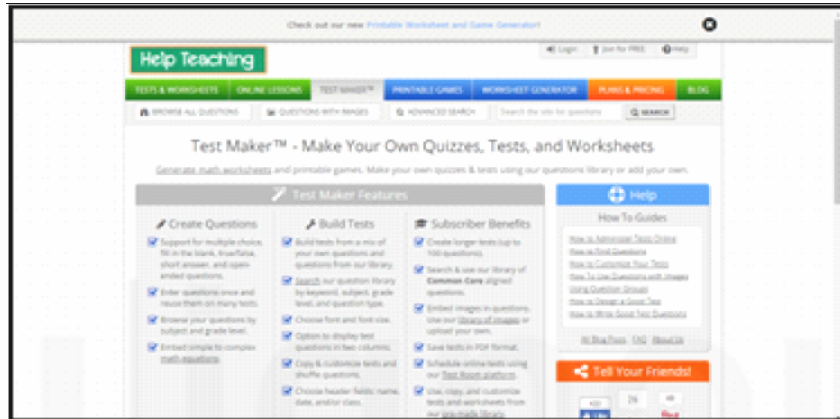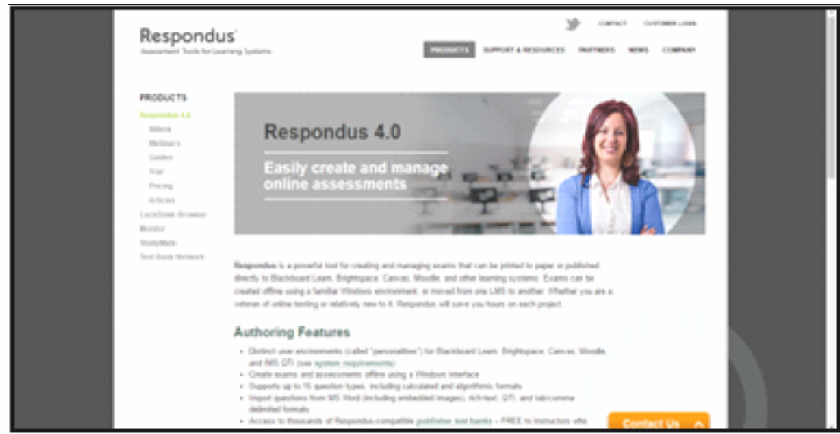| Commercial website	Free website
https://www.classmarker.com/	http://www.fiftysneakers.com/
http://www.contentgenerator.net/	http://funnelbrain.com/
http://www.easytestmaker.com/	http://web.uvic.ca/hrd/hotpot/
http://elearningbrothers.com/	http://www.kubbu.com/
http://www.ecomscotland.com/	http://www.cram.com/
http://www.exambuddy.com/	https://quizslides.com/
https://www.igneon.com/	https://www.studyblue.com/
http://www.mygradebook.com/	http://www.btl.com/surpass/
https://www.onlineexambuilder.com/	http://www.schoolsonic.com/
http://www.proprofs.com/quiz-school/	http://www.learnclick.com/
http://www.questiontools.com/	http://www.mobilestudy.org/
http://www.questionwriter.com/	
https://www.quia.com/web	
http://www.quiz-builder.com/	
http://ww17.quizlet.cm/	
http://www.respondus.com/	
http://www.helpteaching.com/	
http://www.qarbon.com/	
http://eng.smartlite.it/	
https://yacapaca.com/	

Figure 8.8 : Commercial and free websites for ICT tools

In conventional tests, answers are written on papers while ICT tests employ mainly computers. A teacher has to devote lot of his/her time to conduct conventional examinations. In the case of computer made tests, the time devoted is considerably less. Let us understand as to how the time gets reduced in organization of the ICT based of tests. To substantiate, let us pick a free tool from the list given in Figure 8, http://www.learnclick.com/. Learnclick is free application software that provides choices in developing tests of varied nature. You will find options, such as Blank Boxes and Dropdowns, Generated Dropdowns, Drag and Drop, Matching, Multiple Choice (one answer correct), Checkboxes (several answers correct), Essay (open-ended question), Description, etc. As you browse 'learnclick', the homepage will appear where you will find sub-links to create tests. The screen shot is given in Figure 8.9:
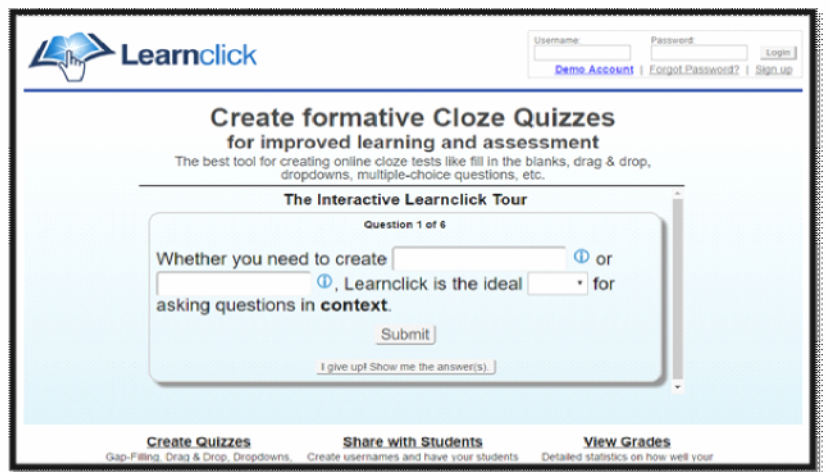
Figure 8.9 : Learnclick – A free application software

(***Source:***   *http://www.learnclick.com/)*

After opening the home, click on the sub link "Create Quizzes, which will take you to another window as shown below.
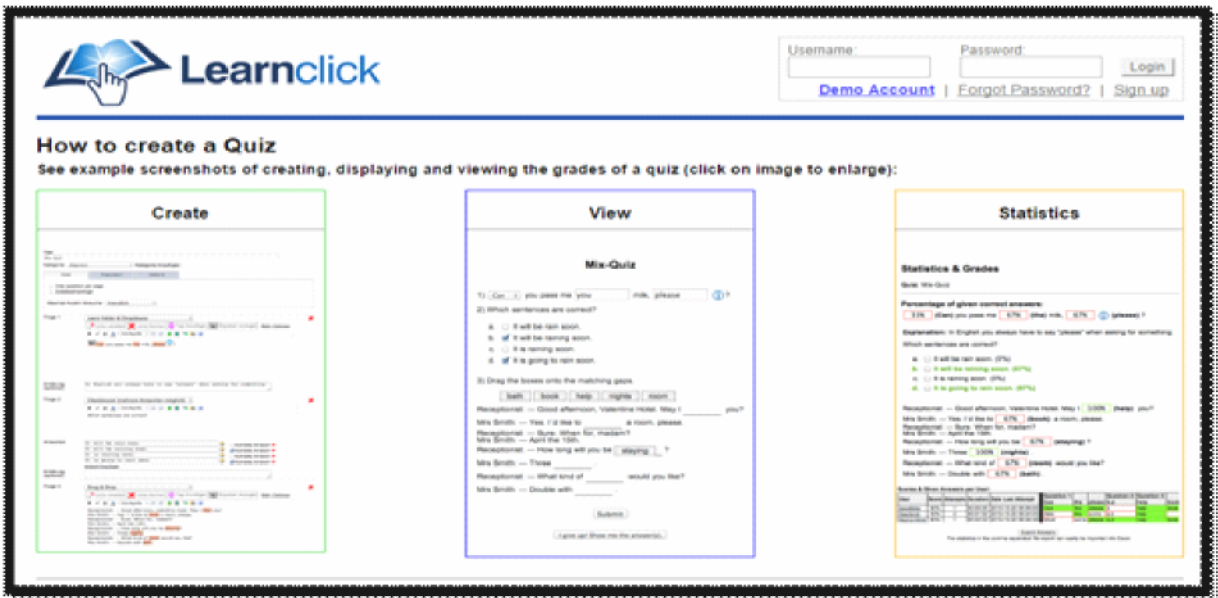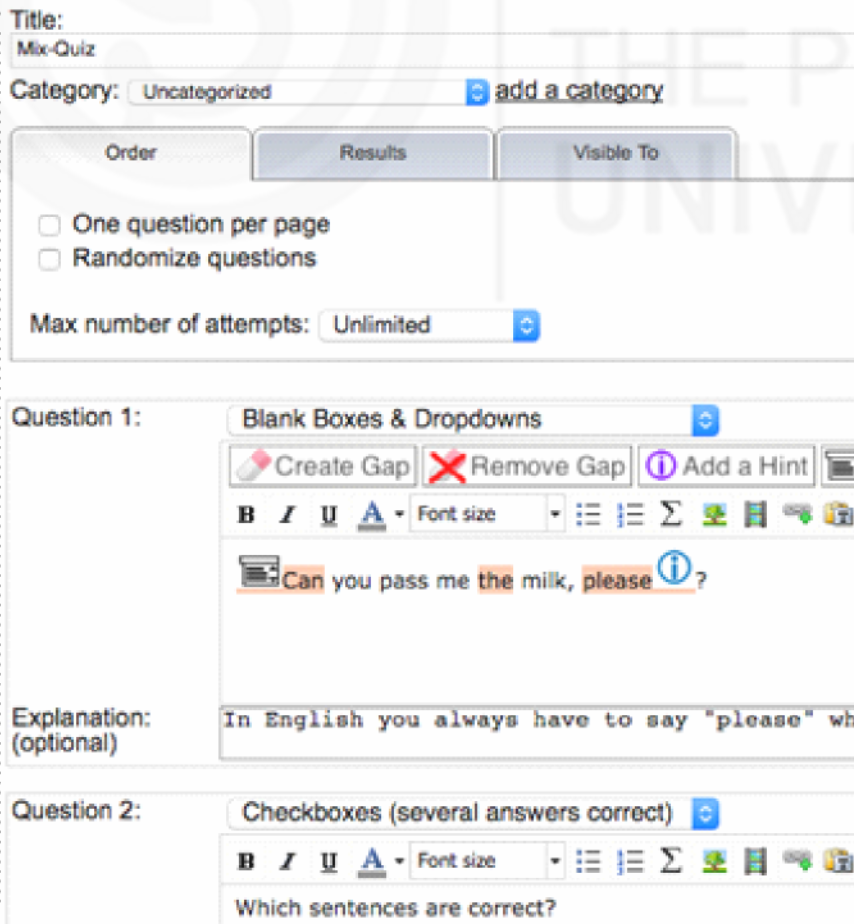
Figure 8.10 : Sub-link of Learnclick website to create quiz

Thereafter, you may insert questions and it may be distributed to your students for answering. After attempting the test, the system itself will evaluate and test results will be displayed. Closer screen shots are given below:
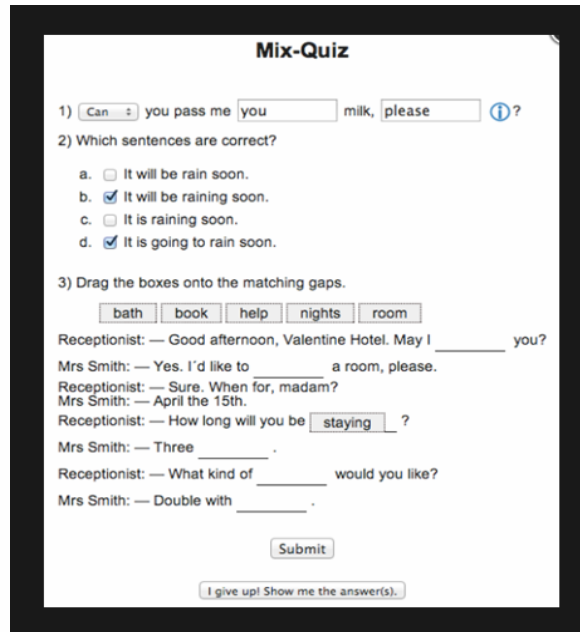
Figure 8.11 : Closer screen shots of Mix-Quiz

Those tools mentioned above are specially designed to create tests/questionnaires/ quizzes. In addition to these Microsoft Word, PowerPoint, etc. can also be used in developing tests. Many e-learning softwares, Learning Management Systems (LMS), survey tools, e-contents, blogs, wikis, e-books, chat services, virtual classrooms, e-mails, PDF readers, concept maps, etc. also have many inbuilt features to test performance of the students. Let us discuss few of them.

You are aware that Microsoft Power Point is extensively used software for presentations. Let us see, how PowerPoint is used to develop tests. Think of a classroom situation. Ms Aparna, a secondary school teacher is teaching the concept of 'force'. After teaching the concept, she tried to check 'whether children have grasped the concept or not. For this, she prepared a multiple choice test in Power Point. The screen shots of the test are given below. The test is prepared in three slides. The first slide contains the question and four options which the student has to pick out the correct option by pressing the button given against each option. As the button is pressed, it will then jump either to second or third slide which will further indicate whether the option chosen by student is right or wrong. If it is wrong, the children have to press the button given in the slide, which will resume to the very first slide.
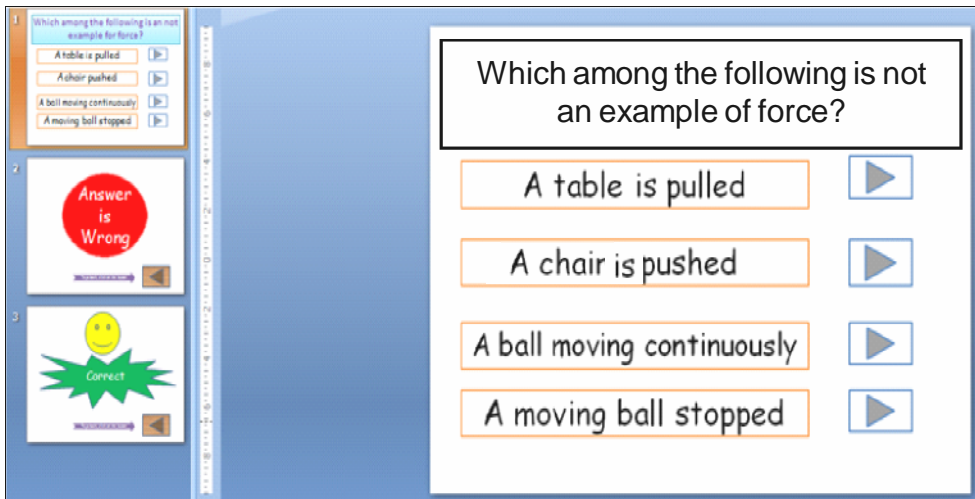
is

Figure 8.12 : Multiple choice test item in power point

Activity 6

i) How will you use blogs for evaluating your children's performance? Suggest a plan.

ii) Develop a test in PowerPoint software and execute it in your classroom. Prepare a report on the difficulties faced by you in executing the test in your class.

....

....

....

....

....

8.9 LET US SUM UP

As you have studied, use of ICT in the process of teaching-learning and assessment is not an isolated activity but it is highly appreciated when ICT becomes an integral part of the process. NCF, 2005 also highlighted such aspect that ICT should be an integral part of teaching-learning process and assessment (refer Unit1 and 2 of Block-1 of this Course).

In this Unit, we have discussed the importance of ICT in assessing students' performance and also role of the teacher in technology-based evaluation. You have also learnt in this Unit about the concept of online/e-examination and its use. E-portfolio and e-rubrics are used as assessment tools in our school system. In this Unit, the concept and the process of developing e-portfolio and e-rubrics are also discussed and the online sources for developing portfolio and rubric have been highlighted. The Unit ends with a discussion on various application softwares/tools/websites such as content generator, learnclick, respondus, help teaching, etc. which are useful for the creation of online tools and tests. As a teacher, you can try to develop the online tools on any subject and topic of your choice as per the list of online sources given in this Unit.

8.10 REFERENCES AND SUGGESTED READINGS

Chen, H.L, and T. Penny Light, (2010).*Electronic Portfolios and Student Success: Effectiveness, Efficiency, and Learning*, Washington: Association of American Colleges and Universities.

Clark, J. Elizabeth, (2010). "The Digital Imperative: Making the Case for a 21st CenturyPedagogy," *Computers and Composition*, 27.1 (2010): 27-35.

Susan M.B.,& Anthony J.(2008).*Assessment and Grading in Classrooms* (p. 201).Upper Saddle River, NJ: Pearson Education.

Websites Referred

http://ftp.jrc.es/EURdoc/JRC76971.pdf

http://www.teindia.nic.in/e9-tm/Files/ICT_Documents/Alternative_paradigm-%20ICT-integrated-assessment-constructivist-perspective.pdf

http://www.icbl.hw.ac.uk/ltdi/implementing-it/using.htm

https://elearningindustry.com/free-testing-tools-for-online-education

https://www.classmarker.com/

https://www.onlineexambuilder.com/

http://speedexam.net/

https://exambuilder.com/

http://www.quiz-creator.com

http://www.flubaroo.com/ ?utm_campaign=Listly&utm_medium=list&utm_source=listly

http://www.rcampus.com/index.cfm

http://www.pebblepad.co.uk/

https://mahara.org/

http://www.portfoliogen.com/

http://rubistar.4teachers.org/index.php

http://www.assessmentfocus.com/rubrics-rubric-makers.php

http://www.rcampus.com/indexrubric.cfm

http://www.teach-nology.com/web_tools/rubrics/

http://www.contentgenerator.net/

http://c4lpt.co.uk/

http://www.learnclick.com/

http://www.respondus.com/products/respondus/index.shtml

http://www.helpteaching.com/create.htm

http://www.btl.com/surpass/

https://moodle.org/

https://www.edmodo.com/

http://anz.blackboard.com/

https://classroom.google.com/h

https://hotpot.uvic.ca/

8.11   ANSWERS TO CHECK YOUR PROGRESS

1. Because it is time saving, economical, interesting, and motivating.

2. Diagnostic evaluation is conducted to know the learning difficulties and accordingly to provide remedial instructions. Formative evaluation is used to know the learning progress whereas summative evaluation is used for grading and certification of the students.

3. Develop an e-content on any topic. Include multiple choice questions in it. Let students attempt it. (You may use e-content developing software like eXe, etc)

4. To select appropriate technology considering weightage to pedagogy and content and standard of teaching

5. Allotting user ID and password, attempting questions, saving answers, submitting answers, etc.

6. E-portfolio is digitized collection of artefacts including demonstrations, resources, and accomplishments that represent an individual, group, or institution. Collection of digital work of any individual/student, images, pdf, audio-video, etc. For the rest of the answer, refer Figure 8.5.

7.   Rubric prepared in digital format is called e-rubric. It helps students to asses their learning.

8. Select learning activity, identify the criterion for that learning activity, assign levels for each criterion, develop rubric template, Use rubric to assess children's learning for that particular activity.