
UNIT 8 MANAGEMENT OF EVALUATION (ASSESSMENT AND EXAMINATIONS)

Structure

- 8.1 Introduction
- 8.2 Objectives
- 8.3 Evaluation — Meaning and Place in the Educational Process
 - 8.3.1 Nature and Concepts of Measurement and Evaluation in Education
 - 8.3.2 Place of Evaluation in the Educative Process
- 8.4 Purposes and Functions of Evaluation
- 8.5 General Defects in the Present Evaluation System
- 8.6 Characteristics of a Good Test
 - 8.6.1 Validity
 - 8.6.2 Reliability
 - 8.6.3 Usability
- 8.7 Evaluation Techniques
 - 8.7.1 Evaluation Tools/Devices
 - 8.7.2 Classification of Written Tests/Examinations
- 8.8 Construction of an Achievement Test
- 8.9 Some New Developments in Evaluation
 - 8.9.1 Grading
 - 8.9.2 Question Banks
- 8.10 Use of ICT
- 8.11 Management of Evaluation
- 8.12 Case Study
- 8.13 Unit-end Exercises
- 8.14 Let Us Sum Up
- 8.15 Points for Discussion
- 8.16 Suggested Readings

8.1 INTRODUCTION

Assessment of the progress of the pupils in education is useful to both - the teachers and the students. Teachers can know the effectiveness of the learning experiences provided in the school. Students can feel the reward of their efforts in learning. So knowledge of the various techniques of evaluation and the preparation of valid tests to be used with the pupils is essential for us as teachers or administrators of the school. The principal or headmaster plays a key role in guiding teachers and in planning various schemes for the development of the school. The head should be well acquainted with the various aspects of education including evaluation. He/she should also keep abreast of new developments and reforms in evaluation methodology. This unit deals with the process of evaluation and some of its related aspects.

8.2 OBJECTIVES

The main purpose of this unit is to understand the evaluation process. After having gone through this unit, you should be able to:

- distinguish between measurement and evaluation;
- describe the features and functions of education;
- identify some defects in the present day system of examinations;
- describe various types of evaluation devices;
- describe the characteristics of a good test;
- understand and develop an achievement test;
- explain the system of grading in evaluation; and
- describe the meaning and utility of question-banks.

8.3 EVALUATION — MEANING AND PLACE IN THE EDUCATIONAL PROCESS

In this section, we shall try to understand some terms used related to evaluation and their relationships with the objectives and learning experiences in education.

8.3.1 Nature and Concepts of Measurement and Evaluation in Education

Everyone of us is well-aware of the examinations or tests conducted in our educational institutions. These are means to measure pupil's achievements in various subjects e.g. Mathematics 35/50, English 30/50. Examinations are normally meant to measure scholastic achievements.

Some terms have been used to describe the abilities or performance of individuals like measurement, evaluation assessment, examination etc. These may mean same thing to a layman, but technically, they are not synonymous in education. Let us try to understand them.

Measurement is assigning some numerals according to certain rules for one's abilities or traits. It is thus a quantitative aspect of the abilities. Suresh's 70% score in Mathematics is a measure of his achievement in Mathematics. Laxmi answered 5 questions (as change of number/gender) correctly and got 5/10 marks in a test of English. Ramesh is 160 cms in height. These are examples of measurement.

Evaluation is relatively a new technical term that has recently emerged in education and has gained wide acceptance. It is wider in scope and includes measurement as well.

To understand it, we take some examples. We analysed the list of Mathematics test of the class of Suresh. We found that he scored 70% and stood second in the class. We also found that Laxmi got 5/10 marks and is average in the class of English. Ramesh is considered to be of short stature in his group with a height of 160 cms. These descriptions are evaluation of abilities or traits. These are quantitative as well as qualitative ranks or positions in the group characterise the qualitative aspect.

Gronlund (1971) explains this fact as: Evaluation = Measurement + Value Judgement.

Evaluation characterises value or worth of ability/phenomenon. The comparison can be made with social, cultural and scientific standards. We may use various adjectives

for comparison or judgement like good language/poor, good/bad, bright/normal/poor, long/average/short, rank etc.

Assessment is usually used to mean evaluation. It may be an estimate of how teachers assess their pupils and this may be through tests and sometimes on the basis of general observations in class.

The term 'Examination' is generally used in the context of a pupil's performance in various subjects in school or college. It is a process to measure his/her academic or scholastic achievements.

Tests are devices to gather information through some specified tasks presented to individuals. We come across various types of tests in education e.g. oral, written and practical; group or individual; performance etc. Sometimes 'examination' is also used to mean a device or a test.

Examination is a process which involves planning, construction/selection of tools, administration and scoring; and compilation of results. It is an aspect of evaluation and uses various tools to gather information about pupils.

Evaluation has a broader meaning than measurement and examination. Some significant features of evaluation are as follows:

- i) Evaluation should be comprehensive in nature. It covers all aspects of pupil's personality - physical, mental, social, moral.
- ii) Evaluation should be a continuous process of assessment and forms an integral part of the educative process. Even in the early stages of the session, a teacher may evaluate pupils to know their capability for learning and may help them accordingly.
- iii) There is an inter-relatedness between the objectives (ends), learning experiences (means) and evaluation (evidence).
- iv) Evaluation has quantitative as well as qualitative aspects.
- v) Evaluation is a co-operative process. It involves teachers, peers, parents and students in appraisal of various aspects of pupil's personality.
- vi) Evaluation employs various devices and sources of secure valid evidence of desired change or achievement of objectives.

We often talk of continuous comprehensive evaluation but evaluation can be in intervals and in parts too.

8.3.2 Place of Evaluation in the Educative Process

Education is imparted in the schools or colleges to promote some aims and values in relation to the individual and society. The educative process is considered a tri-polar process by various educationists like Bloom. This is represented in Figure 1.

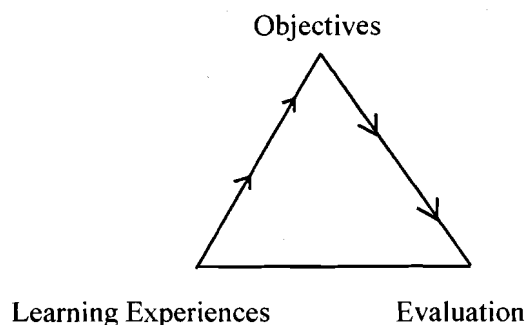


Figure 1

Firstly the objectives are decided and expressed explicitly in expected behavioural changes in the learner. The second step is to plan and organise the learning experiences to achieve the decided objectives. These may be in the form of face to face teaching, self-learning through books or other media, discussions etc. The third step is evaluation using various tools or devices. These devices may be oral, written, practical, performance, observation, interview, pupil's diary etc. These give evidence about changes in education of learner (objectives).

The both way arrows indicate their inter-linking and inter-dependence.

Objectives help to select the type of the learning experiences and the evaluation tools. If we aim to understand spoken English, we should converse and listen to speeches of persons or radio/T.V. and similar other programmes in English. Our evaluation process may also assess comprehension of spoken English. An individual should be able to answer or explain situations presented through English speech. For learning to speak English, conversation or participation in discussions might be used. Similarly, visiting a historical place may be helpful to enhance our knowledge and interest in History.

The type of learning experiences organised in the class explains the achievable objectives. Learning experiences determine the form or mode of evaluation process. Programmes based on listening to spoken English help us to achieve similar objectives (comprehension of spoken English). The evaluation technique, then, will be based on the presented situations or experiences.

The evaluation process also explains the achieved objectives and success of learning experiences. It helps to know which, and to what extent, objectives have been achieved. It also enables on to examine the success of organised experiences. Spoken English and oral Mathematics are not testable through paper-pencil tests.

The above illustrations explain the mutual dependence and relationship between the tri-polar process - objectives, learning experiences and evaluation.

As a school Head you should be aware of this tri-polar process. It can help you to plan, guide and implement the various programmes.

8.4 PURPOSES AND FUNCTIONS OF EVALUATION

Evaluation is an essential aspect of the educational process. Its main purposes and functions are briefly taken up for discussion as below:

- i) *Assessment of abilities:* Schools and school education Boards (like CBSE) conduct examinations to assess the achievements and abilities of students in different subjects and classes.
- ii) *Diagnosis:* Tests help teachers to know students' weaknesses or lacuna in their subjects. Some diagnostic tests enable us to know specific weak points.
- iii) *Guidance:* On the basis of achievement in various subjects or other abilities, students can be helped to plan for educational courses or jobs for future.
- iv) *Grading:* Students can be categorised into Divisions (I, II, III etc.) or grades (A⁺, A, B⁺, ...)
- v) *Selection:* Some schools and colleges conduct tests to select students for various courses. Tests for engineering or medical colleges serve this purpose. Banks recruit clerks etc. through the tests.

- vi) *Prediction:* Some tests are planned to serve the purpose of predicting probably success in a particular job or course. The examples of entrance tests for engineering, CAT (Common Admission Test) of IIMs, B.Ed. admission test etc. fall under this category.

The above is only a brief list. Evaluation gives feedback to teachers to improve their teaching. It motivates pupils to excel and compete with others. It provides the basis for remedial teaching, action research, and to search and select resources to improve the efficiency of the learning experiences.

8.5 GENERAL DEFECTS IN THE PRESENT EVALUATION SYSTEM

Evaluation is an essential part of the educational process as discussed earlier. At present, evaluation is done only by academic examinations. But the present examination systems are unable to fulfil the desired objectives. These have been severely criticised. The University Education Commission (1948-49) mentioned that if a single reform in the education is to be suggested, then it should be the improvement of examinations. As such, examinations have earned a bad name and have been labelled as blood suckers, a dead hand of education, a necessary evil, an enemy of real education and similar others.

Some of the main defects as follows:

- i) Examinations lack specificity of objectives and purposes. School or college examinations test achievement in subjects. But they are used by employers to select candidates for various jobs as well.
- ii) Examinations test the crammed information and fail to assess the examinee's interests, creativity and understanding in depth.
- iii) Examinations partially test the knowledge of course content. Normally pupils have to answer 5 - 6 questions out of 8 - 10.
- iv) Examinations do not give any credit for regularity of work, participation in discussion in the class, etc.
- v) Students scoring equal marks are considered equal for various purposes although they attempted different questions in the same of different subject(s).
- vi) Examinations are found unreliable, especially the essay type ones. This may occur due to ambiguity of language of the question, lack of instruction for marking and subjective approach of the examiners.

Obviously, our examinations system suffers from many defects. Some efforts have been made to improve by using the semester system, internal/sessional work assessment. Curriculum is sometimes divided into 5-6 sections and at least one question is required to be answered. This ensures wider coverage of the syllabus for the testee.

Check Your Progress 1

Note: Write your answers in the space given below.

- 1. Distinguish between measurement and evaluation.

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

2. Mention three main defects of the present day system of examinations.

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

3. Give an example to show how the objectives determine the type of learning experiences and selection of evaluation technique.

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

4. Describe two main functions of evaluation.

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

8.6 CHARACTERISTICS OF A GOOD TEST

The main function of education is to bring about desirable changes in the education of learners (objectives). We conduct tests/examination to assess the progress in achievement of set objectives. A good test should enable us to measure the achievement or ability accurately and economically. Here we take three main characteristics of a good test - validity, reliability and usability.

8.6.1 Validity

A test is said to be valid if it measures what it is supposed to measure. Cronback defines validity as "the extent to which a test measures what it purports to measure".

Oral communication ability in a language can be properly assessed through speech or discussion techniques. So these are valid tools here. But written examinations cannot be a valid tool for assessment of oral communication. We cannot say that one who writes good answers in the examination would be efficient in oral communication. A class test in arithmetic can be valid to know the achievement in arithmetic and not for diagnosis of weakness of the pupils. Practical and theoretical examinations conducted in science subjects judge the different types of abilities and objectives. One cannot be used for the other. They are valid for their related content and objectives.

Normally teachers prepare achievement tests for their classes. A test should cover the given content and the specified objectives (like knowledge, skill and application) for accurate measurement. If these two conditions are satisfied, then it is said to have content validity.

There are some other types of validity also. One easily understandable is predictive validity. This indicates predications of success of an individual in future in accordance with the criterion of a test. National Talent Search Tests for school students are conducted in India. These are supposed to select students who can excel in specified fields in future. Selected ones are awarded scholarships. Banks also conduct test to select clerks on the predictive validity concept.

We, as teachers, are generally concerned about content validity of achievement tests conducted by the school or state school boards of education. So such tests should try to cover the prescribed syllabus and the specified objectives. Their proportionate

weightage is also given sometimes. If not, the teacher may decide weightage (assigning marks) to each selected topic and the each objective.

8.6.2 Reliability

A test is reliable if it gives consistent results of measurement on its administration at different occasions. Pupils should score similar scores if the test is repeated. Its administration may be after some gap of time and the condition is that pupils do not receive any further training of related test content.

Reliability may be affected by vague wording of questions/items, general attitude of particular teachers, lack of instructions about marking, circumstances of examinee (as ill health) etc.

Some precautions may be taken to increase the reliability of a test. Length of test increases reliability and reduction in size decreases it. Language of instructions in items should be clear. This helps the testee to know the specific scope and desired length of the answer. "Write a short note on population problem" is not a good item. It can be improved as, "Write a short note of 200 words on the problem of over-population in the Indian context".

Further, in an objective question, multiple choice type, there should 4 -5 alternatives which resemble the correct answer. True-false type should preferably be avoided. Items of varying difficulty level should be included but a large number of items should be of moderate difficulty.

One relation between validity and reliability is worth noting. A valid test is reliable. High reliability does not necessarily imply validity.

Estimates of Reliability

We consider three simple methods to estimate reliability of tests.

- i) Test-retest method
- ii) Parallel form method
- iii) Split-half method

Test-retest Method

In this method, a test is conducted twice on the same group of pupils after a gap of some time. Thus each individual gets two score. Here we calculate Pearson's correlation coefficient between two sets of scores of the individual. This is known as test-reliability or measure of stability. This is useful for the tests which contain a large number of items to minimise the carry-over effect of memory. The reliability of some intelligence and personality tests can to be calculated by this method.

Parallel Form Reliability

An equivalent test (or parallel form) is also prepared in addition to the original test. These two tests are administered to the same group of pupils on the same day. Then Pearson's coefficient of correlation between these pairs of scores of individuals gives the reliability measure. This is called parallel form or equivalent reliability. It is commonly employed for teacher-made or achievement tests.

Split-half Method of Reliability

A single test is divided into two equivalent parts or sub-parts, say A and B. A may contain odd numbered items (1, 3, 5, ...) and so B should contain the even numbered ones (2, 4, 6...). The test is administered only once. The total scores for half part A and total score for half part B are separately noted for each pupil. A correlation coefficient is calculated for these sets of scores. This gives the reliability of half test

(r_{hh}). The (r_{hh}) is used to calculate reliability of the full test (r_t) using the Spearman-Brown proficiency formula as follows:

$$(r_t) = \frac{2 \times \sqrt{r_{hh}}}{1 + \sqrt{r_{hh}}} = \frac{2 \times \text{reliability of half test}}{1 + \text{reliability of half test}}$$

Some other methods of estimating reliability are also available using Kuder-Richardson and Cronback alpha formulae. But these involve some complex statistical calculations. The teachers and administrators interested in intensive study of the topic are advised to consult books like Educational and Psychological Measurements and Evaluation by Stanley and Hopkin or some other ones on the subject.

8.6.3 Usability

Usability refers to practicability - ease and economy in terms of energy, time and money. Usability means a test should be:

- easy to construct;
- easy in administrating;
- easy in scoring;
- easy in interpreting; and
- economic in terms of money and time.

The School Head and the teachers should take care that cost, economy and facility in administration should not dilute the accuracy, purpose and reliability of the test.

8.7 EVALUATION TECHNIQUES

In this section we have given a brief list of commonly known techniques to assess abilities and characteristics of pupils. We would also try to understand various types of written tests mostly used in our schools.

8.7.1 Evaluation Tools/Devices

Evaluation is a comprehensive process. It covers various types of educational changes and different aspects of the personality. Different kinds of tools or techniques are used by teachers to judge the progress of their pupils. Some of them are listed below:

- i) Oral test
- ii) Written test/examination
- iii) Practical examination
- iv) Interviews of the pupils, peers and the parents
- v) Pupil products like models, paintings, charts etc.
- vi) Observations of teachers

Some special types of tests and techniques like intelligence test, aptitude test, adjustment test, projective type test are also available. These are usually used by counsellors and psychologists.

The proper use of various devices enables teachers, administrators and parents to have comprehensive and accurate assessment of the abilities and characteristics of the pupils.

8.7.2 Classification of Written Tests/Examinations

We are quite familiar with the various types of items/questions set in our school and college examinations. These are long answer, short answer and objective types. These are taken up for brief discussion here.

i) *Long Answer Type Tests/Items*

These types of questions need more time to answer and so carry more marks say between 7 - 10 (depending on the subjects). These assess the examinees' ability to analyse, evaluate, organise and synthesise ideas. Examinees are able to express themselves creatively and originally. But long answer or essay type questions do not cover the whole syllabus in an examination (and hence are not always valid). These make the test unreliable due to subjective marking and lack of objectivity. It is also difficult to frame good essay type questions.

A good essay type question could be made objective e.g. - Describe any three factors affecting the climate of a place, giving examples.

Short Answer Type

These types of items have shorter answers than the longer ones given earlier. They may carry 4 - 6 marks as specified in various subjects or by examination boards. They assess language, creative expression and organising power of pupil. Unlike essay type questions, they can cover more of the subject matter.

Short answer type items should be specified e.g. Name any three factors affecting the climate of a place.

ii) *Objective Type Tests/Items*

Sometimes, objective type items are also called new type or very-very short type ones. Their answers are very short - one word or a sentence. One correct answer usually gets one mark. Such questions are exact, objective, more valid and reliable. These can cover more of the syllabus. But they fail to test the full real knowledge, creative expression and presentation. There is a guess chance in giving answers in these items. A commonly used formula for obtaining corrected scores to eliminate effect of guessing is

$$S = R - \frac{W}{n-1}$$

Here S : The corrected score

R : Number of rightly marked items

W : n wrongly marked items

n : Number of choices in an item

In true-false items, n equals 2 and in four-choices item, n equals 4. Let us assume a student gives 7 correct responses and 3 wrong ones in a test where there are four alternative choices. His or her corrected score(s) would be as

$$S = R - \frac{W}{n-1} = 7 - \frac{3}{4-1} = 7 - 1$$

= 6.

Some examples of the objective type items are as follows:

1. True-false or Yes-No Type test: Mark (✓) the correct-answer

Example - New Delhi is the capital of India ... True/False

The Taj Mahal was built by Akbar. Yes/No

2. Multiple choice items:

Example: Mark the correct answer (tick as ✓)

Question - The year in which India got independence:

(a) 1942 (b) 1945 (c) 1947 (d) 1950

3. Matching type test items:

Example: Match the correct answer by choosing from among a, b, c or d against each item in space provided ().

1. Meter is a measure of () (a) Area
2. Square centimetre is a measure of () (b) Volume
3. Kilogram is a measure of () (c) Weight
(d) Length

4. Completion type test:

Examples: Fill in the blanks:

- i) is called the father of the nation.
- ii) is known as iron man in India.

5. Analog: type items/test:

The pupil has to find some relation or similarity in the situations:

Examples: Supply the missing items:

1. 36 : 49 :: : 7
2. Meter : Length :: Litre :
3. Uttar Pradesh : Rajasthan :: : Jaipur

8.8 CONSTRUCTION OF AN ACHIEVEMENT TEST

Teachers teach the students to enable them to develop some abilities, skills and attitudes. After teaching, the learners need to be evaluated. It may be through monthly, half-yearly or yearly examinations. The teachers construct the tests to assess the achievement of students. The following steps are taken to develop these tests:

- Decision on units of content and their weightage
- Identification of objectives and their weightage
- Deciding types and number of items

- Preparing Blueprint
- Preparation of test items/questions
- Arrangement of items
- Estimation of time
- Preparing scoring key

Decision on Units of Content and their Weightage

Depending upon the time (half-yearly or yearly), we decide the topics of the syllabus and also fix their weightage in terms of marks. More topics are included in the yearly or pre-board type examinations. In quarterly or half year examinations, less topics are included. So each topic would be given more marks in comparison to the yearly examination. Sometimes the specifications of marks for topics are given in the syllabus and some topics are put within a section.

Identification of Objectives and their Weightage

Normally, School Examination Board examinations are designed to test knowledge, skill, understanding and application objectives. We should consult the guidelines of the examining board or school. If such guidelines do not exist, the teacher may decide them on his own by consulting other teachers or seeing the pattern of previous years' examination papers. For example, in a secondary class X examination in social sciences, the allocation for knowledge, understanding, application and skills was 35%, 40%, 15% and 10% respectively.

Deciding Types and Number of Items

An analysis of previous year's examination question papers or model test papers of examining bodies may enable us to have an idea about the different types of items (essay, short answer or objectives), their numbers and marks for each type. It is only an estimate and should not be taken rigidly.

Let us take two examples here.

In an English model paper, the long answer, short answer and objective type items were 5, 8 and 35 respectively and their respective total marks were 38, 27 and 35. Essay type of questions carried 6-10 marks each, short answer 2-4 marks and objective types 1 mark respectively.

In a social sciences test paper, there were 7 essay type items (each of 6 marks), 9 short-answer types (6 items of 4 marks each and the remaining 3 of two marks each) and 8 parts of a question (one mark for each part). The word limit for answers was specified. Answers to questions of 2 marks, 4 marks and six marks were to be given within a range of 30, 60-80 and 100-125 words.

The types of questions, number, marks and limit of words for answer can be thus properly decided and it should be considered tentative.

Preparing Blueprint

A blueprint is a two-dimensional chart showing different types of items with marks for each topic/unit and each of the objectives. It should follow the respective weightage of marks for the different objectives, and topics and various types of items as prescribed by the school or in the syllabus. These specifications have been discussed in the earlier steps of planning of this test.

An outline of the blueprint is explained in Table 8.1.

Table - 8.1

Blueprint

Examination - Matric (X)		Max. Marks - 100									Time - Three hours						
S.No.	Objective	Knowledge			Understanding			Application			Skill			Total			Total
	Content units	E	SA	O	E	SA	O	E	SA	O	E	SA	O	E	SA	O	
1.	Topic 1		3(2)		6(1)		1(2)		3(2)					6(1)	3(4)	1(2)	20
2.	Topic 2	6(1)		1(5)		3(1)*	3(2)	1(3)		1(2)		3(-)*		6(1)	3(3)	1(10)	25
3.	Topic 3		3(1)	1(3)		3(1)		6(1)*			6(-)*			6(1)	3(2)	1(3)	15
4.	Topic 4	6(1)*				3(2)		3(2)	1(2)	6(-)*				6(1)	3(4)	1(2)	20
5.	Topic 5		3(1)	1(2)	6(1)	3(1)*		3(2)			3(-)*			6(1)	3(4)	1(2)	20
	Sub-total	6(2)	3(4)	1(10)	6(2)	3(7)	1(5)	6(1)	3(6)	1(4)	6(-)	3(-)	3(-)	6(5)	3(17)	1(19)	
	Total	34			38(14)			28(11)			10(-)						100

Note: 1. Marks for one question are put outside brackets and number of questions within brackets.

2. * (Star) denotes an item covering two objectives of a topic and so a number of questions have been left blank as (-)

3. Adjustment of total scores of skill has been done in starred common questions and so not included in total of 100.

4. No. of questions: Essay type (E) = 6
 Short Answer type (SA) = 17
 Objective (O) = 19

Preparation of Test Items/Questions

Test items form the very basis of testing. A test constructor should have good knowledge of the subject. He/She should consult related literature of the subject, model test papers (of CBSE or others) and question banks (if available). The test items should be clear, unambiguous and according to the objectives. Different types of items - essay, short-answer and objective types - should be written out in sufficient numbers. Items of varying difficulty should also be prepared. Experienced teachers are able to estimate it by their judgement. Some items from model test papers or question banks can be taken up.

After collection, a review of items is done on the basis of blueprint requirements and on quality of items. Only unambiguous and specific objective type items are retained.

Arrangement of Items

Some examination boards and schools have divided the syllabus into parts/sections. Questions are to be set and answered from each section. Some within - question choices are allowed. After preparing items, they are arranged section-wise. Similar types of items - essay, short answer and objective - are put together and in increasing order of difficulty. If there is no sectional division, then we can take up objective, short answer and essay type items in a sequence and in increasing order of difficulty.

Estimation of Time

For teacher-made achievement tests, only the experience of teachers should be enough for the estimate of time. We should try to analyse and estimate the time for different types of questions. Normally, in Board Annual Examinations, it is three hours. But for a class test it may be 40 - 60 minutes.

Preparing Scoring Key

The test constructor should give proper instructions for marking. Objective type tests have exact answers. Their answers and their marks should be given. Short answer questions are also quite specific in nature and possible points or ideas in answers should be mentioned with their marks for each.

Essay type questions are lengthy and need specificity for uniform marking. Important steps or points of answer should be explicitly mentioned along with their corresponding marks.

The above guidelines for marking questions from first to the last one should be able to make our testing more reliable.

8.9 SOME NEW DEVELOPMENTS IN EVALUATION

As discussed earlier, our examination systems have been severely criticised. Some steps have been initiated by various education boards (CBSE and others), universities, state education departments and the Human Resource Development Ministry. Reforms, two of which are worth knowing in detail are Grading and Question Banks.

8.9.1 Grading

We have been giving marks to our pupils like 70 out of 100 in mathematics and 65 out of 100 in English for one student. Are they really assessing his levels of achievement in the two subjects? No. If we examine his ranks in two subjects in his class, we find he stands second in English and tenth in Mathematics. So our judgement in different subjects on the basis of marks is erroneous and faulty. Such an error is due to subject variability.

A second type of error is due to examiner variability. Two examiners may assign 57 and 62 marks for the same answers (even if proper guidelines for marking are given).

The third error arises due to larger width or range of scale as 0 - 100. Precise assessment or measurement is not possible. It is difficult to discriminate between 55, 56 and 57 marks. Our unit of one mark ($1/100^{\text{th}}$ part of scale) is a very small (micro-type) unit for discrimination.

Some boards of educations and universities are now awarding grades. The grading system has some merits as under:

- i) Different grades are assigned for same marks for different subjects e.g. 70% in English and Mathematics may imply A and B grades (5 points A - E).
- ii) Examiners can be more accurate and judicious as the units (grades) of measurement are small like grading on 5, 7 or 9 points of scale.
- iii) Equal or nearly equal achieving pupils are put in the same grade. It eliminates chance factor in measurement.

Methods of Grading

Sometimes, we as teachers are assigning grades to our students in class tests. This technique is based on our experience. Being subjective, grading may be lenient or strict. The following two methods are often used to moderate subjectivity.

Norm-reference Grading

Norm-reference grading is used when the students are taking the same test say mathematics test of Class X of CBSE. It is already decided that a certain fixed percentage or proportion of students are to be allotted within each specific/chosen grade level. If we choose the three tier grade system (A, B, C), we may place the top 16% in A, 68% in B and 17% (lowest achieving) in C. Students get gradings according to their relative positions/ranks in the group. Usually a normal probability distribution of scores is assumed and the specified percentage of pupils are placed in the chosen grades.

Let us take two other examples of grading on 5 and 9 points of scale. The achievers are placed in grades in decreasing order of their achievement in a subject. 7% (top achievers), 24% above average, 38% (average group), 24% below average and 7% (lowest achieving) are assigned A, B, C, D and E grades respectively.

For a nine point scale, there would be 4% (best achievers), 7%, 12%, 17% above average, 20% (average cases), 17% below average, 12% and 4% (lowest achieving). They can be named as A-I. Sometimes, other notations are also used, as A⁺, A, A⁻ (The NCERT has published 'Grading in Schools' in 2000 related to this subject) which teachers will do well to study.

Grading assumes symmetrical normal distribution of scores around a mean. Sometimes it is not so in some subjects and for some selective group of students (talented or backward).

Criterion Reference Grading

Criterion reference grading compares the achievement of the students with reference to some chosen criterion or standard, say Division or Intervals of Marks is decided earlier for placement of students according to their performance related to the criterion. For example, we normally place students in distinction, first division, second division, third division and failure class according to ranges of marks (in percentage) as 75 - 100, 60 - 74, 50 - 59, 33 - 49 and 0 - 32 respectively. These can also be graded as A, B, C, D, E.

Another classification may be 90 - 100, 80 - 89,, 30 - 39, 20 - 29 and 0 - 19 for the nine point scale. They may be given letter grades A to I with grade values 9-1 (numerical values).

Such grades enable us to know how many students have answered correctly (or learnt successfully) 50 - 59% or 60 - 69% of the content taught. We should not assume that grading can remove all evils of examinations. The system has some demerits also.

For an intensive study of the 'grading' system, teachers may consult some other references.

IGNOU's MES-4 (Evaluation in Higher Education), 'Grading in Schools', NCERT 2000, Monograph on Grading for Universities, AIU, New Delhi (1977) etc. may be found useful.

8.9.2 Question Banks

This is an effort to improve the setting of question papers. An ordinary teacher may not be able to construct good questions or items of different units. So a collection of good items of different types and of varying difficulty is made. This collection of items with some technical information about them is known as a Question/Item Bank. Teachers can use them for selecting questions for tests, improving their knowledge and teaching competency. Students can also use it to assess their own levels and know the possible types of questions. These are very large in number and of various forms-essay, short answers and objective types and not possible to be crammed. Such collections should continue to be updated and random selection of items from them should be made for a certain examination.

For every item, the following information is given:

- i) Topic of the item and class level
- ii) Objectives tested by the item
- iii) Type of the item
- iv) Marks specified for item
- v) Difficulty level
- vi) Discriminating index
- vii) Time for answering
- viii) Key to the question (answer and with step-wise marks)

Question banks can reduce the burden of teachers and improve their competence in preparing valid test. These are useful for preparing parallel or equivalent forms of tests as required by certain examining bodies such as CBSE.

8.10 USE OF ICT

ICT is playing a significant role in the management of evaluation. There are various modes of ICT like VCD, Over Head Projector, slides and Internet etc.

The use of computers in evaluation has probably risen due to the enhanced time, money and effort spent. Technology can be used for assessment purposes at multi-stages, from the management of the assessment information to a fully automated assessment system. Online and webbased evaluation has increased frequency of evaluation that performs both formative and summative functions. Apart from generating the test, computers can do many other things in the offline mode, e.g.

analysis and recording of the awards received by the students. In online mode, thousands of students can be assessed, in offline mode you have to restrict it to a classroom.

Computers can be used for psychometric examinations designed to explore the values, attitudes and thinking patterns of individuals.

All this can be possible when School Head procures various forms of ICT in the school. More over he/she should see that there should be timely training of the staff in using ICT in order to supplement the text book knowledge of the students.

8.11 MANAGEMENT OF EVALUATION

It is essential that the head of the institution draws up a plan for effective and systematic evaluation with the cooperation of the faculty. For this, meetings of different subject teachers may be arranged. They should discuss the objectives and schemes of evaluation/examination. They should be trained by internal or external resource persons to develop tests and use other techniques.

An evaluation committee should be formed in the school. It should comprise teachers from different disciplines, academic and co-curricular fields. The committee should arrange conducting of various examinations in consultation with the School Head/Principal. To make evaluation comprehensive, assessment in other activities or other characteristics of the pupils should also find a place in it. Cumulative record cards should be entrusted to the class teacher.

The Headmaster's democratic leadership, initiative and proper planning will make the evaluation process effective and useful. Such skilful management is able to get better cooperation of staff.

8.12 CASE STUDY

(Low achievement in a particular subject)

A School Head notes that many bright students of class VIII are not achieving well in social science in comparison to other subjects in the half yearly examinations.

He/she wants to know the reasons and to improve the situation.

Here is an example a situation which involves analysis of the problem and its reasons and then the programme of action.

Analysis of the Problem

The Head discussed the above observation with the faculty. They took some specific cases of pupils and studied achievements in various subjects. It was found that many bright students scored quite well in English, Mathematics, Science... but lagged behind in social science. After identifying the low achievement in social science as a problem, it was decided to find out its reasons.

Knowing the reasons involved question-wise analysis of answer looks of the examinations and discussion with the teachers and students. The steps were as follows:

a) *Question-wise analysis of scores in the social science test of 40 pupils*

Some forty above-average students were randomly chosen, their answer sheets were analysed on a two dimensional chart - pupils v/s. question's score. It gave score of each student in each question. This led to the following results:

- i) The Map/filling question was wrongly/partially attempted by 60% students.
- ii) Some geography units were either omitted or not well attempted by 50%.

b) *Discussion with the staff and students*

The Head had discussed the problem with the social science teachers and as well as the students. To main conclusions were:

- i) Teachers partially or hurriedly completed the syllabus.
- ii) Students did not do practice and home-work regularly.
- iii) Some teachers did not have geography in their graduation level and so could not justify teaching of this portion.

Programme of Action

The School Head discussed the problem, the reasons and suggestions in the meeting. The programme of action was planned and implemented as follows:

- i) Internal seminars for social science teachers were conducted by internal faculty or from faculty of available/nearby institutions.
- ii) Extra classes for Class VIII were arranged by effective and efficient teachers.
- iii) Some teaching aids were devised or procured.
- iv) Division of syllabus of the subject into units was discussed and planned and then the learning progress was monitored by the Head.

A test of the specified portion of the social science syllabus was conducted. It was found that the bright students as well as the normal ones showed improvement.

What would you have done to solve the problem, had you been the School Head there?

Check Your Progress 2

Note: Write your answer in the space given below

1. Select the correct answer and tick mark (✓) it:
 - a) A valid test is reliable. Yes/No
 - b) A reliable test is always valid. Yes/No
 - c) Reliability of teacher made test is often judged by test-retest method. Yes/No
 - d) Increasing number of items in a test increases its reliability. Yes/No
2. Explain test/item validity in 25-30 words. (In the space given below.)
.....
3. Tick the correct answer:
 - a) Long answer type tests are more reliable than the objective type tests. Yes/No
 - b) Objective type tests can test creative expression. Yes/No
4. Enlist the main steps of construction of an achievement test.
.....
5. Describe in 20-25 words how question banks are useful to teacher.
.....

8.13 UNIT-END EXERCISES

1. Take one question paper of the high school (class X) examination of your board or CBSE of any subject:
 - i) Analyse to what proportion (in percentage) the objectives of knowledge, understanding/comprehension, application and skill are being tested by this question paper.
 - ii) Also identify the content units (topics) of syllabus covered in the question. Mention weightage given to each unit.
 - iii) Do you consider that the coverage of syllabus is satisfactory?
2. Prepare an achievement test for class VIII/IX/X for any subject.

8.14 LET US SUM UP

In this unit, we explained that educational evaluation is quantitative as well as qualitative. We also described its inter-relatedness with the objectives and learning experiences. We discussed various types of tools and the qualities of good test which are validity, reliability and usability. A practical way of developing an achievement test was also taken up. We also tried to introduce you to the concepts of the grading system and question banks. Further, we explain how a knowledge of evaluation and the related concepts is essential for the teachers and the heads of institutions for efficient managing of evaluation system in the institution.

8.15 POINTS FOR DISCUSSION

1. Suppose you are the Head of your institution -
 - i) How do you plan to make examination papers of your school valid, reliable and objective? (Suggest activities as training teachers, moderation of paper.) Discuss in the meeting with teachers.
 - ii) Form subject-wise groups of teachers. Ask these groups to identify the main mistakes committed by students in their subjects and discuss in group/staff meetings. Also discuss the ways and means to improve the situation.

8.16 SUGGESTED READINGS

Grondlund, N.E. (1981): *Measurement and Evaluation in Teaching*, New York, Macmillan.

IGNOU: *Instruction in Higher Education - MES - Block 4 (Evaluation in Higher Education)*, New Delhi, IGNOU.

Sharma, R.A. (2002): *Measurement and Evaluation in Education and Psychology*, Meerut, R. Lall Book Depot.

Stanlay, Julian C. and Hopkin, Kenneth, D. (1978): *Educational and Psychological Measurement and Evaluation*, New Delhi, Prentice Hall of India.