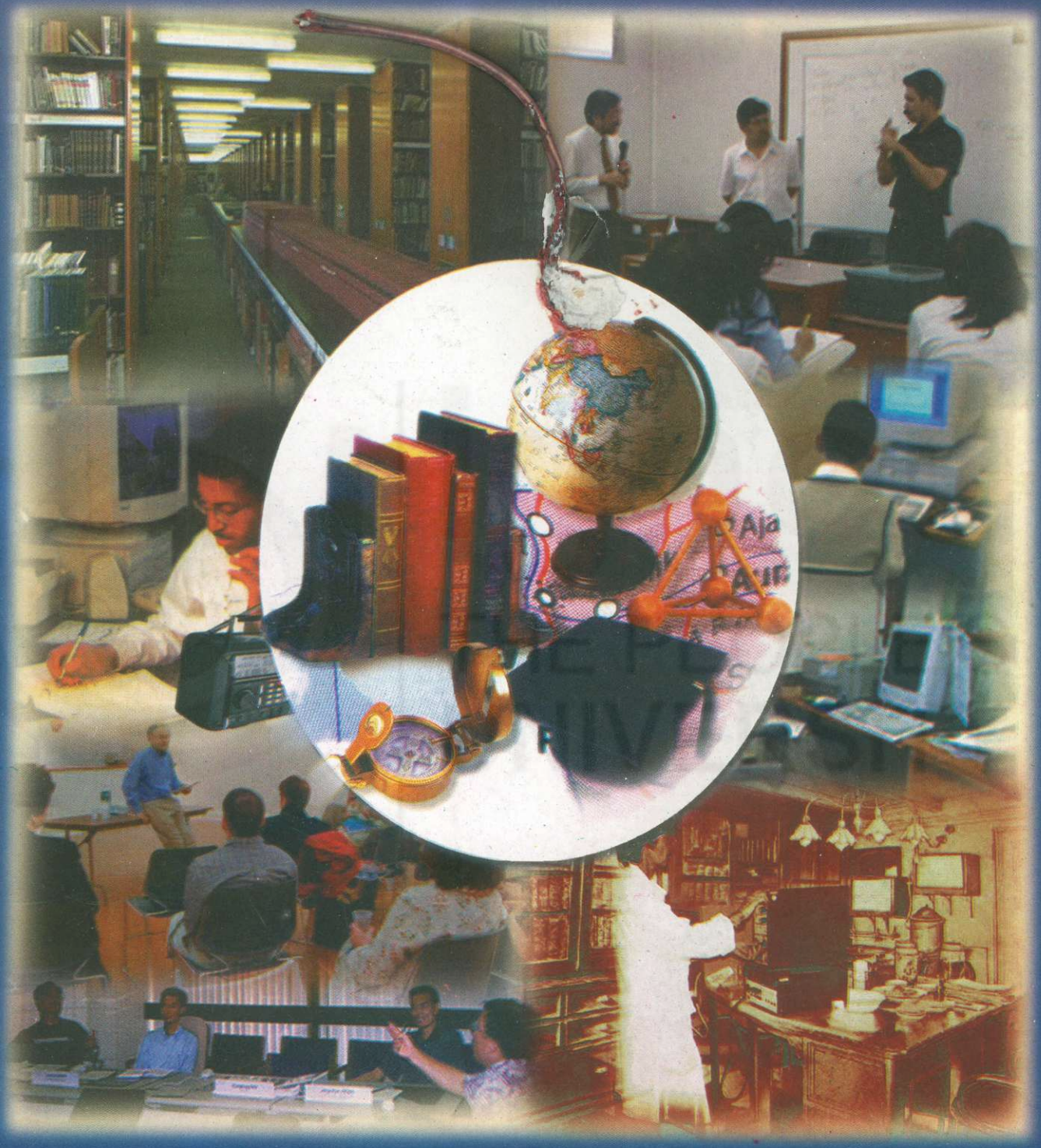
Indira Gandhi National Open University
School of Education

MES-102
Instruction in
Higher Education



Evaluation Perspectives in Higher Education

3

"शिक्षा मानव को बन्धनों से मुक्त करती है और आज के युग में तो यह लोकतंत्र की भावना का आधार भी है। जन्म तथा अन्य कारणों से उत्पन्न जाति एवं वर्गगत विषमताओं को दूर करते हुए मनुष्य को इन सबसे ऊपर उठाती है।"

—इन्दिरा गांधी

"Education is a liberating force, and in our age it is also a democratising force, cutting across the barriers of caste and class, smoothing out inequalities imposed by birth and other circumstances."

— **Indira Gandhi**

Indira Gandhi
National Open University
School of Education

MES-102
Instruction in Higher Education

Block

3

EVALUATION PERSPECTIVES IN HIGHER EDUCATION

UNIT 11

Status of Evaluation in Higher Education-I 5

UNIT 12

Status of Evaluation in Higher Education-II 17

UNIT 13

Evaluation Situations in Higher Education-I 31

UNIT 14

Evaluation Situations in Higher Education-II 46

COURSE TEAM

Course Contribution

Prof. V. Natrajan
Former Director
Research Association
of Indian Universities,
New Delhi

Prof. S.S. Kulkarni
Narayan Ashram
Koloshi, Sindhudurg
Maharashtra

Prof. S. B. Menon
Dean, CIE
University of Delhi
Former Director
SOE, IGNOU

Dr. N. K. Dash
SOE, IGNOU

Coordination

Dr. N. K. Dash
(Course Coordinator)

Prof. S. B. Menon
(Programme Coordinator)

Prof. B. N. Kaul
(Programme Coordinator)
Fomerly Director
SOE, IGNOU

Prof. M. B. Menon
(Programme Coordinator)
Former Director, SOE
IGNOU

Editing

Prof. S. S. Kulkarni
(Course Editor)

Prof. S. B. Menon
(Associate Editor)

Dr. N. K. Dash
(Content Transformation
and Format Editing)

Ms. Jayshree Raveendran
(Language Editing)
Formerly with IGNOU

COURSE TEAM (REVISED EDITION)

Course Contribution

Prof. Pritam Singh
Former Head,
Department of Measurement
and Evaluation, NCERT
New Delhi

Coordination

Dr. N. K. Dash
(Course Coordinator)
SOE, IGNOU

Ms. Poonam Bhushan
(Programme Coordinator)
SOE, IGNOU

Format Concept and Design

Ms. Poonam Bhushan
SOE, IGNOU

Prof. S. V. S. Choudhary
Director, SOE, IGNOU

Dr. N. K. Dash
SOE, IGNOU

Print Production

Mr. B. Natrajan
DR(P)

Mr. Jitender Sethi
AR(P)

Mr. Sunil Kumar
SO(P)

July2010 (Reprint)

ISBN : 81-266-0954-0

Further information on the Indira Gandhi Nastional Open University courses may be obtained from the University's Office at Maidan Garhi, New delhi-110 068

Printed and published on behalf of the Indira Gandhi National open University, New Delhi, by Registrar, MPDD.

Printed at : Young Printing Press 2626 Lane No. 7 Behari Coloney Shahadra Delhi 32

MES-102 Instruction in Higher Education

Block 1 Instruction in a Systemic Perspective

Unit 1 Instructional System
Unit 2 Input Alternatives-Teacher Controlled
Unit 3 Input Alternatives-Learner Controlled
Unit 4 Evolving Instructional Strategies
Unit 5 Unit and Topic Planning

Block 2 Communication Skills

Unit 6 Teacher Competence in Higher Education
Unit 7 Skills Associated with a Good Lecture
Unit 8 Skills Associated with the Conduct of Interaction Session
Unit 9 Skills of Using Communication Aids
Unit 10 Emerging Communication and Information Technologies

Block 3 Evaluation Perspectives in Higher Education

Unit 11 Status of Evaluation in Higher Education-I
Unit 12 Status of Evaluation in Higher Education-II
Unit 13 Evaluation Situations in Higher Education-I
Unit 14 Evaluation Situations in Higher Education-II

Block 4 Evaluation in Higher Education: Mechanics and Processing

Unit 15 Mechanics of Evaluation-I
Unit 16 Mechanics of Evaluation-II
Unit 17 Processing Evaluation Data-I
Unit 18 Alternative Evaluation Procedures
Unit 19 Online/Web-based Student Assessment

Block 3 Evaluation Perspectives in Higher Education

Introduction to the Block

We have already discussed the input and process aspects of instruction in Block 1 and 2 respectively. Hence, the thrust of our discussion in Block 3 and 4 will be on product and output aspects. As you know the output aspects of instruction relate to the students' achievement in terms of instructional objectives. This involves evaluation of students' performance in terms of competencies specified in the form of instructional objectives. Block 3 and 4 will, therefore, deal with evaluation in higher education.

Block 3 deals with an overview of the status and practices of evaluation in higher education. We have tried to present to you with the concept of evaluation, the types of evaluation being practiced in higher education, and reforms being undertaken in the examination system.

Unit 11 of this Block, we talk about the nature and scope of examination, measurement, evaluation and other related forms and the role of evaluation in the teaching-learning process. Besides giving a brief historical retrospect of examinations and the reforms suggested from time to time, we also attempt to appraise the present day examination system and how this determines teaching.

In Unit 12, we discuss the major examination reforms suggested by the University Grants Commission in 1973. In this context, five major issues namely, syllabus, examinations and question papers, question banking, internal assessment and grading have been dealt with. Another innovation in the field of testing and evaluation i.e., National Testing Service has also found a place in our discussion.

In Unit 13, we classify testing and evaluation from three different angles. And on the basis of interpretation of test result, we, divide tests into criterion-referenced and norm-referenced tests. As for the purpose and timing of test the role of formative, diagnostic and summative evaluation is described. We have also classified tests as cognitive and non-cognitive, from the point of view of measurement of cognitive and non-cognitive outcomes of learning.

In Unit 14, we deal with the evaluation of some commonly used techniques of instruction, i.e., laboratory work, seminar participation and dissertation. Tne merits and demerits of internal as well as external examinations have also been highlighted in this Unit.

UNIT

Status of Evaluation in Higher Education-I

11

Introduction

"Evaluation" is a word that came into educational parlance more recently. However, the word is used in many different contexts and hence it's meaning is used with different connotations. "Evaluation" may be said to be the process by which value judgments about the educational (or training) status or achievements of students, are formed. To evaluate then, is to form judgment on the level of achievement and this presupposes that there is a predetermined level available. Since we are confined to educational testing and evaluation, let us hereafter use testing and evaluation in relation to field of education only. At the outset, we will provide you, in brief, with the historical background of examination reform and the development of testing and evaluation. The examination system with which you are involved, is also a process related to evaluation. This unit will provide you with a detailed information on the present status of examination and how an examination determines the nature and content of teaching. You would also be able to understand the role of evaluation in teaching-learning process. Since terms like tests, tools, techniques, measurement, examination, evaluation etc. are used interchangeably by many educators, an attempt is made in the unit to define the nature and scope of each of the basic concepts related to educational evaluation.

Learning outcomes

After going through this unit, you will be able to:

- recall the historical background of examinations and examination reforms;
- discuss the nature and scope of basic concepts related to educational evaluation;
- describe the role of objectives and evaluation in the teaching-learning process; and
- explain how examination determines teaching.

Historical background of examinations and examination reform

"Evaluation" in education, is as old as man's first attempt to teach someone something. However, until recently this process was not very formal and systematic. Before the advent of paper and printing technology, evaluation was mainly limited to **oral testing, recitation of texts** or **oral answers** to questions asked. In our country in the later half of the 19th century, educators began to realize certain shortcomings in depending solely on oral testing and recitation, and started using **written tests** to supplement them. However, oral examinations are still in use for appraising certain types of achievements, together with written tests. Written tests even date back to 2200 B.C. when Chinese are believed to have had a national system of examinations to select their public servants. First educational tests were like present performance tests used in primitive tribes to test young men for their bravery, endurance and learning skills. In medieval times oral examinations were

used in the University of Bologna (1219 AD) and University of Paris by the end of 13th century. Perhaps for the first time educational use of written examinations was made at Cambridge, England, in 1702.

The introduction of standardized tests

At about the turn of the century, the most significant development was the introduction of the **Standardized tests**. The French Psychologist Alfred Binet constructed "Intelligence Tests" (Known as Binet-Simon Scale, in 1905 and as Stanford-Binet Scale in U.S.A. In 1916). Subsequent to this, there were quite a few intelligence and general scholastic aptitude test that were published. These are still being used in many schools and colleges. The impact of these on the teaching-learning process was that they enabled the teachers to learn more scientifically, about the capacities of individual students, their individual differences and also to diagnose their learning difficulties. The early part of this century also saw the significant contribution to achievement testing by J.M. Rice (1894), the real inventor of comparative tests, which became the basis of objective measurement. During the 19th century, popularity of written examination over oral testing, ensuring more objectivity in measurement and comparing performance of bigger groups were major break through. During twentieth century two major trends observed are the broadening concept of measurement and increasing concern for making measurement more and more scientific. These efforts culminated in what is sometimes referred to as the ***testing movement.*** Many tests of achievement and intelligence were constructed and a whole science of psychometrics got developed involving systematic procedures of test construction particularly applicable to objective type tests.

The testing movement

The Testing Movement had two major misuses. First, teachers were so much enamoured by standardized tests, that they began to administer them without considering either the purpose for which they were prepared or the sample on which they were standardized. Second, many teachers became so absorbed with objective type of testing that they used them as the sole method of refinement of the evaluation techniques.

These two misuses culminated in a counter movement that started in the early 1930s, and is often referred to as the **Evaluation Movement**. The main theme of the movement was, and is, that evaluation must be considered as an integral part of the educational process rather than an end of the course activity. If an achievement test is to be valid and appropriate for evaluating students, it should provide a measure of the extent to which instructional objectives have been achieved. A major emphasis of the evaluation movement is that the first step towards planning and programme of evaluating pupil achievement is to prepare a careful and very specific statement of instructional objectives. Another progressive aspect of this movement is that scope of assessment of pupil growth was extended from scholastic to co-scholastic aspects and for that a variety of methods of collecting data regarding student's achievements, like practical tests, rating scales, inventories, observations, etc., are to be used in educational evaluation. The effort would now be to find out as to how far the objectives of instruction in the teaching-learning process are being achieved through our existing evaluation system.

The reform movement in India

The University Education Commission (1948-49), popularly known as Radha Krishnan Commission, expressed that 'if we are to make a single recommendation for the whole university, it should be that of improvement of examinations as they are alarmed by the dominating influence on the teaching-learning practices. The Secondary Education Commission (1952-53) emphasized the need for reducing

external examinations, credit to internal assessment, use of symbolic marking rather than numerical making, etc. Bhopal Seminar (1956) was a landmark in history of examination reform. It recommended 20% weightage to sessional work, inclusion of short answer and objective questions, establishment of examination research bureau, etc. With Dr. Benjamin, S. Bloom's visit (1957-58) to India, guiding philosophy of examination reform emphasized integral relationship among educational objectives, learning experiences and evaluation procedures. An action plan developed in this regard was approved at First Conference of Chairmen and Secretaries (1954). This led to establishment of pilot examination unit (1958), state evaluation units, NCERT examination reform unit, emphasis on comprehensive internal assessment and de-emphasising external examination. Use of grades in place of marks, improving quality of evaluation instruments, question banks, making certification of students more meaningful, etc. were emphasized by the Education Commission (1964-66), The University Examination Reform Action Plan (1972), The National Policy of Education (1986) and in NCERT National Curriculum Framework (2000).

Educational evaluation in the teaching learning process

Educational evaluation relates to any procedure that appraises the extent to which specified educational objectives have been achieved. This definition may be taken to include both the emphases of the evaluation movement namely, specification of instructional objectives and, the variety of tools and techniques used to judge the extent to which the prescribed instructional objectives are achieved.

Basic concepts in educational evaluation

It is commonly observed that many of the terms relating to the field of evaluation are in vogue and used by authors and teachers inter-changeably. To overcome this difficulty, it is considered worthwhile to get an understanding of those terms which form the basic concepts in evaluation. Use of these concepts in right context with definitive connotation is necessary to avoid confusion and their misuses in explaining various concepts. Brief connotation of each is given below:

a) *Evaluation* is a process of collecting information on the basis of which judgements are formed, which in turn are used for taking decisions. Therefore, three phases of evaluation are:

   i) Information gathering (about achievements, competencies developed, instructional objectives, etc.)

   ii) Forming judgements (about attainment of level, in terms of intended learning outcomes, etc.)

   iii) Making decisions (about remedial teaching, grading, promotion, certification, etc.)

b) *Measurement* is a process of assigning numerals to objects or an attribute according to certain rules. It deals with the comparison of quality with an appropriate scale for the purpose of determining numerical value on the scale that corresponds to the quantity to be measured. Attribute may be achievement, physical efficiency, intelligence or any personal or social quality.

c) *Examination* is a process of collecting evidences about students' achievement or growth. It involved planning, construction of instruments, administration, scoring of scripts, compilation of results, etc. Thus examination process included in the first phase of evaluation is concerned with obtaining information. For obtaining information, we use various techniques of evaluation like testing observation, inquiry and analysis. Each of these techniques involves use of various types of evaluation tools.

	Techniques	Usable Tools
i)	Testing	Tests like unit test, term test, annual test, diagnostic tests question papers, record cards, etc.
ii)	Observation	Observation schedule, anecdotal record, rating scale, naked eye, lens microscope, etc.
iii)	Inquiry	Questionnaire, interview schedule, check list, opinionnaire, etc.
iv)	Analysis	Proforma, result sheet, analysis sheet, score card, etc.

Measurement deals with how much of a thing while evaluation goes a step further to tell us, of what value is that measurement. The relationship between the two is depicted by Gronlund, N.E (1971) as under:

Evaluation = Measurement + Value judgement.

Measurement, therefore, is a pre-requisite for evaluation but not synonymous with it.

d) *Assessment* is teacher's estimate of a value against a standard. It is an estimate, not an accurate observation of an attribute, and therefore, suits well in the context of internal assessment, where as assessment of affective outcomes or co-scholastic aspects of students' development is generally approximated and can not be measured accurately like cognitive abilities.

e) *Appraisal* is generally defined as a formal and accurate valuation of some characteristics against some criteria usually made by persons familiar with such values. In a way it is more refined way of estimating the value of a thing by using such criteria, supposedly developed by knowledgeable people in that field.

Measurement alone is not equivalent to evaluation. The Sanskrit term *"Mapan"* for measurement and "Mulaya-Mapan" for evaluation make this distinction clear. Evaluation is more comprehensive in that it invariably includes **valuation** on the basis of some criteria of the measures (scores) obtained.

It is expected that the terms described above, will be used with correct connotation and not inter-changeably. Against each statement given below, write in brackets provided, the term (evaluation, measurement, examination, technique, tool, assessment, appraisal, etc.) that suits.

1. P got 55 marks in biology in paper-I. ( )
2. Q got 55 marks in physics and topped in his class. ( )
3. Teacher's observation about student-X being truest was correct. ( )
4. Which techniques is usable for getting views of experts on internal assessment? ( )
5. Is analysis sheet a tool or technique? ( )
6. Is examination a technique or process? ( )
7. For co-scholastic traits, which of the two terms, evaluation or assessment, suits better? ( )
8. Is testing a tool or technique? ( )
9. Formal accurate valuation against criteria ( )
10. Observation schedule is a technique/tool ( )

Role of objectives and evaluation in teaching learning process

Observe the basic model of teaching learning process depicted below diagrammatically and note that:

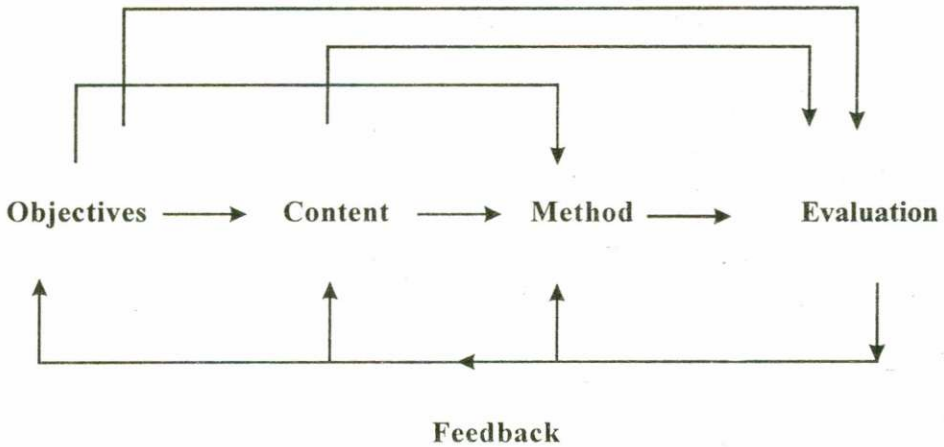


Figure 1 The basic model of teaching learning process

a) Objectives provide basis for content selection, teaching learning strategies and evaluation of students' performance.

b) Content forms the basis for methodology of teaching and evaluation.

c) Objectives, content and methodology of teaching form the basis of evaluation.

d) Evaluation provides feedback on efficacy of objectives, content and methodology of teaching.

e) Evaluation is integral part of teaching learning process and not an end of the course activity.

Instructional objectives in a subject reflect curriculum intention in the form of Intended Learning Outcomes (ILOs) which become the focus of teaching and testing. Let us learn more about them. These objectives can be classified into cognitive, affective and psychomotor domains, emphasizing, intellectual, emotional-social and practical skills outcomes of learning.

a) ***Cognitive Domain:*** Taxonomy of educational objectives (cognitive domain) by Bloom, B.S. (1956) categorizes objectives in hierarchical order of complexity: I) Knowledge, ii) Comprehension, iii) Application, iv) Analysis, v) Synthesis and vi) Evaluation. Indian researches in this area by NCERT and other researches have developed on empirical basis a more realistic and usable three tier taxonomy comprising of I) Knowledge, ii) Understanding and iii) Application objectives arranged in hierarchical order of increasing complexity and reflect almost all the objectives of Bloom's taxonomy, when considered from the point of view of competencies implied in these three objectives Knowledge (K) Understanding (U) Application (A). Each of these objectives (K, U, A) is specified in terms of competencies that are supposed to be developed when an objective is achieved. Such competencies reflect development of abilities like recall, reproduce, recognize (K); translate, relate, compare, classify, interpret extrapolate (U) and analyse, hypothesize, select, suggest, conclude, predict, judge (A) as implied in Knowledge (K)→ Understanding (U)→ and Application (A) objectives.

b) ***Affective domain*** includes objectives which relate to personal social qualities in the form of I) habits, ii) interests, iii) attitudes, iv) values, v) character building qualities. For this commonly used techniques and tools are observation, interviews, inventories, attitude scales, check lists, rating scales etc.

c) ***Psychomotor domain*** objectives which are concerned with the development of psychomotor or practical skills related to process of performance and product of performance. Such skills could be physical fitness skills, artistic skills, observational skills, manipulative skills and other experimental skills.

Tests and Examinations

The word "test" can be defined in a variety of ways. A test provides a situation or a series of situations (questions/items), to students to perform in a particular way to demonstrate their abilities. We may be testing their intelligence, their aptitude or their achievements. The performance may be verbal or manual or a combination of both. Such a test or measuring device, may be a standardized test or a teacher-made test.

In India, we use several words for these measuring devices like Unit Test, Final Examination, Diagnostic Test, Aptitude Test, etc. The Sanskrit term '*Pareeksha*', refers to all these words. And, whether it is a unit test or final examination, it provides an occasion for an outsider- a teacher-to observe the student (***"Eeksha"*** to observe and ***"Pari"*** means comprehensively, hence ***"Pareeksha"*** means comprehensive observation). All such tests can be further classified according to the time or according to the purpose. For example, when a test is given every day or week after a certain small unit of a long duration course is complete, it can be called a unit test or a weekly test, etc. But if the test is given at the end of a term, it is a term test. It could be at the end of a semester at the end of the prescribed course after a year or two and called a final examination.

Another way of classifying tests could be on the basis of the **purpose** with which the test is constructed. As in the final examination, the purpose could be to evaluate a student on all attributes that he has learnt or developed during that year and then possibly grade and promote or certify his achievement and could be labeled as a summative test. But if a teacher wants to assess whether the students have learnt what they have been taught in a unit and get feedback for adapting teaching learning strategies and provide feedback to students, then it is a formative test. When a test is given to poor performers in a unit test to identify the causes of poor performance and specific difficulties they face, it is a diagnostic test, which is usually followed by remedial teaching. You will know more about formative and summative test in Unit 12 of this Block.

Self-ssessment

1. *Suppose after a six-month course in stenography two learners, say 'X' and 'Y', achieve a speed of 100 words per minute. However, when X Joined the course, he had already achieved a speed of 70 words per minute, while Y at the time of joining the course, had a speed of 50 words per minute. Now, which of the three terms* ***'evaluation or measurement' or tools/techniques*** *will you use, to describe the following statements about the achievement of learners? Tick the answer in the appropriate box:*

Statements related to the three Terms	*It is a statement of Evaluation*	*It is a statement of Measruement*	*It is a statement about Tools/ Techniques*
a) The progress made by 'X' is average with reference to his class	☐	☐	☐
b) 'X' could not appear in the inter-school common test	☐	☐	☐

c)	*The progress made by 'Y' is extraordinary in chemistry*	☐	☐	☐
d)	*'Y' could not identify the given slide under the microscope*	☐	☐	☐
e)	*'X' has recorded a progress from 70 w.p.m. to 100 w.p.m.*	☐	☐	☐
f)	*'X' and 'Y' attended the interview conducted by the Principal*	☐	☐	☐
g)	*'Y' got more marks than 'X' when teacher examined them on their academic aptitude*	☐	☐	☐
h)	*At the end of the course, 'X' and 'Y' performed at the rate of 100 w.p.m.*	☐	☐	☐

A careful choice and formulation of educational objectives will lead the teacher on to the construction of effective learning experiences. This will no doubt depend on proper specification and statement of objectives, understanding of how students learn, and comparative idea of the effectiveness and utility of the different communication channels available. If teaching-learning (more properly, the latter), are to be effective and progressive, it is vital on the part of the teacher not to stop with the first two activities but to proceed to the third, involving the evaluation of the extent to which the objectives have been achieved. It is obvious, therefore, those objectives for evaluation and objectives of instruction are identical. Teaching and testing are complementary components of teaching-learning process and should not be treated in isolation.

Evaluation plays an important role in the teaching-learning process. From the foregoing discussion, we have seen that evaluation serves a number of specific purposes and functions. However, we have given below the most common purpose and the related **functions** for which evaluation is used within educational systems.

Diagnosis: To identify the causes of **weaknesses/lapses** in learning process on the part of the learner, and monitor learning progress with a view to suggesting remedial teaching and adapting instructional strategies.

Prediction: To predict probable **future success** on certain types of tasks on the basis of present achievement of related tasks.

Selection: To select suitable persons for admission to a course in a class or a university. The entrance tests to different courses are devised for this purpose.

Guidance: To assist students in **making decisions** for the future, namely, their choice of subjects, courses and career, etc.

Certification of achievements: To assist students' achievement in various subjects in terms of ILOs and certify meaningfully for better comparison of students across the institutions.

Evaluation of a Programme: To monitor the **effectiveness** of teaching in a particular course; to find out the relevance of the objectives and efficiency of methods used, it is necessary to evaluate a programme. A programme is evaluated on *rational* basis before it is being introduced. During the programme *formative* evaluation as integral part of the programme goes on to get and provide feedback on the basis of which teaching learning strategies are adopted according to the stated objectives. At the end of a programme *summative* evaluation is undertaken to find the discrepancy between the intended and the observed outcomes.

Values of testing and evaluation to student and teachers

A) **Use to students**

i) Test scores when properly derived, show the student's progress in the course.

ii) Tests help students to identify the points of the course which the student does not know well enough.

iii) Tests cause the students to review work which help them to organize and retain knowledge.

iv) Tests enable students to distinguish between the relevant and the irrelevant answers.

v) Tests give the student a better conception of the objectives of the course.

vi) Tests motive students to learn more and develop regular study.

vii) Tests help students to self-evaluate with reference to class performance.

B) **Use to teachers**

i) Test scores indicate to the teacher whether or not his/her mode of presentation of the subject matter was sufficiently effective to accomplish the objectives of the course/unit.

ii) Test scores enable the teacher to compare the effectiveness of different teaching methods, s/he adopts.

iii) Test scores reveal to the instructor which parts of the course he could not efficiently teach or make clear, thus enabling him/her to go back and clarify or re-teach.

iv) Test scores enable the teachers to locate the weaknesses of individuals/ groups of students on the basis of which remedial teaching can be undertaken or correctives can be applied.

v) Test scores of the learners help the teacher to judge the efficacy of the unit objectives, he formulated.

vi) Consistently poor scores of the whole class on specific concepts or competencies implied in an objective help the teacher to review his/her instructional strategies.

Self-assessment

2. *Given below is a list of tests/ examinations. Identify the function each of them serves and write the answer in the blank space provided:*

   a) *Half yearly examination in the college* .....

   b) *UGC's JRF Examination* .....

   c) *Graduate Aptitude Test in Engineering* .....

   d) *Bank recruitment examination* .....

   e) *Unit test conducted by a teacher* .....

   f) *In-text exercise provided for students* .....

   g) *Questionsnaires to assess the teachers' performance* .....

   h) *Use of unit test for poor performers* .....

   i) *Public examination at the end of degree course* .....

   j) *Use of oral test at the end of a lesson/unit* .....

Examination as the stumbling block for qualitative assessment

After studying the what and why of testing and evaluation, it is worthwhile to look into the examination system which has been accepted as the means through which we evaluate the performance of our students, teachers, educational institutions and the overall educational system. For quite a long time now, performance in examination has been used as a quantitative indicator of the educational system development. The present examinations, however, do not reflect the qualitative improvement. We value the contribution of the present day examinations in terms of the number of students passed in an educational institution, and the number of students getting high percentage of marks in the examination. Examination is being treated as a certification process in which certificates indicating divisions, percentage of marks/grades have become more important than the actual behavioural changes occurring in the learners as a result of instructional interventions by teachers. Behavioural changes here mean changes in the positive direction reflecting development of desirable competencies, which are usually specified in the stated objective at subject level or delineated by teachers at unit level. These factors are, to a considerable extent, responsible in hindering the efforts towards qualitative improvement, which demand competency based teaching and competency based testing. Purpose of examination is limited to measurement of students' achievement while their role in improving their achievement remains un-appreciated.

Defects in present day examinations

Various educators and Education Commissions appointed by the Govt. of India have pointed out the defects of the present day examination systems in our universities. Some of these are briefly summarized below:

i) The dominance of the final end of the year examination gives the impression to students that they do not have to study throughout the year. There is no continuous assessment which demand diagnostic and formative testing. Hence, many a students cram in the last month of the academic year a certain content area of the syllabus on which the questions are likely to be asked.

ii) Most of the questions asked in these examinations test factual information and hardly assess students' critical thinking or their ability to apply knowledge. The questions do not represent the higher level instructional objectives of the curriculum.

iii) In most of the examinations, the question asked, are essay type (apart from practical-laboratory based tests), with alternative choices/ options. These do not allow wide coverage of the course objectives and lead to selective teaching and selective learning. The assessment of essay type questions is also not reliable because of subjectivity in marking, which give rise to many complaints of favouritism, bias, lack of objectivity in marking, impressionistic marking, etc.

iv) In most university examinations English is the dominant medium of expression. Students whose mother tongue is different, find it difficult to express their thoughts and feelings in English, even if they have understood the concepts and mastered the content of the subject.

v) Because of free options (any 6 out of 9 or 10 types) comparison of students' scores is fallacious since all the students do not attempt the same questions (i.e. different tools of testing) in a test, which makes it difficult to compare performance of two students unless both of them have attempted the same questions.

Because of these and many other defects, the examination systems has corrupted the system of education vis-à-vis the teaching-learning process.

Examinations dominate teaching

Teaching-learning process, now-a-days, to a great extent, is dictated by the examination system. Classroom teaching has become more information oriented. It hardly gives any regard to the cultivation of logical and critical thinking, the spirit of inquiry, ability to apply decision-making skills, etc. The teachers are neither bothered about giving regular tests to students nor are they interested to get the feedback on their own teaching. Teachers fashion their teaching according to the requirements of the examination. They prepare the students by way of dictating notes, providing guide books and readymade learning materials, all geared up towards the final examination. At times, they do not even cover the entire course, instead, they choose some chapters which are important from the examination point of view. Student's involvement in teaching-learning process is hardly appreciated for better learning.

Mostly the teacher prefers a passive lecture i.e. one way communication to an active interaction in the form of discussions and active participation by students. It seems that in college classrooms, a teacher is alienated from students. Teaching-learning process which should be a two-way process turns out to a one-way process. Teaching gets synonymous with teacher talking and students as passive listeners.

Thus examinations have a negative influence over the teaching-learning process. They have made teaching a monotonous, disinterested and passive process in which lessons are just like stones pelted at students and students are like dead specimens of a museum (Rabindranath Tagore). All this leads to a question. Should we abolish the examination system? The answer is certainly 'No'. We can not totally wipe out examination system, but what we can do is to improve the existing system. In this context, we will discuss the examination reforms suggested and recommended by different educational bodies, committees and commissions, like the University Grants Commission, etc., in the next Unit (Unit 12).

Self-assessment

3. *Consider any external college/university examination for a course you are familiar with. What purpose(s) does this fulfill for each of the following beneficiaries?*

   a) The students taking the examination.

   b) The teachers who had taught the course to the students.

   c) The college attended by the students.

   d) Employing agencies likely to employ the students on their leaving the college.

4. *Suppose on the other hand, for the same course, there is no external college/university examination and the students have to be evaluated by you. Would this change your earlier answers to the above questions in any way? If yes, what would be your answer?*

Summary

In this unit, the main focsus of discussion was the role of testing and evaluation, and the status of the present day examination system. In the beginning, we started the discussion with the history of the development of testing and evaluation. Then, we defined the concept of evaluation and other related terms. We distinguished between measurement and evaluation and said that evaluation is measurement with value judgement. We have also described the role of objectives and evaluation in the teaching-learning process. Our discussion then shifted to the analysis of the present day examination. Towards the last part of the unit, we examined the fact that examination determines teaching and hence is in need of reform. In the next unit, we shall describe the various examination reforms.

Unit-end activities

1. Add some points of your own to the list of defects and advantages of the present day examination system, mentioned in this Unit.

2. Make a list of objectives which are tested in public/university external examination and which are more amenable to internal assessment.

Points for discussion

1. In spite of the several criticisms of the present day public examinations systems, by various Education Commissions, it still continues to hold its sway. What could be the reasons for persons ascribing it more credibility than internal assessment system?

2. If you are appointed a principal of an autonomous college, what kind of an examination system would you evolve to met the requirements of both the students and the teachers more efficiently.

Answers to self-assessment

1)	It is a statement of Evaluation	It is a statement of Measurement	It is a statement about Evaluation tools and Technique
a)	√		
b)			

c) √

d) √

e) √

f) √

g). √

h) √

2) a) Grading b) Selection c) Prediction

d) Selection e) Guidance/ diagnostic f) Evaluation

g) Judge instruction efficiency h) Diagnostic and remedial action i) Certification of students' achievement

j) Review of lesson/unit

3) a) The students taking the examination get certificates, which indicate their performances in the examination.

b) The teacher who taught the course to the students get feedback on the overall performance of their students.

c) The college attended by the students gets feedback on the performance of both the teachers and the students which help to improve teaching learning practices.

d) Employing agencies use the certificates for selecting students to different posts.

4) a) Students taking the examination come to know about their progress in academic achievement and accordingly improve their performances.

b) The teachers who taught the course to the students can diagnose the learning difficulties of the students and provide remedial teaching to them.

c) The college is able to monitor the progress of the students as well as the teaching performance of the teachers more efficiently.

d) Employing agencies may use the comprehensive record on various types of recorded achievements for each student, for the purpose of recruitment.

Suggested readings

Ebel, Robert. L. (1979) *Essentials of Educational Measurement.* (3rd Edition), New Delhi : Prentice Hall of India Private Limited.

Gronlund, N.E. (1981) *Measurement and Evaluation in Teaching.* (4th Ed.), New York: Macmillan.

Lindeman, R.H. (1971) *Educational Measurement.* (First Indian Reprint), Bombay: D.N. Taraporeval a Sons & Co. Pvt. Ltd.

Mehrens, W.A. and Irvin J. Lehman (1984) *Measurement and Evaluation in Education and Psychology.* (3rd Ed.), New York: Holt, Rinehart and Winston.

Nunally, J.C. (1972) *Educational Measurement and Evaluation.* (2nd Ed.), New York: McGraw Hill.

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

UNIT

Status of Evaluation in Higher Education-II

12

Introduction

In the previous unit, we discussed the present day examination system in India in the context of the broad framework of evaluation and testing. An analysis of the present day examination system reveals that it has become very obsolete and outdated and it hardly leaves any scope for innovations in teaching, etc. Hence there is a strong need for examination reforms. In this unit, we shall discuss syllabus reforms, question paper reform, question banks, internal assessment and grading, in the context of examination reforms. Towards the end of the unit, we will discuss the objectives and functions of the National Testing Service.

Learning outcomes

After having gone through the unit, you will be able to:

- describe the significant aspects of examination reforms such as reformulation of syllabus, improvement of question papers, and development of question bank;
- appreciate the need for internal assessment for making evaluation more continuous, comprehensive and integral part of instruction;
- discuss the concept of grading and its advantages over numerical system of marking; and
- enumerate the objectives and the role of the National Testing Service in

Examination reforms — Significant aspects

In the earlier unit, we have seen that examination has been identified as the major stumbling block in qualitative improvement of instruction in higher education as well as at other levels of education. In order to revamp the examination system, various committees and commissions have suggested certain reforms. In this context special mention may be made of University Education Commission wherein it was mentioned that if we have to make a single recommendation in the whole system of education it would be to improve existing system of examinations. The Education Commission (1964-66) says, that 'examination reform has become crucial to all progress and has to go hand in hand with the improvements in teaching'. Moreover, the University Grants Commission, in a monograph entitled, "Examination Reform – A Plan of Action" (1973), has made specific recommendations which relate to four major aspects of examination reforms, viz., internal evaluation, grading system, question banks and national examinations. We shall discuss these reforms along with other reforms, in the subsequent sections.

Reformulation of syllabus

As we say earlier, the examination should reflect the instructional objectives. At present, in the context of universities, the only document that represents a statement of instructional objectives, is the syllabus, which is generally prepared and approved in the **Board of Studies** of the respective universities. This document-syllabus suffers from many defects. Some of these are:

i) It is usually prepared casually – mostly based on the available text books rather than on the basis of a **comprehensive analysis** of the curriculum content/ subject matter and the intended skills and abilities, which the students develop at the end of the course. In most cases, specific instructional objectives or intended learning outcomes are not stated (nor even deliberated on), by the Board of Studies.

ii) In most cases, even the content areas are not defined or described fully. The syllabus becomes just a list of themes/topics without reference to the scope and depth at which they need to be dealt with.

iii) The document do **not** provide for possible **learning experiences,** except that in some natural science courses, a list of practicals is given. This absence of any indication of nature and type of learning experiences lead a teacher to make a course dominated by lectures, without ensuring students' involvement to develop the needed competencies and mastery over the skills implied in that course.

One of the implications of a new system like the 'Semester System' for courses, is that the *'syllabus'* should be more scientifically and comprehensively prepared to specify clearly the:

i) Unit-wise content clusters to be taught.

ii) Competencies students have to develop.

iii) Information to be acquired and skills to be developed by students.

iv) Activities the students are supposed to undertake.

v) Linkage with other subjects wherever essential.

vi) Breadth and depth of treatment of concepts, understandings, principles, etc.

vii) Practical examples of applications of subject matter in daily life and industrial world.

viii) Teacher's methods of teaching and instructional strategies.

ix) Usable teaching-learning aids for efficient teaching.

x) Time allocation for different units of the syllabus for instruction.

xi) Mode of classroom management and techniques of lesson planning.

xii) Mode, frequency and type of tests to be used for evaluation of students.

All these aspects shall have to be taken into consideration while undertaking revision of syllabus in any future qualitative improvement programmes of the universities.

Importance of instructional objectives in testing

If learning is not given direction, meaning and organization by clear-cut instructional objectives, testing remains limited to memorized factual information, the relative importance and the ultimate utility of which, is not only unknown, but remains uncertain. And the same gets reflected in examinations. Hence, if objectives are to serve as guide to the teaching-learning process and to evaluation, their statement in explicit way is of great importance. It follows, therefore, that the most important step towards qualitative improvement is to get the syllabi revised to take care of this aspect. This is to be done as a part of syllabus framing perhaps by the Board of Studies, whereby, statement of instructional objectives in the prescribed syllabus becomes a necessity for better teaching and testing.

Nature of examinations and question papers

Quality of question papers (achievement test) is one of the most important concern for improving the present day examinations. In this context, we shall discuss the nature of examinations, which include.

- frequency of examinations
- setting of question papers
- language of examinations
- reliability of examinations
- **Frequency of examinations**

It is desirable to 'spread out' examinations through out the year besides the ones that are conducted at the end of every semester or a session. This is now universally favoured that students concentrate on smaller courses followed by taking examinations, thereby making students' assessment more continuous which, in turn, leads to more regular study habits among students and improved learning. More frequent feedback help teachers also to adapt their instructional strategies.

It is true that the examination conducting body of a university finds it difficult to efficiently conduct increasing number of examinations every semester. Some of these examinations could, however, be organized at the college level in a suitable manner. Semester examinations at the end of every consecutive semester (1st, 3rd, 5th, etc.), may be undertaken by college staff will also acquire expertise in this area, which is desirable, when they become autonomous. **Another alternative** is to get the first 2 semester examinations conducted by colleges in the first year, and the next year's (2nd year) semester examinations may be conducted by the university on the second year. Gradually the dependence on the public examinations, conducted by university or college should be reduced by introducing internal assessments-teacher made tests/ practicals, as is done in Indian Institute of Technology and such other organizations like Boards of School Education in India.

- **Setting of question papers**

While preparing a question paper, the teacher has to take care of certain important aspects. These are the cognizance of instructional objectives, design of question papers, types of questions to be asked and options to be given to the students. Questions paper is the main evaluation tool used all over the world for assessment of students on cognitive learning outcomes. Therefore, it is essential to ensure the quality of a question paper. Either the teacher sets or a university gets it set for different examinations. But what is a good question paper? In fact, examination improvement requires knowledge of :

a) Criteria of a good question paper

b) Technology of setting better question paper

c) Management of examination processes like scheduling, conducting, scoring, processing of results and certification

We are concerned with the first two aspects.

a) **Criteria of a good question paper**

A good question paper is one that is:

i) *Valid* which means that it tests what it is supposed to test, i.e., it serves the intended purpose in terms of relevance to content of testing (adequacy of sampling of content) and the assessment objectives (competencies being tested).

ii) *Reliable* which means that it yields the same scores consistently and accurately if administered again. Therefore, accuracy of measurement and consistency in score are necessary.

iii) *Usable* of practicable in the given context, i.e., it should be easy to administer, score and interpret.

b) **Technology of setting better question paper**

i) **Designing of question paper-** It refers to the framework that the Board of Studies, University or any examination agency should provided to the paper setters in external examinations (while at college level it is the teacher himself who designs the question paper) Whosoever develops the design of question paper should take the following steps:

- Identify the assessment objectives (knowledge, understanding, application) and give weightage to each objective (e.g.) 50%, 30% and 20%)
- Identify major units of syllabus or follow the prescribed units.
- Give proportional weightage to units of syllabus.
- Indicate scheme of sections if any in the question paper
- Indicate scheme of options to be followed.

ii) **Preparing the blueprint** (by paper setter) On the basis of design the paper setter is to prepare the blue pint. Design remaining the same there can be many prints on the same design.

A sample blueprint is given below. Blueprint may differ from subject to subject and from class to class and even from teacher to teacher or paper setter to paper setter. A blueprint indicates the distribution of questions (along with marks) over various units of the **syllabus** to test various **objectives** besides reflecting the **options** as stipulated in the design. It would look like the following 2-dimensional chart showing entries related to **assessment objectives content areas/units and form of questions**.

Objective→	Knowledge			Understanding			Application			Skill			Total
From of Question → ↓ Content	E	S	O	E	S	O	E	S	O	E	S	O	
Unit/ Theme-1	-	3(1)	2(2)	-	-	2(2)	-	3(1)	-	-	-		10
Unit/ Theme-2	-	3(1)	1(1)	10(1)	-	2(2)	-	-	1(1)	-	3(1)		20
Unit/ Theme-3	-	3(1)	1(1)	-	6(2)	1(1)	-	3(1)	1(1)	-	-		15
Unit/ Theme-4	10(1) 10(1) (a)	-	1(1)	-	3(1)	1(1)	-	-	2(2)	-	3(1)		20
Unit/ Theme-5	7(1)	-	4(4)	-	-	2(2)	-	3(1)	1(1)	3(-)	-		20
Unit/ Theme-6	-	3(1)	2(2)	-	3(1)	-	-	6(2)	-		-	1(1)	15
Sub-total	17 (2)	12 (4)	11 (11)	10 (1)	12 (4)	8 (8)	-	15 (5)	5 (5)	3 (-)	6 (2)	1 (1)	100 (43)
Total	40			30			20			10			100

mmary: Internal option (a) in Essay type questions in Unit-4 under Knowledge
jective.

tries made are only to illustrate. Paper setters have to decide about these while
eparing the blueprint. However, it must confirm to weightages indicated in the
sign to the various objectives (40, 30, 30 10), content units (10, 20, 15, 20, 20, 15)
d form of questions E=30%, S.A.= 45% O.T. 25%) as reflected in this table of
ecification or blueprint.

Types of questions used

paper pencil test or written examinations there are mainly three types of questions:
say type, Short Answer type and Objective type. All these types can be used in
st of the written examinations. In many college/university examinations all the
ee types are combined. However, the patterns of weightage differ from one
ject to another and from one level to another. Some rather typical combinations
as presented below:

uestion type	Marks	Time
o 3 say type questions	10 to 30 Marks	20 to 60 mts
to 20 ort answer type questions	30 to 60 marks	40 to 80 mts
to 30 jective type	20 to 30 marks	20 to 30 mts

should be noted that while the short-answer questions with average three marks
ch and requiring four minutes each, the marks and the time need not be uniform
all the questions. Some might warrant only one mark, some two, and others three
four. In case of multiple-choice questions, usually one mark each is allotted,
suming that it takes about a minute to solve each question. But this presumption
uld be wrong. Some multiple-choice questions involving higher level objectives or
mpetencies could require more time.

important feature of this pattern is, that compared to the present practice of
e or six essay type questions, this increases the number of questions, quite
nificantly. With a large number of questions, say forty questions, it is possible
cover the syllabus adequately and assess all the important competencies implied
assessment objectives, which is not possible when students answer only a few
estions. Choosing a particular type of question depends on the assessment
ectives and the nature and scope of the intended response. For example, essay
estions, if properly constructed, are useful when purpose is to test the students'
lity to develop a plan, an argument, or a sequence of ideas, express their ideas
appropriate language, summarize, integrate and suggest measures or an action
n. Every effort should be made to develop questions which measure higher order
lities/intellectual skills, as distinguished from memorization and factual
ormation. Increase in number of questions, with introduction of short answer
d objective type questions, helps to improve validity and reliability of the test
uestion paper).

Optional questions

e provision of choice in the present question papers ranges from papers which
lude ten questions out of which the students may answer any five or six, to those
which choice is limited to certain sections of the paper and those in which some
estions may be compulsory. In the combined pattern of questions shown earlier,

it is suggested that an internal choice be given in essay type questions only (i.e. include two questions, out of which the students may answer any one). The alternative question should, however, be framed with a deliberate attempt to make it nearly as comparable in terms of content area, objective tested, form of question (essay type), level of difficulty and require the same amount of time to answer as the other question.

(You will know more about designing a question paper in Unit 15 of Block 4.)

b) **Language of examinations**

This is a hot issue in many universities. While there are great pressures from the political and social forces for allowing the students to answer in their regional language/mother tounge, the academic community is rather reluctant to introduce the change for some sound reasons. The issue is tied up with the question of instructions and with the availability of textbooks/learning materials in that language. The argument is also that university students must have high ability in English so as to compete in the national/international job market. However, it must be conceded that learning is facilitated if both instruction and examination are conducted in a language familiar to the learners; and that, in most states where high school/higher secondary education is provided in languages other than English, there is a large body of students (may be a majority) who do not find it easy to cope with university instruction/examination if these are conducted in English.

Keeping these difficulties in mind, most of the universities now do conduct examinations in more than one language, giving the students an option to choose the medium that they are more comfortable with. This, however, raises the problem of preparing tests in different languages, which are equivalent without becoming ambiguous, and also the problem of finding examiners who are well versed in more than one language.

- **Increasing reliability of examinations**

  i) Those who set question papers and those who mark the answer books must be thoroughly familiar with the subject as it is taught in the colleges. In fact the need for training of paper setters and examiners in evaluation cannot be over emphasized. It is desirable that a team of 2 to 3 teachers set the question papers.

  ii) Any paper setter or teacher who writes an examination question, should also provide a model answer. For multiple choice questions, key should be given for each question. For short answer question, there should be a complete answer. For an essay type question it should be an outline giving the value points to be included in the answer and the suggested allotment of marks for each value point. This information is useful in judging the quality of the questions and in increasing the objectivity and uniformity in scoring the answer scripts.

  iii) Another way to increase the reliability of the examination is the centralized evaluation of answer books. The examiners should work together at designated centers under the guidance of chief examiners. They should discuss the answers to questions until they have a common understanding of standards for valuing them. They should work full time at the center until the valuing is complete. This will serve to increase the uniformity of judgements of examiners, and will also reduce the time spent on the marking of answer books. This would ensure more uniformity in scoring of scripts thereby improving reliability of scores.

Self-assessment

1. *Suppose you are given a question paper, in your subject area, which has been prepared by an outside paper setter for the college final examination. You are asked by your Head of institution to see whether the question paper is valid and reliable none.*

   a) How do you judge the validity of the question paper?

   b) How do you judge the reliability of the questions paper?

Question banks

A question bank is a library of questions or test items pooled through cooperative efforts of teachers and item writers, pre-validated, coded and classified as per design of the question bank developed for use as an open or confidential bank, depending upon the use to which it has to be put. A Bank or collection of questions/items can be built for use in university examination and also for classroom tests which are a part of periodic continuous assessment. Certain basic principles in forming such question banks are:

i) A large number of questions/items must be stored. For example, in every paper (course), there are to be, say, 50 objective items, 10 short answer questions and 3 essay/problem solving questions. A bank that can be OPEN to teacher/ students and others, should have about 40 to 50 times the number i.e., 1600 to 2000 objectives items; 400 to 500 short answer questions and 120 to 150 essay questions. The possibility of an answer bank for this is remote, and if the students do memorise the answers to this large number of questions and come prepared, it will serve the purpose of educating them, since thorough coverage of syllabus is ensured.

ii) Teachers on the other hand, should not feel obliged to provide answers to those questions from the bank, instead of teaching/ instructing students on the concerned topics. They should point out to the students that the bank is after all only a sample and an examiner is not bound to select questions only from the bank. However, teachers can use these questions in periodical tests.

iv) If the Controller of Exams/Board of Studies decides to keep the Bank CLOSED (Confidential) to teachers/students, well, the number of questions/items can be limited to say 800 objective items; 200 short answer and 60 essay type questions. They will be constructed, pre-validated and stored in secrecy and only a few items/questions are distributed to students/teachers as specimen. The specimen questions are meant to illustrate what is expected of students and not to bind the examiner. Sample items on all competencies implied in assessment objectives can be given as exemplars.

Question banks should thus help in setting better examination papers as well as classroom tests. You will learn more about question banks in Unit 16 of Block 4.

Internal assessment

a) Internal assessment-that is, periodical assessment of students' learning by their own teachers in the college is gaining more and more acceptance. Teachers speak of it as 'desirable' or 'essential' despite the fact that they recognize the difficulties in introducing such a programme. Internal assessment includes day to day assessment as continuous assessment which is an integral part of teaching also, unit-wise testing of periodical testing or even end of the session or semester testing by college teachers themselves are all examples of internal assessment. Some students are opposed to internal assessment because of the fear of biased teacher ratings. On the other hand, some teachers and students favour internal

assessment as a device to compensate for the fall in marks obtained in external assessment. Hence, one preferred mode is to separate the marks obtained and to show them separately in the final record/certificate. Students accept the idea of having class tests, laboratory work, field work, etc., which are separately assessed and shown in the final record separate from the marks obtained in the final external written examination. Moreover, regular, continuous and comprehensive assessment covering assessment of even non-cognitive learning outcomes, carried out by college teachers themselves, improves the validity of assessment of students.

b) Since evaluation should be an integral and essential part of teaching, an attempt to make internal assessment a regular feature of the programme of the colleges is desirable. However, such an attempt should be undertaken with careful planning and a through preparation of teachers for their combined roles as guides and examiners.

c) The first and the foremost thing should be to develop consensus on the design or basic framework of internal assessment for use in all colleges with provision for some flexibility in different subjects. Another point is to ensure a degree of uniformity in the assessment policy of the colleges, because teachers in the colleges lack a clear understanding of what is really expected of them. It is believed that each university must take a lead in working with the colleges to develop their policies. The primary responsibility in this task can be assigned to the Examination Reform Unit. Obviously, the principals of colleges must be involved. It is also necessary that representative groups of teachers and students be involved in framing the policy. A questionnaire may be used to elicit opinions from all teachers teaching different subjects. Consensus about the design of internal assessment in every subject should be worked out within the broad framework or design worked out at the central level e.g. Association of Indian Universities (AIU).

d) Now, what are characteristics of the students, which should be assessed? It is recommended that right from the beginning the concept of continuous and comprehensive assessment should be emphasized. Assessment should not be limited only to academic achievement in the curricular subjects, but efforts also should be made to develop appropriate tools for assessing co-scholastic aspects like personal social qualities, interests, attitudes, values, etc. It is possible to evaluate these if the traits are adequately defined and procedure of assessing them are developed. Criteria of assessment be listed and weightage to different aspects of assessment be given and relevant tools and techniques have to be developed for collecting evidences about students' attainments.

e) A more gradual, phased approach of starting internal assessment is also possible. Such an internal assessment programme could be stated in all colleges but in only a few subjects, or in a few subjects, or in a new college as a pilot project, or at one level only, it might be advantageous to try internal assessment initially with postgraduate students, before extending it to lower levels.

f) Teachers should be trained through workshops/seminars/follow up activities, to conduct internal assessment programmes. They should be helped to prepare design and blueprint of appropriate tests, both written and performance type, unique to their subject/discipline.

g) It is recommended that the internal marks should be reported along with external examination marks; but the two should not be combined. It is believed that a student's record gives more meaningful information about the student, when both sets of marks appear separately, rather than, when they are combined.

h) It is suggested that the university too develops a Student Record Book, similar to the S.S.L.C. Mark Book currently used in the high schools, for use in

colleges. The student's marks on all college tests and examinations, together with assessment on personal traits, interests, attitudes, and the rank order of his standing in each subject, should be indicated. In the same book, the university should enter student's examination marks or grades. On the cover of the Record Book, the name of the university and the name of the college should appear prominently. Wherever and however the internal marks of students is reported, the identity of the college, which awarded them, the marks, should be known. You will learn more about internal assessment in Unit 14 of this Block.

Grading

Reporting the results of university examinations (internal assessment and/or the final (external) exam), in terms of grades, instead of the familiar numerical marking scale, is now widely advocated in India. For example, a committee of the Central Advisory Board of Education said, that for all public examinations, the results should be declared subject-wise and furnished in the form of grades. The 'raw' marks, even when given by the examiners, should be converted into letter grades before being made available to students.

It is natural that any proposed substitute for the time honoured marking system would be resisted by many. The present system is familiar to teachers and students, to employees, and to the public, and all of them think they know what it means. The trouble is that much of what they know is not true. The numerical marking system is highly misleading.

The traditional marking system on a zero to hundred scale appears to be an absolute scale indicating the exact level of a student's achievement. But it is, in fact, a relative scale. A mark has real meaning only when the numbers are compared to those assigned to other students. And these numbers are by no means free from errors. In all examination results, there are certain standard errors of measurement. A study conducted in Guwahati University reveals that generally, in university examinations in India, the error is about 5 marks and often it is as high as 7 (Taylor, 1964a) and even more in certain cases. Methods are being proposed to make the marking system more accurate. But by and large, all these methods are not adequate to compensate for basic human weaknesses especially when it comes to making precise objective differentiations when the scale is as large as zero to one hunered. Hence many studies reveal that the same examiner gives different marks to the same paper if it is examined on two different occasions, or that a paper which is scored as getting 61% (second class) on another occasion by another examiner. We are also not sure what precise or absolute level of achievement 61 or 54 indicates, although numerically there is a lot of difference between the two numbers.

A grading scale does not attempt to make such fine distinctions and thus it is less misleading. It groups students into a smaller number of categories say five, seven or nine, each representing a range of marks, and we can be more confident that the range within which a student is placed gives an accurate indication of his standing and misclassification of students is minimized.

A university is not justified, and is deceiving itself, when it declares that a fixed number of marks is required for passing an examination. It does not have the evidence that a mark of 35 in this year's examination is equivalent to last year's 35. And in fact, it departs at times, from its own standard. If too few students pass, the marks are moderated throwing away the pretence of the sanctity of the passing mark. It should not have had that sanctity in the first place. The university would do better to admit that the proportion of students who pass, is a more stable base for judgement than a required minimum mark.

Most of us are opposed to moderation of marks. It is, however, believed that even with a grading system, there are borderline cases which ought to be carefully

examined on their individual merit to be sure that injustice is not done to deserving students. Moderation makes no distinction between those who deserve higher marks and those who do not.

It is recommended that to begin with, at the Post-Graduate level, a system of 'direct grading' be adopted. It is possible to work out a paradigm of criteria for graded answers to every question (by a committee of valuers) as to what should be an 'outstanding' answer to a question, an 'average' answer to the same question and so on. Teachers have to be trained in the science of direct grading. Some People, however, are of the view that to begin with grading via marking may be better before introducing direct grading. This will make switch over more easily acceptable to teacher, students and the public at large.

- ***Adopting the grading scheme***

Before adopting the grading scheme, certain preparations are to be made:

i) Past examination results in every subject in the last three years have to be systematically analysed to find statistical indicators like, Mean, Standard Deviation, Reliability Co-efficients, Standard Errors, etc.

ii) Standard Tables for conversion of marks into grade will have to be prepared for every paper (course) and used as an interim measure.

iii) Direct grading should be resorted to, after training in this area has been given to teachers and other.

iv) Information giving details of the system, and implications for teachers, students, parents, and other concerned, must be given to them to highlight the desirability of this change into system of grading and its advantages thereof.

You will know more about marking and grading in Unit 17 of Block 4.

Preparation of teachers and students

It is suggested that, as soon as possible, the university prepares and publishes a manual to be distributed to all teachers in colleges, to give them essential information on the intended changes in examinations. It is desirable that this manual may be given to students as well, who are likely to be affected by the changes.

This manual may contain the following information:

i) A general description of changes, being introduced.

ii) A separate section for each subject providing:

   a) objectives of instruction in terms of intended learning outcomes.

   b) illustrative charts of question papers indicating basic framework or design of question papers.

   c) sample questions based on different competencies implied in assessment objectives (knowledge, understanding, application, skill, etc.) for use as exemplars.

iii) Implications of grading and other changes for teaching, learning and testing.

It would probably be also desirable to organize short term training courses for teachers who have not attended the workshops. There could be 'short courses' conducted in the colleges. The suggested booklet might serve as a hand book for discussion in these courses.

Self-assessment

2. *How do question banks help teachers in the instructional process?*
3. *Point out the drawbacks which exist in the present system of numerical marking.*
4. *In what way the grading system would overcome the drawbacks of numerical marking?*

National testing service

Different Committees and Commissions, set up from time to time, have strongly emphasized the education-employment linkages, the delinking of degrees from jobs, the vocationalisation of education and the preparation of manpower to suit developmental requirements. 'National examinations' was mentioned as a reform in the 'Report on the Examination Reform- A Plan of Action (1973). This idea later developed into the National Testing Service put forth by the National Policy on Education (1986). The policy has stressed on the need for creating a demand for vocationally trained manpower in the employment sectors. This Policy also envisages the establishment of a National Testing Service (NTS), to conduct tests on a voluntary basis, to determine the suitability of candidates for certain specified jobs and to pave the way for the emergence of norms of comparable competence across the nation. Its primary objectives is to allow skilled and proficient people even without the degree qualification, to qualify for a variety of jobs that have traditionally been limited to graduates. The policy also envisages that the administering of a single test on a National basis at the + 2 level can help the students to qualify for admission to courses conducted by different universities, without appearing in multiple entrance tests organized by various agencies. Suitable common entrance tests can be devised for various vocational courses like medical or engineering, at different levels, including that of masters level. Such a national level body would ensure uniformity of standards across which performance standard of students from different states can be compared. In the field of engineering and medicine such efforts are already initiated on similar lines. National Evaluation organization is visualized in NPE (1986) and POA (1992). It will act as a quality control mechanism to organize nation wide tests on a voluntary basis so that norms can be evolved for comparability of performance and also for conducting of independent tests. As a non-profit organization, it will perform developmental activities and conduct in the areas having bearing on students' evaluation.

Self-assessment

5. *Which of the following purposes does the National Testing Service serve?*
   a) *education-employment linkages*
   b) *delinking of degrees from jobs*
   c) *vocationalisation of education*
   d) *preparation of person-power needed for development need*
   e) *all of the above*
6. *Which of the following is not a function of NTS?*
   a) *test development and administration*
   b) *test scoring and computer services*
   c) *certification of eligible candidates*
   d) *appointment of candidates to different posts*
   e) *quality control and monitoring*

Summary

In this unit, we started our discussion with examination reforms/new alternatives and went on to deal with syllabus reforms and question paper reforms. In question paper reform the teacher has to take care of the instructional objectives to be achieved, the design of question papers, blueprint, types of question to be asked and the type of options to be given in the question papers. We introduced you to three major reforms out of four suggested by the University Grants Commission, i.e., question banking, internal assessment and grading. Lastly, we briefly touched upon the objectives of National Testing Service (an extension of the fourth reform of UGC) and discussed the various functions to be carried out by it, besides mentioning the role of National Evaluation Organization as envisaged in the National Policy of Education (1986) and Programme of Action (1992).

Unit-end activities

1) Examine the syllabus which you are involved in. identify the defects as you have observed. Suggest various steps you would take to rectify those defects.

2) Have you ever involved yourself in the centralized evaluation of answer books? If not, associate yourself in such a type of evaluation and evaluate its appropriateness in the examination system.

3) Develop ten questions/items testing different competencies/abilities, which can be used as exemplars for developing question banks.

Points for discussion

1) What steps should one take to increase the credibility of internal assessment?

2) To what extent can internal assessments include students' self-assessment (i.e., students assessing their own progress against a model answer key) peer assessment and teachers' assessment of students of another section of the same class. How can this be integrated with examination reforms suggested in this unit?

Answers to self-assessment

1. a) **Bases for judging validity**

   i) Adequate coverage of instructional objectives

   ii) Due weightage given to assessment objectives

   iii) Proportional weightage to different units of the syllabus

   iv) Adequate sampling of the content of the syllabus

   v) Avoiding free options if given in essay type questions

   vi) Framing of questions reflecting clearly the nature and scope of intended response

   b) **Bases for judging reliability**

   i) Increase in number of questions

   ii) Increase in short answer and objective questions

   iii) Decrease in essay type questions

   iv) Precise and unambiguous language of questions to clarify scope of expected answer

   v) Use of detailed marking scheme

2. a) The question banks help the teachers by providing a list of readymade, standardized questions, which the teachers may use in their question papers. Thus, the teachers do not get pressurized to construct appropriate items at the eleventh hour.

   b) Moreover, the quality of questions being good, they yield more dependable evidences about students' achievement.

   c) A good question bank has questions with calculated facility and discrimination indices, which are useful in setting question paper.

   d) Teachers can select question of their choice depending upon their requirement in terms of assessment objectives and intended difficulty level of question paper they set in different subjects.

3. There are two major drawbacks, which exist in the present numerical marking system.

   a) First, when the same examination paper is evaluated by more than one teacher, score varies from examiner, i.e., inter- examiner differences.

   b) Second, in numerical marking, the value of scores differs from subject to subject, e.g., a score of 60 in English is not equivalent to a score of 60 in Mathematics. While in the former, the score is more valuable, in the latter is not so. These drawbacks happen due to 101 point scale followed in marking. But in grading, since a narrow-range scale is used, these drawbacks are overcome, e.g., the value of grade 'A' in English is equal to the value of grade 'A' in Mathematics. Similarly, there will be hardly any inter- examiner differences in grading.

4. a) Grading takes care of measurement errors and the 3, 5, 7 or 9 grades provide better judgement in place of 101 point scale.

   b) Mis-classification of students is minimized because of grouping of students into much less number of bands (grades) taking into consideration the measurement errors.

   c) Provides more valid comparison of students' performance across the subjects.

5. Answer- e

6. Answer- d

Suggested readings

Association of Indian Universities (1970) *Monograph on Question Banking.* New Delhi.

Association of Indian Universities (1977) *Monograph on Grading for Universities.* New Delhi.

Association of Indian Universities (1977) *Monograph on Assessment for Universities.* New Delhi.

Government of India (1964-66) *Report of the Education Commission.* New Delhi: Government of India Press.

Government of India- *National Policy on Education (1986) and Programme of Action* (1992). New Delhi: Ministry of Human Resources Development.

NCERT (2000) *National Curriculum Framework.* NCERT, New Delhi.

Singh, Pritam (1996) *Towards Better Questions.* New Delhi: Navraj Publication.

Taylor, H.J. (1964a) *An Examination of Examiners.* Guwahati University: (UGC Financed Project).

University Grants Commission (1971) *Principles and Mechanics of the Semester System.* New Delhi.

University Grants Commission (1973) *Examination Reforms - A Plan of Action.* New Delhi.

UNIT

Evaluation Situations in Higher Education-I

13

Introduction

In the preceding unit, we discussed the concept of evaluation and its present status, especially the present day examination system and the different examination reforms that are in vogue in our country today. In this unit as well as in the next, we shall discuss the different methods and techniques of evaluation and the approaches to be followed by college teachers. As practitioners in the teaching profession, you are often engaged in the task of evaluation. In institutions of higher education, evaluation is done in many ways. The most common periodicity of evaluation is the half-yearly and the annual examination. Teachers, however, do evaluate the students' performances by providing unit tests, unit quizzes, weakly tests, monthly tests etc. Teachers also evaluate the students' performances through projects, seminars, symposia, brainstorming sessions and such other interactive situations. All these evaluation techniques serve different purposes. Hence, These techniques can be classified in several ways depending upon the purposes they serve. Tests, which are given to students, may be called as criterion-referenced and norm-referenced, formative and summative, and cognitive and non-cognitive, etc. depending upon interpretation of scores and the aspect of assessment or intended learning outcomes. The main thrust of discussion in the present unit will be based on the purpose of testing.

Learning outcomes

After you have worked through the unit, you will be able to:

- define norm-referenced testing (NRT) and criterion-referenced testing (CRT);
- compare the role of NRT and CRT in the instruction and evaluation;
- distinguish between nature, scope and purpose of formative and summative tests;
- appraise the role of formative and summative tests in the teaching-learning process;
- distinguish between cognitive and non-cognitive objectives and their behavioural outcomes; and
- list assessment tools and techniques for cognitive and non-cognitive objectives.

Norm-referenced testing and criterion-referenced testing

Nature, purpose and scope

In colleges, teachers generally perform two important functions, i.e., teaching and testing. Testing is done to measure the student's performance after teaching with a view to finding out whether the learning has taken place. Students' performances are generally reported in terms of marks or grades. But these marks or grades are meaningless, unless these are interpreted. So the common practice prevailing in the colleges is to compare the performance of a student with a certain norm, i.e., local, regional or national. Let us now analyse the following statements:

- Raman stood first in the class.
- Malhotra's score is the lowest in the group.
- Panda secured 10th position in H.S.C. examination.
- Mukherjee scored at the 76th percentile.

For arriving at a meaningful conclusion to the above statements two criteria are needed. They are: a) the specific learning outcomes that were measured by the test and b) the level of achievement of the group of students who have taken the test? Here, the group of students refers to the *norm group*. This is the group that took the test. If the test is standardized, as distinguished from a classroom test, such a group is a carefully selected and well-defined group, representing different age levels on a national, regional, state or local level. The classroom tests and standardized achievement tests that report test performances in terms of an individual's relative position in some group, are called *norm-referenced tests*. Most of our modern testing theories are based on this norm, wherein a good spread of marks is expected so that selection/classification of students can be done.

Contrary to norm-referenced testing, many college teachers interpret the test performance against a prefixed standard or criterion-defined level of achievement-knowledge and/or skills in that subject.

Let us take a look at the following statement:

i) Given 50 minutes, out of the 100 objective items, Hari did 90 items correctly.

ii) Mathur can type 60 words per minute with not more than 6 errors.

iii) Given 10 numerical problems on calculation of facility indices students should be able to work out at least 8 correctly within a period of 20 minutes.

If we analyse the above statements, we find that here the students are being described in terms of a predetermined standard of performance. Take the first statement, the teacher might have expected every student to do at least 75 items correctly in 50 minutes. Since, in this case, Hari has answered 90 correctly, the performance of Hari is far better than the standard kept in mind by the teacher. Here, the performance of the student is described in behavioural terms without any reference to the marks obtained by the other students in the group. The level of performance that is acceptable is predetermined, based on the importance of the unit/task concerned. Taking this as the absolute standard, each student's performance is compared. Since a criterion/standard, rather than a relative position in a norm-group, is used for describing test performance, such interpretations are called *criterion-referenced*. Hence, the interpretation of the test results reveals whether a particular test is criterion-referenced or norm-referenced. Let us study a little more about these two types of testing.

Why norm-referenced testing (NRT) and criterion-referenced testing (CRT)?

It was generally accepted by the teachers that the test results are not uniform for all the learners. Some students' scores are very high, while others' scores are very low. Teachers believe that individual differences in learning exits because all students are not equipped with the same intellectual abilities to benefit from teaching. This assumption gave birth to the norm-referenced measurement of intellectual abilities. The aim of the norm-referenced testing is to rank order the students in a class, rather than assess their attainment of specified curriculum objective. Hence, in norm-referenced measurement, a particular student's achievement is evaluated in terms of the comparison between his performance and the performances of the other members of the group. It hardly gives any information regarding the attainment of the intended objectives by the individual learner.

The researches on efficient strategies and the growing realization of the impact of environment on learning, led educators to think that student's performances, irrespective of their innate abilities, can be increased through an appropriate teaching-learning environment. And this gave rise to the concept of mastery-learning. The basic idea behind this was that any student can attain mastery if he is provided with adequate time and an appropriate teaching-learning environment. In the context of mastery-learning, a new method of measurement, known as the *criterion-referenced* testing, emerged. This type of testing emphasizes that teachers should interpret the test results, not in comparison to the performance of the group, but in relation to a specific performance criterion. The word criterion in criterion-referenced tests denotes:

a) an instructional objective;

b) an expected post instructional learning outcome;

c) an intended level of students' performance;

d) an acceptable level of learners achievement;

e) a desired standard of product of performance.

Individualized instruction and CRT

Alongwith the mastery-learning concept, there grew the movement of Individualized Instruction-Programmed Learning or Self Instruction. Here, the emphasis was on preparing instructional materials in such a manner that a learner would learn at his/her own pace and time. Experts from the Programmed Learning movement propagated the idea of CRT. According to Robert Glaser (1963) 'Criterion-referenced testing is deliberately constructed to yield measurements that are directly interpretable in terms of specified performance standards; W. James Popham adds "A criterion referenced test is used to ascertain an individual's status with resect to a well-defined behaviour domain". Here, "well-defined behaviour domain" refers to a class of learner behaviour. With this background, let us have a look at the advantages of both NRT and CRT.

Advantages and disadvantages of norm-referenced testing advantages

a) **Advantages**

i) Norm-referenced tests serve in classifying students for selection purposes. The purpose of NRT is to differentiate among the students.

ii) Parents are provided with information regarding their children's progress vis-à-vis other students.

iii) They help in certifying students for employment, indicating their attainment level vis-à-vis other students level.

iv) They help the teacher to provide educational and vocational guidance to students to improve their achievement.

v) They help in curriculum construction and also for curriculum validation.

b) **Limitations of NRT**

i) One major disadvantage of nor-referenced testing lies in the inaccuracy of measurement. An NRT measurement hardly indicates the difference between the performance of A and B who have secured 60 and 59 per cent respectively in the same test. The difference in performance between two such pupils is not easily demarcated because of some measurement errors, otherwise both have performed rather alike on the test.

ii) Scores on NRT do not indicate how far instructional objectives have really been achieved

iii) Classification of students made on the basis of NRT as 1st, 2nd or 3rd divisioners is misleading.

Advantages and disadvantages of criterion-referenced testing

a) **Advantages**

i) A properly constructed CRT, which reflects clearly the instructional objectives in behavioural terms, helps a teacher to find out the extent to which his students have achieved these objectives.

ii) It encourages individualization of instruction for improvement of students' learning.

iii) It motivates teachers to improve upon their instructional strategies and preparation of teaching-learning material.

iv) It provides teachers better evidence to diagnose students' levels of achievement to the parents.

v) The progress reports of students provide a clear picture of students' levels of achievement to the parents.

vi) It encourages students for mastery learning and the teachers for mastery testing.

b) **Limitations of CRT**

i) In a selection situation, where only a certain number of candidates are to be selected a CRT may not differentiate amongst the group, while a NRT, can better discriminate for ranking purpose.

ii) Since the focus of CRT is on mastery testing, it would demand a lot of time on remedial teaching, which may not be practicable.

iii) Retesting for judging the mastery after each remedial programme on intended learning outcomes has to be undertaken, which demands unaffordable additional time from teachers.

Self-assessment

1. *How do you differentiate between norm-referenced testing and criterion-referenced testing? List two advantages of each.*

2. *Given below is a list of statements. Write NRT against the statement, which reflects norm-referenced testing and CRT against the statement, which reflects criterion-referenced testing.*

   a) *Patel stood second in the final B.Sc. examination* .....

   b) *Reddy cycles at a speed of 20 kms per hour* .....

   c) *Yadav Scored at the 55th percentile rank*.....

   d) *Mishra Scores 5th position in CBSE Examination* .....

   e) *When asked Bose spells 8 words out of 10 words correctly in 5 minutes* .....

Formative and summative tests

In the previous section, we discussed one way of classifying tests and examinations as NRT and CRT, from the viewpoint of interpreting the results. We will now classify the tests and examinations from the viewpoint of the sequence in which they are likely to be used in classroom instruction. In colleges, teachers evaluate the performance of student even while they continue with the instruction, as also at the end of instruction. Very often, teachers put a few questions at the end of every class to review what is taught or when the teaching of a certain skill or a unit is complete. In colleges such tests are mostly teacher-made tests. This type of evaluation is called formative evaluation and the tests used for this are called *formative tests*. College teachers also administer tests to students at the end of a semester or course of instruction. This means that evaluation is done at the end of a course of instruction. Such a type of evaluation is called summative evaluation and the tests used for this are called *summative* tests. Let us now discuss in detail the need for formative and summative tests.

Need for formative and summative tests

Right from the beginning of formative assessment, there have been many attempts to adopt formative assessment to enable learners to understand their progress in learning, their weaknesses and to take the necessary remedial measures to correct learning. The need arises from the fact that there are individual differences in learning and in the quality of learning. Also, the pace of learning is different and so, the time within which a certain amount of learning is to be expected varies from person to person. The need for a formative assessment is to be seen in this perspective that in the matter of learning, a student needs to be assessed in order to provide information to him as to the quantum, quality, pace and amount of learning he has acquired, so that this acts as a feedback to him. The formative assessment is thus a feedback to the student. This feedback is to be seen in terms of cognitive, affective and psycho-motor characteristics included in his learning. Formative assessment also serves as a feedback to the teacher as to which part of the course is not clearly taught, and which objectives not clearly achieved. It is this feedback that makes formative assessment a useful mode of assessment and helps the teachers adapt their instructional strategies for improving students' achievement.

We have seen that formative assessment is to serve the twin purposes of feedback to students in monitoring the progress of learning and to the teachers on the efficacy of their teaching. The feedback to students is mainly to monitor the progress of learning when the learning takes place, either at an assumed rate of learning or at their own rate of learning. The problem of measurement in this context is to take care of the setting of standard level of achievement, against which 'students' own levels are compared. The formative tests, therefore, measure the students' learning levels against this standard set up in the test. The decisions that are taken, amount to the success of failure of any individual to come up to this standard. The strength of testing lies in revealing what has been learnt correctly and precisely. To get this information on the strength or weakness of the learning, on the part of an individual, is precisely the major purpose of formative assessment. Formative assessment, therefore, is not very meaningful in terms of marks. It becomes slightly more meaningful when grades are given and it becomes extremely meaningful when the results of the test are discussed with the students and the weaknesses of the learners are identified and remedial measures suggested. Traditionally, such formative assessment has taken the format of oral test at the end of every class or a test used at the end of module or unit or even a test for testing the mastery and acquisition of a certain ability. This may take the form of a 'paper-pencil' test, a skill test or an oral test. Whatever be the form, its main purpose is to draw attention to the weaknesses of learning and the remedial measures with which to improve upon. One of the modern formats of formative evaluation is the use of a flow chart that

includes instructions, safety checks, diagnostic questions and remedial loops, all in one package easily comprehensible to students. In such a formative testing the emphasis is on diagnostic questions and responses of the students to those, so as to bring forth the weaknesses of learning.

In the comprehensive learning package mentioned above, remedies for a set of assumed weaknesses of learning are also suggested. Another format is to use a programmed instruction (learning) approach, which consists of a self-paced learning material component in addition to what is contained in the flowchart. In the programmed instruction (learning) sequence, formative evaluation is built in; so not only does the learning take place in steps, but it is also checked in steps, and remedied in steps. (See section on programmed learning (PLM) in Unit 3 of Block 1.)

Formative-diagnostic oriented assessment in any format requires a plan- a table of specifications indicating the level of abilities and the specific content area.

Need for summative tests

Summative tests are usually given at the end of a course a semester. It could be used at the end of a unit of instruction when purpose is to grade the students. Unlike the formative test, the aim here is not to monitor the progress of the students during instruction but to assess and grade the over all performance of the students at the end of a term, semester or the course in order to determine the extent to which the instructional objectives have been achieved. This also helps the teacher to do course evaluation.

The second important objective of summative evaluation is to assign course grades or certify pupils in terms of the intended learning outcomes (ILOs). This ensures the students' promotion to the next task or course. And this in turn facilitates in classifying students' according to their abilities, and their selection for different purposes, such as admission to a new course, selection for a particular job, etc. the techniques used in summative evaluation are determined by instructional objectives, but they typically include teacher-made achievement tests, ratings on various types of performance (e.g. reports). Though the main purpose of summative evaluation is the certification of pupil achievement, it also provides evidences for judging the appropriateness of content (syllabus), the course objectives and the effectiveness of the instruction. It helps to validate the whole teaching-learning process.

Distinction between formative and summative tests

i) **Purpose of evaluation**

The main purpose of formative evaluation is to determine the degree of mastery of a given learning task, and also that part of the task that has not been mastered. Thus the emphasis is not on grading or certifying, but on helping both the learner and the teacher to improve the teaching-learning process. Summative evaluation, on the other hand, is directed towards a much more general assessment of the degree to which the larger and more general outcomes of learning have been achieved over the entire courses or a substantial part of it. To cite an example, at the end of a first semester in Physics, (electricity and magnetism) summative evaluation will have the major objectives of determining the degree to which the students have learnt to interpret the various principles and concepts, and also to apply them to new or unfamiliar situation. The purpose, therefore, is to grade the students and report to the administrators. On the other hand, the purpose of the formative evaluation, in one unit dealing with electromagnetic induction, will be to determine whether a student has learnt to define, interpret, translate and apply the principle of electromagnetic inductions. If the learner is not successful in doing this, he is given further support material and resource to learn to do such things. Certainly here, the major purpose is not to grade him or certify his learning but to get and provide feedback to learner, and improve teaching-learning strategies. A teacher can and should use both formative and summative evaluation for their appropriate purposes.

ii) **Periodicity of evaluation**

Another distinction is that tests for formative evaluation are given at regular and frequent intervals, while for summative evaluation , they are given at the end of considerable learning over a fairly long period of time. In a course that extends over a period of sixteen weeks, a test at the end of every week is a formative test, while the test at the end of sixteen weeks of learning, will be summative. A test at the end of a teaching unit/topics formative while an annual test, a semester end or course end tests are summative since they are used for grading/promotion/certification.

iii) **Level of generalization-transfer of learning**

Perhaps the level of generalization will be the factor that differentiates formative from summative more sharply. While in formative evaluation one may be interested in assessing knowledge, skills or the ability to apply a given principle in an unfamiliar situation, in a limited context, in the summative evaluation, one may try to assess the acquisition of knowledge, skill or the ability to apply it in a wide variety of situations that integrate several principles covering bigger chunk of syllabus. Focus of formative evaluation is on assessment of students' competencies related to a restricted sample of content or selected chunk of syllabus. While in summative evaluation, the population of items is large due to much bigger chunk of syllabus to be covered. Therefore, There is a higher probability of sampling error. Hence the marks/ grades obtained in summative evaluation suffer from the sampling limitations.

Planning summative evaluation, therefore, requires a careful process of sampling not only the content, but also the abilities specified under various assessment objectives. This is usually done through a table of specifications. Such a table should not only indicate the topics separately but also their integration wherever desired. The validity well as the assessment objectives and the proportional weightage given to different content areas and abilities (implied in assessment objectives) to be tested in an integrated form. Content validity and curricular validity are thus, built into framework of summative evaluations.

Types of formative and summative evaluation

The types of formative and summative evaluation tools can take the form of:

i) written or oral quiz (announced or unannounced)

ii) short duration objective type tests consisting of items of Constant Alternative (CA), Multiple-choice (MC), Multiple Facet (MF), Matching (M), Multiple Completion (MC) Rearrangement ® and Combined Response (CR) variety

iii) fixed response-open ended questions

iv) short answer questions

v) short essay type questions (restricted response essay questions)

vi) practical exercises.

In formative testing as pointed out earlier, the emphasis is on checking learning, locating the weaknesses of learning and therefore suggesting remedial steps for improving the learning. The summative evaluation can take the from of a judicial combination of objective type items, short answer questions and long answer questions for better sampling of syllabus. Such a combined form may look similar to the one described in the earlier unit (i.e. Uuit 12). In a nutshell focus of formative testing is on diagnosis, feedback, remediation and improvement of students' achievement whereas summative testing is judgemental in nature with focus on measurement of achievement, grading and certification of achievement.

Self-assessment

3. *Classify each of the following by indicating whether it rejers to formative evaluation or summative evaluation:*

   a) *A Sociology studies term test used to assign grades.....*

   b) *A test used for observing and recording reading errors .....*

   c) *A unit test in science used to certify mastery of concepts .....*

   d) *A short quiz in mathematics to determine learning progress of the learners.....*

   e) *A unit test to identify nature of errors students commit in a Mathematics .....*

   f) *A language test to provide feedback to the learners .....*

   g) *A Biology test on respiration in plants to get feedback on instructional effectiveness.....*

4. *List two bases on which we can classify tests as formative or summative?*

Cognitive and non-cognitive assessment of learning outcomes

Still another way of classifying evaluation activities, is on the basis of aspects of human development or learner behaviours to be measured and evaluated.

Some of these behavarious are related to the student's intellectual development, i.e., his knowledge, understanding and ability to apply the knowledge in the new situations. These behaviours are sometimes described as non-cognitive behaviours or learning outcomes.

A teacher or educational planner, while specifying the content and the related learning experience for an education programme i.e., while developing a curriculum plan should describe both cognitive and non-cognitive learning outcomes as desirable behaviours of that programme, i.e., the intended learning outcomes.

It is, however, true that most universities' syllabi at present refer only to content- the list of themes/topics, which only partially indicate the desired cognitive outcomes to the total exclusion of non-cognitive outcomes. The list of contents- the syllabi gets further narrowly interpreted as to indicate that the only desirable cognitive outcome is the memorization of the content or factual information (limited to knowledge objectives) while higher order abilities are relegated to the background.

A good teacher, however, always tries to go beyond such a narrow interpretation and attempts at helping his students to achieve higher level cognitive learning outcomes, as well. Only a few teachers bother to develop specified non-cognitive outcomes and also tries to assess both these types of intended learning outcomes. The obvious reason is the absence of non-cognitive outcomes in external examinations. Their assessment even in college as a part of internal assessment would remain a neglected aspect unless some weightage to such outcomes is given meaningful entries in the certification of achievement and taken cognizance.

Objectives of assessing cognitive and non-cognitive outcomes

There are many purposes that such an assessment would serve. Some of these are listed below.

From the teachers' perspective:

i) to monitor learning progress and locate weaknesses of learners in order to suggest remedial action;

ii) to monitor the effectiveness of teaching;

iii) to make decisions for the future, regarding choices of subjects, courses and career;

iv) to predict the future successes of the learners;

v) to choose the suitable candidates for a course or a class;

vi) to assign students with exercises, reading materials, etc., based on his/her progress up to that point.

From the students' perspective:

i) to build a realistic self-picture;

ii) to set educational, vocational, personal and social goals;

iii) to determine the emphasis to be put on different areas of instruction;

iv) to determine areas- subjectwise or workwise, which would provide satisfaction;

v) to ascertain areas of interest;

vi) to identify changes in character/value pattern;

vii) to assess the curricular experiences of pupils;

viii) to get accurate and objective information regarding their aptitudes, achievements, interests, values, etc.

Difficulties in assessing non-cognitive outcomes

It is a common observation that in formal examination setting, teachers and examiners do not attempt to assess the interests, attitudes or the character development of students. There are many reasons behind this apathy. Some of the important ones are;

i) Non-cognitive abilities are not easily observable as the cognitive abilities. Some of them are not at all observable. Therefore, they seem to be non-assessable

ii) Non-cognitive abilities like interest, attitudes, beliefs and values are generally regarded as private matters. For example, some teachers think that attitudes towards God, home and family are private matters and this privacy should be respected and hence, no attempt to disclose them should be done.

iii) Non-cognitive abilities develop slowly and become visible only after a long period of time, perhaps even several years, hence they are not considered as the concern of the teachers. Teachers think that their duty is just to assess the academic achievement and not the other aspects of development.

iv) It is difficult to assign quantitative values to the non-cognitive abilities as it is difficult to differentiate the degree of their intensity, etc., based on the unstructured indictors of these outcomes.

v) Most of the non-cognitive abilities are not very clearly defined. This vagueness in definitions also creates apathy towards assessment in this area.

vi) Development vis-a-vis assessment of non-cognitive behavioural outcomes are not reflected or given any weightage in external examination and therefore, considered non-essential.

Identification of areas in cognitive and non-cognitive abilities

a) **Cognitive domain:**

When we talk of identifying or specifying cognitive and non-cognitive outcomes, it is very difficult to pinpoint the exact list of learning outcomes under both the categories. However, on the basis of work done in the field of taxonomy of educational objectives and its use all over the world and in India, we can categorise them into a classificatory scheme. Although classification appears to be arbitrary, yet it serves a number of useful purposes. It highlights different categories of cognitive and non-cognitive outcomes, which the teachers may consider as desirable learning outcomes and hence worthy of assessment.

Description of cognitive domain taxonomy (By Bloom)

This taxonomy is based on four basic postulates.

i) Behaviours designated are cognitive.

ii) These are hierarchical in nature.

iii) These are cumulative in nature.

iv) Behaviours are learned behaviours.

The following objectives in the cognitive domain have been identified by Bloom and his associates (1956). The six objectives have been listed hierarchically ranging from the simplest to the most complex. Each objective is further defined into sub-sections.

1) Knowledge
   - Knowledge of specifics
   - Knowledge of universals and abstractions in a field
   - Knowledge of universals and abstractions in a field

2) Comprehension
   - Translation
   - Interpretation
   - Extrapolation

3) Application
   - Ability to use a theory, a principle or method to solve a problem involving a new or unfamiliar situation.

4) Analysis
   - Analysis of elements (identifying assumptions, and logical fallcies)
   - Analysis of relationships
   - Analysis of organizational principles

5) Synthesis
   - Production of a unique communication
   - Production of a plan or set of operations
   - Derivation of a set of abstract relations

6) Evaluation

- Judgement in terms of internal criteria
- Judgement in terms of external criteria

b) **Non-cognitive domain**

In the beginning, we should not go in for a very ambitious programme for assessing all the areas listed in the following table. Initially we should restrict ourselves to a few relevant categories, concentrate on them, prepare feasible scheme for providing learning experiences and assessment, and execute them. We can then review and modify these schemes with a view to improving them. And only then we proceed to next categories. It would be desirable to prepare semester-wise or year-wise schemes for assessing these learning outcomes.

Table 1: Non-cognitive aspects related to higher education

Non-cognitive category	Specific areas to be measured at the higher education level
i) Habits	Regular and distinctive habits related to academic activities. (e.g. study habits and work habits)
ii) Interests	Classroom interests, laboratory interests, institutional interests, other curricular and co-curricular interests and national interests.
iii) Personal qualities	Regularity in work, perseverance, originality, initiative, open mindedness, keenness, imagination, persuasiveness, inventiveness, creativity.
iv) Social qualities	Empathic (participating in feelings of others) qualities, Social maturity, cooperation, and leadership.
v) Adjustments	Individual, classroom, institutional, curricular and co-curricular adjustment.
vi) Attitude	Rational attitude towards-self, college, teachers and university activities and examination, national activities and programmes.
vii) Sociometric	How s/he relates to the other members of the student community? Is s/he popular, isolated, neglected? Is s/he a member of gang/clique, etc,?
viii) Appreciation	Does s/he appreciate social/ aesthetic aspects of life-e.g. art forms, role of objectives in testing.
ix) Values	Values related to faculty and institution, personal, economic, social, national, moral, universal and spiritual values etc.

Tools and techniques for assessment of cognitive and non-congnitive outcomes of learning

Different tools and techniques are available for the assessment of cognitive and non-cognitive aspects of learners' growth. So far as cognitive domain is concerned, they are generally tests and examinations. As regards non-cognitive behaviours, they include formal observations, personality and interest inventories, attitude scales and tools like the projective tests designed primarily for use by clinicians. Let us, now, discuss these tools and techniques both for cognitive and non-cognitive abilities.

a) **Tools and techniques for measuring cognitive abilities**

Since we will discuss in detail the various modes of assessing cognitive abilities in Unit 15 of Block 4, in this unit let us study more about assessing non-cognitive domain.

b) **Tools and techniques for measuring non-cognitive abilities**

In educational situations, we may use the following tools and techniques for measuring non-cognitive learning outcomes.

i) **Tools and techniques based on observation**

a) Anecdotal record

b) Checklist

c) Rating scale

d) Observation schedule

ii) **Tools and techniques based on product assessment**

a) Drawing and painting

b) Direct questioning

c) Written tests

d) Autobiography

iii) **Tools and techniques based on sociometric status**

a) Sociograms

b) Social distance technique

c) Guess who technique

iv) **Tools and techniques based on self-reported inventories**

a) Biographical data bank

b) Personality inventory

c) Interest inventory/rating scale

d) 'Q' techniques

e) Attitude scale

v) **Tools and techniques based on ideas of projection**

a) Inkblot technique

b) Picture presentation technique

c) Presentation of verbal stimuli

d) Sentence completion technique

e) Play technique

vi) **Tools and techniques based on school records**

a) Attendance register

b) Learner's history

c) Cumulative records

Some of these tools and techniques will be discussed in details in Unit 18 of Block 4. These tools and techniques may also be classified in the following manner.

Table 2: Assessment tools and techniques for non-cognitive abilities

Subjective Methods	Objective Methods	Projective techniques
Anecdotal record; Observation; Direct question; Autobiography; Interviews; Questionnaries; Project and field work; Assignments	checklist; Rating scale; Study of physiological changes; Situational tests; Sociograms; Attitude scale	Inkblot; Picture projection; Presentation of verbal stimuli; Sentence completion; Role playing; Drawing test

Self-assessment

5. *Given below is a list of statements. Tick (√) that statement which is a measure of cognitive outcome and mark a cross (×) against the statement, which is a measure of the non-cognitive outcome.*

   a) *A history teacher asks the students to write down the landmarks in the Indian Independence Movement.* ( )

   b) *The Principal of a college assesses the attitude of the teachers towards organizing co-curricular activities in the college.* ( )

   c) *A language teacher assigns the task of composing a patriotic song to his students.* ( )

   d) *The teacher in charge of N.S.S. found that the students were taking keen interest in community participation.* ( )

   e) *A science teacher observed that the students had followed the right procedure while preparing oxygen.* ( )

   f) *The class teacher selected Ravi as the Leader of the class.* ( )

Summary

In this unit, we discussed the three major ways followed by college teachers while evaluating the performance of students. We presented three different ways of classifying tests and examinations. The first line of classification is on the basis of interpretation of the test results. Thus, we have *norm-referenced testing* (NRT) and *criterion-referenced testing* (CRT). We examined the merits and demerits of both NRT and CRT. The second line of classifying tests and examinations is on the basis of the sequence and purpose for which evidences are used either for feedback or for judgemental purpose. Accordingly we have *formative* and *summative tests*. We discussed the need of both formative and summative testing in instructional process. Towards the last part of the unit, we classified tests on the basis of the nature of learning outcomes to be evaluated. Thus, we have evaluation of *cognitive* and *non-cognitive outcomes*. We identified the different areas in cognitive and non-cognitive domains to be evaluated. We also listed a number of tools and techniques to be used for measuring the cognitive and non-cognitive outcomes besides listing tools and techniques for non-cognitive outcomes. In the next unit, we shall learn more about the other methods used to evaluate the students' performance in institution of higher education.

Unit-end activities

1) From your past college experiences, list some tests or other evaluation instruments, which according to you, are inappropriate. For each of them, describe how they should have been handled?

2) In an undergraduate college (students' age group ranges from 18 to 22) of Arts and Commerce, what non-cognitive aspects of behaviour would you like to develop and then evaluate? Prepare a detailed scheme to develop and evaluate them.

3) Keeping in view the Indian context in what way do you think advantage of norm-referenced and criterion-referenced tests can be taken for more valid and reliable testing and for certification of students.

Points for discussion

1) What are the reasons that most colleges, at present, do not provide enough attention to the non-cognitive development of the students? Are there any ways of overcoming these difficulties/

2) What are the agencies in your region, which use tests for selection purpose (e.g. State Public Service Commission)? What kind of tests do they use? What type of cognitive and non-cognitive abilities of the testes do these tests measure?

3) Develop a chart of your own listing cognitive and non-cognitive outcomes which you think must be made a part of the syllabus for comprehensive assessment of learners.

4) What difficulties do you visualize in using various types of tools and techniques described in this unit?

Answers to self-assessment

1. We differentiate between norm-referenced testing and criterion-referenced testing on the basis of the interpretation of test results. In norm-referenced testing, the students' performance is evaluated in terms of the comparison between his performance and the performances of the other members of the group. whereas in criterion-referenced testing, a student's performance standard or criterion, usually defined in terms of intended learning outcomes/ competencies.

   Advantages

   (NRT)

   a) NRT is useful in ranking students and a student's performance can be judged with reference to his class, district or state.

   b) These tests are helpful in making selections across the institutions of the states and from the basis for reflecting and setting performance standards.

   (CRT)

   a) CRT is useful in determining the level of mastery attained with reference to criterion behaviour or intended learning outcomes.

   b) Use of CRT provides feedback on attainment level of students on desired or intended competencies and becomes the basis for remediation.

2. a) NRT, b) CRT, c) NRT, d) NRT, e) CRT

3. a) Summative, b) Formative, c) Summative, d) Formative, e) Formative, f) Formative, g) Formative

4. The purpose for which they are used:

   a) The sequence in which evaluation procedures are likely to be used in classroom, e.g. readiness test, review test, unit tests, diagnostic tests are used in instructional sequence, purpose being to get feedback on students' learning or progress.

   b) Periodicity or frequency of the tests in use, e.g. unit-wise tests, review tests etc. arte frequently used and are formative but a semester, annual, sessional or course end tests are summative because purpose is to grade, promote or certify students' achievement.

5. √ b) × c) √ d) × e) √ f)

Suggested readings

Bloom, B.S., et. al. (1956) *Taxonomy of Educational Objectives: Handbook 1, Cognitive Domain.* New York: David Mckay Co.

Bloom, B.S., Hashings, J. Thomas, Madaus, George, F. (1971) *Handbook of Formative and Summative Evaluation of Student Learning.* Mc Graw-Hill Book Company.

Glaser, Robert (1963) *Instructional Technology and the Measurement Learning Outcomes.* American Psychologists, Vol. 18, PP. 519-21.

Grondlund, N.E. (1981) *Measurement and Evaluation in Teaching.* (4th Ed.), New York: MacMillan.

Gronlund, N.E. (1973) *Preparing Criterion-referenced Tests for Classroom Instruction.* New York: MacMillan.

Popham, W.J. (1973), *Criterion-referenced Measurement: An Introduction.* New Jersey: Englewood Cliffs; Educational Technology Publications.

Singh, Pritam (1970). *Criterion Referenced Testing-A Selected Readings.* NCERT: New Delhi.

UNIT

14 Evaluation Situations in Higher Education-II

Introduction

As we have learnt earlier, the objectives of an instructional situation determine the choice of instructional techniques, methods or media. Similarly, it is the objectives of an instructional situation that determine the nature of evaluation techniques to be made use of. In the most commonly found instructional situations of higher education, student-learning is mostly confined to acquiring knowledge or other cognitive abilities which manifest in terms of a final product . Such learning, therefore, is tangible for assessment through a conventional written test or term paper . However, there are certain **instructional situations such as laboratory work, field, project work, seminar, etc., where the emphasis is on the students' learning certain process-oriented psychomotor skills and related competencies.** The kind of approach one takes up for evaluating student-performance in such instructional situations depends essentially on these situations that form the basis to serve the intended instructional functions. In all such instructional situations, which involve process-oriented learning, student-achievement should cover both the process and the product of performance. Evaluation of student-performance on process skills and related competencies can be based on the behaviour manifested by the students while they are undergoing the instructional process, which the situation puts them through. Thus, evaluating the process dimension in such situations has a utility as an assessment exercise as well as a diagnostic device and remedial exercise, as is the case with most formative evaluation situations. In this Unit, we will discuss evaluation of laboratory work, seminar, project work and dissertation. We will also talk about internal and external evaluations.

Learning outcomes

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

- prepare a checklist to assess the process-oriented aspects of learning through laboratory assignments' (requiring yes or no, right or wrong, true or false etc.);
- develop a 5 point rating scale to assess qualitative aspects of process oriented learning outcomes of laboratory experiments;
- observe a seminar and assess the students' performances in the seminar with the help of an observation schedule;
- assess any of the dissertations you have guided or come across as an external evaluator using an assessment scheme;
- compare the limitations and advantages of internal assessments with those of external year end examinations; and
- work out a plan of introducing internal assessment in your own institution to reduce teacher's subjectivity to the minimum.

Evaluation of laboratory work

Why do we have laboratory work as part of the instructional programme? What objectives does it help to realize? As we have seen in the Unit on alternative components (Unit 3 of Block 1), laboratory work is essentially a means to develop in students, certain intellectual processes and skills, like hypothesizing, identifying variables, controlling variables, selecting relevant procedures, establishing cause and effect relationships, interpreting data, inferring and predicting. It also serves to develop skills of observation, dexterity with which instruments are used and specimens are handled, skills of processing and presenting data, etc. It has the potentiality to develop certain affective attributes, such as those associated with scientific attitude or scientific temper, an appreciation of the processes of scientific inquiry, a positive attitude towards compliance to safety rules, unbiased outlook, open mindedness, intellectual honesty, etc. The skills mentioned above usually manifest in terms of student behaviour during the laboratory session. The affective outcomes are more difficult to discern but, even they can be seen manifested in the student behaviour, if observed systematically over a long period of time, such as, the duration of an entire course.

The only way perhaps to evaluate students' performance in the laboratory to systematically observe them while they are performing their laboratory tasks involving various processes of performance. The evaluation will not, of course, be complete, unless one scrutinizes the laboratory reports (practical records) in which the data collected are presented and the final results reported. These aspects pertain to both the process and product of performance. What happens many times in the actual laboratory situation in our undergraduate classes, is that evaluation is confined to merely the scrutiny of the laboratory report (practical record), which represents the product aspect of laboratory work and the process-aspect is altogether ignored. The major reason for this is that our college teachers are often unaware of the ways of systematically observing the process aspect of laboratory work.

What does the systematic observation of laboratory work entail? It involves anticipating the various categories of students' behaviours that constitute laboratory work, and recording on the basis of observation, whether or not a student is manifesting each of those behaviours. Perhaps rating on the basis of observation as to how well a student manifests each behavior indicates his performance. Therefore, what needs to be developed is a checklist, a rating scale or also an observation schedule and a scoring sheet for systematic observation, recording and reporting.

It is not a highly technical job to develop one's own rating scale or a checklist. One does not always have to go through elaborate procedures of standardizing and validating the instrument. For day-to-day use, the most pragmatic arrangement is to develop a teacher-made observation schedule. The observation schedule may be in the form of a list of things we expect the student to do, with a provision to check whether s/he does each of those things listed or not. This is nothing but a 'checklist' of observable processes representing intended outcomes of laboratory work.

The checklist that a Physics teacher develops for evaluating the student performance in a 'practical' class may take the following shape (see Box-1), which has been developed by Garg and Ghosh (1990) when they had a typical experiment performed at the undergraduate level, with Physics laboratory in mind.

Please examine this critically and see if you can develop a similar one that suits your specific needs.

Box–1

Measurement of 'Y' by the the Method of Flexture

Checklist

(Put a tick ($\surd$) against those steps enumerated below, which the student performed correctly.)

i) a) Determine : The vernier constant of the callipers

b) Measure : The breadth of the beam

ii) a) Determine : The least count of the screw gauge

b) Measure : The depth of the beam

iii) Set the beam properly between the knife-edges to have the desired length under study.

iv) Measure the vernier constant of the trowelling microscope(TM)

v) Place the frame containing the indicator (pin) properly on the beam (i.e. it should be at the center perpendicular to the beam).

vi) Focus the pin at the TM.

a) Measure the height of the tip of the pins from the surface of the table and bring the operating part of the TM at the same height.

b) Use one hand to place the paper (or anything easily identifiable) a little away from the outer end of the objective lens of the TM. See through the microscope and use the other hand to rotate the knob till the object is sharply focused. Now measure the approximate distance of this object from the objective lens. Let this distance be 'X'

c) Adjust the knob of the microscope so that the front side of the operating tube is approximately at a distance 'X' from the pin. The pin should get focussed by this arrangement. If not, then give slight movement to the operating tube in the possible directions and you will get the sharply focussed image of the pin. Make sure that there is no parallax between the tip of the pin and the horizontal cross-wire.

vii) Take load vs. depression data

viii) Draw load vs. depression graph

ix) Take any one point on the graph, other than the experimental point to calculate 'Y'

x) Point out the sources of error.

On the other hand, an observation schedule can also take the form of a list of things that the teacher expects the student to do during the laboratory work with a provision for the teacher rating the performance of the student on each of those things listed. The teacher may rate the performance on each item on the list as to whether it is good, average or poor (3 point scale). The rating can also be as to whether the performance is very good, good, satisfactory, fair and poor (5 point scale) in certain cases. It can also be as to whether the performance is 'adequate' or 'inadequate' (2 point scale). Whatever be the nature of rating, this kind of instrument is a 'rating scale'. A Biology teacher may think of developing a rating scale for the purpose of evaluating the performance of undergraduate students in a laboratory class like the one developed by N. Kapoor (1990). This has been cited in Box 2. Please examine

this and see if you can develop a similar rating scale for specific teaching situation that you are in.

In case of 5-point numerical scale (5-1), it is assumed that the student would get a minimum of 1 mark. But in certain laboratory experiments the student may not be able to undertake a step or omit it. In that case he cannot be given 1 mark for a wrong or omitted step and deserves zero mark, not one. In such cases rating scale can be a 5-point scale (4-0) scale which is more appropriate. Depending upon the nature of aspects, skills or activities to be assessed we can use 3, 5 or 9 point scale. As for 2 point scale it is indeed a checklist as the evaluator has only to tick mark (√) or cross mark (×) against each statement or step, to indicate whether the answer, skill or step undertaken by the student is right/wrong, adequate/inadequate, true/ false, yes/ no etc.

Box–2

An example of a rating scale

Dissection of Rat to expose Arterial System

Steps followed in the	Weightage	Rating	Weighted
a) Keeping the animal ventrally in the dissection tray filled with 1/3 level of water	03%	5	0.15
b) Fixing the pins on the arms and legs of the rat	05%	4	0.20
c) Giving deep incision from the center of thoracic region thereby removing the fat bodies and exposing the internal organs	07%	3	0.21
d) Puncturing the veins	15%	3	0.45
e) Exposing the arteries (Note: Here it is important to see the number of arteries, e.g., internal and external carotid)	35%	4	1.40
f) Display of arterial system by putting blackpaper strips or photographic film strips	25%	3	0.75
g) Diagram of arterial system	10%	3	0.30
			3.46

Self-assessment

1. *Mentione below are two statements which are followed by four alternatives. Choose the correct one:*

   i) *Evaluation of laboratory work involves evaluation of products of students' performance.*

   ii) *Evaluation of laboratory work involves evaluation of process skills of students' performance.*

   a) *Statement (i) is correct and (ii) is incorrect*

   b) *Statement (i) is incorrect and (ii) is correct*

   c) *Both the statements (i) and (ii) are incorrect*

   d) *Both the statements (i) and (ii) are correct*

2. *For assessing the following skills which of the two, checklist or a rating scale is appropriate?*

   a) *Drawing of diagram of digestive system*

   b) *Labeling of different part of the diagram correctly*

   c) *Exposing the different arteries in dissection of a rat*

   d) *Drawing of right inferences or conclusions.*

Evaluation of students' perfrmance in seminars or similar group-controlled learning situations

We have learnt in Block 1, Unit 3, the place of seminar and other such group-controlled instructional methods, and the specific instructional objectives that they help to realize. Subsequently, in Unit 8 Block 2 on skills associated with organizing interaction sessions, we learnt the important competencies of the teacher ought to have, in drawing out the best out of such group situations. In both these units, we understood that a seminar provides opportunities for a student to present a paper that s/he has developed, or report a project that s/he has undertaken. There are certain abilities related to organization and synthesis of ideas, presentation of ones' views and arguments besides developing affective attributes like, tolerance, openness to others' view points, communication skills etc. that one tends to develop while going through the exercise. This, in fact, constitutes one of the main purposes for which a seminar as an instructional method, is used. Alternately, a seminar facilitates a group of students to discuss the paper presented. Through this, not only does the chairperson or the speaker tend to develop certain intellectual abilities related to analysis and critical evaluation, but s/he also develops certain attributes like tolerance, openness to others' points of view ability to argue, moderation, restraint, etc.

Thus, the goals attached to the seminar as an instructional technique, pertains chiefly to the performance of the student who presents the paper, and those of the students who participate in the discussion thereafter. Evaluating the extent to which this instructional technique is effective would mean evaluating the performance of the speaker and the participants of the seminar. Now, what are the problems involved in evaluating seminar performance: Let us see.

Similar to laboratory work, performance in a seminar also involves several process skills. Thus, evaluating a seminar performance is essentially process evaluation. However, unlike laboratory work the process skills involved in a seminar performance do not manifest in concrete, discernible and overt behaviour, which can be registered in a relatively value- neutral manner. On the other hand, the processes involved in a seminar performance are mental and attitudinal, which sometimes manifest, in overt behaviours like a spoken intervention, or in gestures, or even in silence. Therefore, discerning and registering such behaviour and subjecting them to evaluation, is quite a value-mediated exercise. However, there is no alternative to observing such behaviours and assessing them on the attitudes that they reflect. Even when it comes to intellectual skills of coherence, logical cohesiveness, analytical abilities and clarity of expressions in the intervention that the students make, it is extremely difficult to make evaluative judgements. It is one thing to carefully observe and record every behaviour that the students manifest, and quite another, to deduce a meaning from these behaviours and come out with a correct assessment on the skills, higher cognitive abilities and affective attributes that they stand for. Casual observation may not enable one to register the right kinds of behaviour for one may not be sure about the kind of behaviour to look for. What is imperative is a systematic observation and this calls for an instrument, which helps the observer to anticipate behaviours and to interpret depending on the context in which they are manifested.

For this purpose, there is a need for fairly comprehensive observation tool with which one may register the behaviours manifested by the various members of the student group with the help of which one can interpret them for all the covert skills, attitudes and other attributes that they reflect. One such tool has been developed by Menon (1984) meant for postgraduate classes.(See Box-3).

Box - 3

Discussion Observation Schedule

No. of questions and observations made
No. of point discussed
No. of supporting observations made
No. of supporting questions raised
No. of high level observations made
No. of high level questions asked
No. questions asked for clarification
No. of irrelevant questions and observations made
Speaking confidently
Communicating coherently
Conveying views precisely
Concentration on the discussion
Sharp comprehension of ideas/ issues
Openness to others' ideas/ views
Objective outlook/ unbiased
Elaborating of ideas put forth
Integration of different ideas
Initiation of ideas
Facility of languages/ Communication
Self-contradiction making contradictory statements
Getting excited easily
Getting impatient

For judging the effectiveness of a seminar, workshop or any other programme it is necessary that it may be evaluated with reference to:

a) the context in which seminar or other similar programme are conducted;

b) input given in terms of facilities, resources (academic preparation, material, financial etc.);

c) progresses undertaken to prepare, organize, conduct, coordinate, participants' involvement etc.;

d) products or outcomes or attainment of objectives of the seminar, quality of papers, report, material produced etc.;

e) impact made on the participants, institutional learning environment, teaching-learning practices, etc.;

While evaluating a seminar or seminar or similar programmes we have to find out the discrepancy between the intended outcomes of the seminar and the actual or observed outcomes which indicate the effectiveness of the seminar.

(i) _____

(ii) _____

(iii) _____

(iv) _____

(v) _____

(vi) _____

(vii) _____

(viii) _____

Categorise: Student names Quantitative Qualitative

ME : To a maximum extent

GE : To a great extent

FE : To a fair extent

LE : To a less extent

NA : To the least extent (Not at all)

(**Source:** Menon, M.B. (1984) *Evolving Multimedia Approach to Teaching at Post-Graduate Level;* Doctoral Thesis: CASE, Baroda, P. 301.)

Activity 1

See Box-3 and solve the questions given below:

1) How many dimensions/ categories of behaviours are included in this scale?

2) How many of these are quantitative?

3) Would you like to add to these dimensions? Which are those? (e.g. obstructing others' expressions).

4) Can you transform any of the qualitative dimensions into a quantitative one? (e.g. the number of times he elaborated on other's ideas).

5) Would you like to develop similar instrument (may be shorter version, which is more appropriate for your instructional situation? Develop such an instrument and discuss it with your colleagues and then try it out on your student group.

Evaluation of project work and dissertation

One of the major objectives of higher education is to develop in the students, the capability to undertake on their own, such activities as, study, research, exploring the literature available in a field, undertaking direct observation of reality and so on. It is in order to realize these objectives that instructional techniques such as project

work, assignments and term papers, research resulting in a dissertation, etc., are commonly used in good institutions of higher education. While the use of such techniques, as we have seen in some of the earlier units, by the proper evaluation of these instructional techniques that one can ensure proper development of intended abilities, psychomotor skills and higher cognitive skills besides cultivating desirable interests and attitudes.

Suppose, you are planning to give a project work dissertation to your students. The first thing that you need to be clear about is the specific objectives or intended learning outcomes that you want your students to develop through the project. Once you delineated the objectives you can go ahead with the specification of these objectives in terms of the competencies and specific skills that you would look for in the students while they are doing the project, as well as the attributes that you would look for, in the final project report which will indicate the extent to which the students have developed the competencies that were aimed at. In other words, what you need to do is to develop a list of process and product attributes of competencies that you want to observe in the performance of the students.

There are certain aspects, processes or products which you need to know whether they exist or not, present or absent, true or false, correct or incorrect. Their assessment can be done with the help of a checklist on the basis of which you go on tick ($\surd$) marking or crosses ($\times$) against each item at conclusion about such attributes, skills, competencies, etc. – a quantitative assessment. There are also aspects, attributes and other things on which you will want to rate the performance of the students in qualitative terms which can be rated on 3, 5, 7, or 9 point scale and yes or no, right or wrong . If you are the supervisor, your job as an evaluator starts at the very beginning of the project. A lot of rigour that has gone into the conduct of the project will not reveal itself in the project report; you get an idea about it only if you have observed the student at every step of the project (formative evaluation) to get feedback and provide feedback. On the other hand, if you are external examiner, many a times you will be compelled to restrict your evaluation to the final project report (summative evaluation). However, you must realize that this is rather an incomplete source of information on the way the student must have carried out the project. In fact such an evaluation is restricted to product evaluation unless external examiner observes all processes.

The customary practice of evaluating project work and dissertations even by internal evaluators, unfortunately, is to give a global grade or mark indicating their judgement about the quality of the work. But. one does not adequately realize that many different kinds of errors creep into this global judgement. It makes a sounder evaluation practice if one breaks up the global judgement into judgements about the student's performance on specific process and product aspects. For instance, if you are guiding a post-graduate student for his/ her dissertation work, and if in that capacity you are his/ her internal examiner, it makes much better sense, if you break up the entire process of conducting the research work leading to the dissertation into smaller steps and judge the student's competence from the way s/he accomplishes each step ranging from identification of problem to the submission of report of dissertation.

If, on the other hand, you are his/ her external examiner, then again you could avoid giving global grades or marks, and instead, assess the way each part or step of the research project has been presented in the report, giving separate weightage to each aspect/ step/ dimension and then giving separate grade on each dimension/ aspect/ attribute. Most commonly usable grading is 5 point grade and therefore, A, B, C, D, and E can be used as letter grading corresponding to 5, 4, 3, 2 and 1 or in certain cases it may be 4, 3, 2, 1and 0 when certain aspects require awarding of zero mark as in case of correct or incorrect use of formula, correctness of results etc. On the basis of weightage of each aspect and grading weighted grade point or score can be calculated.

Box–4

Illustration of an assessment scheme for a dissertation

Name : XYZ

Class : M.Ed.

Title of Dissertation : A Study of Relative Efficacy of Different Instructional **Media** in the Teaching of Physics at the Higher Secondary Level.

S. No.	Dimensions/ aspects of assessment	Weightage	Grade*	Grade Point	Weighted Grade Point
1)	Need and significance of the problem identified	5%	D	2	0.1
2)	Proper Statement of the problem reflecting answers or conclusion	5%	A	5	0.25
3)	Relevance and adequacy of Review of related studies	10%	C	3	0.3
4)	Justification of the hypothesis formulated or research questions raised	10%	A	5	0.5
5)	Appropriateness of the research design used	10%	B	4	0.4
6)	Rigour of methodology and procedure adopted	10%	C	3	0.3
7)	Relevance and quality of data gathering devices	15%	A	5	0.75
8)	Appropriateness of data analysis techniques used	10%	B	4	0.4
9)	Correctness of the interpretations made	15%	B	4	0.6
10)	Accuracy, precision and logical sequence in reporting	10%	B	4	0.4
Overall Grade Point Average Overall Grade					4.0 B

*Numerical equivalents of Grades : A (Outstanding) = 5; B(Good) = 4; C (Satisfactory) = 3; D (Fair) = 2; E (unsatisfactory) = 1

Activity 2

1) List the new items that your would like to add or the ones that you would like to delete from the above mentioned scale.

2) Would you like to add sampling process as one of the dimensions? If yes, at what place in the list?

3) Think about any one of the dissertations that you have read (on experimental, descriptive or historical research) and develop a rating scale to evaluate that, bearing in mind both the process and product dimensions.

Self-assessment

3. *Fill in the blanks:*

   a) *Evaluation of a student's performance in a seminar or a similar group controlled learning situations means .....*

   b) *Cognitive abilities, as well as ..... are developed through seminars and other group interaction techniques.*

   c) *Process evaluation in case of a dissertation is carried out by ..... evaluator.*

   d) *Evaluation of project work means both the evaluation of ..... and the process skills involved in it.*

Internal assessment versus external examination

The terms "Internal Assessment" is often used as if it were synonymous with continuous assessment. It may often be so, but it is not absolutely necessary that it should be. The term "internal" refers to the examiner – the teacher or other experts – who can closely observe the students during the teaching programme and not an external referee who examines the student at the end of the teaching programme. The term " Continuous" on the other hand, refers to the time dimension. An assessment, which is carried out during the programme at different stages – from the beginning to the end at progressive phases of the programme, is continuous assessment. It is carried out regularly and periodically at the end of a chapter or unit of teaching whereby the teacher gets regular feedback about students learning. This is distinguished from the assessment, which is conducted once at the end of a programme, term, session or a semester, namely, summative evaluation whose focus is on grading, giving judgement or certification of achievement .

What is continuous internal assessment?

Operationally thus, continuous internal assessment would mean the following:

i) It represents continuous awareness by the teacher of the development and knowledge of his students. It is a process, which extends over a period of time.

ii) It enables a teacher to look for signs, which indicate acquiring of developmental objectives, e.g., the growth of thinking process, and the development of those varying abilities/ competencies, which are implied in the instructional objectives and towards which teaching is directed.

iii) It is used by a teacher as a diagnostic device to plan his remedial material/ exercises on the basis of diagnosis of the students' difficulties or inadequacies as revealed from time to time in unit or topic testing.

iv) Its focus is on formative evaluation getting feedback on the attainment of intended learning outcomes on the basis of which the teacher can adapt his instructional strategies.

Merits and demerits of both systems

There is no doubt about it that both the "External end examination" and "Continuous internal assessment" are equally important. Therefore, it is essential that teachers should desire the adoption of continuous internal assessment as a method of measuring attainment within the context of a nationally recognized external examination. Keeping in view the merits and demerits of both the external end examination and continuous internal assessment it is essential to use both for the

specific purposes for which they are more appropriate. Some of the selected merits and demerits of both the systems are given below:

a) **Merits of external examination**

This is the most important and universally accepted means of assessment. The following are the merits of this system of evaluation:

i) It provides an independent assessment of those whose results are most likely to be accepted by the users of examination results.

ii) It provides uniformity of practice and standards with all the students being assessed by common qualities and by common criteria of assessment.

iii) It concentrates on the resources and expertise for paper setting and marking.

iv) It eliminates the possibility of the rise of any conflict between the teacher as a teacher and the teacher as an assessor.

v) It provides an assessment procedure unaffected by personal relationship between the teacher and the students.

vi) It provides better basis for setting performance standards.

b) **Demerits of external examination**

i) It is usually more time consuming and cumbersome for administrators to plan, organize, conduct and process the results.

ii) Its focus is on measurement of achievement, not on improvement of achievement through immediate feedback.

iii) It has strangulating influence on teaching learning practices which are dominated by the mode and nature of external examination.

iv) It has low validity and reliability which cast doubt on certification of students' achievement.

v) It is limited to assessment of cognitive outcomes and ignores non-cognitive outcomes of students' growth.

vi) Use of free options (any 6 out of 10 type) in question paper leads to selective teaching and selective learning besides difficulty of comparing scores of students who attempt different questions.

c) **Merits of continuous internal assessment**

i) Internal assessment provides more opportunities to students to improve their achievement because of more frequent assessment and regular feedback about their progress.

ii) During internal assessment a wide variety of tools and techniques of assessment may be used covering a wide variety of competencies, thereby making the assessment more valid and reliable.

iii) Internal assessment requires the teachers to make assessment more comprehensive in terms for assessment objectives, which include cognitive, affective and psychomotor outcomes of learning.

iv) Internal assessment enables students to know their progress and inadequacies in learning by getting continual feedback on theoretical and practical work, both on the process and product of performance.

v) Internal assessment provides better opportunities to make evaluation an integral part of teaching learning process. Being a formative assessment it helps to diagnose students' weaknesses and provides basis for remedial instruction.

d) **Demerits of continuous internal assessment**

i) Internal assessment can sometimes endanger the relationship between a teacher and a student because of subjective opnion in assessment and favouritism of teachers.

ii) It makes it difficult to relate and compare the standards between institutions/ colleges. i.e., inter- institutional comparison of students' performance becomes difficult.

iii) It may give rise to suspicions in the mind of users on the validity and reliability of the results of assessment.

iv) There is likelihood of dilution of standards of performance if teachers use lower order assessment objectives to show better results.

It is, therefore, obvious that both external examinations and continuous internal assessments are equally important sides of the same coin. Internal assessment should not be taken as something with which the external examination can be replaced. Both have a part to play in a well planned evaluation system. Very often the annual examination is to be taken as a part of the internal assessment. There is growing concern to de-emphasis external examinations and make teacher based internal assessment as important complement of the total assessment system and make certification of students more meaningful.

Special features of internal assessment

i) The introduction of internal assessment, in contrast to the present point-in-time examination, is an important step towards the integration of the trinity of functions namely teaching, learning and testing. At present, the performance of the student is judged entirely at the end of the year/ semester/ course examination, which is usually of three hours duration wherein they are expected to produce answer to stereotyped questions and which involve a great deal of rote memory work . By such simple repetition of stereotyped questions, various abilities involving higher cognitive operations like application, analysis, synthesis and evaluation are neglected in favour of simple memory based factual information. The introduction of internal assessment will enable a teacher to assess certain intellectual abilities and skills that are not tested by external examinations especially of the affective domains' personal social qualities.

ii) The main purpose of introducing internal assessment is to integrate teaching and evaluation, and to test those skills and abilities, which cannot be tested through a written examination at the end of a course. In internal assessment, very low weightage should be given to testing simple recall of knowledge and information and more attention should be given to higher order assessment objectives.

iii) The data obtained through internal assessment should not be used just to rank the students, but more often to improve the programme and for improvement of students' learning. Students, of course, should be informed about their progress and inadequacies in learning. This helps the teacher to provide remedial instruction to the students where necessary to improve their achievement. This also helps the teacher to revise his programme and adapt his instructional strategies to make his instruction more effective.

Weightage to internal assessment

The question of giving weightage to internal assessments and external examinations is important. In a unitary university/ autonomous college, this question does not arise, since a design of internal assessment can be worked out including the end of session/ semester/ course examination. In an affiliating type of university, dealing with a good number of colleges, the question of fixing weightage for internal assessment and external university examination does arise. This is being tackled by the following ways.

i) Internal assessment and external university examination scores or grades are shown separately, on the ground that the scores on external examinations reflect assessment of cognitive abilities (Psychomotor also in case of practical subjects) whereas coverage of assessment objectives is comprehensive (cognitive + affective + psychomotor) in case of internal assessment.

ii) Various percentages are fixed mostly on arbitrary criteria to give weightage to various assessment objectives/ intended learning outcomes.

Some teachers feel that because the end examination involves an external observer, it may be more impartial (less biased), therefore, this should be given more weightage: but some others feel that the end examination only tests the memory and does not reflect on any process dimensions and hence, should be given less weightage.

The dilemma can be solved not by classifying the examinations as internal and external, but by referring to its dimensions – the competencies, skills, etc., involved in a particular examination and giving a weightage to that dimension in the context of total terminal behaviour. Any objective- based assessment demands that the intended outcomes of learning in a particular subject must be spelt out first in terms of observable and testable behaviours. This is to be followed by selection of suitable, usable, valid and reliable evaluation tools to match these abilities and skills that yield needed dependable evidences on development of listed competencies corresponding to stipulated assessment objectives.

Various types of evaluation

Some of the tools of evaluation, suitable for measuring these learning outcomes are given below:

1) Quizes (announced)

2) Quizes (unannounced)

3) Oral tests for review of a chapter (5 to 10 minutes) in the classroom or a formal oral test on a specified theme, unit for assessment of specific abilities.

4) Short duration objective type tests (10 to 15 minutes for periodical assessment of cognitive abilities).

5) Short answer questions (15 to 30 minutes for periodical assessment of cognitive abilities).

6) Long answer essay (30 to 60 minutes for periodical assessment of higher order cognitive abilities)

7) Class assignments/ Tutorials (once or twice a for developing higher order abilities and checking on them)

8) Home assignment (once or twice a week for developing higher order abilities and checking on them)

9) Guided individual projects (Given twice in a semester or 4 times an academic year, to bring in new meaningful and relevant problems that would work better

in relation to the subject matter) besides know how of planning, development, execution and evaluation of projects.

10) Guided group projects (in addition to the above, to develop leadership, qualities, team work, discipline, etc.)

11) Laboratory/ Field / Practical work (periodically, to develop and assess practical skills in handling instruments, experiments, reports, etc.)

12) Observation (periodically to assess personal qualities of regularity , hardwork, inventiveness, initiative, originality, etc.)

13) Checklist (periodically to assess presence or absence, correct or incorrect, yes or no, right or wrong of development of such abilities or skills, attitudes or values which can be judged on a two point scale like health habits, study habits, work habits etc.)

14) Rating scale (periodically to assess personal qualities of regularity, hardwork inventiveness, initiative, originality etc.)

15) Rating scale for interest (3 times in a semester or 6 times in a year, to rate his interest in the subject).

16) Rating scale for attitude (3 times in a semester or 6 times in a year to rate his attitude).

17) Group Discussions (once a month to assess his originality, creativity, initiative, communication skills, etc.)

18) Seminar (once a semester or twice a year, to develop and assess originality, creativity, discipline, communication skills, initiative, etc.)

19) Thesis/ Dissertation (once a year, to develop skills of scientific method, critical thinking, analysis, synthesis, evaluation, performing skills, communication skills, etc.)

20) Projects-individual and group projects (once a year to develop insight and training in process and product attributes and skills related to planning, execution, evaluating and reporting of findings)

The ground realities about internal assessment

In spite of the theoretical superiority of continuous internal assessments and their being recommended by the various education commissions, the system is not, at present, being implemented and is not given its due weightage in most of the Indian colleges and universities .Still external year end examinations reign supreme and carry quite higher if not entire weightage in our scheme of evaluation.

The reasons for this are:

i) Students and/or their guardians exert pressures on teachers for awarding high marks, irrespective of their deserved score. Teachers even when given autonomy by the institution, feel that they are under threat to express their real assessments of the students. This is true not only in rural areas but also in urban areas.

ii) Teachers feel that it is too much hard work to continuously examine, score and record evidences/data about their students and sometimes feel that they do not have enough time for teaching.

iii) Some teachers are too lenient and others are too strict in awarding marks on internally conducted examinations, which have adverse effect on students in certain subjects and favourable for students in a different subject.

iv) Comparability of students' performance across the institutions is not possible due to non-comparability of the evaluation tools used and subjective scoring.

The dilemma is genuine, and it can not be resolved overnight . But over the years, we shall have to get over these difficulties, to make our educational system more effective. Some of the ways presently suggested to get us out of this problem are as under:

a) **At micro level**

i) Internal assessment/ Unit test/ class assignments, etc. should not be used for ranking purposes, but should be used primarily for guidance and for helping a student to find his strengths and weaknesses (diagnostic orientation) and providing basis for teachers to undertake remedial teaching.

ii) Tests are to be used as an integral part of instruction/ teaching – that is to say - to be sometimes used in the beginning of a session to start a discussion or during the session where students score their own records (self or peer scoring tests). These may take the form of readiness tests, formative tests, review tests, quizzes or assignments

iii) To design tests which are objectively scorable (objective type items) – and can be assessed by computers or by the students themselves.

iv) To use more and more short answer questions, which are scorable easily and help in better sampling to content to ensure more validity.

v) To develop open book tests, and quizzes wherein students develop questions etc., where just memorization to facts would not help.

vi) To encourage cooperative testing among neighbouring colleges to minimize disparity and improve comparability of students' performance across the institutions.

b) **At macro-national level**

i) Delinking degrees from job-specific testing.

ii) An agency like National Testing Service at the national level to design objective testing programmes based on which student profiles on different dimensions can be prepared and provided to the students and / or employing agencies to find the comparative strengths and weaknesses on certain specific dimensions.

iii) State level evaluation organizations may be encouraged to develop high quality question banks for which U.G.C. may undertake training courses for item writers and develop guidelines for cooperative testing to minimize disparity in inter–institutional standards of students' performance.

iv) U.G.C. may work out a common basic design or framework (with some flexibility) for internal assessment in colleges in collaboration with state level educational agencies to reduce disparity in subjective modes of assessment in colleges.

Self-assessment

4. *Below is a list of statements. Write T if the statement is true and F if it is false, against each statement in the bracket provided.*

   a) *Internal assessment is the same as continuous assessment.* ( )

   b) *Internal assessment provides scope for innovations in examination practices.* ( )

   c) *External end examination encourages the linking of jobs with degress.* ( )

   d) *External end examination is meant for diagnostic evaluation.* ( )

   e) *The triangle of teaching, learning and testing is reflected in internal assessment.* ( )

   f) *In internal assessment it is not possible to test higher cognitive abilities like application, analysis, synthesis and evaluation.* ( )

   g) *The weightage to internal and external assessments must be based on the coverage of assessment objectives in a particular subject, by each type of assessment.* ( )

   h) *Students do like internal assessment because these can be used to know their inadequacies in learning and improve their achievement.* ( )

   i) *Continuous internal assessment of students provides regular feedback to teacher about effectiveness of their instructional strategies.* ( )

Summary

In this Unit, an attempt has been made to introduce some of the evaluation techniques, which you may use while evaluating laboratory work, seminar participation, project work and dissertation. Unlike the prevailing situation, we laid stress on evaluation of process skills involved in all these instructional techniques. We provided examples and illustrations of checklist, rating scale and the observation tools to measure the process skills involved in the above mentioned instructional techniques. We also suggested some guidelines to be used in the evaluation of a laboratory record/ project report/ dissertations.

In the second section of the Unit, we had a detailed discussion on continuous internal assessment versus external examinations. We pointed out the merits and demerits of both continuous internal assessment and external examinations. We opined that the weightage to be given to both types of examinations should depend on coverage of intended learning outcomes of a particular subject. We mentioned various types of tools that may be fruitfully employed by you while undertaking continuous internal assessment. Lastly, we discussed some ground realities, related to the unpopularity of continuous internal assessment and suggested some measures to overcome them.

Unit-end activities

1) Prepare a checklist to assess study habits of your students and administer the same and summarise your findings.

2) Develop a rating scale for a laboratory experiment, which you have assigned to your students and judge their performances through it.

3) Organise a seminar or visualize the seminar, which you have organized earlier and make an assessment of students' performance in the seminar with the help of the observation tool given in this text.

4) Consider any of the dissertations you have guided or come across as an external evaluator, and assess the same by developing your own, assessment scheme reflecting a) different aspects/ components of assessment. B) assigning weightage to each aspect, c) giving rating using 5 point scale, d) indicating weighted grade point and e) overall grade point average.

Points for discussion

Suppose you are the principal of an autonomous college and have invited for a meeting of the teaching staff to discuss introduction of internal assessment in your college. What objections do you expect from a) the teaching staff and b) the administrative staff? What steps you are likely to take to overcome the problems or issues raised by them.

Answers to self-assessment

1. Answer – d

2. a) Rating scale b) Checklist c) Rating d) Checklist

3. a) Evaluation of the performance of both the speaker and the participants of seminar.

   b) Affective attributes

   c) Internal evaluator

   d) Project report

4. a) false, b) true, c) true, d) false, e) true, f) false, g) true, h) true, i) true

Suggested readings

Association of Indian Universities (1977) *Monograph on Question Banking*. New Delhi.

A.I.U. (1977) *Monograph on Grading for Universities*. New Delhi.

A.I.U. (1977) *Monograph on Internal Assessment of Universities*. New Delhi.

A.I.U.(1977) *Monograph in Practical Examination for Universities*. New Delhi.

Armstrong, J.M.A. (1950) *Projects and their Place in Education*. London: Pitman.

Garg, S.C. and C.K. Ghosh (1990) Measurement of 'Y' by the Method of Flexture with the Help of Checklist, School of Sciences, Indira Gandhi National Open University, Mimeograph.

Government of India (1986), *National Policy on Education*. New Delhi: Ministry of Education.

Joseph. T. (1977) *Reforming Laboratory Training*, Technical Teacher, Annual.

Kahlon, A.S. (1977) *Student Evaluation and Internal Assessment*. University News: 15(3); Feb., 56-7.

Kapoor, N. (1990) Developing a Rating Scale to Evaluate the Performance of Undergraduate Students in Dissection of Rat to Expose Arterial System, School of Sciences, Indira Gandhi National Open University, Mimeograph.

Popham, W.J. (1973) *Criterion-Referenced Measurement: An Introduction.* Englewood Cliffs,New Jersey: Educational Technology Publications.

Singh, Pritam (1983) *A Monograph on Improving Practical Examination.* NCERT: New Delhi.

MPDD-IGNOU/ P.O.1 T/July2010 (Reprint)

ISBN : 81-266-0954-0