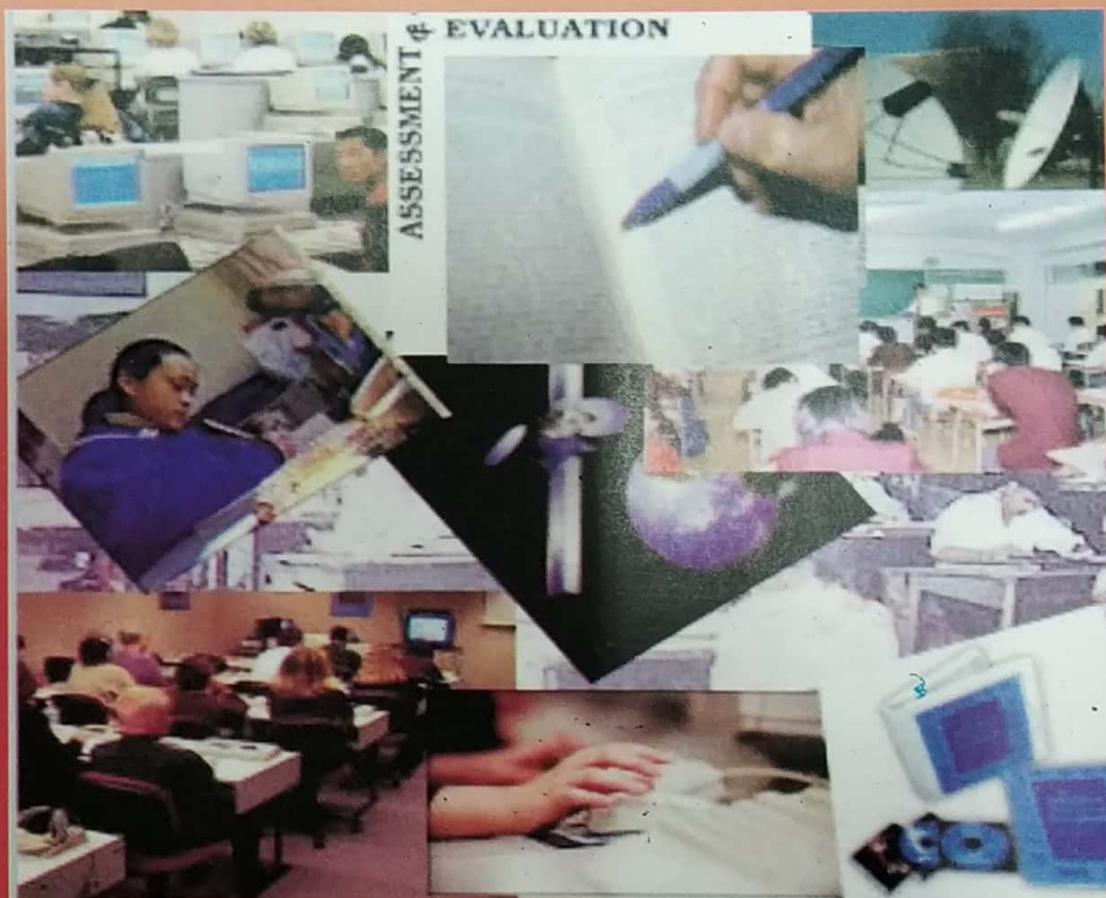


Assessment and Evaluation in Distance Education



**Staff Training and Research Institute of Distance Education
Indira Gandhi National Open University
Maidan Garhi, New Delhi - 110068**



STRIDE HANDBOOK 4

Assessment and Evaluation in Distance Education

Prabir Biswas and Basanti Pradhan

Editor:
Santosh Panda



Staff Training and Research Institute of Distance Education
Indira Gandhi National Open University
Maidan Garhi, New Delhi-110068, INDIA
<http://www.ignou.ac.in>

June 2002

STRIDE HANDBOOK 4

ASSESSMENT AND EVALUATION IN DISTANCE EDUCATION

(Revised and up-dated version of earlier Handbook-4 entitled "Assessment of Assignment Responses" prepared by **B.N.Koul** and **Santosh Panda** in 1989)

Designed and Developed by

Prabir Biswas, Ph.D.

Senior Lecturer

&

Basanti Pradhan, Ph.D.

Reader

Staff Training and Research Institute of Distance Education
Indira Gandhi National Open University

Format and Content Editing

Santosh Panda, Ph.D.

Professor of Distance Education

Staff Training and Research Institute of Distance Education
Indira Gandhi National Open University

Language Editing

Eisha Kannadi, Ph.D.

Lecturer

School of Education, IGNOU

Type set and CRC Preparation

Shashi Surendranath
STRIDE, IGNOU

Cover Design

G. Mythili, Analyst
STRIDE, IGNOU

Print and Production

T.R.Manoj, SO (Publication), STRIDE, IGNOU
C.N. Pandey, STRIDE, IGNOU

© Indira Gandhi National Open University, New Delhi
June, 2002

ISBN 81-266-0487-5

All rights reserved. No part of this work may be reproduced in any form, by mimeograph or any other means, without permission in writing from the "Indira Gandhi National Open University".

Further information about Staff Training and Research Institute of Distance Education (STRIDE) and the Indira Gandhi National Open University may be obtained from the University office at Maidan Garhi, New Delhi-110068, India.

Published by Prof. P.R.Ramanujam, Director, STRIDE, on behalf of Indira Gandhi National Open University, Maidan Garhi, New Delhi-110068.

Printed at New Chanab Offset Printers, New Delhi-20

FOREWORD

Starting from correspondence courses, a few decades ago, teaching and learning through non-conventional and non-traditional modes have undergone very dynamic and effective changes to improve the quality of education through these new systems. These emerging systems and trends are quite open, flexible and widely accessible to large target groups of learners. In view of these unique and distinctive features of these systems, they are now identified as the open and distance learning systems (ODL). It is now very well realized and understood that for effective and successful handling of the ODL systems, very well trained faculty and other professionals, and distance educators are all required to work together. The system has gained considerable importance and acceptance to attract significantly large number of learners seeking education and training through the ODL systems. It is in this background that the Indira Gandhi National Open University (IGNOU), which was set up in the year 1985 has already emerged as one of the largest universities in the world today. A more recent realization of the effectiveness of the ODL systems is its unique capability of handling programmes for trainers and learners especially those related to empowerment and awareness programmes involving substantially large groups of trainers and learners. Never before, we had witnessed such a great impact of any technology than that of the contemporary developments in Information and Communication Technology (ICT) on the system of education. It is becoming very apparent that in recent years the ODL has very well and meaningfully absorbed the ICT for the betterment of the system, to make it more cost effective and accessible so as to bring in equity in education.

The ODL has been growing at a very fast rate and a massive human resource trained in the Open and Distance Learning System is required to handle the system efficiently. Keeping in view the above requirements of human resource development, the Division of Distance Education of IGNOU prepared 7 Handbooks on 7 themes in 1989 to meet the above requirements of the ODL System. The Division of Distance Education grew into the Staff Training and Research Institute of Distance Education (STRIDE) in the year 1993 and 10 Handbooks on different themes of ODL have already been brought out. These Handbooks have found to be very useful for the teachers, administrators and other functionaries of the ODL. As of today, the STRIDE has successfully revised the 7 Handbooks and has also added 3 more Handbooks on 3 different themes relevant to emerging trends in ODL. I appreciate the efforts made by STRIDE and hope that the Handbooks will be immensely useful for different functionaries of the ODL System in the country and abroad.



(H.P. Dikshit)
Vice-Chancellor
IGNOU

Contents

Page

Foreword

iii

Acknowledgements

v

Introduction

9

Section 1

Assessment: An Overview

11

- Objectives
- Meaning of assessment
- Explaining assessment
- Purposes of assessment
- Assessment in open and distance education
- Who needs assessment?
- Reliable and valid assessment
- Methods of assessment
- Computer based assessment

Section 2

Evaluation: An Overview

27

- Objectives
- Meaning of evaluation
- Difference between assessment and evaluation
- Purposes of evaluation
- Evaluation in an educational programme (EIEP)
- Evaluation of an educational programme (EOEP)
- Evaluation in distance education
- Reliability and validity in distance education

Section 3

Two Way Communication in Distance Education

37

- Objectives
 - Types of two way communication
 - Level of two way communication
-

Section 4**Assessment of Assignment Responses****43**

- Objectives
- What is an assignment?
- Types of assignments and designing assignments
- What is an assignment response?
- Commenting on assignment response
- Tutor comments
- Importance of tutor comments
- Writing comments: Some important steps
- Types of tutor comments

Section 5**Grading in Distance Education****58**

- Objectives
- Marking vs. grading
- The Advantages and disadvantages of marking
- The advantages of grading
- Bases of grading
- Relative grading standard
- Absolute grading standard
- Grading mechanisms

Section 6**Question Bank****71**

- Objectives
- Purpose of a question bank
- Planning a question bank
- Development of question bank
- Screening of questions
- Coding of questions
- Filing and storage of questions
- Using the question bank
- Conclusion

Section 7**Programme Evaluation****82**

- Objectives
 - What is a programme?
 - Concept of programme evaluation
 - Programme evaluation at IGNOU
 - Conclusion
-

Introduction

Over the last decade, open and distance education institutions have increasingly recognised the need for and the benefits of systematic training and staff development. Thus an array of induction courses, refresher programmes and workshop-cum-training programmes were designed and conducted for the staff members of open and distance institutions. The purpose of such training programmes for teachers/academics is to enhance their knowledge, understanding, confidence and competence concerning the processes and practicalities of distance learning, teaching, assessment and evaluation. However, the recent developments in technology have prompted a question as to what kind of training and development is to be offered and indeed required to undertake for distance educators in different aspects of distance education.

During 1997, an important turning point was the publication of Higher Education in the Learning Society (the Delor's Report). This report made far reaching recommendations regarding the quality of provision of higher education and provision of training for faculty in matters relating to learning, teaching and assessment.

Assessment is a critical focus of attention in any programme for university teachers, not simply because of the considerable time and effort it demands, but also because of the dilemmas it poses in trying to reconcile the tensions between the summative purposes of *assessment for grading* and the formative purposes of *assessment for learning*. A distance educator during assessment process has to consider the potentially powerful impact of the tasks set, the judgements made, and the feedback given, on what distance learners learn and on how they go about their own learning. Assessment procedures often operate at several institutional levels, and there are the technicalities of different types of assessment, of scaling and of aggregating marks or grades. Moreover, there are substantial variations between different disciplines and schools in the criteria for grading that are likely to be seen as the most appropriate. Considerable differences exist in the course contexts within which productive feedback strategies are to be devised and implemented. The challenge of assessment is all the greater, however, when the training and development are being targeted not at

established, mainstream academics, but at the growing number of part-time tutors or counsellors who are responsible for some aspects of assessment.

This handbook aims to help the distance educators/trainers to develop professional practices related to the area of assessment and evaluation. This book would serve as a toolkit for staff development programmes for academics of open and distance education institutions. There are VII sections in this book. Each chapter is self-sufficient and prefaced by some objectives or intended learning outcomes, which tell more about the specific purposes of the section. Besides, we have presented the content with the help of activities. The important points are presented in boxes for quick reference during a training programme.

SECTION 1

Assessment: An Overview

Assessment is an essential component of the teaching learning process. Teachers in schools, professors/lecturers in universities/colleges and trainers in training institutes are inevitably involved in assessing learners and/or trainees. Every teacher/lecturer/trainer should, therefore, have a clear knowledge of designing and carrying out assessment in their respective fields.

Objectives

After going through this section you would be able to:

- ☛ define and explain the term assessment;
- ☛ discuss different purposes of assessment;
- ☛ describe what is reliable assessment; and
- ☛ discuss various methods of assessment.

Meaning of Assessment

The term ‘*assess*’ originates from the Latin word ‘*assidere*’, which means ‘*to sit by*’ (in judgement). According to *The Oxford English Dictionary* (1998) the word ‘*assess*’ means (i) estimate the value of (a property) for taxation etc., (ii) fix the amount of (a tax etc.) and impose it on a person and community, (iii) fine or tax (a person, community etc.) in or at a specific amount, (iv) estimate the size or quality of (something).

The first three meanings are related to fines or taxes. The last meaning has some relation with the teaching and learning process.

Since the middle of the twentieth century educationists have been using the word ‘*assess*’ in the field of education and psychology. The latest use of its meaning is: *to judge the extent of students’ learning* (Freeman and Lewis, 1998).

Here, we would like to introduce and explore some key terms and issues in assessment. We have looked at the issues from the open and distance learning perspective.

Explaining Assessment

There are many definitions and explanations of assessment in education. Let us look at the definitions stated in Box 1.1 of Rowntree (1977) and Erwin (1994).

Box 1.1

Assessment in education can be thought of as occurring whenever one person, in some kind of interaction, direct or indirect, with another, is conscious of obtaining and interpreting information about the knowledge and understanding, or abilities and attitudes of that other person. To some extent or other it is an attempt to know that person.

(Rowntree, 1977)

Assessment is a systematic basis for making inferences about the learning and development of students... the process of defining, selecting, designing, collecting, analysing, interpreting and using information to increase students' learning and development.

(Erwin, in Brown and Knight, 1994)

The above definitions emphasise a couple of significant common themes.

They are:

- Assessment is a human activity which involves interaction aimed at seeking to understand what learners have achieved. Rowntree implies that assessment may occur in formal or informal ways, and *it may be descriptive than judgemental in nature*.
- The role of assessment is to increase students' learning and development, rather than simply to grade or rank student performance. Naturally, one cannot grade student performance without first assessing it, but it is implied that grading is a secondary activity to the primary goal of helping learners to diagnose problems and improve the quality of their subsequent learning.

Activity 1.1



Please note down a few lines about your own understanding of assessment in open and distance learning.

.....

.....

.....

.....

.....

.....

From the above definitions, we could reconceptualise assessment as the engine that drives and shapes learning, rather than simply an end of term examination that grades and reports performance.

Purposes of Assessment

Assessment has different purposes. The main purposes related to the teaching learning process are:

- i) **Selecting the students:** Assessment helps in selecting the students for a course (general, professional, technical and so on).
- ii) **Certification:** Assessment helps in certifying that a student has achieved a particular level of performance.
- iii) **Stimulating learning:** Assessment can stimulate learning in different ways, e.g. motivating the students, providing information/feedback, suggesting practice, and so on.
- iv) **Improving teaching:** Assessment information helps to review the effectiveness of teaching arrangements.

Rowntree (1987) and Freeman and Lewis (1998) have identified a number of purposes of assessment such as selection; certification/accreditation; maintenance of standards; description; improving learning; and improving teaching. From the above statement, it can be stated that the purpose of assessment varies as per the need of the target group and the context higher/open and distance learning/training. But, one can realize that there are two distinct interpretations of ‘assessment’. They are:

- (i) Assessment is interpreted in terms of the routine tasks that students undertake in order to receive feedback on their learning and a mark or grade signifying their achievement.
- (ii) Assessment as applied to processes at the institutional level, for example, programme evaluation or course evaluation.

As a teacher/trainer in open and distance education system, we come across different meaning of assessment and assessment strategy and use of appropriate assessment vehicles in a variety of teaching contexts i.e. in terms of individual learning, independent learning, group learning and self assessment. But the most consistent message we receive from the teachers/trainers is the *need for evidence* to judge the extent of student's learning and the extent of the competencies developed by the trainees. For example, "Student learning research has repeatedly demonstrated the impact of assessment on learner's approaches to learning..... Ask them to explain the physics and chemistry of muscle contraction, but test them on the names of the muscles, and they will 'learn' the names but not be able to explain how contraction happens. Ask learners to discuss narrative perspective in the novel but test them on the author's background and they will describe a lot about the author and little about narrative perspective. (Nightingale 1996)

Evidence of the extent of students' learning comes from their behaviour. Students' behaviour may be specific to a course or more general one. It may include various types of activities like written, verbal or practical work. The students may 'produce' a report, an assignment response, a story and so on; and they will follow some 'process' for doing these. So, two things are involved: a *product* and a *process*. Both of these constitute the 'evidence' on which the judgement may be based. In short, assessment may be defined as the judgement of a sample of relevant student behaviour (Freeman and Lewis, 1998).

There has been a change in the assessment scenario. We observe that there is a change from testing conceptual or disciplinary knowledge to testing the application of that knowledge in different contexts, together with a wider appreciation of or sensitivity to the boundaries and changing nature of knowledge.

Now a days the tests are developed in more flexible and contingent ways so that not only the disciplinary knowledge is tested but the general skills and abilities of students are also tested within the disciplinary contexts.

Lifelong learning skills have become an important and explicit feature of the open and distance education curriculum and the need for evidence to judge the 'knowledge, understanding, transferable or generic skills and attributes in learners (for example oral and written communication, teamwork, self management, creativity, inquiry and problem solving etc.) emphasises more demands on assessment.

The need to demonstrate understanding and application of knowledge, together with generic skills, has moved assessment practice towards new vehicles. It is clear that while a handwritten, three or four hours essay type examination may be able to legitimately test something, it cannot measure all, especially the generic skills that students have developed. This has led to integrated assessment. In such assessment, complex simulations, case studies, role plays or multi faceted projects are used to assess a range of knowledge, skills and attitudes in the single assessment (Nightingale, 1996, p.03). In situations where self assessment is important, reflective journals or reflective moments in portfolios have become popular. So, testing what learners can actually learn and can do, and how they have developed the skills and attitudes and need to continue their learning, marks a significant shift in approach to assessment. At the same time, sometimes assessment faces a few risking challenges such as cheating, plagiarism and a variety of assessment problems. So, while designing assessment tools considerable care has to be taken and ensure that the purpose of assessment has been achieved.

Assessment in Open and Distance Education

In face to face teaching, you get to know your learners through lectures, tutorials and individual consultations. The students in face to face settings have a range of opportunities to demonstrate their learning which are not confined to formal assessment tasks. Their interest, motivation, questioning and interactions are all on display throughout a learning encounter. As open and distance learners rarely enjoy these varied opportunities to communicate their learning, they are much more dependent on formal assessment tasks. They have also less opportunity in which to diagnose their own errors or mistakes before they go for a formal assessment task. For example, it may not be until midway through a course, when a student's first assignment is returned, that a simple error is discovered, to the detriment of the student's final grade.

While face to face learners can often rectify these problems long before they submit an assessment task, open and distance learners do not have the same kind of opportunities to check their understanding of an assessment task, or to compare approaches or methods with other students. So, distance learners are more dependent upon effective, early communication of assessment requirements, together with well designed and cohesive assessment tasks, useful and timely support, and a transparent marking scheme that explains how *judgements* are to be made. They are also more dependent on rapid turn around of assignments, so that the feedback can contribute to subsequent efforts and help maximise the valuable formative functions of assignments. Assessment activities do not have the same level of flexibility as they do in face to face settings. They must be thoroughly planned, communicated and managed.

Who Needs Assessment?

There are a number of stakeholders in the assessment process and it involves the students, teachers, trainers, institutions and the community at large.

Nightingale *et al* (1996) categorise these needs into four groups. They are:

1. Students' needs

- to know how they are progressing with their studies/courses
- to know whether they are achieving the required standard
- to gain certification of a level of achievement

2. Teachers' needs

- to know whether students are attaining the intended learning outcomes
- to know whether course materials and distance teaching activities are effective
- to be able to certify that students have achieved standards or met requirements.

3. Institutions' needs

- to provide evidence of achievement of institutional aims
- to know whether programmes are effective in their stated aims
- to certify that learners can practise in specific vocational areas
- to make judgement about admission to courses/programmes

4. Community needs

- to know whether institutions and teachers are effective and deserve continued funding
- to know whether students are adequately prepared for their careers
- to know whether education is being geared to meet the broad, long term needs of the society.

Reliable and Valid Assessment

Reliability and validity are terms used for educational assessment. In this context validity refers to what we really measure and what we are supposed to measure as per the detailed objectives. For validity to be high:

- The assessment must analyse student's performance on each objective.
- The assessment should provide the appropriate situation possible for measuring the specific abilities being measured.

For example, if the objective is to test learner's practical skills, we should *not* set a question for oral development measuring skills. Reliability is about the consistency or precision with which the assessment item measured the desired objective.

Whenever we make an assessment, it may or may not be reliable. We should try our level best to make it reliable.

Reliability or consistency operates at two levels:

- individual assessor; and
- more than one assessor.

As a reliable assessor, you should make the same decision on a particular assignment response whenever you grade/mark it. It means whenever you go through an identical response/solution of problem, every time you give the same grade/mark. Then you will be called consistent or reliable.

The key component in determining reliability of an assessment is consistency in marking. There is always the opportunity for human error, particularly when more than one assessor is assigned to a group of students assignment responses. If all the assessors make the same judgment — then reliability is achieved.

It is not possible to attain complete reliability due to various factors. We may try to make assessment more reliable with the help of a few guidelines as mentioned in Box 1.2. The most important way is to create, communicate and use clear criteria against which student performance is measured/assessed. The good criteria are those which are explicit, understood and agreed by all assessors and also by the students (Freeman and Lewis, 1998).

Box 1.2

To achieve more reliability in assessment

- cover all objectives (learning outcomes);
- create and use clear assessment criteria;
- ensure that the assessors understand the criteria;
- monitor assessors' marking; and
- use moderator where necessary.

Methods of Assessment

There are various methods of assessment. The main methods are the following:

Objective questions: Here, the word 'objective' refers to the method of marking the questions. This marking is a simple mechanical process. This can be done by an individual or by a computer. There are various types of objective type questions like, true false, fill in the blanks, multiple choice, matching, and so on.

Short answer questions: Short answer questions may be of different types like one sentence answer, completing a table/diagram, preparing a list, writing a paragraph, and so on. (For example, write the features of self learning material within 300 words).

Long answer questions: Long answer questions also may be of various types like essay, reports, dissertations, etc. (For example, prepare a self learning unit of 1200 words of a theme of your choice).

Performance methods: These methods do not require a separate assessment device. We may assess the presentation itself. Examples of

performance methods include making a presentation, acting, dancing, singing and playing a musical instrument (Freeman and Lewis, 1998).

Presentations: Presentation is a popular method of assessing. When presentation is used to assess presentation skills, the method has high validity, if it is assessed by a group. To assess presentations, suitable assessment criteria may be developed. These may include: *appearance of presenter, introduction of self, introduction of presentation, content of presentation, logic and order of presentation, eye contact, audibility, handling questions, use of visual aids*, and so on.

Formative assessment: This method of assessment helps to form and develop student learning. It provides support and feedback to learners to improve their *ongoing* learning. This comprises of all those activities designed to motivate, to enhance understanding and to provide learners with an indication of their progress. These are in the form of self assessment questions and tests given in the SIMs that help the learners monitor their own progress feedback from the assignments or from peers or mentors or counsellors and interaction with the teachers and tutors.

Summative assessment: It provides a total feedback or a report on what the distance learners have already achieved, whether this be a grade or a written assessment. The purpose of summative assessment is to record or report an estimate of students' achievement. These will often take the form of end of course examinations, course work assignments that contribute to a final grade or mark, supervised practical demonstrations, the project reports, etc. Depending on the setting and goals, it might entail awarding marks, grades, written reports or achievement or recognition of acquired competencies.

In most open and distance learning contexts, assessment usually involves both formative and summative component. Assessments are often designed in such a way that one assignment builds upon the next, with formative feedback from the first contributing to the next, and so on. This is referred to as continuous assessment. Marks are awarded for each assignment, which taken together, form a final grade. Continuous assessment plays a key role in open and distance education. We have presented a few more points pertaining to role of continuous evaluation in box 1.3. An *end of course examination or term end examination* which has a summative function might also be included in a final grade.

Box 1.3

In open and distance education context, the role of continuous assessment is to:

- provide some structure to distance learning;
- break down the assessment load into manageable chunks;
- encourage, motivate and develop confidence in the distance learners;
- work as a source of ongoing dialogue or two way communication between teacher and learners or institution and learners; and
- provide insight for learners into their progress, including their understanding and mastery of the subject.

Criterion referenced assessment: This method uses clearly stated criteria and performance standards against which each student's achievements are judged. If the criteria and standards are met, the student achieves the corresponding grade, irrespective of how others in the group have performed or how many others have achieved the same grade. Although criteria may be achieved in norm referenced assessment to assist in marker's judgements, criterion referenced assessment provides a clearer focus to assessment, for both learners and evaluators, and a clear description of what learners have achieved and the standard of achievement.

Norm referenced assessment: This method uses the achievement of group of students to set the standards for specific grades awarded to learners. Initially, learners are ranked using a scale (0-100) in order of achievement among the members of the group. A predetermined formula is used to nominate which percentages of students achieve the top grade, the lowest grade, and the range of grades in between. This method of assessment provides a description of where a student's achievement lies in relation to others in the group, rather than the particular qualities or competencies that students individually achieve. This method promotes competition between students for limited grades, rather than a detailed description of students' progress or abilities.

To sum up, criteria referenced and norm referenced assessment refer only to the method of interpreting the learners' performance on test and other evaluation instruments.

Activity 1.2



Explain how norm and criterion referenced assessment are different in their approach. Give an example from the distance learning situation/context.

.....

.....

.....

.....

.....

.....

As educators and trainers, on the one hand, we are working with learners to review and support learning, and on other hand, we are making judgement regarding their merits and their achievements. In the context of open and distance learning, we have embraced self assessment, continuous assessment and terminal evaluation. In distance education, teacher assessment is no longer pre eminent and there is no immediate feedback and reinforcement. The learner has to assess her/himself and for the purpose the distance teaching materials provide questions and answers which enable the learners to evaluate their progress frequently and provide immediate feedback. This is carried out through self assessment questions. The functions of self assessment are to help the student to check that the content is understood; to reinforce memory or understanding and stimulate the student to go forward in his or her thinking.

The continuous assessment in open and distance education is through tutor marked assignments (TMAs) and computer marked assignments. Assignments carry, weightage of 25% to 30% to pass an examination. Terminal evaluation carries 70% to 75% weightage in the final results. In case of IGNOU the terminal evaluation or term end examinations of various courses and programmes are held in the months of June and December every year. Distance learners are free to appear at any of these examinations either for specific courses or for the whole programme provided that the minimum period of study prescribed for the relevant course/programme is completed.

Computer Based Assessment

Computer technology is an important device for assessing students' learning in different subjects as it opens up opportunities for developing innovative assessment tools in distance education. The nature of computers as information processing tools, the role of computer technology in user friendly interactive learning environments, and the possibility of designing instructional tools to meet individual needs of distance learners, make computers potentially *powerful tools for assessment*. Computer based assessment applications are used in different subjects, such as computer based and Computerized Adaptive Testing, Figural Response Item Testing, Computer Simulations and Anchored Assessment etc. Computer based tools are found to have a positive impact on student's attitudes in learning.

Guidelines for computer based testing

In testing and assessment applications, computer has changed the ways in which tests and assessments are developed and administered. *Computer based tests are defined as tests or assessments that are administered by computer in either stand alone or networked configuration or by other technology devices linked to the Internet or the World Wide Web (www)*. In the face of the rapid growth of computer based testing, the Association of Test Publishers (Ref: <http://www.isoc.org/oti/articles/0500/olsen.html>) sponsored the development of formal, written guidelines to ensure high measurement quality of computer and internet based tests and to provide direction for the principles and procedures used for developing and administering those tests. Guidelines for Computer Based Testing are intended to supplement, extend, and elaborate on the recently published Standards for Educational and Psychological Testing (Joint Standards) as they apply to computer based and internet based testing and assessment.

All computer based tests must be used in accordance with fundamental measurement *standards* for test fairness, including fairness in testing and test use, in the right and responsibilities of test takers, in testing individuals with diverse linguistic backgrounds of distance learners and in testing individuals with disabilities. So, the purpose of statement of 'standards' is to provide criteria for the evaluation of tests, testing practices, and the effects of test use.

The *Standards* applied to computer based testing are:

- 1) The rationale and supporting evidence for computerized adaptive tests should be documented. The documentation should include procedures used in selecting subsets of items for administration, in determining the starting point and termination conditions for the test, in scoring the test and for controlling item exposure.
- 2) Instructions to test takers should clearly indicate how to make responses. Instructions should also be given in the use of any equipment likely to be unfamiliar to test takers. Opportunities to practice responding should be given when equipment is involved, unless use of the equipment is being assessed.
- 3) If a test is designed so that more than one method can be used for administration or recording responses — such as marking responses in a test booklet, on a separate answer sheet, or on a computer keyboard — then the manual should clearly document the extent to which scores arising from these methods are interchangeable.

The following six steps are presented to illustrate the essence of the guidelines and the standards:

Step I

Planning and Design A wide variety of computer based tests can be designed and developed to meet different purposes. The test specification for computer based tests should include: the test purpose, the content domain definitions, the content structure for the test items, required response formats for the test items, sample test to be developed and administered, scoring and reporting formats and procedures, and test administration procedures. The test specification should be thoroughly documented.

Step II

Test Development The test delivery environment should be evaluated before item authoring begins. It is important to make sure that items being created can be properly displayed in the test delivery environment and that test taker input and results can be collected, aggregated, and reported. For example, graphics constraints need to be identified so that item writers do not create items that have too many colors or require a screen resolution that is too high for the current test delivery environment.

Step III

Test Administration The test sponsor should provide test takers with clear and concise information regarding procedures to register for an examination, obtain an authorization for testing document, and scheduling a test appointment.

Step IV

Scoring and Score Reporting The accuracy of computer scoring algorithms should be established prior to implementation of the computer based test.

Step V

Psychometric Analysis Determine appropriate reliability indices if different test takers are given different items or exercises or pictures.

Step VI

Stakeholder Communications Developers of computer based tests should provide sufficient information concerning the test purpose, and test content specifications to test users, prior to when the test is available for widespread administration. This test information /instructions should be kept accurate and as up to date as possible.

The above six guidelines show that the guidelines provide supplemental and elaborative information on the standards for individuals and organizations seeking to develop computer based or internet based tests and assessments. (Source: Association of Test Publishers, 1201 Pennsylvania Avenue, Suite 300, Washington, DC, 20004).

The guidelines can be appropriately used by a wide variety of audiences. They are:

- *Test development organizations* — for specifying procedures for designing, developing, field testing, and validating computer based tests.
- *Test publishers and administrators and test delivery organizations* — for establishing common institutional guidelines for communication of test items, examination scores, and item response information to and from computer based testing locations.

- *Test delivery organizations* — for providing information about how to achieve high quality delivery of computer based tests.
- *Test takers* — for providing information about the types of test items, tests, and test score interpretations and test orientations they might encounter when they take a computer based test.
- *Research and evaluation specialists* — for providing information on current and expected future uses of computer based tests.
- *Teachers at educational institutions that administer or use computer based tests* — for providing information about interpreting test scores about examinations and using the test scores appropriately, improved psychometric methods, and expanded or up dated question bank or item bank etc., as well as about helping students to prepare appropriately for computer based tests.

Designing computer based testing

All computer based tests should be designed by using the fundamental *standards* identified in the *six* technical areas. These areas are:

- 1) Test construction, evaluation, and documentation;
- 2) Reliability and errors or measurement;
- 3) Test development and revision;
- 4) Scales, norms, and score comparability;
- 5) Test administration, scoring, and reporting; and
- 6) Supporting documentation for tests.

The advantages of computer based testing

There are a number of reasons why computer based testing is more helpful for a distance teaching institution. First, computer based testing can be more responsive to the needs of both the test provider or the institution and the distance learners. For example, where “on demand” testing of examinees is needed, the use of computer based testing is more useful.

- Computer based assessment works on a completely different model than paper and pencil administration and offers benefits the latter can not match.
- Smaller numbers of candidates can be tested throughout the year rather than larger numbers several times a year. Candidates can register just two days before the test, as opposed to weeks in advance for paper and pencil testing. Centers can offer different tests at the same time, because

exams are delivered on Personal Computers (PCs) using a Local Area Network (LAN).

- The computer selects test questions from a pool, so candidates taking the same exam will not be answering identical questions, which would, enhance test security. Computers eliminate the need for test booklets and answer sheets, that increase the security levels.
- Self paced tutorials show candidates how to use a mouse and other testing tools, ensuring that even those without computer experience are comfortable. Candidates can use either paper and pencil or word processing for essay questions, depending on their own preference.
- Computer based assessment allows for a diverse range of question types, which is a better test of a candidate's competency. Technology based assessment improves the link between instruction and assessment, providing a profile of candidates' strengths and weaknesses, and matches questions and the order in which they are presented to the ability of each test taker.
- Computer based tests are scored immediately or shortly after administration. The paper and pencil test results usually take four to six weeks to process. Scores are sent electronically to universities or licensing and certifying agencies. So, fast scoring helps for publishing results quickly.

Disadvantages of computer based testing

The disadvantages of computer based testing are all related to the resources required. These resources are a sufficient numbers of computers, a room to install them, appropriate software, and adequate technological expertise.

To sum up, in computer based testing the computer continuously re evaluates the ability of the distance learner resulting in a test that is tailored to each individual distance learner and computer technology is a viable tool for performance assessment in different subjects and a potentially powerful tool for replacing traditional product oriented paper and pencil tests.

In this context, Craig Mills, Executive Director of Examinations at the American Institute of Certified Public Accounts, quote "There will be an explosion of new item types and testing methodologies. So, we must be ready with good tools and approaches to manage this explosion." (Ref.: <http://www.isoc.org/oti/articles/0500/olsen.html>).

SECTION 2

Evaluation: An Overview

We have presented, in the previous section, an overview of *assessment*. In this section, we will present a brief overview of *evaluation*. Though ‘assessment’ is very often used interchangeably with ‘evaluation’, these two terms have different meanings in the context of educational practice. As stated by Thorpe (1988), evaluation is not synonymous with assessment, because evaluation is the procedure of assigning *values* to the learning outcomes during and at the end of a course. Assessment focuses on the learning of students and the results (marks/grades) may be used as a source of information for evaluation.

Objectives

After going through this section you would be able to:

- ☛ define evaluation;
- ☛ differentiate between assessment and evaluation;
- ☛ discuss the purposes of evaluation; and
- ☛ discuss evaluation in an education programme.

Meaning of Evaluation

Evaluation is an integral part of the instructional process. It involves three steps. They are:

- i) identifying and defining the intended learning outcomes,
- ii) constructing or selecting tests and other evaluation tools relevant to the specified outcomes; and
- iii) using the evaluation results to improve learning and teaching.

The above steps emphasise that evaluation is a continuous process. It is essential in all fields of teaching and learning activity where judgments need to be made. The teaching and learning process involves a continuous and inter related series of instructional decisions to promote student learning. Our main contention here is that the effectiveness of the instruction depends

to a large extent on the quality of the evaluation data/information on which the decisions are based.

Generally, most of the people engaged in the educational system (mainly teaching) are interested in ascertaining the outputs of an educational programme. Output is counted in terms of test results. Naturally, the results are expressed in quantitative indices, say scores or marks. For obtaining those scores a device consisting of a set of tasks, called a test is used. Tests may consist of term-end question papers, assignments, interviews, group discussions, projects and so on. On the other hand, an act of measurement is done when we award marks, say 60 out of 100 (or 60%) to an answer paper or a project report. Thus, we obtain the measurement of an individual's objectivity in quantitative or numerical indices when we administer a test. Sometimes confusion arises when another term evaluation is used along with the term measurement for the same process. But there is a clear cut difference between these two terms: measurement and evaluation.

In case of measurement we express an individual's ability in numerical indices or scores, in evaluation we express it in qualitative indices (say, good, excellent). By doing so we attach a value judgment to our measurement. When we say Ms. Angela has secured 75% marks in term end examination we talk about 'measurement'. But we 'evaluate' her ability by saying that she has done very good (value judgment) and stood first in the examination. We provide a qualitative description to Angela's ability.

The above example distinguishes evaluation as qualitative descriptions of Angela's performance from "measurement" which is quantitative description (75% marks).

So, from instructional point of view evaluation may be defined as a systematic process of determining the extent to which instructional objectives are achieved by learners. [Gronlund (1981)]

There are many definitions and explanations of evaluation in education. Let us see the definitions stated by Thorpe (1993) Cronbach (1980) in box 2.1.

Box 2.1

“By the term evaluation, we mean systematic examination of events occurring in and consequent on a contemporary programme – an examination conducted to assist in improving this programme and other programmes having the same general purpose.” (Cronbach et al 1980)

“Evaluation is the collection, analysis and interpretation of information about any aspect of a programme of education and training, as part of a recognised process of judging its effectiveness, its efficiency and any other outcomes it may have” (Thorpe, 1993).

Wottawa and Thierue (1990) made an attempt to explain the concept of evaluation which can serve different purposes. They are:

- Evaluation has something to do with valuation.
- Evaluation serves to help in planning and deciding and thus has something to do with assessment and valuation of alternative ways of acting.
- Evaluation is oriented towards aims and purposes. It primarily has the aim of checking practical measures, of improving them or of making decisions concerning them.
- Evaluation measures reflect the current state of techniques and research methods.

Difference between Assessment and Evaluation

The terms ‘assessment’ and ‘evaluation’ are not synonymous. Assessment is distinct from evaluation in the following way:

Assessment	Evaluation
<ul style="list-style-type: none"> • Focuses on the learning of the students. • Focuses on the performance of the students (grading or marking). • Assessment results may be used as a source of information for evaluation. 	<ul style="list-style-type: none"> • Focuses on the way the various components of a course perform — e.g. the syllabus, the teacher, the resources, and so on. • Focuses on the performance of the provider and the provision. • Evaluation results have no direct bearing on students’ assessment.

Purposes of Evaluation

The main purposes of evaluation are the following:

- *Proving*: to demonstrate conclusively that something has happened as a result of learning or training and that this may also be linked to judgements about the value of the activity; whether the right thing was done, whether it was well done, whether it was worth the cost, and so on.
- *Improving*: to ensure that either the current or future programmes and activities would be better than they are at present.
- *Controlling*: to use evaluation data to ensure that an individual learner or trainee is performing upto the standard or that subsidiary learning/training establishments are meeting targets according to some centrally determined plan.

Evaluation in an Educational Programme (EIEP)

The modern concept of evaluation has assumed wider meaning and scope than the traditional role of evaluation that was concerned only with the end of the course measurement of the quantum of learning possessed by the learners. The modern role of evaluation encompasses a wider scope in the whole educational programme (see Figure 1 below).

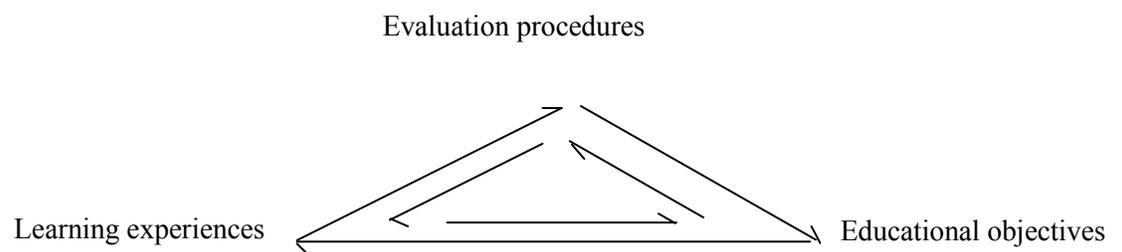


Figure 1: Evaluation in an educational programme

Any educational programme consists of three components:

- educational objectives;
- learning experiences; and
- evaluation procedures.

These three components must support one another to achieve the desired effectiveness of the programme. The function of evaluation procedures in this triangular relationship is dual:

- to examine to what extent the pre stipulated learning objectives of a programme have been achieved in terms of learners' academic achievement; and
- in an on going process, to provide continuous feedback, after evaluation (i.e., the above function), to the planners of the programme to suitably revise/modify either the programme (educational objectives), or the programme activities (learning experiences) or even both for achieving improved effectiveness of the programme.

Evaluation procedures have their respective functions at the beginning, during the process and at the end of the programme or course. At the initial stage, the programme planners use pretests to study the existing knowledge of learners about the proposed programme before they enter it, so that learning experiences, appropriate to the concerned learners can be suitably designed and effectively presented. At the mid process stage, they help monitor the progress of the learner and the teacher, and provide feedback to improve upon the process of learning. And, at the end of the programme, they help in the measurement of the learning outcomes.

Within the process-product relationship in an educational programme, the process of evaluation may help in:

- diagnosing weakness in learning;
- predicting learner's aptitudes and abilities;
- selecting suitable persons for a course or career through entrance tests;
- grading learner's abilities through tests/examinations;
- providing guidance for course choice and subject choice within a course; and
- evaluating the effectiveness of the whole programme, i.e. programme evaluation.

Evaluation of an Educational Programme (EOEP)

Evaluation within an educational context or training programme can be described as a systematic activity involving the analysis or documentation of programme related components and processes, the measurement of variables associated with the programme, and the elaboration of recommendations based upon the collected information. This activity provides opportunities for feedback, accountability, and cost-benefit information.

Evaluation in Distance Education

Evaluation has significant implications for distance education both in terms of evaluating learner's progress and attainment, and evaluating the effectiveness of the whole programme. The former refers to "evaluation *in* distance education", and the latter to "evaluation *of* distance education programmes".

In distance education, we depend on both the *continuous assessment* of learner's progress on the course with the help of both commenting and grading the assignment responses, and the *term end examinations*. In this context, assessment can be both formative and summative.

Box 2.2

Going through the assignment responses of the learner and writing comments on them is *formative assessment*, for it not only identifies the weaknesses and strengths in learning, but also helps him/her to improve upon both the process and the attainment of learning.

Assessment is also *summative* because of the fact that at the end of the course or the term, the attainment of the learner is graded on the bases of both the continuous assessment and the term end examination.

Grading or scoring may be done in formative assessment too. So, grading or scoring at both formative and summative assessment lead to award of final grades and certification of what has been achieved by the learner. Assessment in terms of both commenting and grading the assignment responses of the learner gives him/her the necessary feedback to know not only the weaknesses in his/her learning but also how to overcome them. Further, they suggest the relative position of the learner in the peer group. The knowledge of the relative position in his/her group of learners helps him/her considerably to do self assessment which is an important means of learning at a distance. Moreover, assessments (both continuous and terminal) in terms of grading lead to final award of a certificate/diploma/degree to the learner. In an open and of distance education system, the individual learners proceed at their own pace, and sit for examinations when they feel they have successfully learnt what is to be learnt in a course. This may create problems to the distance teaching institution in making arrangements for individual assessment/evaluation. Besides this, because of heterogeneity in the test taking behaviour of the distance learners, there may

be problem of constructing different items/questions for different learners. These problems call for a large question/item bank (with reliable and valid questions or items) from which questions/items are picked up to constitute tests as and when needed. (For a detailed discussion on this issue please see section 6 of this handbook).

Reliability and Validity in Distance Education

Reliability and validity assume greater significance in distance education with special reference to commenting and grading the assignment responses. Comments on assignment response are considered reliable if the distance teacher/tutor has similar reaction to the same response and writes nearly similar comments when he/she goes through that particular assignment response for the second time or even for the third time. Not only that, the contexts/occasions of commenting by one tutor must also be nearly similar to those of the second tutor if both go through the same assignment response (see box 2.2). Similarly, in case of grading the assignment responses and answer scripts, there should be higher inter reliability. The same logic is also applicable to the validity of both commenting and grading (see box 2.3).

Box 2.3

Reliability of a Test

Reliability of a test refers to the “accuracy” or the “consistency” of measurement of what the test purports to measure. To understand it more clearly in terms of our day-to-day experiences of life, perhaps a good example may be the weighing machine at the railway station. Let us suppose that the machine for all time to come weighs constantly 2 kg. less than the actual weight of persons. So if a person gets his/her weight as 62 kg in the first trial, also gets the same weight in the next trial the next minute, the weight is constant at each trial. Though the actual weight of that person is 64 kg (because the machine due to some defects always weighs a constant of 2 kg less than the actual), the machine can be relied upon because of its consistency in measuring weight.

In academic situation, a test is reliable if a learner, for instance, takes the same test at two occasions (say, with a gap of fortnight) and secures more or less the same score at both the occasions. In this context, Gronlund (1981)

points out that it is more appropriate to talk of the reliability of the test results rather than the test itself, for a test may have different reliabilities at different occasions/situations with different groups of subjects. Further, the obtained scores on a test are said to be reliable on the basis of certain criteria. These are called the estimates of reliability. Generally there are five methods to estimate reliability coefficients on the basis of which reliability may be expressed. They are:

- i) stability,
- ii) equivalence,
- iii) stability and equivalence,
- iv) internal consistency, and
- v) scorer reliability.

- Stability refers to the consistency of test results over two occasions. Test-retest method is used to estimate stability. The same test is administered twice to the same group of subjects with a time interval of, say, a fortnight or a month, and the two sets of scores (obtained at two different occasions) are correlated. A higher coefficient of correlation indicates greater stability, and vice versa.
- In case of establishing the estimate of equivalence of a test, two different versions of the same test equivalent in terms of content and the level of difficulty are administered to the same group of subjects on the same day with very little time interval. The two sets of scores obtained on the two versions of the same test are correlated, and the coefficient of correlation gives the measure of equivalence.
- The measures of stability and equivalence include two purposes:
 - i) to predict the long-term reliability of the test, and
 - ii) to infer one's knowledge in a subject area.
- The split-half method (i.e., correlating the two sets of scores obtained on the odd and even items of the same test separately) is used to establish the estimate of internal consistency.
- Variations exist in the scoring/marking/grading patterns of more than one scorer/examiner/evaluator. Two examiners may award different marks or scores to the same answer script. To estimate the reliability between evaluators/scorers, product moment correlation and Spearman-Brown prophecy formula are used.

Box 2.4

Validity of a Test

Validity of a test refers to the “duty” or “truthfulness” of the test. Does the test measure what it is supposed to/purports to measure? Before establishing the “truthfulness” of its measurement or assessment, it is necessary to define “what” it is going to measure/assess. It is very difficult to achieve the situation where the test “fully” measures whatever it purports to measure. So there is always a gap between the purposes of a test and the extent to which they may be achieved.

There are mainly four types of validity corresponding to four aspects of establishing test validity:

- i) **content validity** — related to the content of the test items;
- ii) **concurrent validity** — related to the extent to which the test under consideration matches with other tests with similar purposes;
- iii) **predictive validity** — related to its prediction of other related variable(s); and
- iv) **construct validity** — related to the extent of its inference of the constructs of the theory on which it is based.

In the above discussion we have focused upon the definition of evaluation in distance education, role of reliability and validity in distance education. Even though the definition is comprehensive enough to include different approaches, concepts, tasks and methods of evaluation, we thought of presenting, in this section, a variety of evaluation tasks in distance education. These tasks are:

- Tasks of Evaluation
 - Student assessment/evaluation
 - Course evaluation
 - Evaluation of the effectiveness of media
 - Student services
 - System evaluation and social impact analysis
 - Social needs analysis
 - Policy evaluation

Assessment and Evaluation in Distance Education

- Organizational evaluation
- Market analysis

- Evaluation Methodology

The variety of evaluation purposes and the above tasks correspond to the diversity of the methodological approaches, procedures and methods that are used in evaluation. They are:

- Qualitative approach
- Rating procedures
- Retention tests
- Survey method
- Face to face interviews
- Methods of observation
- Methods of content analysis
- Cost benefit analysis
- Experimental studies

Activity 2.1



Do you think evaluation in distance education is a difficult task? Give reasons in support of your answer.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

SECTION 3

Two Way Communication in Distance Education

Unlike conventional classroom teaching and learning where communication takes place through face to face interaction between the teacher and the students, in distance education, because the distance teacher and the learner are physically separated, two way communication takes place at a distance.

Communication at a distance may be of various levels. As a distance teacher you should have a clear understanding of two way communication in distance education.

In this section, we will discuss various types and levels of communication between the learner and the distance teacher.

Objectives

After working on this section you would be able to:

- ☛ discuss two types of communication between the learner and the distance teacher; and
- ☛ describe three levels of communication between the learner and the distance teacher.

Types of Two Way Communication

Communication at a distance may be two types:

- i) *real* as in the case of writing to the learners, commenting on assignment responses, telephone contact, computer assisted teaching, face to face counselling (that is kept at a low key), and so on,
- ii) *simulated* in which the self learning materials are written in such a conversational style that when the learner interacts with or goes through the pre produced course (i.e., the SLM) a constant conversation takes place between him/her and the author of the self learning materials.

The latter type (i.e. simulated two way communication) is called “guided didactic conversation”. For our present purpose **we are concerned with the former, i.e., real two way communication** by taking into account only the type of communication that takes place through the medium of assignment responses.

Comments written by the distance teacher on assignment responses help the learner to sustain his/her interest in learning, to get motivated and activated, and to facilitate learning. Besides, they assess the progress of the learner in comparison with the peer group (Holmberg, 1985). So it is worth mentioning here that the comments written by a distance teacher on assignment response(s) should necessarily be teaching type, otherwise, they might defeat the purposes of learning teaching at a distance.

This non face to face communication between the distance teacher and the learner is, otherwise, known as *non contiguous two way communication*. Distance learning teaching is more successful when this two way communication is frequent and constructive. Like the conventional teachers, the distance teachers are often carried away by the notion that in an answer script/or assignment response they have to find out and correct what is wrong in that, and finally award a score or a grade. This is, though not totally wrong, only half the truth.

Besides locating and correcting what is wrong in the assignment response, the distance teacher has to write ‘constructive and meaningful comments’ that suggest to the learner his/her learning style, level of performance, limitations and how to overcome those for better performance. This is the crux of two way ‘real’ communication in distance learning teaching.

Level of Two Way Communication

Communication between the learner and the distance teacher may further be categorised into three broad levels:

- Academic level
- Personal level
- Supplemental level

Let us briefly discuss each of them separately.

Communication at Academic Level

So far as the academic level of communication is concerned, the medium is the assignment response or answer script or a project or log book which the distance teacher reads and on which he/she writes comments and awards a grade. Whatever is being communicated between the distance teacher and the learner is at the academic level only and concerns only the course. The duty of the learner is to go through the course units sent to him/her and write and send back responses to assignment questions supplied from time to time throughout the course. The duty of the distance teacher, at this level of communication, is to write meaningful or teaching type comments on the assignment responses so that they may help correct and guide the learner, to accept and encourage good aspects, to assess his/her level of achievement and to explain it with keys, and to further improve the response/performance. For an elaborate discussion see Section IV of this Handbook.

Communication at Personal Level

Distance learners belong to a heterogeneous group of various socio economic status (family background, occupation, income, caste, community), academic status, experiences, personality types, and the like. So, when they react to the course units in terms of their responses to assignment questions these background characteristics are reflected in their assignment responses. It may so happen that, at different times, the learners may hide their weaknesses, flatter the distance teacher, show off their abilities, and express negative attitudes towards distance education in general and the course, the distance teacher and their own modes of learning in particular. This is what we mean by the communication at personal level. Differently, we call them tutor comment at non academic level. These comments are written in such a way that they reflect the fact that the teacher is there to help the learner in his/her learning and in other non academic/personal matters related to his/her study.

Communication at the personal level tries to break the isolation of the learner and provide extra incentive/motivation and help, other than academic communication, not only to sustain him/her on the course but also

help him/her go ahead smoothly to achieve the stipulated learning objectives. Therefore, both the distance teacher and the learner should be genuine and objective in their personal communication, though in most cases the learners' communications are dominated more by emotion than reasoning. Let us consider the following personal correspondence sent by one learner to the distance teacher.

“You are so good, intelligent and dedicated that I cannot stop appreciating you. I was so thrilled to get and go through your scholarly papers that I cannot think of a better teacher than you. And I am sure you can be the best leader and caretaker of the system of education of our country through your lighthouse of knowledge and perseverance”.

On the other hand, we may look into another personal communication (given below) from another learner written to the same distance teacher on genuine reasons and strong motivation.

“I could not submit the fourth set of two assignments in time though the previous three sets were submitted in time. By now I have also got them back with useful comments. My younger sister got married, and naturally in the absence of my father I was busy full time. I could not touch my study materials even once in the past thirty days or so. I fear I may not be able to make up and keep pace. The final exam is also approaching nearer. I don't know what to do...”

In response to this second communication from the learner concerned, the distance teacher finds interesting and genuine points to write on and that helps the latter in independent learning. In case of the first one, not only that it is difficult to respond to the learner, but also the distance teacher may lose interest in such communications (because of unnecessary and unwanted flattery) and so he/she may not write at all.

Communication at Supplemental Level

Supplemental communication is what arises out of either academic communication, or personal communication or both. It may be initiated either by the learner or the distance teacher. When a learner reacts strongly to the comments written by the distance teacher or even challenges the grade

awarded to him/her, the latter has to explain his/her position clarifying the issues. Similarly, the initiation may start from the distance teacher through, for example, model answers sent to learners when he/she finds that a particular assignment question is difficult to answer.

Communication at supplemental level provides “feedback” to the objectivity and functions as a “check” on the carelessness of the distance teacher. From an operational point of view the probability of subsequent enquiry/question from the learner related to the comments written or grades awarded on his/her assignment response compels the distance teacher to be alert and do justice to the task of distance teaching. But the task becomes more complex due to heterogeneity in the characteristics and needs of distance learners. While diverse learner backgrounds demand more insights from the distance teacher so far as the “personal” level of communication is concerned, communication at “supplemental” level requires equanimity, objectivity and scholarship because of the diverse “scholastic” abilities of the learners.

Here is an example (see box 3.1) of such communication between the distance teacher and the learner:

Box 3.1:	
Distance teacher’s comment on the assignment response	
<p><u>Comment</u></p> <p>This is fine. However, as the second group stressed on the learner, the former group highlighted the human aspect of teaching, i.e., the teacher....</p>	<p><u>Assignment Response (AR)</u></p> <p>Both Holmberg and Baath gave stress on the two way communication between the teacher and the learner. But Moore and Wedemeyer described learner’s freedom as the important criterion in distance learning...</p>
<u>Learner’s supplemental communication:</u>	
<p>Dear Sir,</p> <p>I received back the third AR in time. Your comments have been very useful to me to be in right track of learning. But, as I could not understand your comment on the second para (page 3), I would like to know whether the first group stressed on human element. I think it is Sewart who talked of this...</p>	

Distance teacher's clarification/supplemental communication:

Dear Mr. X,

Kindly refer to your last letter related to your reaction on my comments on your third assignment response. You are right that David Sewart argues for human element or 'continuity of concern for students learning at a distance'. But what I had pointed out in my comment is that while the latter group perceived DE from learner's point of view, the former group of two thinkers looked at it from a different angle, i.e., teaching/teacher/teaching organisation. I brought in the human aspect in the former group to point to the notion of guided didactic conversation....

At this stage it must be stated clearly that any administrative system which negates any of these three levels of 'real' communication should not claim to be concerned with effective two way purposeful interaction. Negating any one of these levels amounts to 'undoing' the system.

SECTION 4

Assessment of Assignment Responses

Distance learners' responses to the assignment questions sent to them by the distance teaching institutions form an important instrument of teaching learning at a distance. Long spatial distance may exist between the distance learners and the distance teaching institution, but non contiguous communication through assignment questions, responses to these by the learners, and subsequent comments by the distance teacher on those responses minimise this spatial distance and increase academic and psychological closeness between the two. Meaningful and constructive suggestions through comments written on assignment responses can effect successful distance learning. So, assessment of assignment responses of distance learners as a part of continuous assessment in terms of both commenting and grading is an activity that needs careful analysis, discussion and understanding on the part of the distance teachers.

Thus, in this section we will discuss about various types of assignments used in distance education system and various types of tutor comments (teaching and non teaching types).

Objectives

After working on this section you would be able to:

- ☛ discuss various types of assignments used in distance education system;
- ☛ discuss the importance of tutor comments; and
- ☛ describe various types of tutor comments.

What is An Assignment?

Assignment is an important device through which two way communication takes place in a distance education system. This is a learning task. This enables the learners to ensure that they have learnt what they are expected to

learn from the course materials and their response to it give their distance teacher an opportunity to help them by commenting on their performance.

Types of Assignments and designing assignments

Generally, two types of assignments are used in distance education system:

(i) Tutor Marked Assignments (TMAs) and (ii) Computer Marked Assignments (CMAs).

i) Tutor Marked Assignments (TMAs)

These assignments are marked/evaluated by the distance teacher/tutor. Thus, these are called tutor marked assignments. These assignments comprise a variety of questions such as essay type, short answer type, problem solving exercises etc. These assignments are generally used to assess the higher order cognitive objectives such as analysis, synthesis, judgement, comprehension, application, etc.

Box 4.1

Examples of TMAs

1. Essay type: Answer the following in about 1500 words.

Imagine that you have been asked to plan and organise a training programme for the academics in a newly established open university in your country. In which area will you give training first, and why? Prepare an outline of your training programme.

2. Short answer type: Write short notes of about 250 words each of the following:

- i) The wild life of Himalayas
- ii) Seasonal marketing

3. Practical type exercises:

Answer the following in about 1200 words.

Prepare a self learning unit on a theme of your choice incorporating a unit structure, two sections, SAQs, Summary and check your progress possible answers.

Sometimes these assignments comprise of objective type questions also. There are generally seven major types:

- a. True false
- b. Fill in the blanks
- c. Completion of sentences
- d. Matching
- e. Multiple choice
- f. Sequencing
- g. Graphical/pictorial type.

We assume, all of you have either designed or evaluated these kinds of questions and their responses in several occasions. So, here we will not discuss these in detail.

ii) Computer Marked Assignments (CMAs)

These assignments are marked/evaluated by the computer only. Thus, these are called computer marked assignments. These assignments are generally comprised of objective type questions, preferably the multiple choice questions, which are most versatile amongst all objective type test items. (See box 4.2).

Box 4.2

Example of CMAs

Note: Every question has 4 alternatives. Of these only one is correct. Select the right answer and write in the appropriate box in the answer sheet.

- i) When was the Planning Commission set up in India?
 - a. 1948
 - b. 1950
 - c. 1952
 - d. 1954
2. The earliest form of organisation was
 - a. Family
 - b. Army
 - c. Department
 - d. Society

Note: We have presented a sample response sheet for computer marked assignment in page no. 56.

Tutor marked assignments (TMAs)

A few criteria for designing useful tutor marked assignment questions are:

- Is the assignment question clear and unambiguous?
- Does the question directly relate to stated objectives?
- Does the assignment as a whole test the main things the learner needs?
- Does the assignment begin with fairly easy questions?
- Does the assignment include a marking scheme?

Computer marked assignments (CMAs)

In a computer marked assignment the distance learner responds to multiple choice questions. The writer of computer marked assignment is the best person to decide exactly what message should reach the distance learner. So human skills and experience are needed for designing CMAs.

What is an Assignment Response?

Assignment is a learning task. The learners make an attempt to respond to the assignment after going through the study materials. This he/she does according to the instructions written on the assignment. Thus, the answer written by the learner with respect to a particular assignment is called *assignment response*.

Commenting on Assignment Response

Passing comment is a natural habit of human being. You have observed your friends, relatives, colleagues passing comments on several occasions. You may have also done the same in several occasions. But here, we are talking about 'comments' which have special significance in teaching learning. These comments mean your reactions to the assignment response the learner expressed through written words that promote his learning. These comments should have a special significance while you are commenting on an assignment response of a distance learner. These written words should provide academic support to the isolated learner. As a teacher of a conventional system of education you use spoken words, gestures, etc. to affect classroom/face to face learning, similarly as a distance teacher you should use *written words with varied styles and in different ways* to enhance learning at a distance.

Tutor Comments

You have learnt that some assignments are marked/evaluated by the distance teacher/tutor which are called tutor marked assignments. Similarly, comments written by a distance teacher/tutor on assignment response are called 'tutor comments'. This word is popular in distance education system. A 'tutor' is a person who helps the distance learner to learn on his own. He is a facilitator. He facilitates learning by providing academic counselling, guidance, etc. and also writing comments on assignment response of the learner. In other words, writing comments on assignment response is primarily the task/responsibility of the 'tutor' or 'counsellor' in a distance education system.

Importance of Tutor Comments

As a distance teacher/tutor you should know the importance of tutor comments first.

- i) The distance learner is an independent learner. He/she may need continuous feedback to sustain and/or increase his/her motivation. Feedback in terms of fruitful comments increase the motivation of a distance learner.
- ii) The distance learner is an isolated learner. He/she does not get frequent occasions to interact with peers and evaluate his/her position among the peers. Comparison with peer group provides indirect feedback that increases the competitive spirit of the learner. He/she can assess his/her mode of learning and make necessary modifications for better performance. But the distance learner does not have the facility of regularly meeting the peer group. Occasionally, he/she can meet the fellow learners at the study centres during face to face contact sessions. Through fruitful comments on assignment response, the distance teacher can remove the learner's feeling of isolation, and can also bring him/her closer to the peer group by making him/her see clearly his/her achievement, drawbacks etc. in relation to those of the peers.

- iii) Through written comments the distance teacher can provide guidance, counselling and suggestions to improve the study habits of the learners, if necessary.
- iv) Through written comments the distance teacher can clarify the ambiguities, if any, or difficult portions of a course unit.

In short, the tutor comments are a very important tool in distance education as they are both the *content* and *vehicle of communication* which effect learning.

Writing Comments: Some Important Steps

While writing comments on any assignment response you may follow the steps mentioned below:

- i) You may go through the assignment question critically, understand what the assignment asks for and then build in your mind, the ‘ideal response’ which may be from your point of view, the best response to that question.
- ii) You may identify the weakness, if any, in the assignments question and course unit(s) on which the particular assignment is based. Then, you may match these weaknesses with the pre determined ‘ideal response’. This will help to reduce the bias in the assessment of assignment response.
- iii) Next, you may go through the assignment response and evaluate its organisational aspects, viz., introduction, body or mid part, ending/conclusion, etc., accuracy of information, content density, logical and critical analysis of concept, clarity in expression, language and so on. For assessment you may assign precise value to these organisational aspects of the response. While doing these activities you may write pedagogically meaningful comments in the *margin* of each page of the response. At the end you may write the *global comments*.
- iv) While writing global comments you may award the *grade or mark/score* to the particular assignment, as the global comments, among other things, justify the grade or marks awarded to that assignment response (Koul, 1991).

Types of Tutor Comments

If you randomly select a few evaluated assignment responses of any open university, you will find various types of comments written by the tutors. Some comments may be very effective for distance teaching, some may not. Because it is not so easy to write precise and pedagogically purposeful comments. Only the *oriented or trained distance teacher/tutor* can write pedagogically purposeful comments. Before we discuss the types of tutor comments we will suggest you to do the following activity.

Activity 4.1



First, you may go through the comments written below. Then mention in the blank space which comments in your opinion are *effective for distance teaching* (write the numbers only).

- i) “Your English is very poor. You should not have taken admission in this course”.
- ii) “You have presented an elaborate discussion. You could have discussed this more technically with the help of diagrams, illustrations and so on”.
- iii) “No, no, you are not right. You see section 2 of unit 1. You will get the right information”.
- iv) “All rubbish. Try again, you can do better”.

.....

.....

.....

.....

.....

.....

.....

.....

Here, we will discuss various types of comments generally written by the tutors under three different sub heads: i) **Comments: must be written for distance teaching**, ii) **Comments: may be written with caution for distance teaching**, and iii) **Comments: must be avoided**.

1. Comments: must be written for distance teaching

You must write those comments which help the distance learner to learn successfully. Consider the following comments:

- a) *“Your explanation with regard to information processing theory is very good. I appreciate your diagram number one.”*
- b) *“You could have discussed the factors of environmental pollution with examples and illustrations”.*

We have presented above two types of comments which are *must* for distance teaching. These are **positive comment** (a) **constructive comment**. We will discuss these very briefly.

Positive comments

The first comment (a) approves of what the learner has written. Thus, it is called positive comment. It motivates the learner and encourages him/her to improve his/her future performance. It increases the frequency and improve the pedagogic strength of two way communication. So, you must support the right information, explanation, illustration, example etc. written by the learner. These comments are generally written in the margin against or opposite to the text of the assignment response you are commenting upon.

Constructive comments

The second comment (b) offers constructive suggestion as to how the answer could have been improved. So, this may be called ‘constructive’ comment. This comment is very helpful in bringing about purposeful didactic communication. You must write this kind of comments for effective distance teaching. This should be written in the margins beside the relevant portions of the answer scripts.

Global comments

This is another type of comment which must be written for increasing further effectiveness of distance learning. This is comprised of overall

remarks made on an assignment response. This should be written on a plain paper or in a table/sheet specially prepared for writing such comments. The detailed comments on various aspects of answer or response are put together, and the inter related overall issues are highlighted in the global comments. Here you may mention about spelling, neatness, handwriting, etc. But do not forget to explain and justify the grade you have awarded to the learner. These explanations will help the learner identify both his/her weaknesses and the ways to improve the performance (see page 57).

2. Comments: may be written with caution for distance teaching

There are some types of comments written by some distance teachers which cannot function as successful teaching comments. Only with suitable additions and modifications these can be made effective distance teaching comments. Consider the following comments:

- a) *“The first part of your answer is not clear. The second part is not relevant”.*
- b) *“?”, “=”, “??”*

The first comment mentioned above is called **negative comment**. The second comment (zero meaning, non verbal remarks) is called **null comment**. We will discuss these very briefly.

Negative comments

Through negative comment, the distance teacher/tutor points out the wrong presentations and interpretations in concepts, facts, explanations, examples, relevancy of content, etc. Such comments indicate whether the answer is incorrect, incomplete or inadequate, but do not indicate what is correct and how to make it complete and adequate. Thus, this type of comments do not serve the purpose of distance teaching. The distance teacher needs to indicate categorically what can be added or changed to make the answer correct/complete/adequate.

So, while evaluating an assignment response you may point out the *weaknesses* (which is necessary), *but do not forget to suggest how these weaknesses may be removed.*

Null comments

Many teachers have the habit of putting non verbal remarks like question marks (?), underlining () etc. in the answer script of the learners. While going through the assignments some distance teachers/tutors tend to put this kind of signs and symbols.

Comments expressed through these signs and symbols do not convey any meaning to the learner. The distance teacher needs to indicate in writing what these signs and symbols mean.

So, while evaluating an assignment response you may put ‘question mark’ or ‘cross’, but *do not forget to write a few words explaining why you have put these signs and symbols* and also guide the learner to write the correct answer.

3. Comments: must be avoided while distance teaching

Sometimes, distance teachers write some comments which do not suggest anything meaningful to the learner, but on the other hand misguide the learner. Such comments disturb and even hurt the distance learner.

Consider the following comments:

- a) *“Your answer is so so. You can improve your presentation. You may spend some more time in writing this answer”.*
- b) *“I am not going to tell you in which unit you will get this answer. Better you go through the whole block once again”.*
- c) *“Are you writing English for the first time? So many grammatical and spelling mistakes? How do you expect pass marks”?*

Above we have mentioned three types of comments.

Do you think any one of it can be used for effective distance teaching?

Probably your answer is ‘No’. You are right. These comments are not at all helpful in effecting didactic communication. The first type of comments may be called ***hollow comment***, the second type is ***misleading comment*** and the third is ***harmful comment***.

Hollow comments

Some comments are only ‘word salads’. These are written words which do not suggest anything meaningful to the distance learner. The learner does not get any benefit out of these comments. Consider comments (a) above. It says that the learner can improve presentation, but how? The learner can spend some more time in writing the answer, but what will be the benefit? He may commit similar mistakes. The learner will not academically gain anything from these comments. So, while evaluating assignment response you must avoid this kind of ‘hollow comment’.

Misleading comments

Some comments may misguide or mislead the learner. Consider comments (b) above. The comment reads “...go through the whole block again”. The learner may go through the whole block again, but will it solve his/her problem? He/she needs some specific instructions and guidance to answer the assignment properly. He/she may be guided accordingly. He/she may be told the unit no. where he/she will get the answer. He/she should not be misled.

So, never write this kind of comments which may mislead the learners.

Harmful comments

Some comments are ‘rude’ by themselves. They build barriers between the distance teacher and the learner and consequently, all possibilities of ‘communication’ get blocked. Consider the comment (iii) in the previous page. This kind of comments may disturb, hurt and demotivate the distance learner. This may compel the distance learner to withdraw from the course. So, never write any comment which may cause harm to the learner.

4. Comments: to break the ‘isolation’

As a distance teacher you may break the isolation of the learner and can provide extra incentive and help through communication at the personal level. The learner has no way of assessing his/her performance in relation to that of the other learners on the course. Thus, he/she may suffer from ‘isolation’. Breaking this isolation is much more difficult than writing

comments for effective teaching at a distance. You may write as many *personal comments* as possibly you can, but your comments should be properly worded. Consider the following *personal comments*.

- “Your answer is better than other answers I have gone through till today”.
- “I appreciate your style of presentation. It is something new. No other student has presented the diagram on information processing theory. Carry on please. You will definitely be able to achieve learning objectives. Write to me whenever you feel necessary”.
- I think you have faced some difficulties while responding to this assignment. Don’t worry, most of the students have faced the same problem. Actually, you should have answered this based on your personal experience only. You have done well. While some students have done well, some haven’t”.

Activity 4.2



Five types of comments are mentioned below. Identify the types. You may discuss with your colleagues/resource persons in a workshop.

1. “Your assignment response contains many unnecessary materials.
2. “Very good. Your argument here is quite appropriate”.
3. “This example is not correct”.
4. “A person with minimum intelligence will not write like this. Do you understand?”
5. “Besides your discussion on these issues you could have made your response more authentic by mentioning the following points:
 - i) population of Bombay,
 - ii) Bombay as a business city, andii) Climate of Bombay”.

You may use this space for writing your response.

- 1) _____

- 2) _____

- 3) _____

- 4) _____

- 5) _____

Box 4.3

Broad Classification of Tutor Comments

Various types of tutor comments as discussed in this section may be broadly categorised as follows:

Teaching Comments

- Positive comments
- Constructive comments
- Global comments
- Personal comments

Non Teaching Comments

- Harmful comments
- Hollow comments
- Misleading comments
- Null comments
- Negative comments

Note: Null comments and negative comments can be made effective distance teaching comments with suitable additions and modifications.

**Assessment and
Evaluation in
Distance Education**

**Assessment of
Assignment
Responses**

SECTION 5

Grading In Distance Education

Both the traditional practices of ‘marking’ and the relatively recent innovation of ‘grading’ are concerned with reporting and interpreting examination/test results. While marking of learning outcomes is measured and expressed with the help of numerical or quantitative indices (say 40, 60, 65, etc.), grading is related to the quality of performance or expression of learning outcomes in qualitative terms (say, first/second, higher/lower, etc.). Let us discuss both the practices, marking and grading, and examine to what extent the latter is an improvement over the former.

Objectives

After going through this section you would be able to:

- ☞ discuss the advantages of grading;
- ☞ describe relative and absolute grading; and
- ☞ discuss mechanisms of grading.

Marking vs. Grading

In our examination system, marking has been followed for a long time and most of the educational institutions/systems depend on the same practice. In the practice of marking, the examination results of individual papers are expressed on a scale ranging from 0-100. It is assumed that whatever the student has learnt over a year can be expressed in terms of his/her performance on a three hour test and the performance is measured in numerical scores. But in grading the 0-100 scale is reduced to a five point scale. Grading provides an overall estimate of human ability which is more reliable. Unlike marks, grades are not influenced by the variations in subjects/disciplines.

The Advantages and Disadvantages of Marking

The **advantage** of marking is that it is convenient to express the measurement of learning outcomes/performance in numerical indices or

scores. Another advantage is that the easy comparison of the scores of a number of students helps the educational authorities to select and admit them into various courses. Though this process is easy, it may not be valid and in fact it is not. And by tradition/convention we didn't/don't question this practice. Though it has many disadvantages, we will concentrate on three issues.

- The foremost disadvantage of marking is its unattainability of the assumption on which it is based. There are varied and at times different, learner tasks and it will be wrong to analyse these tasks into exactly hundred constituent 'bits'. Moreover, in a 101-point scale, the difference of 1, for example, between 1 and 2, between 40 and 41, and between 76 and 77 may not have the same connotation.
- The second disadvantage is that the hundred constituent 'bits' (in the scale of 0-100) are spread differently (i.e., in different proportions) in the three domains of learning, viz., cognitive, psychomotor and affective. For example, a score range of 30-45 measuring learning outcomes in the cognitive domain is different from the same score range measuring learning outcomes in the psychomotor domain. Moreover, it is difficult to express human abilities, especially at affective domain, in numerical indices or scores.
- The third, but very important, disadvantage is the 'error' that is involved in the practice of marking. This error creeps in due to 'marker variability' and 'subject variability' in marking the answer scripts. Let us briefly discuss errors in marking due to these two types of variations.

i) Marker variability

Reliability of scoring objective type/fixed answer items/questions is guaranteed even if the same answer script is checked and marked/scored by a group of examiners. But scoring of essay type/free response items by more than one examiner does not provide objectivity. In fact, it can't be possible because of the fact that it is impossible to predetermine rigid 'bits' constituting a response or an answer, and it is difficult on the part of one individual marker/examiner to consistently distinguish the variations in one response/answer given/attempted by a group of learners/students.

There is no justification for awarding a particular score (neither more or less) to an answer. The correctness of the decision of an examiner to award 55 and not 56 out of 100, and vice versa can always be questioned.

Moreover, the probable ‘error’ in awarding one score, say 55 in this case, to one answer script by the same examiner varies from 5 to 7%, and in 50% of the cases of answer scripts checked, this error is more than 5%. So the true mark of a student who has been awarded 55 may well fall between 55 ± 5 , i.e., 50 to 60. To put it precisely, besides inter marker variability (variation between examiners) in scoring the same answer script, there is intra marker variability (variation in one examiner at different occasions) in scoring many answer scripts at the same period of time. So, the decision of labelling 39% marks as “fail”, 40% marks as “Pass”, and 60% marks as “First Division”, and so on has no sound theoretical basis. Also, it is wrong to declare a student scoring 55% as “Second Division”, for if we allow an error (of scoring) of 7% the same student would have secured either 48% (“Third Division” or “Pass”) or 62% (“First Division”).

ii) Subject variability

In the traditional practice of marking, most of us are used to the 0-100 scale. If you look at the ‘range’ of marking or the ‘spread’ of marking in various ‘disciplines’ including yours, you may find that the lower and the upper limits of this spread/range vary from discipline to discipline. In Mathematical Sciences the whole range from 0-100 is utilized in marking/scoring, while in Literature the range may fall approximately between 15-60. Obviously, then, a score of 60 in both Mathematics and English Literature may not have similar connotations. Likewise, a student of Science securing 60% marks in aggregate may not be equaled with a student of English Literature securing the same percentage of marks in aggregate.

So, because of these variations in the range of scores across subjects/disciplines, it becomes difficult to compare the scores of students from different subjects/disciplines. Moreover, equalisation of aggregate marks across disciplines, without taking into consideration the respective spread or ranges of scores, at the time of admission to courses and recruitment to jobs may be fallacious. For example, a student securing 65% marks in Statistics may not necessarily be considered superior to a student securing 55% marks in Sociology because, for instance, the score-range of students of Statistics may vary from 45% to 85% while that of Sociology from 30% to 65%. In that case both the Statistics and the Sociology students may be considered nearly equal in their respective abilities/knowledge in their respective disciplines.

The Advantages of Grading

As against marking, the acceptance of the principle of grading is an honest confession of our inability to be so precise in assessing human qualities. Unlike marks, grades are not influenced by the variations in disciplines/ subjects. If almost all human qualities and achievements can be measured quantitatively, only then scoring/marking is justified. But this is not always the case.

Given below are a few advantages of grading.

- Grading is more ‘precise’ and ‘reliable’ than marking. It is ‘precise’ because the 0-100 scale is reduced to a five point, or seven or even nine point scale. As we have seen earlier, marking (with its numerical scale of 0-100) ensures a false precision in the measurement of human abilities. Grading, on the other hand, provides an overall estimate of human ability which is more realistic and so ‘reliable’.
- The accuracy in assessment cannot be fully ensured, whether it is marking or grading. But the chances of error of judging the ability of one individual to be better or worse than that of the other are much less in grading. At the same time, because of a very precise scale with a short range, error due to marker variability in grading the same assignment response is minimised.
- In all subjects whether Literature or Mathematics there is a greater possibility of utilising the whole grade range in the system of grading. Therefore, the error due to subject variability is lower because grading minimises the disparity in the value of scores/grades in different subjects.
- Another reason why grading is more realistic and reliable is that it is based on two standards.

Let us briefly discuss these two basic standards of grading before examining the techniques of grading assignment responses.

Bases of Grading

Grading is based on two standards — *relative* and *absolute*.

A *standard* is an external criterion or a set of external criteria. While grading, the standard or the model is kept in view against which a grade to a particular assignment response is awarded.

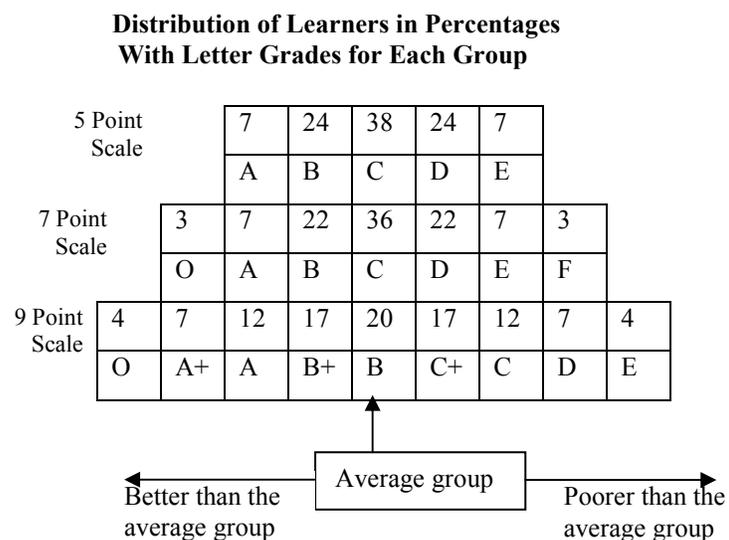
For example:

- When the relative grading standard is followed in grading the assignment responses of a group of learners on a given question or task, the criterion is the performance of the whole group of learners on that task.
- In case of absolute grading standard the criterion may be the ideal performance on the given task, irrespective of anybody's performance.

Both relative and absolute grading standards are elaborated in the following pages.

Relative Grading Standard

As said earlier, in relative grading standard the performance of the whole group of learners is kept in view as the standard against which grades are to be awarded. Before assessing the assignment response, a decision is made regarding the number of learners to be allocated/put under various categories within the range of the scale, say five point, seven point, nine-point, etc. The preferred mode of distribution determines the number of learners to be put in various grade categories. For example, if 500 learners take one assignment question and prepare assignment responses, the distribution in a five point scale would be even and balanced on either side of the central figure, i.e., 38% in C grade, and 24% and 7% in B/D and A/E grades respectively. You will find three distribution modes corresponding to three different point scales, (i.e., 5, 7 and 9) from Figure 3.



You might notice from Figure 3 that the averages or the middle portions in the 5 point, 7 point and 9 point scales are 38% (C grade), 36% (C grade) and 20% (B grade) respectively and the remaining ranges in both the sides of the middle portion are equally distributed. This can be practically presented in a tabular form (both in case of 5 point and 7 point scales) when the assignment responses of 500 learners on a given assignment question are to be evaluated. (See Table 1)

Table 1: Distribution of “population” (N = 500) on 5-point and 7-point scales

5 Point Scale			7 Point Scale		
Grade	% of students	No. of students out of 500	Grade	% of students	No. of students out of 500
A	7	35	O	3	15
B	24	120	A	7	35
C	38	190	B	22	110
D	24	120	C	36	180
E	7	35	D	22	110
			E	7	35
			F	3	15

Table 1 suggests that if a 5 point grading scale is to be followed, the number of learners out of 500 to be awarded grades A, B, C, D and E would be 35, 120, 190, 120 and 35 respectively. The middle portion (i.e., grade C) has 190 learners and the rest 310 learners are equally distributed on both the sides of the scale. Similar principle is followed in case of 7 point scale. (See the right hand portion of the table)

Activity 5.1



Before you proceed, you may take up the following activity.

- i) List the limitations of relative grading standard.

.....

.....

.....

.....

.....

Though very easy to use and handle, the limitations of relative grading standard are many.

- First, the grades do not tell the teacher exactly “how much” of “what” the learners have mastered in a subject. This standard may be helpful to, for example, Recruitment Boards of Universities which make select/reject or pass/fail decisions respectively for their purposes.
- Second, grading a learner’s ability in relation to the overall performance level of the class/group does not convey much meaning to him/her unless accompanied by a clarification on group performance. A learner who has secured ‘B’ grade in one class/group may secure ‘D’ grade in another class/group. This inconsistency challenges the reliability of grading learners’ performance.
- Third, in a 5 point scale for example, at least 7 % of learners are to be given ‘E’ grade (which means “poor” or “unsatisfactory”) even if they have written good assignment responses. If these 7 % learners do better next time, they may come up to ‘D’ grade, which further means that 7 % of learners from D grade have to come down to ‘E’ grade. This not only causes injustice to learners but also creates misunderstanding among them.

Now let us see in what way “absolute grading standard” overcomes these drawbacks in grading learners’ performance.

Absolute Grading Standard

Unlike relative grading where a fixed number of learners are put in each category of grades, in absolute grading a range of scores is put against each grade. Table 2 presents an example of absolute grading, if, a 7 point scale is followed.

Table 2: Absolute grading: distribution of range of scores across grades

Grade	Range of scores
O	90-100%
A	80-89%
B	70-79%
C	60-69%
D	50-59%
E	40-49%
F	Less than 40%

So, unlike relative grading, there is no prefixed hard and fast rule to put a fixed percentage of learners under each grade. In absolute grading, for example, a situation may arise when no learner gets F grade, and the average learners may fall within 50-59 score range (which is not the mid portion). Though ideal in evaluation and grading learners' performance, it is not easy to follow this standard of grading. Much depends upon the type of item/question/assignment response, which is being used to evaluate learners' ability. Depending on the objectives of a particular programme, one may go in for either absolute or relative grading system.

Grading Mechanisms

Grading assignment responses can be done on different scales suitable for different items like objective type, short answer, essay type, etc. For objective type items/questions having two alternative answers "Yes/No", for example, scoring can be done with the help of a two point scale, say 0-1, and the final score can be converted into a grade from the given table of specification (say Table 2 in this case). For essay type items/questions, the range of scores can be a 7 point scale putting the evaluation of responses into a wide range of qualitative judgments. The interpretation of the scale consisting of seven grade points can be done as given in Table 3.

Table 3: Interpretation of 7 point scale

Grade points	Interpretation
6	Outstanding
5	Excellent
4	Very Good
3	Good
2	Satisfactory
1	Pass
0	Failure

When grading is done, especially of essay type items/questions, certain criteria regarding

- **content** (viz., relevance, adequacy, appropriacy, etc.);
- **form** (viz., clarity, coherence, etc.); and
- **presentation** (viz., imagination, intelligence, resourcefulness, etc.)

of the answers/responses are taken into consideration. The objective items may be “Yes/No” type or “multiple choice” type having only one correct answer, and so they cannot be spread over a 7-point scoring scale. Similarly, essay type items cannot be put within two point scale of right/wrong answer.

You must be interested at this stage to know “how” to grade different item/question types and combine the grades to find out the final grade for a learner. The succeeding sub section realises this objective.

Grading Different Item Types and Combining Grades

Let us consider the grading of answer scripts of learners appearing at the term end/final examination.

- Suppose the final examination “Paper 1” consists of:
 - i) four essay type questions,
 - ii) eight short answer questions, and
 - iii) sixteen objective type questions.
- Suppose again that the instructions say that the essay type questions are to be graded on a 7-point scale (from O to F or 0 to 6), the short answer questions are to be graded on a 3-point scale (from A to C or 0 to 2), and the objective type questions have two alternatives (Yes/No) and so have 2-point scale (0-1).
- In the examination, for example, Mr. Y, has secured grades C,B,D,E in the four essay type questions (on the 7-point grading scales), grades A, A, B, C, C, A, A, C in the eight short answer questions (on the 3-point grading scale) and has given 11 correct answers to objective questions (on the 2-point grading scale).

Now let us calculate grades for each of the three items (question types) and find out the final grade.

Step 1

$$\begin{aligned}
 \text{Sum of grade points on essay type items (on 7-point scale)} &= \frac{\text{Combined grade for essay type items} \times 6}{\text{No. of essay type items} \times 6} \\
 &= \frac{[C + B + D + E] \times 6}{4 \times 6} \\
 &= \frac{[3 + 4 + 2 + 1] \times 6}{4 \times 6} \\
 &= \frac{60}{24} = 2.50 \text{ or Grade 'C'}
 \end{aligned}$$

(Where 6 is the maximum grade point on a 7-point scale)

Step 2

$$\begin{aligned}
 \text{Combined grade for short answer items (converted into 7-point scale)} &= \frac{[A + A + B + C + C + A + A + C] \times 6}{8 \times 2} \\
 &= \frac{[2 + 2 + 1 + 0 + 0 + 2 + 2 + 0] \times 6}{8 \times 2} \\
 &= \frac{54}{16} = 3.38 \text{ or Grade 'C'}
 \end{aligned}$$

(where 6 is the maximum grade point on a 7 point scale, and 2 is the maximum on a 3-point scale)

Step 3

$$\begin{aligned}
 \text{Combined grade for objective type items (converted into 7-point scale)} &= \frac{11 \times 6}{16 \times 1} \\
 &= 4.13 \text{ or Grade 'B'}
 \end{aligned}$$

(where 6 is the maximum grade point in a 7 point scale, and 1 is the maximum on a 2-point scale)

We can find out the overall grade for this question paper consisting of three item types (essay, short answer, and objective type) by calculating the average of the grade points of the three item types.

Step 4

$$\begin{aligned}
 \text{The overall grade for Mr. Y is} &= \frac{\text{Sum of grade - points on different subsets}}{\text{Total number of subsets}} \\
 &= \frac{C + C + B}{3} = \frac{3 + 3 + 4}{3} = \frac{10}{3} = 3.33 \text{ or Grade 'C'}
 \end{aligned}$$

In certain cases, the subsets (essay type, objective type, etc.) in a question paper may carry different weights for which weighted grades are to be calculated.

For example, in the present case if subset I (essay type) has a relative weight of 2, subset II (short answer type) a relative weight of 3, and subset III (objective type) a relative weight of 1, then the weighted grade and the overall grade will be as given in Table 4.

Table 4: Tabulation of scores on subsets of different weights

Subsets	Relative weight	Grade-point obtained	Weighted grade
I	2	2.50	5.00
II	3	3.38	10.14
III	1	4.13	4.13
Total	6		19.27
Overall grade = $\frac{19.27}{6} = 3.21$ or Grade 'C' (Good) (see Table 5)			

For converting GPA or grade point average to corresponding grades and verbal description of those grades, a conversion table may be prepared that might help in easy calculation of the overall grade for individual learner. Table 5 is an example of such a conversion table for a 7-point scale of grading.

Table 5: Conversion table of GPA into letter grade

Grade range	Grade point	Letter grade	Verbal description
5.50 and above	6	O	Outstanding
4.50 to 5.49	5	A	Excellent
3.50 to 4.49	4	B	Very Good
2.50 to 3.49	3	C	Good
1.50 to 2.49	2	D	Satisfactory
0.50 to 1.49	1	E	Pass
Less than 0.50	0	F	Failure

From Table 5 we notice that each grade point has a corresponding grade range. The grade range for a particular grade point is $\pm .50$ to that grade point. For example, the grade point of 3 has the lower limit of 2.50 and the upper limit of 3.49. Any calculated GPA falling in between the range of 2.50 to 3.49 is converted into the letter grade of C, which means “Good” (GPA of 2.49 is Grade D whereas 2.50 is Grade C). In the foregoing example, we find that the GPA of Mr. Y is 3.33 which falls in the grade range of 2.50 to 3.49 and is represented by a letter grade of ‘C’, meaning

that the performance is “Good”. Similarly GPA falling in between the range of 3.50 to 4.49 indicates “Very Good”, 5.50 and above indicates “Outstanding” and 0.50 and below indicates “Failure”.

The Indira Gandhi National Open University follows a letter grading system of 5 point scale (A, B, C, D & E). The student is required to obtain at least a “D” grade in assignments and terminal examinations of each course.

The evaluator as per his/her judicious estimation awards grades to the questions in accordance with their national values while assessing the assignments relating to Diploma and Certificate Courses.

The following tables indicate the qualitative levels and grade points corresponding to various letter grades, average grade points of assignments and term end examinations and over all course grade point range and also the percentage of course.

Table-I
Qualitative values of letter grades and their equivalent point grades and point grade range of diploma and certificate programmes having weightage of 30:70 in continuous evaluation and term end examination

Letter Grade	Qualitative Value	Grade Point	Assignment Weightage Range		Direct P.G. In T.E.E.	Point Grade Range	Overall P.G. Range of C.E. & TEE
			15%	30%			
A	Excellent	5	0.75	1.35 & above	3.50	3.15 & above	4.5 & above
B	Very Good	4	0.60	1.05 & below 1.35	2.80	2.45 & below 3.15	3.5 & below 4.5
C	Good	3	0.45	0.75 & below 1.05	2.10	1.75 & below 2.45	2.5 & below 3.5
D	Satisfactory	2	0.30	0.45 & below 0.75	1.40	1.05 & below 1.75	1.5 & below 2.5
E	Unsatisfactory	1	0.15	Below 0.45	0.70	Below 1.05	Below 1.5

Table-II
Qualitative values of letter grades and their equivalent point grades and point grade range where the proportion is 25:75 in continuous evaluation and term end examination

Letter Grade	Qualitative Value	Grade Point	Assignment Weightage Range		Direct P.G. In T.E.E.	Point Grade Range	Overall P.G. Range of C.E. & TEE
			12.5%	25%	75%	75%	
A	Excellent	5	0.625	1.125 & above	3.75	3.375 & above	4.5 & above
B	Very Good	4	0.50	0.875 & below 1.125	3.00	2.625 & below 3.375	3.5 & below 4.5
C	Good	3	0.375	0.625 & below 0.875	2.25	1.875 & below 2.625	2.5 & below 3.5
D	Satisfactory	2	0.25	0.375 & below 0.625	1.50	1.125 & below 1.775	1.5 & below 2.5
E	Unsatisfactory	1	0.125	Below 0.375	0.75	Below 1.125	Below 1.5

SECTION 6

Question Bank

The number of questions needed for a course/programme which runs for a number of years is likely to be very large in a distance education institution. Course team members, in open and distance learning system, therefore, might consider gradually introducing a system whereby they could develop questions, question types, and a number of different types of questions based on specific objectives. All the questions would be vetted by course team members for quality, then they would be indexed and banked. These questions when pooled in a bank, is called a question bank. Such banks allow a great deal of flexibility into assessment practices and distance teaching. For example, 'parallel' questions can be generated from the bank which would allow students who were ill for an examination or due to some urgent work remained absent to take up the examination later. The advantage of question bank is that questions which work out well in practice can be reused on a number of later situations. Thus, new questions do not have to be generated at the same rate from year to year and the quality of questions gradually improves.

The question bank is, thus, a planned library of test items pooled through co operative efforts under the aegis of an institution for the use of evaluators, academics and students in partial fulfillment of the requirements of the teaching learning process. It is, therefore, a systematic collection of a number of questions. The question bank is designed to fulfil certain pre determined purposes. Its efficient functioning demands co operative enterprise. Its clientele is specified depending upon the nature and scope of questions that constitute the bank. It is a utility service with an in built feedback mechanism for improvement of its questions.

Objectives

After going through this section you would be able to:

- ☛ describe the purposes of a question bank;
- ☛ develop questions for a question bank;
- ☛ designing a blue print; and
- ☛ use a question bank.

Purpose of a Question Bank

A question bank can serve two purposes: i) to enrich the instructional aspect, and (ii) to judge the distance learners in terms of instructional efforts. On the instructional side, questions can be used by the teachers at the pre-testing stage for the development of a unit and for revision purpose. These questions can be used for self assessment questions (SAQ) as assignments and for term end examination. A pool of test questions can be used to frame a unit or a topic and to test for formative evaluation which is an integral part of distance teaching. Such a stock of questions can also be made use of in preparing question papers at the end of a term or a session for summative evaluation. Questions for mastery testing can be used for diagnosis of the student's difficulties while going through the learning material.

Keeping in view the place of evaluation in distance teaching, it is obvious that every teacher is supposed to prepare a larger number of quality questions on different topics of the prescribed course/programmes for various purposes. Because of lack of adequate know how and the limited time at his/her disposal, a teacher cannot be expected to develop a question pool of reasonable magnitude and quality. This is true, also, of paper setters who are required to frame question papers within a specified time. In the absence of a stock of ready made questions, the quality of question papers is liable to suffer. That is why preparation of quality questions in different subjects has to be entrusted to experienced teachers who are well conversant with the content and technique of framing questions. Under these conditions, it is necessary that a readymade stock of questions is built up and made available to teachers and testers. Such a pool of test materials can be of immense use if developed according to predetermined objectives.

Activity 6.1



Write down some of the features of a question bank.

.....

.....

.....

.....

Planning a Question Bank

In order to run a question bank efficiently it should be established after due planning. At this stage it is essential that the objectives of such a bank should be clearly visualised. The two major objectives can be:

- a) To increase the value of measurement.
- b) To increase the pedagogical value of evaluation.

Further details in terms of more specific objectives can be worked out by the institutions establishing the question bank.

Where should such a bank be located? To begin with, the Association of Indian Universities (AIU) has established question banks in various subjects. Apart from this, individual universities (state or central) can also have their own question banks. This, of course, would mainly depend on the materials received from the key institutions of the State, besides some materials developed in the institution itself.

Wherever the location of the bank may be, it should be managed properly to provide utility services to all those who are interested. For this, suitable place and equipment for storing, vetting or screening, sorting and classifying of questions are required. For efficient maintaining and up keep, subject wise assistance has to be arranged under the overall guidance of an academic manager who directs the activities of the question bank.

Development of Question Bank

What *type of questions* make up the bank depends entirely on the total framework of reference envisaged at the planning stage whether only the written examination questions, oral examination questions, practical exam questions or the questions of all three are to be kept in mind for developing the question bank. This development will involve the following guidance.

Blue Printing for Developing Question Bank: The blue print in question bank may be thought of as a two dimensional grid:

- i) the behaviour/objective aspect; and
- ii) the content/subject area aspect.

The *objective aspect* refers to the expected learning outcomes in terms of abilities like recall, recognition, translation, extrapolation, application, analysis, synthesis, evaluation and any other abilities. The *content aspect* connotes the unit, sub unit of topic through the medium of which the above mentioned abilities are developed or tested. A good question pool will be one that contains questions on all the topics in a subject testing all abilities. A mere collection of a large number of questions will not constitute a question pool of quality, unless all such items fit into a pre determined structure. A sample of blue print for developing a question bank is given below. Before going through the blue print kindly perform these two activities given below.

Activity 6.2



1. Explain the factors you should consider for establishing a question bank in your area.

.....

.....

.....

.....

.....

.....

.....

.....

.....

2. Design a blue print for developing a question bank of your own subject.

.....

.....

.....

.....

.....

.....

.....

.....

In accordance with the blue print, questions are written or collected from various sources as under:

- Ready made questions may be collected from old question papers set in various examinations, from standardised tests and some from the review exercise of good text books. Such questions may need pruning from the point of view of their format before they are accepted as such.
- In the case of new questions, these may be invited from experienced teachers, examiners and paper setters. Such questions should, invariably, be accompanied by a key and outline answers, besides indicating the objective and the content area.
- Apart from the above techniques, it is still better to get the questions prepared by practicing teachers, invited to a workshop for the purpose. Such questions should be prepared by the participating teachers in their area of excellence and discussed in details before finding a place in the bank.
- Help of sister agencies may also be sought in enriching the bank as well as expertise.

Screening of Questions

After the questions are written, the question sheets are passed onto other members of the group for their comments. The comments are further passed onto the author of the question who in consultation with two or three participants finalises the question. Alternatively, the individual questions are written on the black board by the author followed by a discussion by the participants. Though it is a time consuming process, it is educationally more potent and pedagogically sound. Not only the quality of questions improve but it also provides good training to the participants for framing good questions.

The second level screening may be done with the help of a group of three subject experts, all conversant with the technique of test construction. Such a group may consists of subject specialists who are in a better position to pass judgement on the authenticity of the subject matter. A teacher who has the experience of teaching that particular class for which the question is written, should also be associated. S/he is in a better position to judge the suitability of the question for a particular grade level. The third person may be an evaluation expert who can help in improving the format of the

question in the light of the objectives to be tested. Further refinement is possible only after try out of the question. One of the main purpose in trying out of the question is to obtain valid, reliable and useful information about a question. This involves determining what is to be measured and then defining it in such precise terms that test items can be framed that call forth the desired performance. It also involves specifying the achievement domain in such a manner that the sample of test tasks will be representative of the total domain of achievement tasks and that the results will be appropriate for the intended instructional uses.

The possibility of preparing valid, reliable and useful questions is greatly enhanced if a few basic steps are followed. They are:

- Determining the purpose of testing
- Developing the test specifications
- Selecting appropriate questions
- Preparing relevant questions
- Assembling the test
- Using the results

There are certain factors which influence the validity, usability and reliability of test items. These factors are:

- Clear instructions to respond to the questions
- Sentence structure of a question

Test items appropriate for the outcomes being measured.

- Ease of scoring
- Ease of interpretation and application
- Cost of developing items.

Nevertheless, the questions are now ready for transfer to question cards for proper record.

Activity 6.3



Why should we screen the question while putting them in a bank?

.....

.....

.....

.....

.....

.....

.....

Preparing The Question Cards

The finalised questions may now be transferred to cards made of thick chart paper cut into a post card size. Cards with different colours may be used for different objectives. For example, we may use the following colour scheme in preparing cards for different objectives and various forms of questions:

Objectives Form of Question/	Knowledge	Comprehension	Application	Skill
1. Essay type	Dark Yellow	Dark Green	Dark Pink	Dark Blue
2. Short answer type	Yellow	Green	Pink	Blue
3. Objective type	Light yellow	Light Green	Light Pink	Light Blue

In case we are interested in having one colour, we may use white for all cards with top margin of about one inch coloured strip on the same pattern. Such a colour scheme facilitates the identification and sorting of questions arranged in cabinets. Each card will carry the question data on it. A sample question card is given below:

QUESTION CARD (FRONT SIDE) OBTVERSE

Objective.....	Topic...	Form...	Estimated difficulty level....	S.No...
Specification:	Sub topic....	Org level.....	Estimated time...	Code No.
Questions:				

QUESTION CARD (BACK SIDE) REVERSE

Class for which suitable.....	Date of accession	Date of try out
Difficulty level	Discrimination value.....	Place of try out
Key/outline answer:		

Estimated difficulty level may be indicated on a three point scale, e.g., A for difficult, B for average and C for easy. The actual difficulty value can be calculated only after the try out of the questions and, therefore, may be entered after the try out.

Activity 6.4



Prepare a question card which will be helpful for your subject.

.....

.....

.....

.....

.....

.....

.....

Coding of Questions

After the cards are prepared, they are to be coded and classified in a manner that facilitates their use. This is necessary for quick location and sorting of cards just like books in a library. It is also advised that a system like that of

Dewey's Decimal System of cataloguing the books is followed here to handle thousands of cards containing the questions.

Filing and Storage of Questions

A drawer/almirah/file cabinet with four or five cabinets each may be procured to hold the cards. It is easier to get the question cards typed as and when these are finalised. In case of unit tests and question papers, the questions may be typed on sheets of paper or even cyclostyled when more copies are needed for circulation. When a pool of unit tests full of question papers, diagnostic tests, standardised tests etc. are to be stored, almirahs of desired specifications may be procured to file such test materials. All question cards and other materials may be stored in duplicate so that one copy of that may be issued to the users as and when required. Original copy may be kept intact for official use and reference. *Computer facilities are the latest additions to these aspects* (MS Access).

Using the Question Bank

When a sufficient number of questions, unit tests and question papers are ready in the pool, these can be used for the following purposes:

- A judicious selection of questions can be made for instructional purpose. Different types of questions selected from a question bank may be used for pre testing, development, review and revision of a lesson.
- In the preparation of textual material a question pool can be utilised for preparing review exercises in text books. Of course, the caps are to be filled in by the authors. Likewise, the preparation of teaching units or resource units also involve the use of evaluation materials which may be picked up from the question bank.
- For evaluating pupils' progress the question bank can be used most efficiently. Individual questions can be stored and grouped for use in topic or unit testing in periodical test. Individual questions, unit tests and question papers can be profitably used by the examining agencies by making the question bank available to their paper setters.
- When question banks are established in institutions, students can use them for self evaluation in their spare time. As an outline answer or key is provided in such questions, students can check their response against

such keys and answers, especially in case of short answer and objective type questions. When questions on all topics of the prescribed syllabus are available pupils can revise their lessons. Even teachers can make use of such cards for quick revision. Like quiz cards, these cards can be used for an inter section/class competition by apportioning a particular subject area for students to learn through healthy competition.

Activity 6.5



Describe how you can utilize a question bank.

.....

.....

.....

.....

.....

Conclusion

A question bank may become a store of outdated material after some years, if not evaluated at regular intervals. At least once in three years, the question pool must be screened to discard the obsolete questions. Dead wood, therefore, must be continually removed and new material added. Curricular changes also put new demands from time to time. Enrichment of questions by updating, replacing, discarding, modifying, adding new questions, regrouping and classification is to be an ongoing process to give the question bank a dynamic look.

SECTION 7

Programme Evaluation

In section 2 of this handbook we have noted that the aim of evaluation in the case of an open and distance education institution is to support the institution, in achieving its goals. One of the dimensions of such an exercise is evaluation of various programmes of study. In this section we discuss the evaluation of distance education programmes. In the context of IGNOU, the term ‘programme evaluation’ is used to mean evaluation of an academic programme as a whole or a part thereof or a training imparted or service rendered. It also includes both print and electronic media. Before going to discuss what is programme evaluation, let us see what does a programme mean.

Objectives

After going through this section, you would be able to:

- ☞ define the concept of programme evaluation;
- ☞ describe an approach to programme evaluation in distance education; and
- ☞ explain the process of programme evaluation at IGNOU.

What Is a Programme?

An academic programme in open and distance learning consists of a few courses offered at a distance. It might involve print, audio and video materials, other modern technologies, assignments with tutor commenting and grading, project work, extra reading materials, face to face academic counselling, workshops and seminars, training, learning from resource centres, home kits, students’ own study, term end examination, etc.

The programme has certain broad objective(s). For example the PGDDE programme of IGNOU has the following objectives:

- Promoting awareness about the concept and utility of open and distance education in India and other developing countries; and

- Developing the much needed human resources for the existing open universities and correspondence/distance education centres, and many more that may come up in the near future.

The PGDDE programme is designed and developed to provide education and training which is of benefit to both distance learners and the distance educators. The PGDDE programme is evaluated to help the institution in meeting both internal and external needs for information, feedback and insight into the nature and quality of the teaching and training provision.

So, the term programme evaluation is intended to encompass the evaluation for developmental or appraisal purposes of a programme. Please see box 7.1 for a definition of programme evaluation.

Box 7.1

“By programme evaluation I mean evaluation which focuses on programmes of study. It is at this level that the pedagogic, management and often the financial responsibilities lie in education and training. It is usually here that the responsibility for the detailed issues of quality and accountability have to be exercised.” (Calder, 1994).

We shall discuss in this section the issues involved in the process of evaluation of printed coursewares, evaluation of media materials, evaluation of student support services, evaluation for accreditation, evaluation for financial backup and evaluation to assess social impact of the programmes.

Concept of Programme Evaluation

“Programme evaluation is necessarily concerned with the specific programme that has already taken place, i.e., it fairly deals with the retrospective findings. On the other hand, the results of such evaluation certainly feed to the long term decisions of the programme and looked this way, a programme evaluation obviously contributes to planning of the programme. Therefore, this is necessarily an on going process and there is certainly a starting point for both the first planning and the first evaluation of a programme. What we are concerned here with is suggesting and describing the first evaluation of a programme; both processes and products, which obviously contribute to certain aspects of planning of the programme. And, this is a continuous process.

Secondly, such evaluation might adhere to either an *accountability perspective* (i.e., to examine especially the efficiency of a programme so as

to report to the funding agency/authority) or a *managerial perspective* (i.e., to assess the effectiveness of the programme so as to provide feedback to the programme manager or the programme team regarding the effectiveness of programme delivery and management, and the programme itself). In the former situation, the focus is on the ultimate objectives of the programme. The evaluation methodology is scientifically objective. The data gathered are mostly quantitative. And the purpose behind such an exercise is to decide whether to retain the programme or reject it altogether. On the other hand, the latter type of evaluation is concerned with the immediate or intermediate objectives. The methodology followed is rigorous to the extent that sound *decisions* can be derived. Collected data are both quantitative and qualitative in nature. And the purpose of such evaluation is to improve the programme along with its delivery. In our discussion on programme evaluation, both the perspectives have been thought as necessary and therefore have been considered, though the stress has been on the managerial perspective so as to increase the effectiveness of the programme”.

(Extract from Panda, 1991)

Based on the works by Jenkins (1978), Gooler (1979), Rumble (1981), Feasley (1988), an evaluation perspective/approach has been formulated as given below.

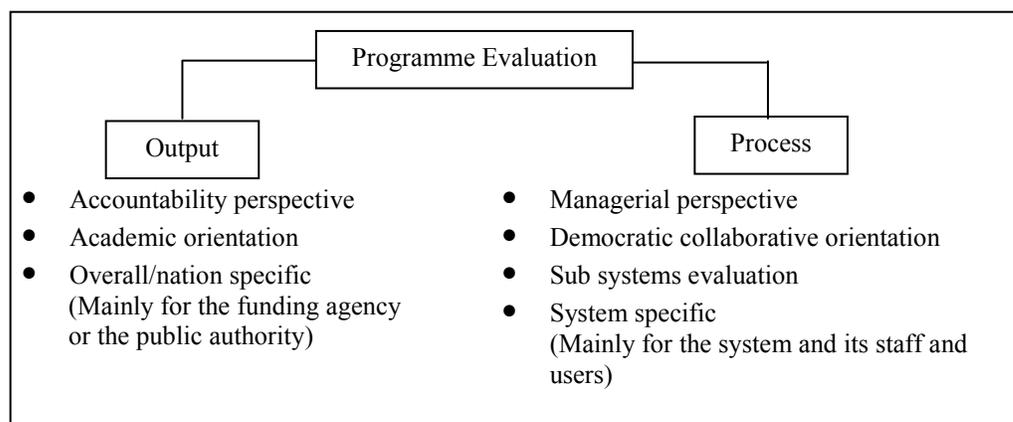


Fig.7.1: A perspective/approach to programme evaluation in distance education

Source: S. Panda (1991) ‘Programme Evaluation in Distance Education: A Perspective and Proposed Agenda of Actions’.

“Output

This is concerned with the overall evaluation of the system that has more cultural or national connotation with an accountability perspective (accountable to the government or some funding agency like the public in general/the parents of the students in particular) where the evaluator acts as an adviser and has control over the entire process of evaluation. Some of the dimensions/parameters at this level include the following:

- Equality of educational opportunity.
- Access or quantity.
- Student grading/marking.
- Relevance to needs and expectations.
- Impact on other open universities/distance learning systems, and traditional learning systems.
- Overall efficiency: cost not only to the system but importantly to the nation because many of the cost components like broadcasting, infrastructure at regional and study centres, etc. include expenditure by public authority and/or private institutions.

Process

In fact, this includes both the process, the input, and a part of output put together (i.e. operation of the sub systems) that has more system connotation with a managerial perspective (with the intention/objective of improving the effectiveness and efficiency of the functioning of the subsystems so as to increase the overall system effectiveness and efficiency) where the evaluator acts within a democratic collaborative style at every stage of the evaluation exercise(s). The following are some of the dimensions/parameters at this level:

- Quality of programme(s)
- Generation of knowledge (teaching methods and learning processes in general)
- Student learning (pace, style, strategy, attitude, satisfaction, student feedback, etc.)
- Curriculum development and implementation
- Instructional design and development; course design and self instructional materials print and non print (audio/video), and their openness
- Assignments (two way communication, commenting, grading, etc).
- Support system (the system, academic counselling, other supports) perceived by the students, the ACs, the faculty, etc.
- Course team
- Student entry characteristics (age, sex, residence, caste, previous educational experience, economic status, study skills, language proficiency, course needs, attitude to the system, occupational background, spare study time, etc.)
- Student dropout/dropdown: background characteristics, social integration, academic integration, goal commitment, institutional commitment, course/personal/socioeconomic problems, etc.
- Subsystem(s) efficiency: costs; subsystem(s) effectiveness
- Student administration and related student affairs
- Material production and distribution
- Admission system
- Evaluation system (with special reference to course objective vis à vis testing and grading/marking)

Assessment and Evaluation in Distance Education

- Quality control and decision making subsystems; evaluation of the *line managers* and their efficiency and decision making; evaluation of decision making processes/mechanisms, etc.
- The logistical systems: personnel, finance, establishment, administration, etc.
- The coordination system, especially among the important instructional functionaries like course writers, academic counsellors, media producers, paper setters, evaluators, etc.
- The system of staff development: orientation and training
- Staff attitude towards and satisfaction of teaching (materials, tutorials, commenting and grading, etc.)
- Academic freedom and its utilisation
- Users' perception of data base
- Staff progress report vs. actual performance/programme management
- Graduate placement/employment
- Employers' perception of the programme.”

(Extract from Panda, 1991)

The above model might be misinterpreted that if the output (for example, the students) is up to expectation, there is no need to study the system operation and, therefore, only when output is not satisfactory against certain criteria (id at all pre fixed), then only one may go for evaluating the system operation. But it is construed that this may not always be the case. What if the students at entrance to the system were very weak or dull? If they do not succeed, the system may not necessarily be non functional. The argument can be that if the courses/programme(s) are based on learner needs and characteristics, the question of non functionality may not arise, though this ideal is rarely achieved.

Therefore, what we are concerned with is evaluating not only the output but also the process/subsystems even if the output is very satisfactory. So in this model of programme evaluation, both the final product as well as the process of teaching learning including the logistics are simultaneously evaluated on a continuous basis so as to increase and maintain subsystem efficiency. (Panda, 1991)

learning as far as the student is concerned. These cost factors help in estimating whether a course is cheaper than the other, etc., and thus establish their relative cost effectiveness.

However, it should be mentioned here that IGNOU is still at the experimental level with regard to evaluation of curriculum planning.

Box 7.2

For evaluation of the materials produced, a checklist of the following type is generally used:

- Is the course content adequate?
- Is the language used appropriate to the level concerned?
- Is the material adequately self instructional?
- Is the media utilisation pedagogically rational?
- Is the material easily accessible?
- Is the course material considered useful by academics outside the university and are they using it in their systems?
- Does the material conform to the institutional norms of format, size, etc., as required by the university?
- Does the material help the student to get good grade in the examination?
- Does the material facilitate and help in effective and active student learning?

This discussion should suggest that programme evaluation is a two tier operation — the first consisting of those elements of evaluation which are a part of the general management process. For example, the rationale behind and the outline of a particular programme/course is looked into by an expert committee; modifications brought in at this stage are again discussed and reviewed with the course writers; in the process of course writing, the materials go through the hands of a language editor, a content editor, a format editor, etc., depending on the need, and at the post production level informal/unsystematic feedback is obtained from students, counsellors and others who matter. All these stages are the components of programme/course evaluation and a part of the management process — the establishment has not to put in any separate resources for this level of programme/course evaluation.

However, the second tier of course evaluation may or may not fall within the general management process. It may have any one or more of the following forms:

i) *Piloting of programmes/courses:* In this scheme the materials are tried out with the first batch of students and revisions brought in subsequently. This is a well known approach and very often the management takes to it almost without giving any thought to its utility in relation to its cost. The following are strong reasons why IGNOU should not depend on this approach for each and every course:

- Given the constraints on the resources (financial as well as human), it will not be possible for all courses to follow this approach and then offer an improved course to the second batch of students.
- It is also not within educational ethics to use the first batch of students as guinea pigs always and ever.
- Correspondingly, the credibility of a course will be in question, should different batches of students get different kinds of courses and yet assessed by the same examination procedures.
- Operationally, bringing in revisions immediately after the first launch adversely affects distribution and support services.
- The cost of course design will be prohibitive and the rate of course production also gets adversely affected by this approach.
- If the first version of each and every course is prepared hastily, the image of the university will be affected adversely in the long run.

It is not, however, suggested that piloting a programme/course is entirely a futile task. Better ways of achieving similar results, nevertheless, should be explored. For example, the concept of ‘quality assurance’ may be inculcated in the production processes.

ii) *Routine evaluation:* This is best done as a part of the management process. The advantage of keeping this evaluation within the management process is that in doing so it will not demand high resource allocations. Besides, it will keep the institution in constant touch with the students and the materials, resulting in appropriate and timely feedback to the academic schools, course writers and academic counsellors. It will eventually feed into the process of course maintenance which can be carried out through supplementary materials and thereby saving the resources. Routine evaluation, thus, implies

- evaluating each course within the year it is launched; and
- evaluating assignment question and the term end examination question papers, in addition to course materials.

- **Evaluation of non-print media materials**

A few learning objectives, especially those concerned with acquiring skills and/or attitudes can not be realized by print materials alone. The use of electronic media within the total media mix of a programme is useful to achieve objectives pertaining to psychomotor affective domains.

To determine the worth of non print materials, we need to go for the description of its attributes. The attributes are the opinions and beliefs, needs, results and outcomes. The evaluator/teacher has to describe the physical conditions for conducting evaluation and the psychological attributes associated with a programme. For example, a video presentation of a theme related to specific unit may be described as 20-25 minutes long, in 16 mm format, with high quality colour pictures and sound. It may be described also as having four episodes at its conclusion in which learners are asked to identify *interpersonal skills* that the film purports to teach. The description might document that most counselors using the film think that it is effective, and that the skills it develops are important. It might indicate that 70 per cent of those viewing the film are able to identify at least three of the four skills demonstrated in the concluding episodes.

To describe the video cassette, assessments are made, some of these are *quantitative*, such as recording the time and percentage of viewers who can identify a given number of skills, etc., while others are qualitative such as collecting statements about effectiveness or monitoring individual perceptions, of picture and sound quality. This may be considered as example of how a non print media is evaluated adequately with the help of the descriptions of the attributes.

- **Evaluation of student support services**

Evaluation of student support services includes the following:

i) Tutor evaluation: Tutors/academic counsellors are essentially engaged in the activities of informing, advising and counselling students both in the face to face contexts and at a distance by post; meeting students in face to face situations for a number of sessions during an academic year; assessing assignment responses and helping distance learners to learn how they should learn on their own. These activities are the bases of evaluating the tutors. This evaluation, thus, consists of evaluating

- the nature of communication between them and the students — through letters and/or in face to face situations,
- the quality of their comments on assignment responses,
- the reliability and validity of the assessment of the responses,
- the turn around rate of assessed assignment responses.

ii) Evaluation of face to face sessions: As far as face to face sessions are concerned, IGNOU evaluates the academic counsellors' punctuality, regularity, nature of rapport with students, quality of the conduct of sessions, students' satisfaction, academic counsellors' motivation/enthusiasm, etc., through its regional centers.

iii) Evaluation of continuous assessment: This entails evaluating the validity and reliability of continuous assessment. This in effect means evaluation of assignments for their validity, their reliability and for their correlation with the questions set at the end of the term examination. A certain percentage of the evaluated assignment responses are further monitored by the concerned school faculty (sometimes, the regional center senior functionaries), every year.

iv) Evaluation of support system: Evaluation of the support system itself consists in evaluating the quality and quantity of support available to students in relation to various courses at various study centres. Such evaluation will have implications for recruitment of academic counsellors, providing support in terms of equipment and other support to counter issues relating to local and socio geographical constraints, etc. In general, to evaluate student support services, IGNOU seeks answers to a few questions of the kind given here:

- Do these services cater to the information needs promptly, adequately and convincingly?
- Do these services provide for advice at pre course, on course and post course stages adequately and effectively?
- Do these services provide for academic counselling adequately and conveniently? This includes the effectiveness of tutors with regard to their work on assignment responses and face to face situations.
- Are these services easily accessible?
- Are the staff involved in these services attitudinally tuned to the kind of work assigned to them?
- How do these students rate the support services?

- Do the support services make a pedagogically rational use of various media in operation?
- Do the support services help the learners get a good pass in the examination?

Conclusion

In this section we have presented the perspectives of programme evaluation. We have also attempted to explain the different types of programme evaluation carried out at IGNOU.

- Brown, S. and Knight, S. (1994) *Assessing Learners in Higher Education*, Kogan Page, London.
- Calder, J. (1994) *Programme Evaluation and Quality: A comprehensive guide to setting up an evaluation system*, Kogan Page Ltd. London N19JN.
- Cronback L.J. et.al. (1980) *Towards Reform of Programme Evaluation*. San Francisco: Jossey-Bass
- Freeman, R. and Lewis, R. (1998) *Planning and Implementing Assessment*, Kogan Page, London.
- Gronlund N. (1981) *Measurement and Evaluation in Teaching*, Fourth Edition, Macmillan Publishing Co., INC, New York.
- Holmberg, B. (1985) *Status and trends of distance education*. Lund: Lector Publishing
- Koul, B.N. (1991) *Evaluation at Indira Gandhi National Open University*, New Delhi. In R. Schuemer (ed.) *Evaluation Concepts and Practices in Selected Distance Education Institutions*. ZIFF, Germany.
- Morgan, C. and O'Reilly, M. (1999) *Assessing Open and Distance Learners*, Kogan Page Ltd., London.
- Nightingale, P. et.al. (1996) *Assessing Learning in Universities*, Professional Development Centre, University of New South Wales, Sydney.
- Panda, S. (1991) 'Programme Evaluation in Distance Education: A perspective and proposed agenda of actions', In *Educational Technology, Third Year Book Vol. 1*, AIAET, New Delhi, pp 168-181.
- Rowntree, D. (1997) *Assessing Students: How shall we know them?*, Kogan Page Ltd., London.
- Thorpe, M. (1993) *Evaluating open and distance learning* (Revised Ed.) Longman: U.K.
- Wattawa, H. & Thierau, H. (1990): *Lehrbuch Evaluation*. Bern: Huber

Web site sources

<http://www.isoc.org/oti/articles/0500/olsen.html>