Module 1

Lesson 2:

Evaluation and Assessment

Introduction

Our second learning unit covers the definition, importance, strategies and tools of Evaluation and Assessment in TE/TB learning. It is designed to help practitioners understand the basics of program evaluation and assessment in Technologically Enhanced/ Technological Based (TE/TB) learning environments. In this lesson, we will start with the definition of program evaluation, the points to take into account during evaluating a program and different models of program evaluation. Then we will continue with definition, purpose, strategies and tools of assessment.

Definition, Importance, Purpose of Program / Course Evaluation

In general evaluation is the systematic investigation of the worth or merit of an object. Based on this definition evaluation of an instruction program encompasses systematic investigation of the worth of an ongoing education program thereby identifying its strengths, weaknesses, benefits and drawbacks. Evaluation suggests the attribution of significance or quality of the current status of a particular object and it implies a judgement regarding the information that assessment activities provide.

The purpose of evaluation is to gather information in order to assist in decision making and to as well as assess the value of the education and the degree to which the intended objectives have been achieved. For judging a program's effectiveness, information about the program is collected, analyzed and interpreted as a part of evaluation process. Simply put evaluation is a process whereby information is carefully and systematically collected about an institution, person, activity a program with a view to making an informed decision. Based on the definitions it becomes clear that evaluation is important for making informed decisions on improving education programs. In order to assist informed decision making process, a program evaluation must be methodologically systematic, addressing

questions that provide information about the quality of a program, contribute to a recognized level of value.

In the literature, several criteria are mentioned for determining the effectiveness of an education program. Among these criteria, most common are the following:

- How students are learning from the program
- The difficulties the students encounter at different stages of education program (both administrative and academic).
- The adequacy and appropriateness of the support services
- The need for changing program assessment
- The cost of running the program
- Whether the program objectives have been met and the goals achieved

Types and Purposes of Program Evaluation

For describing the purpose and nature of program evaluation it is needed to refer to three main types of program evaluation for their unique focus and rationale: formative program evaluation, summative program evaluation and developmental program evaluation.

The objective of formative program evaluation is improvement by providing constructive feedback to program implementers and clients. Formative evaluation informs program implementers about ways to improve program quality and services. Formative evaluation design includes implementation evaluation, process studies and evaluability assessment.

Summative program evaluation's purpose is to measure program performance in terms of outcomes and impacts during ongoing operation or after the program completion. It determines the overall quality of the program. It asks whether the program was worth what it cost in terms of time, money and other resources. It informs the implementers whether the program is successful.

Developmental program evaluation aims at informing social program developers who intend to introduce a major change to program models. It intends to facilitate innovation.

Recapping the three types of evaluation, the formative program evaluation gather information for improving the existing programs; summative program gathers information for reporting on a program's merits or its worth to participants; the developmental program evaluation relates to the continuous gathering of information for making adaptations in the course of developing a program.

Course\Program Evaluation Models

Two alternative strategies have been emerged for program evaluation. The first is traditional, positivist approach. The second one represents an eclectic view of evaluation that includes both qualitative and quantitative techniques for evaluation.

In the traditional/positivist strategy, the education program is structured from the start with the requirements of the evaluator. In this quantitative method, after carefully picking the samples and establishing the controls variables are selected for collecting the comparison data.

As opposed to this quantitative procedures for evaluation activities use focus groups, interviews, observations to understand opinions and perspectives of different stakeholders participating in the learning process. Evaluation procedures are of critical interest to trainers and teachers for the purpose of improving education programs and determining the worth of education activities. Within this context three different evaluation approaches will be elaborated the first one is Kirkpatrick's evaluation approach, the second one is the evaluation approach developed by Woodley and Kirkwood and the third one is AEIOU developed by Fortune and Keith.

Kirkpatrick's five level evaluation approach has been traditionally used to evaluate classroom training and teaching:

1. Reactions: This level measures how participants feel about the educational activity
2. Learning attempts to determine the extent to which learners have advanced in skills, knowledge or attitude
3. Transfer this level attempts to determine if the skills, knowledge and attitudes learned are being transferred to the workplace.
4. Results this level attempts to measure the success of the training program in tersm of increased productivity, improved quality, lower costs etc.

Woodley and Kirkwood's model proposes six categories of evaluation information:

1. *Measures of activity* counts the number of events, people and objects.
2. *Measures of efficiency* is concerned with the number of successful students, average workload of students and number of students enrolled in additional courses.
3. *Measures of outcomes* most importantly includes measures of adequate learning. Interview with learners are used for measures of outcomes in addition to course grades.
4. *Measures of program aims* establishes the extent to which the aims of the teaching program are met.
5. *Measures of policy* takes the form of market research. This category includes surveys with prospective students, employers and monitoring current students.
6. *Measures of organizations* contributes the efficiency of an organization by monitoring the process of course development and delivery. This category requires on-site visits and interviews.

Another approach for program evaluation is AEIOU developed by Fortune and Keith (1992) also brings quantitative and qualitative methodologies together in that sense it is eclectic. AEIOU model provides key questions necessary for effective evaluation and consists of five main components:

1. *Accountability* aims at determining if the project's objectives and activities are completed.
2. *Effectiveness* attempts to place some value on the project's activities. Grades, achievement testes, attitude inventories as well as course material reviews are used to determine effectiveness.
3. *Impact* focuses on identifying the changes generated by the project activities.
4. *Organizational Context focuses* on identifying contextual or environmental factors that contribute to the program's ability to conduct activities.
5. *Unanticipated Consequences* aims to identify positive or negative unexpected changes as a result of the program.

	Book: Planning and Management in Distance Education, edited by S. Panda
	e-book: Evaluation in Distance Education and e-Learning

Assessment

Definition, Importance, and Purpose of Assessment

As a key element of an instructional design, assessment can be defined as a process where learners' achievement and progress are measured. By measuring learning gains and progress toward a learning goal, assessment also enables improvement of teaching and learning as it provides feedback on the whole process.

As defined above assessment is core to the instructional design process in the sense that it is a natural consequence of the development of learning objectives and sets a precedent for the development and implementation of instructional strategies. In more concrete terms, assessment activities should be matched to what the learners are expected to get out of the instruction. This will be followed by determining the ways to know whether or

not the learners are successful. In this sense assessment becomes the desirable basis for instruction for it measures what is considered important for the learning objectives. To put it simply, the ideal is that the instruction is based on assessment plans in order to ensure that the instruction does not depart from the original goals.

The importance of assessment manifests in its different uses in the instructional design process in TE\TB education. First of all there are several administrative uses of assessment in evaluation and improvement of program, student placement, justification for funding and scholarship. More significant uses of assessment involve learners and instructors. Basically, learner assessment enables the instructor to assign grades at the end of a course or lesson. This use of assessment indirectly helps to determine how to improve the instruction for future students. Furthermore, learners gain a sense of control and take on greater responsibility for their own learning if they get frequent feedback based on different assessment types. Also by using assessments carefully, instructors can identify the problems -if any- in their lesson design.

Another important use of assessment is to decide whether or not a learner is ready to begin an instructional unit especially in a skills-training environment. Especially effective entry-level tests to determine whether or not students or employees lack entry-level abilities save a considerable amount of time, money and energy.
For a good assessment in a TE|TB course the following principles can be followed:

- Design learner-centered assessment that includes self-reflection,
- Design and include grading rubrics,
- Include collaborative assessment,
- Encourage students to develop skills to provide feedback in peer review activities,
- Use assessment techniques that fit the context of the learning objectives,
- Incorporate student input on how assessment should be conducted.

The Difference Between Assessment And Evaluation

Assessment and evaluation are often used interchangeably but as seen by the definitions so far, they have different meanings and applications. Assessment refers to the measurement of progress toward a learning goal. For instance if we want to measure a student performance compared with a desired level of proficiency this is defined by the term assessment. Evaluation, on the other hand, means that an informed judgement will occur based on the information provided by assessment activities. All in all, evaluation involves judgements on the efficiency, effectiveness, worth of an instructional program and assessment results of learners are among the tools used for the evaluation process.

	ebook: Assessing Open and Distance Learners

Assessment Models

In TE/TB education based on their different purposes four main assessment models can be defined. These are termed as diagnostic, continuous, formative and summative assessment.

In diagnostic assessment models, teachers or tutors identify students' prior knowledge at the beginning of each course using various assessment tools such as tests, essays, presentations.

In continuous assessment, teachers manage various ways of assessments over time to allow students engage in multiple tasks and to gather information about their level of knowledge, understanding and ability. There is a continuous delivering of learning activities or demonstration of skills, and continuous marks and/or feedback on knowledge and competencies acquired. This can include: e-portfolios (demonstrates student's progress over time through showcasing of academic essays, multimedia presentations, digital storytelling, multimedia posters, academic reports, etc.), practices or exercises, continuous assessment activities, professional practices, final project degree, etc.

The goal of formative assessment is to improve teachers' teaching and students' learning process by monitoring student learning with providing ongoing feedback. They help students identify their strengths and weaknesses and improve their knowledge, skills and competencies. In formative assessment activities students prepare essays, reports, multimedia presentations, projects, assignments, etc. Teachers provide feedback to students for their progress in a constructively and rich way for enhancing students' activities. Different from continuous assessment in formative assessment models the works of students are not formally graded.

The goal of summative assessment is to evaluate student learning at the end of an education process by comparing it against some standards. The final grade constitutes the average grade for all the tasks, or midterm exams, delivered and/or the final exam grade which can include computer based individual assessments, computer based testing group assessments, midterm and final exams, written assignments such as essays, reports, portfolio discussion forums, group projects, etc.,

In TE\TB education we can talk about two main reasons for assessing learner's achievements. First one is to determine the level of competence of the learner the second one is to compare or rank the learners' abilities. For instance, the instruments designed to assess whether or not a candidate driver has the skill to be licensed as a driver are termed criterion-referenced assessment instruments. The information provided by this assessment can be used whether or not the skills or knowledge of an individual can be licensed, or whether the learner is ready for further study.

The instruments of assessment designed to enable comparison or ranking of individuals are termed as norm-referenced tests. One example of these tests are university entrance exams or Gaduate Record Exam (GRE). These tests are designed to provide scores that enable admission officers in selecting individuals to higher education. With a clear picture of ranking the admission officers maximize their ability to compare students.

Criterion and norm referenced assessment design models differ in the content that is to be assessed and in the items to include on an instrument. Norm-referenced test designer define the scope of the material for the assessment instruments more broadly than do criterion-referenced test designers. Criterion-referenced test designers use precise objectives of the learning task to assess the selection of skills and knowledge to be obtained.

For norm-referenced instruments, the selected items helps to obtain a spread of scores leading to normal distribution or bell curve of scores. So that it becomes possible to discriminate between the performance of individuals and rank them.

For criterion-referenced tests, designers write items that match a goal or objective. The criterion-referenced assessment designer is less concerned with obtaining a wide spread of scores. The scores do not represent a normal distribution and skewed by many individuals performing well. In other words, if the instruction is effective test results will not form a normal distribution on a well designed test.

Assessment Strategies

For assessing learners achievements two main strategies can be followed. First strategy involves traditional methods of face to face assessment in which students are administered with assessment techniques in person in an authorized exam center. This strategy is time and space bound and has room for very limited or no flexibility.

The second strategy involves both traditional and new methods made available by information technologies such as TETB assessment strategies. As in the case of face-to-face education, the students may take online exam appointments or take the exam on a designated time. Employment of new technologies in education in the form of TETB education allows spatial flexibility. The assessment techniques can be conducted at any place as well as at an authorized exam center. Furthermore, the exams can be conducted with or without the presence of a supervisor.

Assessment Tools

For assessing a student's progress, there are different types of tools for administering assessment. We can classify assessment tools under two broad categories, namely traditional paper-pencil tests and alternative assessment. In traditional paper and pencil assessment tests, there are four frequently used types of questions: multiple-choice, true/false, short answer and essay styles.

Multiple-choice tests are objective assessments that are mostly effective in measuring learning if the objective of learning is at a low level of cognitive effort like recalling a previously memorized information. A multiple choice question consists of a stem and three to five alternative choices. One of the options is the correct answer, and the others are known as distracters. While preparing the options one must be cautious to not to cause any ambiguity.

Another type of objective test item is true/false question. In this exam item students are asked to determine which response is appropriate for a given statement.

Short answer test items can be either objective or subjective assessments. In these items students are asked to provide answer consists of a word or a short phrase.

Essay questions are developed to assess high level of learning such as analyzing concepts. This question type enables students to show their capability to present an insightful well-organized response and to show their level of understanding. For this type of question rubric is essential to establish a level of objectivity. Rubric enables the answers to be judged accurately and rating to remain to criterion-referenced. To assist

grading the answers, a rubric should be developed to ensure the rating remains criterion-referenced.

For evaluating the pros and cons of objective and subjective assessment tools for TE/TB educational environment, objective tools like multiple choice test and true-false questions are easy to administer and score electronically. Short answer items can also be easily created and delivered electronically but scoring is more difficult compared to multiple choice tests. The scoring system must include the desired correct answer and possible alternative acceptable answers and the instructors need to pay attention to any misspellings. Besides its advantages of flexibility and testing for high level of learning outcomes essay questions present several disadvantages in te/tb education. The most apparent disadvantage is scoring the essay responses are time consuming and difficult. Also unlike multiple-choice tests, essay can only cover a very limited amount of material. For all the above mentioned reasons, it is suggested to mixing objective items with short answer and essay items for a successful assessment process.

Besides above mentioned exam items, in te/tb education system alternative assessment tools can also be used such as assignments, quizzes, performance based assessments and portfolio assessment.

Assignment is a crucial assessment tool that helps students to deepen in the subject, improve research skills, encourage team work and enables the instructors to assess a students progress throughout the course. Assignments can be designed in the form of projects, problem based cased studies. Assignments can also be delivered in the form of Authentic assessment in which the students are asked to transfer their skills to unfamiliar situations in real world challenges.

Quiz are implemented to assess learning in progress for a portion of the course content. Quizzes enable the lecturers to give feedback and reflection for improving the knowledge of learner or for assessment. In TE/TB education quizzes can easily be published in LMS

after each learning unit. In that way students can be provided with a way of self-assessment.

In Performance-based assessment, learners are expected to perform skill and may include determining what the learner knows about the skills per se.

Portfolio assessment can be defined as a means of collecting and judging examples of student work. One significant aspect of portfolio creation is that student decides which works to include in the portfolio. Portfolio assessment enables a level of self-reflection, a process through which the student determines the criteria to use in selecting the materials. In other words, as a student identifying what represents her/his best work constitutes a level of self-assessment.

The traditional types of tests are successfully administered in TE/TB education settings. Online tools have brought new opportunities and enhanced features to web-based assessment within the environment of TE/TB education. Testing software offer anytime, anyplace access to quizzes, tests and exercises. In test design and use online assessment resembles traditional types of tests the advantages of online testing lies in its logistical nature. Online testing packages may randomly change the order of the test items, can categorize the test items according to learning objectives. Security features may include browser lock-down to prevent students from accessing other computer resources during the test. Automated scoring can be time and labour saving. However there are disadvantages of online assessment especially in terms of test security. Proctoring is required to ensure that the test taker is who he/she claims to be and additional resources are not being used during the assessment process. Also slow and limited internet access may take some students in disadvantaged position in te/tb education system.

Informal learning groups can also be considered as assessment tools which automatically develops in face-to-face instruction. In an informal learning group, a benchmarking system is unofficially developed which provides feedback to students. Such a benchmark

system also serves for enhancing student motivation to keep up with their peers. However in TE/TB education systems it becomes more restrictive to provide the conditions for creation of such group as a model of ongoing assessment. However in te/tb based learning where informal learning group is not formed, different types of assessments may include in-class exercises, holding students responsible for collaborative work, expecting students to participate online. For many teachers in te/tb education nongraded assessment measure can be part of their course and alleviates the strain of grading massive amounts of student work. Spontaneous question and answer activities students can check their own work and participate in the discussion of correct answers.

	Read more on assessment models, tools and strategies in distance learning.

Feedback in TETB Education

Feedback to students is a critical component of the assessment activities in the partner universities. Especially formative feedback constitutes a great majority in the assessment activities to enable progress of the students in the learning activities. Sometimes formative feedback is provided along with grades. Students are given the marks as well as explanation about their learning activity performance and progress. As well as that, qualitative, quantitative, personalized or dialogic feedback are provided depending on the design of the assessment activities. For instance, qualitative and formative feedback is provided on all written assignments and also on exams. Dialogic feedback, which can be defined as a continuous dialogue in an assessment process to support interaction and enhance positive dialogue among students and teachers, is offered especially on online courses and thesis seminars. In some cases, when personalized feedback is not possible, there are other options such as class feedback, facilitation of examples of exemplar products or exercises and rubrics. Feedback could be provided by the teachers, tutors or by the other students. The students are informed about their learning progress after

completing each learning activity (their results are published in the LMS or they are available offline). In some practices, automatic feedback descriptions are given for the assignments. For some courses, activities are discussed with tutors and fellow students which allows feedback to be given and shared learning. Assessments, either formative or summative are the points at which feedback is provided to students on progress. On many courses a self-assessment is required as well.

Self-assessment activities also offer benefits to learners in TE/TB education.
Especially for adult learners who have a tight schedule, self-assessment activities provide the opportunity to prepare assignments and other materials that demonstrate their level of learning in a flexible timetable.

Validity and Reliability in Assessment Tools

For an effective construction of assessment tools in TE/TB delivery, the tools must meet the conditions of validity and reliability. An assessment instrument is **valid** if it actually measures what it intends to measure. An instrument is valid if:

1. its individual items are consistent with the goals or objectives they claim to assess
2. the items for each objective represent the range of items that are possible to develop for that objective
3. objectives upon which the instrument is based are adequately sampled.

Basically, an individual item is congruent with its goal if the conditions and performance specified in the objective are represented in the item.
To design a valid assessment tool these guidelines can be followed:
Items should be written clearly
The content should directly match the learning outcomes
The required level of thinking should match the learning outcomes
Range of items is wide enough to represent the learning outcomes.

An assessment instrument is **reliable** if it consistently measures what it claims to measure and if we have a high degree of confidence in the scores that it produces. With a reliable

assessment instrument we can confidently say that learners who achieve a high score are competent and learners who score poorly are not. With a reliable assessment instrument, learners achieve basically equivalent score from assessments conducted in different times.

For designing an assessment instrument one should be cautious of a number of factors that can lead to unreliable instrument. For instance, an instrument that cannot be objectively graded may lead to unreliability. In case of essay questions the instructor might grade it differently in different time spans. Objective assessment formats such as multiple choice, true-false and matching are easier to evaluate objectively than constructed responses such as essay questions.

Especially in criterion-referenced assessments it is important to differentiate between competent and incompetent individuals with regards to a particular goal. A test that does not differentiate between a skilled and unskilled learner is not a reliable test. So the assessment designer needs to eliminate the factors that allow incompetent learners to appear competent or competent learners to appear to be incompetent. The reliability of an instrument is also affected by the clarity of the items and the conditions under which the assessment is administered.

Cheating and Plagiarism

In TE/TB education, the use of advanced communication technologies led to new concerns about the security of assessment process. The assessment processes being administered in the virtual environment unintentionally results in a perceived absence of an authority figure. Therefore in TE/TB education new concerns arise about different and new forms of cheating and plagiarism.

To elaborate the issue, it is useful to clarify the difference between cheating and plagiarism. Cheating may take many forms like stolen test questions circulated among students, the use of course materials during the test or tests taken by student other than

the enrolled student. Especially non-monitored online testing it may be quite easy to use course materials during the test. Plagiarism on the other hand is submitting the works of others as their own, may also result from students who don't know how to reference their sources.

TE/TB education programs create an environment that makes it easy for students to cheat and plagiarize. For that reason building awareness of what plagiarism is and how to avoid is the first step of preventive strategies. Additionally ongoing assessment techniques involving students in the design of assessment can also used to prevent cheating and plagiarism. Utilization of plagiarism detection or browser lock-down tools can help avoiding cheating but it must not be forgotten that plagiarism and cheating may be cultural.

References

Braimosh, Dale (2003). Institutional and Programme Evaluation, *Planning and Management in Distance Education*, S. Panda (ed.). Kogan Page: London.

Gunawardena, C. N., LaPointe, D.K. (2003). Planning and Management of Student Assessment, *Planning and Management in Distance Education*, S. Panda (ed.). Kogan Page: London.

Palloff, R. M., Pratt, K. (2009). *Assessing the Online Learner: Resources and Strategies for Faculty.* Jossey-Bass: San Fransisco.

Simonson, M., Smaldino, S., Albright, M., Zvacek, S.(2012). *Teaching and Learning at a Distance: Foundations of Distance Education.* Pearson: Boston.

Smith, P. L., Tillman, J. R. (2005). *Instructional Design.* Jossey-Bass: San Fransisco.

Mackh, B.M. (2018). *Higher Education by Design: Best Practices for Curricular Planning and Instruction.* Routledge: New York.

The Society for Research into Higher Education (2003). *Teaching for Quality Learning at University*. Open University Press: Berkshire.

Secolsky C., Denison, D. B. (eds., 2012). *Handbook on Measurement, Assessment and Evaluation in Higher Education.* Routledge: Oxon.