Lecture 1: Concept of Assessment

Topic no 1: Measurement, Assessment and Evaluation

In this topic, we will discuss about the three main concepts of educational testing and measurement i.e. measurement, assessment and evaluation and the difference about them

At the end of this topic, students will be able to explain

1. What are educational measurement, assessment and evaluation?
2. Difference among educational measurement, assessment and evaluation

1. Measurement

Measurement is the process by which the attributes or dimensions of some object (both physical and abstract) are quantified. The tools used for this purpose may include test, observation, checklist, homework, portfolios, project etc. (the process of changing the person's ability into numbers is measurement). Measurement can be easily understood if we use this word to measure height and distance because these things are physical present in their existence so height can easily be measured by scale. But in the field of education, our variable are not physical and cannot be directly measured e.g. attitude, behavior, and achievement etc. these all are abstract, so there measurement is relatively difficult than those who have physical existence. The tool used for measuring the abstract variables cannot measure exactly like scale (thermometer).

So in this whole course, whenever the measurement word is used it means that tool which will be used for measuring student abilities and then converts it in numerical form.

2. Assessment

It means appraisal of something to improve quality of teaching and learning process for deciding what more can be done to improve the teaching, learning and outcomes.

3. Evaluation

Evaluation is process of making a value judgment against intended learning outcomes and behavior, to decide quality and extent of learning.

Evaluation is always related to your purpose, you aligned your purpose of teaching with what students achieved at the end, with their quality and quantity of learning

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Let's take an example

In classroom situation, as a teacher, when you teach a chapter or unit to a class, first you made the objectives, either you will make it yourself as a teacher or you take it from curriculum document of the particular subject. Objectives are also written at the start of the chapter or book which shows that at the end of the unit what student will be able to do this, that is also referred as Student learning outcomes. So these SLOs can be checked by two aspects i. assessment ii. Evaluation. When you, as a teacher reflect on your teaching on daily basis that what you teach yesterday was a good way to teach or not, what i taught was the students need, did they understand what i teach, and then you can decide changing which things can cause the improvement in the learning process of students. This is ASSESSMENT. In the end of assessment, opinion is not about the individual, it is about the process from which the students or individuals are passing for the betterment of that process.

Evaluation is that when the learning process is complete and you want to see what are your targets or objectives and how much my students achieved those objectives , then tools are made and measures comes from that tell that how much your student learn and what is the quality of their learning.

Difference in measurement, assessment and evaluation

These terms are not different words for same concept but a process serving as prerequisite to each other and having unique purpose.

Every mechanism or process start with the measurement

In the field of education, measurement generate tools i.e. Test, Observation, Quiz, Checklist, Homework, Portfolios. They can give information about student learning. The information we got from the tool is *MEASUREMENT*

Now if that measurement is used for the process of making teaching learning process better than it is *ASSESSMENT*

e.g. if i want to decide that the procedure or way i adopted for student learning i.e. books, activities, field experience is good for their learning and they are according to the needs of student, then it is *ASSESSMENT*

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

And if from these measure i decide that either their learning is getting better or not or how much student learned what i taught them and what is the quality or standard of that learning, then it is *EVALUATION*

Assessment and Evaluation don‘t exist in hierarchy; they both are parallel and different in purpose. Measurement is the source to move towards assessment and evaluation because it provides base and evidence to quantify teaching learning process. The quantified number has no meaning until we do assessment or evaluation. Assessment purpose is to make teaching learning process better so that student learning improve and measurement purpose is to align the learning with purpose.

Topic 2: Classroom Assessment

In this topic students will learn

1. Ways to classify the classroom assessment
2. Types of assessment in each classification

Concept of Classroom Assessment

The process of gathering, recording, interpreting, using and communicating information about a child's progress and achievement during the development of knowledge, concepts, skills and attitudes.

When we are teaching in a classroom, we are doing four things

i. Developing students‘ knowledge

ii. Improving their concepts

iii. Teaching the skills

iv. Making their attitudes.

And while doing so, we collect information about the development of these things, and then we record it, interpret it and use it to communicate about the learning progress of students.

This procedure is called classroom assessment

Classification of Assessment

Assessment can be classified in four ways

1. Nature of Assessment
2. Format of Assessment
3. Use in classroom instruction

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

4. Method of interpreting results

1. Nature of Assessment

i. Maximum Performance Assessment

ii. Typical Performance Assessment

2. Format of Assessment

i. Fixed Choice Assessment

ii. Complex Performance Assessment

3. Use in classroom instruction

i. Placement Assessment

ii. Formative Assessment

iii. Diagnostic Assessment

iv. Summative Assessment

4. Method of interpreting results

i. Norm Referenced Assessment

ii. Criterion Referenced Assessment

Topic 3: Types of Assessment

In this topic, Student will learn

1. Types of assessment by nature
2. Types of assessment by format

1. By Nature of Assessment

i. Maximum Performance Assessment

ii. Typical Performance Assessment

i. Maximum Performance Assessment

Maximum performance assessment determines what individual can do when performing at their best. E.g. assess student in an environment when they exhibit their best performance. Procedure of this type is concerned with how well individual perform when they are motivated to obtain as high score as possible. This type of assessment includes Aptitude tests and Achievement tests.

In Achievement test , student learn by themselves or we teach them and at the end we want to see that how much student learn against our target so in this situation we make a test from which we

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

can determine their best abilities. It is designed to indicate the degree of success in some past learning activity.

Aptitude test, we measure through aptitude test when we want to predict the success in future learning activity e.g. it is used when we want to see the interest of student in a particular field like medicine, sport, teaching. We know that different abilities used for going in different professions, we make a test depending on these abilities and then try to assess, in what abilities the students perform well.

ii. Typical Performance Assessment

The second category is typical performance assessment determines what individual will do under natural conditions. This type of assessment includes attitude, Interest and personality inventories, observational techniques, peer appraisal. Here emphasis is on what students will do rather than what they can do.

2. By format of assessment

i. Fixed Choice Assessment

ii. Complex Performance Assessment

i. Fixed Choice Assessment

Fixed Choice Assessment is sued to measure the skills of people efficiently (means measure more skills in less time) and for this we usually use fixed choice items i.e. Multiple choice question, matching exercise, fill in the blanks and true false. It is called fixed choice because the person who is attempting the paper does not need to write the answer, just need to choice the answer. From these, we can assess student abilities of lower level learning

Fixed Choice Assessment is used for efficient measurement of knowledge and skills. This type of assessment includes standardized multiple choice questions

ii. Complex Performance Assessment

Complex Performance assessment is used for measurement of performance in contexts and the problems valued in their own right. This includes hands on laboratory experiments, projects, essays, oral presentations.

E.g. if want to measure the student ability of writing an essay and this cannot be judged by fixed response items

Topic 4: Use of Assessment in Classroom Instruction

Placement and Diagnostic

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

In this session student will learn Classification of assessment in terms of its uses in classroom instruction

3. Use in classroom instruction

i. Use of Placement Assessment
ii. Use of Diagnostic Assessment
iii. Formative Assessment
iv. Summative assessment

i. Placement Assessment

Placement Assessment determines prerequisite skills, degree of mastery of course goals and mode of learning. Placement assessment is used when we want to assess student‘s prior knowledge so that we can decide what the level of student is. It is associated with student‘s entry level performance to know either student have a sufficient knowledge required for a particular course or not. Through placement assessment, teacher can be able to know that where student should be place according to their present knowledge or skills. It determines the level of student knowledge at the beginning of session and helps teacher plan the lesson accordingly. In the classroom, the teacher can use placement assessment to assess the level of students‘ knowledge and skills and then make lesson plans keeping in mind the level and need of students accordingly.

It also determines the interest and aptitude of student regarding a subject and helps in selecting correct path for future.

Examples

Readiness test: It is a test used to determine the students‘ knowledge or concept about a particular course of instruction or what is the level of students

Aptitude test: It is used for the admission in a particular program

Pretest: It is made according to the course objectives and determines the student present knowledge about them

Self- report inventories: Determines the student level by interviewing or discussion

ii. Diagnostic Assessment

Diagnostic Assessment determines causes (intellectual physical, emotional environmental) of persistent learning difficulties. e.g. if you are having a headache , first you will try to cure it by yourself by taking a medicine and you got a relief but if you didn‘t got a relief by

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

taking a medicine then either you change your medicine or you go to the physician or doctor. At first, doctor prescribed medicines , if you still have headache you again go to the doctor , then the doctor suggest you the tests i.e. blood test, urine test etc. and then finally by seeing the test report the doctors able to recognize the reason or cause of headache. And when doctors know the root of your headache then he will prescribe you the medicine for that cause, this is diagnosis.

Diagnosis doesn‘t start first day, it is for the constant or continuous problems e.g. if a student continuous to experience failure in reading or mathematics or any other subject despite the use of prescribed alternative methods , then a diagnosis is indicated. Teachers‘ try to find out what is the root of students failure.

Topic 5: Use of Assessment in Classroom Instruction – Formative and Summative Assessment

In this session students will learn uses of:

i. Formative Assessment

Ii. Summative Assessment

i. Formative Assessment

Formative Assessment determines learning progress, provides feedback to reinforce learning, and correct learning errors. When we assess student during classroom instruction with a purpose to have a feedback that how can we make our teacher learning process better, that is formative assessment. In this assessment, we are not assessing what students learnt or not rather we assess the process behind the students learning .The process behind the student learning includes teaching method , book, . If we make all these things according to the needs of students then learning will improve.

It is conducted during the academic session or teaching-learning process so that i can get a feedback about my way of teaching and how students are learning and decisions are made on the basis of results immediately. It is and ongoing process to modify teaching strategies on the basis of students need.

It provides feedback to teachers

. About weakness and strength of learning process

. To modify their teaching practices

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

. To improve teacher-learning process

It also helps students to reflect on their weaknesses and encourages them for their successful learning. (When we tell students that their problems is their way of learning rather than their intelligence, then we tell them how to change your way of learning to learn better. With this, students can reflect on their learning process alone, and then what goes wrong with my learning process)

Formative assessment provides feedback to students who are struggling with specific content area or concept.

The main difference between formative and summative assessment is that the in formative assessment improvement is in the process of learning rather than to certify students We use different tools for formative assessment and it includes teacher made tests, custom made tests from textbook publishers, observational techniques

ii. Summative Assessment

Summative assessment comes at the end of the instructional session (course of unit). It is designed to measure the extend of achievement of intended learning outcomes. The primary utility of this type of assessment is to assign grades and certifying the level of mastery and expertise in certain subject. It is usually done through teacher made achievement tests or alternative assessment techniques like portfolio to summarize the overall performance of the student at the end of the session. It is not compulsory to done it at the end of the semester, in semester system there is midterm and final term in one semester, and these both are summative assessment. It usually compares the student learning either with other students learning (norm-referenced) or the standard for a grade level (Criterion-referenced). Summative assessment includes teacher made survey test, performance, rating scales, product scales.

Topic 6: Types of Assessment

Methods of Interpreting Results

4. Method of interpreting result

i. Norm-referenced Assessment

ii. Criterion-referenced Assessment

i. Norm-referenced Assessment

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Norm-referenced Assessment measures the students‘ performance according to relative position in some known group. E.g. ranks tenth in classroom group of 50 or top 5 students in a class. Relative position means a point defined with reference to another position (where a student stands compared to other students)

Rather than reporting student‘s achievement it reports students standing among other students. NRT is utilized to discriminate between a certain groups of students. It‘s never used for certification or issuing grades to students. Position of the student is generally represented by percentile scarce noting the percentage of students achieving same or low scarce in the test. Examples of NRT are NTS or CSS exams. E.g. if you achieved 97% percentile on NTS test, it means there are 96% who scored lower then you.

Norm referenced test includes items with average difficulty and high discriminating power. This provides large spread of scores which makes it easy to declare the relative positions of students. Because the purpose of NRT is not to certify grades, so the test made for it must be of average difficulty means test items should not be very easy or very tough. If all the items are difficult then no student will be able to solve it and we are not able to discriminate who is good students. And if all the items are easy then even the low ability students can solve it then we are not able to discriminate it.

Norm referenced test includes standardized aptitude and achievement tests, teacher-made survey tests, interest inventories, adjustment inventories

ii. Criterion-referenced Assessment

Criterion-referenced assessment describes student performance according to a specific domain of clearly defined learning tasks e.g. adds single-digit while numbers. In this you don‘t compare student performance with other students rather you compare the performance of all students with criteria (in our case that criteria are our learning outcomes). It is most commonly used in schools to report the achievement of learning outcomes against set goals rather than other students. It grades the students to pre-defined criteria and student‘s grades represent their mastery over content. Students with same level of expertise achieve same level of grades. A cut point is determined to distinguish between failed and successful students regardless of score of highest and lowest achiever. It consists of teacher-made tests, custom made tests from the test publishers and observational techniques.

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Lecture 2: Assessment, Testing and National Curriculum

In this unit, students will learn about link between curriculum and assessment. For this purpose, we precede our discussion in reference to National Curriculum of Pakistan 2006.

Topic 7: Role of national curriculum in assessment

In national curriculums of Pakistan, learning of student is classified into four levels.

- Competency
- Standards
- Benchmarks
- Student learning outcome (SLOs)

Competency

It is a key learning area. For example algebra, arithmetic, geometry etc. in mathematics and vocabulary, grammar, composition etc. in English.

Standards

These define the competency by specifying broadly, the knowledge, skills and attitudes that students will acquire, should know and be able to do in a particular key learning area during twelve years of schooling.

Benchmarks

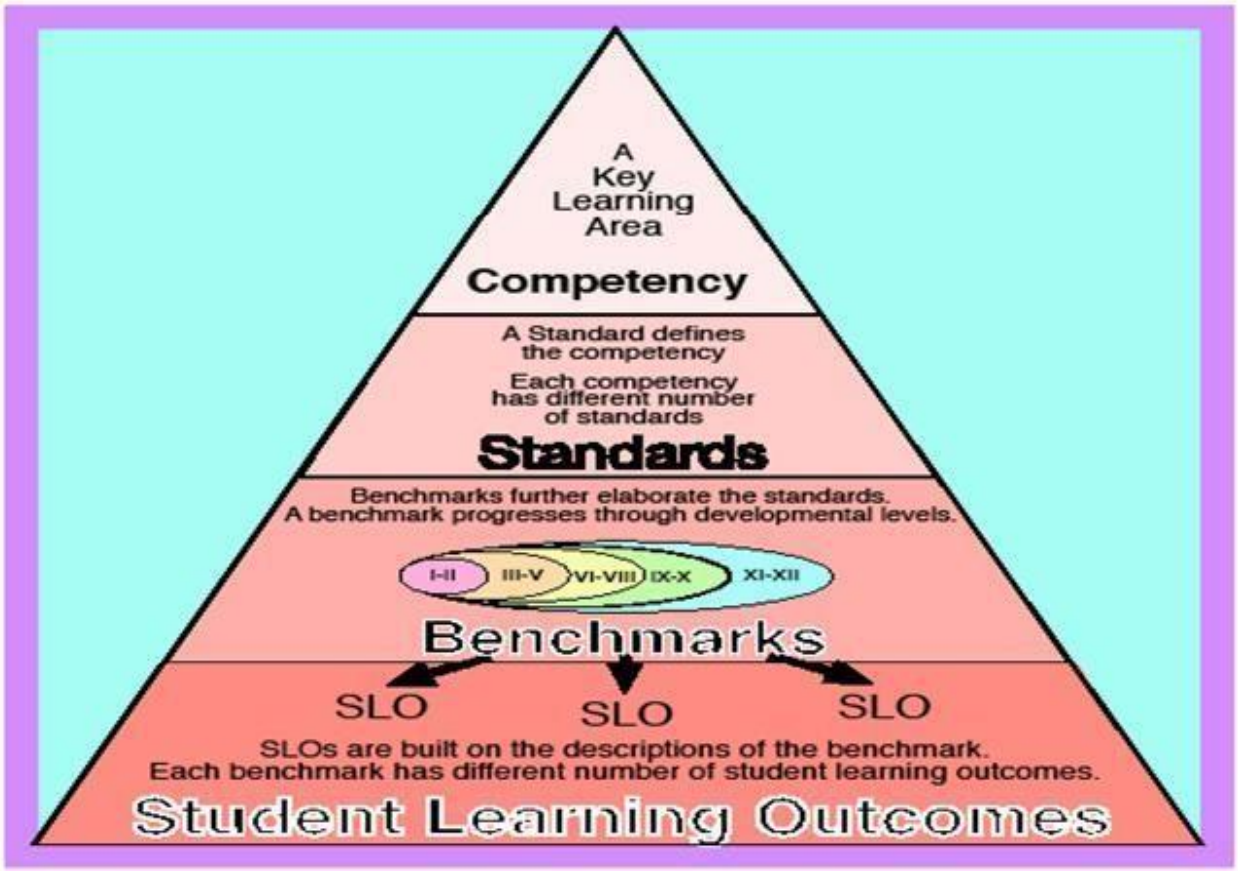
The benchmarks further elaborate the standards, indicating what the students will accomplish at the end of each of the five developmental levels in order to meet the standard.

Student learning outcomes

These are built on the descriptions of the benchmarks and describe what students will accomplish at the end of each grade. It is the lowest level of hierarchy.

Topic 8: Connecting all four levels in curriculum

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3



In the image above, SLOs are at the bottom which is the lowest level. All SLOs combined to make a benchmark and benchmarks convert into standards and then into competency.

Example:

Example is taken from curriculum of English subject.

Competency 1: Reading and thinking skills

Standard 1: All students will search for, discover and understand a variety of text types through tasks which require multiple reading and thinking strategies for comprehension, fluency and enjoyment.

Benchmark 1: use reading readiness strategies.

Student learning outcome:

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

1. Articulate, identify and differentiate between the sounds of individual letters, digraphs and trigraphs in initial and final positions in a word.
2. Identify paragraph as a graphical unit of expansion, ¾ know that words in a sentence join to make sense in relation to each other.

Topic 9: Modes of Assessment in Curriculum

Curriculum document provides specific guidelines for assessment.

Example:

6.1. The two forms of assessment recommended are:

6.1.1. periodic/formative assessment through homework, quizzes, class tests and group discussions.

6.1.2. end of term/ summative assessment through final examination.

Purpose of assessment and curriculum-English 2006

The assessment system for the present curriculum should include

- A clear statement of the specific purpose(s) for which the assessment us being carried out.
- A wide variety of assessment tools and techniques to measure students ability to use language effectively.
- Criteria to be used for determining performance levels for the SLOs for each grade level.
- Procedures for interpretation and use of assessment results to evaluate the learning outcomes.

Form of suitable assessment tools- English 2006

- MCQs
- Constructed response
    - Restricted response
    - Extended response
- Performance tasks

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Lecture 3

Topic 10: Taxonomies of Educational Objectives and Assessment

In this session student will learn:

1. Concept and use of Taxonomies
2. Three popular Taxonomies

Every assessment, regardless of its purposes rests on three important pillars:

1. A model for how students present knowledge and develop competence in the subject domain
2. Tasks or situations that allow the examiner to observe the students‘ performance
3. Inferences from performance evidence about the quality of learning.

In developing a test to assess student learning, taxonomy provides a framework of categories with different hierarchical levels of outcomes

Popular Taxonomies

1. Bloom‘s taxonomy of educational objective
2. Structure of Observed Learning Outcomes (SOLO)
3. Depth of Knowledge (DOK)

Topic 11: Bloom’s Taxonomy and SOLO Taxonomy

In this session students will learn

1. Levels of Structure of Observed Learning Outcomes (SOLO)
2. Levels of Depth of Knowledge (DOK)
3. Levels of Bloom‘s taxonomy of educational objective

Structure of Observed Learning Outcomes (SOLO)

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

The taxonomy of Structure of Observed Learning Outcomes (SOLO) was initially developed by Biggs and Collis in 1982, and then well described in Biggs and Tang in 2007. It carries five different levels of competency of learners

Levels of Structure of Observed Learning Outcomes (SOLO)

1. Pre-structural
2. Uni-structural
3. Multi-structural
4. Relational
5. Extended Abstract

Depth of Knowledge (DOK)

DOK (Depth of Knowledge) was presented by Webb in 1997, giving four levels of learning activities

Levels of Depth of Knowledge (DOK)

1. Recall
2. Skill/Concept
3. Strategic Thinking
4. Extended Thinking

Bloom's Taxonomy of Learning Objectives

Bloom's Taxonomy was presented by Benjamin Bloom in the 1956, consists of a framework with most common objectives of classroom instruction.

Bloom's Taxonomy of Learning Objectives

Those dealing in three different domains and further sub categories in these domains.

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

1. Cognitive
2. Affective
3. Psychomotor

Cognitive Domain

i. Knowledge
ii. Comprehension
iii. Application
iv. Analysis
v. Synthesis
vi. Evaluation

Affective Domain

i. Receiving
ii. Responding
iii. Valuing
iv. Organization
v. Characterization

Psychomotor Domain

i. Perception
ii. Set
iii. Guided Response
iv. Mechanism
v. Complex covert Response
vi. Adaption
vii. Origination

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Topic 12: SOLO Taxonomy

In this topic student will learn

1. Levels of Structure of Observed Learning Outcomes (SOLO).
2. Indicative verbs for each level

Levels of SOLO

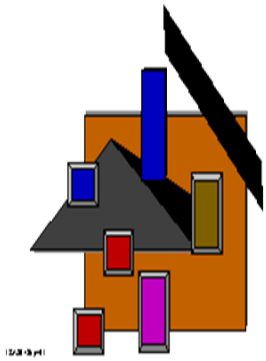
1. Pre-structural
2. Uni-Structural
3. Multi-structural
4. Relational
5. Extended Abstract

1. Pre-structural

Students are simply able to acquire bits of unconnected information and respond to a question in meaningless way. Example of pre-structural level:

Question: What is your name?

Answer: What is your name?

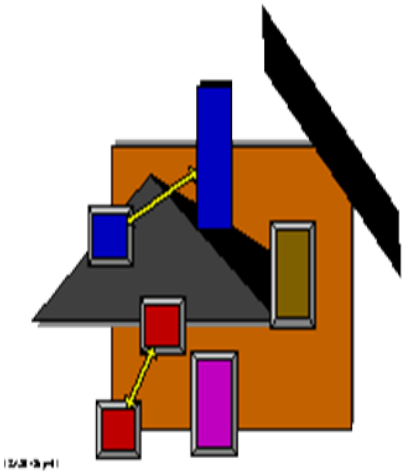


2. Uni Structural

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Student shows concrete understanding of the topic. But at this level is only able to respond one relevant element from the stimuli or item that is provided.

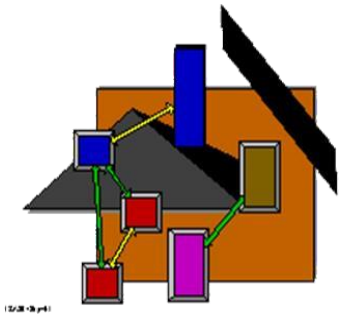
Indicative verbs: identify, memorize, do simple procedure



3. Multi- Structural

Student can understand several components but the understanding of each remains discreet. A number of connections are made but the significance of the whole is not determined. Ideas and concepts around an issue are disorganized and aren't related together.

Indicative verbs: enumerate, classify, describe, list, combine, do algorithms.



4. Relational

Student can indicate connection between facts and theory, action and purpose. Shows understanding of several components which are integrated conceptually showing

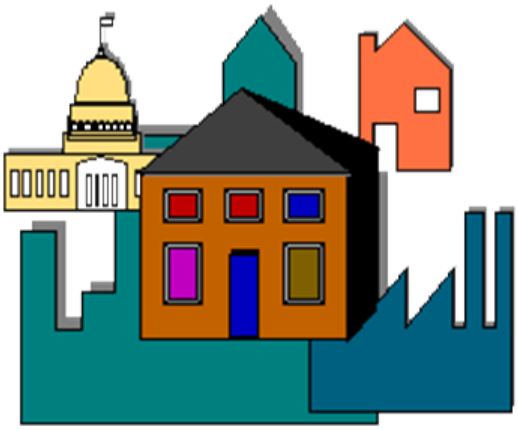
© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

understanding of how the parts contribute to the whole. Indicative verbs: compare/contrast, explain causes, integrate, analyze, relate, and apply.

5. Extended Abstract

Student at this level is able to think hypothetically and can synthesize a material logically. Student make connections not only with in the given subject area but understanding is transferable and generalizable to different areas. Indicative verbs: theorize, generalize, hypothesize, reflect, generate



© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

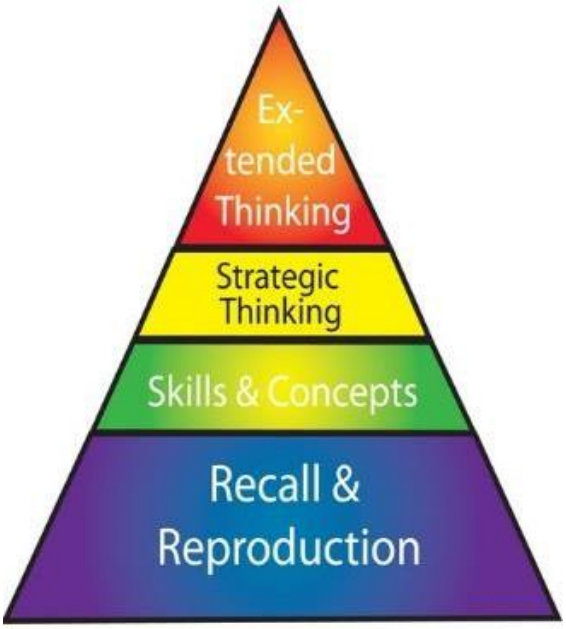
Topic 13: Depth of Knowledge

In this topic, Students will learn

1. Levels of DOK (Depth of Knowledge)
2. Key verbs for each level

Levels of DOK

1. Recall
2. Skill/concept
3. Strategic Thinking
4. Extended Thinking



© 2013 Heidi Butkus www.heidisongs.com

DOK measures the degree to which the knowledge bring about from students on assessments is as complex as what students are expected to know and do as stated in the curriculum.

Levels of DOK (Depth of Knowledge)

Recall

Recall of a fact, information, or procedure. The subject matter at this particular level usually involves working with facts, terms and/or properties of objects.

Key words: list, enlist, name, define etc.

Skill/Concept

It includes the engagement of some mental processing beyond recalling or reproducing a response. Use information or conceptual knowledge, two or more steps, not just recalling

Key words: Graph, separate, relate, contrast, narrate, compare etc.

Strategic Thinking

Items falling in this category demand a short-term use of higher order thinking processes, such as analysis and evaluation, to solve real-world problems with predictable outcomes.

Key words: Argue, critique, formulate

Extended Thinking

Learning outcomes to this level demand extended use of higher order thinking processes such as synthesis, reflection, assessment and adjustment of plans over time.

Key words: Create, Synthesize, Design and reflection.

Topic 14: Bloom's Taxonomy

In this topic student will learn:

1. Bloom's taxonomy of educational objectives.
2. Old and revised taxonomy.

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

3. Key words for each level

Three Domains of Learning

There are three main domains of learning and all teachers should know about them and use them to construct lessons.

- Cognitive Domain
- Affective Domain
- Psychomotor Domain

1. **Cognitive or Thinking Domain**

In 2000-01 revisions to the cognitive taxonomy were spearheaded by one of Bloom's former students, Lorin Anderson, and Bloom's original partner in defining and publishing the cognitive domain, David Krathwohl. One of the major changes that occurred between the old and the newer updated version is that the two highest forms of cognition have been reversed.

Levels of Cognitive Domain

Old Cognitive Domain

Knowledge:

It is defined as the remembering of previously learned material. This may involve the recall of a wide range of facts, procedures principals and generals, the recall of procedures and the processes.

Sample Question: Define the 6 levels of Bloom's taxonomy of the cognitive domain.

Comprehension:

It is defined as the ability to grasp the meaning of the material. individual can make use of the content or idea being communicated without necessarily related it to other content and seeing its fullest implications. Sample Question: explain the purpose of Bloom's taxonomy of the cognitive domain.

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Application:

It refers to the ability to use the previously learned material in new and concrete situations. The abstractions may be in the shape of universal ideas, rules of methods. Sample Question: write an instructional objective for each level of Bloom's taxonomy.

Topic 15: Bloom's Taxonomy (Continued)

In this topic student will learn:

1. Bloom‘s taxonomy of educational objectives.
2. Old and revised taxonomy.
3. Key words for each level

Analysis:

The breakdown of a concept into its constituents parts such that the relative hierarchy of the concept is made easy to understand or the relation between the parts of the concept is elaborated. Sample Question: compare and contrast the cognitive and affective domains.

Synthesis:

There is a collection of the constituents or parts of a concept so as to make a whole. This is a stage when an individual is working with the pieces and assorting them in such a way as to formulate a pattern or structure not clearly there before. Sample Question: Design a classification scheme for writing educational objectives that combines the cognitive, affective, and psychomotor domains.

Evaluation:

It is concerned with the ability to judge the value of the material for a given purpose. Judgments are made on the definite criteria. Sample Question: How far the different BISEs and universities are developing papers using Bloom's taxonomy? Support your answer with arguments.

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Topic 16: Revised version of Bloom's Taxonomy

In this topic student will learn:

1. Old and revised taxonomy.
2. Key words for each level

Revised Cognitive Domains

Levels

Remembering:

Exhibitmemory of previously learned material by recalling facts, terms, basic concepts, and answers.

Key verbs:

Choose, Define, Find, How, Label, List, Match, Name, Omit, Recall, Relate, Select, Show, Spell, Tell, What, When, Where, Which, Who, Why

Understanding:

Constructing meaning from different types of functions be they written or graphic messages, or activities.

Key verbs: Classify, Compare, Contrast, Demonstrate, Explain, Extend, Illustrate, Infer, Interpret, Outline, Relate, Rephrase, Show, Summarize, Translate

Applying:

Solve problems to new situations by applying acquired knowledge, facts, techniques and rules in a different way.

Key verbs:, Apply, Build, Choose, Construct, Develop, Experiment with, Identify, Interview, Make use of, Model, Organize, Plan, Select, Solve, Utilize.

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Analyzing:

Breaking materials or concepts into parts, determining how the parts relate to one another, or how the parts relate to an overall structure or purpose.

Key verbs: Analyze, Assume, Categorize, Classify, Compare, Conclusion, Contrast, Discover, Dissect, Distinguish, Divide, Examine, Function, Inference, Inspect

Evaluating:

Making judgments based on criteria and standards through checking and critiquing.

Key verbs: Agree, Appraise, Assess, Award, Choose, Compare, Conclude, Criteria, Criticize, Decide, Deduct, Defend, Determine, Disprove, Estimate

Creating:

Putting elements together to form a coherent or functional whole; reorganizing elements into a new pattern or structure through generating, planning, or producing.

Key verbs: Adapt, Build, Change, Choose, Combine, Compile, Compose, Construct, Create, Delete, Design, Develop, Discuss, Elaborate, Estimate, Formulate

These categories range from simple to complex and from concrete to abstract level of student‘s learning. It is assumed that the taxonomy represents a cumulative hierarchy, so that mastery of each simpler category is considered as prerequisite to mastery of the next, more complex one.

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Comparison of Bloom, SOLO and DoK

A comparison of hierarchical division in three taxonomies of learning objectives

SOLO Taxonomy	Bloom's Taxonomy	DOK Taxonomy
Level 1: Pre-structural Level 2: Uni-structural	Level 1 Knowledge	Level 1:REcall
Level 3:Multi-structural	Level 2:Comprehension	Level 2: Skill/Concept
Level 4: Relational	Level 3: Application	Level 3 Strategic Thinking
Level 5: Extended Response	Level 4: Analysis Level 5: Synthesis Level 6: Evaluation	Level 4: Extended Thinking

Topic 17: Instructional Objectives

In this session student will learn:

1. Instructional Objectives as Learning Outcome
2. General objectives
3. Specific objectives

Instructional Objectives as Learning Outcome

Instructional goals and objectives are stated in terms of actions to be taken

When viewing instructional objectives in terms of learning outcomes, we are concerned with products rather than process of learning

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Sources for lists of Objectives

- Professional Associations standards
- State Content Standards
- Methods books
- Year books
- Encyclopedia of educational Research
- Curriculum Frameworks
- Test manuals

Criteria of selecting the final list of objectives

Prepare tentative list of instructionally relevant learning outcomes

Review the list for

1. Completeness
2. Appropriateness
3. Soundness
4. Feasibility

General Objectives

- Stating the general objectives is selecting the proper level of generality.
- Objective should be specific enough to provide the direction for instruction but not so specific that instruction is reduced to training
- Stating general objectives in general terms, we provide for the integration of specific facts and skills into complex response

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

- General statements gives teachers freedom in selecting the method and materials of instruction

List of general objectives shows the desired level of generality

- Knows basic terminology
- Understands concepts
- Relates concepts to everyday observations
- Applies principles to new situations
- Interpret graphs
- Demonstrate scientific attitude

Steps for stating General Objectives

1. State each general objective as an intended learning outcome
2. Begin each general objective with a verb
3. State each general objective to include only one general learning outcome
4. State each general objective at the proper level of generality
5. Keep each general objective sufficiently free of course content so it can be used with various units of study
6. Minimize the overlap with other objectives

Specific Learning Outcomes

Each General objective must be defined by a sample of specific learning outcome to clarify how students can demonstrate that they have achieved general objective. Until the general objective are further defined in this manner they will not provide adequate direction for assessment

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Steps for Stating Specific Outcomes

1. List beneath each general objective a representative sample of specific learning outcome that describes terminal performance students are expected to demonstrate
2. Begin each specific learning outcome with an action verb that specifies observable performance
3. Make sure that each specific learning outcome is relevant to the general objective it describes
4. Include enough SLOs to describe adequately the performances of students who have attained the objectives.
5. Keep the SLOs sufficiently free of course content so that the list can be used with various units of study
6. Consult reference materials for the specific components of those complex outcomes that are difficult to define
7. Add a third level of specificity to the list of outcomes, if needed

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Lecture 4: Purpose of Testing

Topic 18: Educational Decisions making

In this session student will learn:

1. Use of testing for decision making
2. Types of educational decisions

Why Test?

In the classroom, decisions are constantly being made. Teachers face huge numbers of dilemmas every day. These decisions can be of following nature

Types of Educational Decisions

- Instructional
- Grading
- Diagnostic
- Selecting
- Placement
- Counseling and Guidance
- Program or Curriculum
- Administrative

These types of decisions are taken at different levels. Some are decided at Board/Administrative level while some are taken at school management level and other are taken in classrooms by teachers

Instructional Decisions

Instructional decisions are the nuts and bolts types of decisions made in classroom by teachers.

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

These are most frequently made decisions. Such decisions include deciding to

- Spend more time on specific units
- Regroup student in class for better management
- Instructional plans

Grading Decisions

Educational decisions based on grades are also made by the classroom teacher but much less frequently than instructional decisions. For most students grading decisions are most influential decision made about them

Diagnostic Decisions

Diagnostic decisions are those made about a student‘s strengths and weaknesses and the reasons behind them. Teachers make diagnostic decisions based on information yielded by an in-formal teacher made test

Decisions of diagnostic nature can also be made by the help of standardized tests (will be discussed in next session)

Selection Decisions

Selection decisions involves test data used in part for accepting or rejecting applicants for admission into a group, program, or institution

Placement Decisions

Placement decisions are made after an individual has been accepted in a program. They involve determining where in program someone is best suited to begin with.

Counseling and Guidance Decisions

Counseling and guidance decisions involve the use of test data to help recommend programs of study that are likely to be appropriate for the students

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Program or curriculum decision

This type of decision is taken at policy level. Where it is decided if a lesion, unit or subject will continue or abandoned for next academic session according to the national objectives of education.

Administrative Decisions

Administrative policy decisions may be made at school, district, state or national level.

Based on measurement data. This includes financial decisions of schools

Topic 19: Types of test

In this session student will learn:

- Types of written tests

How to Measure

In classroom assessment different forms of assessments are utilized. Each form of test has its own benefits and disadvantages. Most common type of assessment used in classrooms is written assessment

Types of Written Tests

- Verbal
- Non-verbal
- Objective
- Subjective
- Teacher Made
- Standardized
- Power

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

- Speed

Verbal

Emphasize reading, writing, or speaking. Most tests in education are verbal tests.

Non-verbal

Does not require reading, writing or speaking ability, tests composed of numerals or drawings is example.

Objective

Refers to scoring of tests when two or more scorers can easily agree on whether the answer is correct or incorrect, the test is objective one. True false, multiple choice and matching tests are example

Subjective

Also refers to scoring. When it is difficult for two scorers to agree on whether an item is correct or incorrect, the test is a subjective one. Essay tests are the example.

Teacher Made

Constructed solely by teacher only to be used in his/her own classroom. This type of test is custom designed according to need and issues related to specific class

Standardized

Test constructed by measurement experts over a period of years. They are designed to measure broad national objectives and have a uniform set of instructions that are adhered to during each administration

Most also have tables of norms, to which a student performance may be compared to determine where the student stands in relation to a national sample of students at same level of age or grade

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Power

Tests with liberal time limits that allow each student to attempt each item. Item tend to be difficult.

Speed

Tests with time limits so strict that no one is expected to complete all items. Items tend to be easy.

Topic 20: Norm- Referenced Assessment (NRT)

In this session student will learn:

1. Characteristics of norm-referenced assessment
2. Questions included in NRT

Why Test?

General purpose of assessment is to gather information to make better and more informed decision. The utility of that information is what differentiate between types of assessments. In earlier session classification of assessment by method of interpreting results was discussed. This session will further unpack the complexity of norm and criterion referenced assessment

NRT

Type of test which tells us where a student stands compared to other students. It helps determining a student‘s place or rank among a group of similar students. Such kind of test is called norm-referenced test

Dimensions

- It provides estimate of ability in a variety of skills in much shorter time.
- NRT tend to be general. It measures variety of skills at same time but fails to measure them thoroughly.
- It‘s hard to make decisions regarding the mastery of student‘s skill in subject.

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

It provides estimate of ability in a variety of skills in much shorter time. NRT are much difficult for students to solve. On average only 50% students are able to get an item right in a test

Topic 21: Criterion Referenced Assessment

In this session student will learn:

1. Characteristics of Criterion-referenced assessment
2. Questions included in CRT

A second type of test tells us about student‘s level of proficiency in or mastery of some skill or set of skills. This is achieved by comparing a student‘s performance to a standard mastery called a criterion. Test that yields such information is called Criterion Referenced Test

Dimensions

- CRT tends to be specific. It measures particular set of skill at one time and focus on level of achievement of that skill. CRT gives clear picture regarding the mastery of student‘s skill in subject.
- It measures skill more thoroughly so naturally it takes more time comparing to NRT in measuring the mastery of said skill
- Items included in CRT are relatively easier. Around 80% of the students are expected to respond item correctly in the test
- CRT compares students‘ performance to standards indicative of mastery.
- Breadth of content sampled is narrow and covers very few objectives

Topic 22: Characteristics of Criterion Referenced Assessment

In this session student will learn:

1. Characteristics of Criterion-referenced assessment
2. Questions included in CRT

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Dimensions

- Sampled content in CRT is much more comprehensive, usually three or more items are used to cover single objective.
- The meaning of the score does not depend upon on comparison with other scores.
- It flows directly from the connection between the items and the criterion.
- Items are chosen to reflect the criterion behavior. Emphasis is placed upon the domain of relevant responses.
- Number succeeding or failing or range of acceptable performance used.
- Example: 90% proficiency achieved, or 80% class reached 90% proficiency.

Topic 23: Difference between NRT and CRT

In this session student will learn:

1. Factors differentiating items of CRT and NRT

Basis of comparison

- Comparison targets
- Selection of items
- Meaning of success
- Average item difficulty
- Score distribution
- Reported scores

Comparison targets

In CRT, the examinee‘s performance is compared to an external standard of competence. While in NRT, examinee‘s performance is typically compared to that of other examinees.

Selection of items

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Items included in CRT are of specific nature and designed for the student skilled in particular subject. In NRT items are of general knowledge nature. Student should be able to answer it but superficial knowledge is sufficient to respond the item correctly

Meaning of success

In CRT, an examinee is classified as a master or non-master. There is no limit to the number of _pass‘ or _fail‘. In NRT, examinee‘s opportunity for success is relative to the performance of the other individuals who take the test.

Average item difficulty

In CRT, the average item difficulty is fairly high. Examinees are expected to show mastery. In NRT, the average item difficulty is lower. Tests are able to spread out the examinees‘ and provide a reliable ranking.

Score Distributions

In CRT, a plot of the resulting score distribution will show most of the scores clustering near the high end of the score scale. In NRT, broader spread of scores is expected, with a few examinees earning very low or high scores and many earning medium scores.

Reported Scores

In CRT, classification of the examinee as master/non-master or pass/fail. In NRT, percentile ranks or scale scores are frequently used.

Topic 24: Formative Assessment

In this session student will learn:

1. Concept of Formative assessment
2. Types of formative assessment

In earlier session classification of assessment by use in classroom instruction was discussed. This session will further unpack the complexity of norm and criterion referenced assessment

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Formative Assessment

Formative assessment provides feedback and information during the instructional process, while learning is taking place, and while learning is occurring. Formative assessment measures student progress but it can also assess your own progress as an instructor.

Types of Formative Assessment

- Observations during in-class activities; of student‘s non-verbal feedback during lecture.
- Homework exercises as review for exams and class discussions)
- Reflections journals that are reviewed periodically during the semester
- Question and answer sessions, both formal (planned) and informal (spontaneous)
- Conferences between the instructor and student at various points in the semester
- In-class activities where students informally present their results
- Student feedback collected by periodically answering specific question about the instruction and their self-evaluation of performance and progress

Topic 25: Functions of Formative Assessment

In this session student will learn:

1. Functions of Formative assessment.

Functions of Formative Assessment

1. Focus of measurement in formative assessment is predefined segment of instruction.
2. Limited sample of learning tasks are addressed.
3. The difficulty of item varies with each segment of instruction.
4. Formative assessment is conducted periodically during the instructional process.
5. Results of formative assessment are used to improve and direct learning through ongoing feedback.

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Topic 26: Summative Assessment

In this session student will learn:

1. Concept of summative assessment
2. Types of summative assessment

Summative Assessment

Summative assessment takes place after the learning has been completed and provides information and feedback that sums up the teaching and learning process. Typically, no more formal learning is taking place at this stage, other than incidental learning which might take place through the completion of projects and assignments.

Summative assessment is more product-oriented and assesses the final product, whereas formative assessment focuses on the process toward completing the product. Once the project is completed, no further revisions can be made.

If, students are allowed to make revisions, the assessment becomes formative.

Types of Summative Assessment

- Examinations (major, high-stakes exams)
- Final examination (a truly summative assessment)
- Term papers (drafts submitted during the semester would be a formative assessment)
- Projects (project phases submitted at various completion points could be formatively assessed)
- Portfolios (could also be assessed during its development as a formative assessment)
- Performances

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Topic 27: Functions of Summative Assessment

In this session student will learn:

1. Functions of summative assessment.

Functions of Summative Assessment

1. Focus of measurement in summative assessment is on course or unit objectives.
2. Broad sample of all objectives is used in summative assessment.
3. This type of assessment uses wide range of difficulty while selecting items for the test.
4. Summative assessment is done at the end of the unit or the course
5. Most important functions of summative assessment is to assign grade, certification of accomplishment and evaluation of teaching

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Lecture 5: Table of Specification

Topic 28: Table of specification (video 28)

In this session student will learn:

1. Concept of two way table of specification
2. Two way table of specification

Table of specification

One of the tools used by teachers to develop a blueprint for the test is called ―Table of Specification‖ in other words Table of Specification is a technical name for the blue print of test. It is the first formal step to develop a test.

Concept of Table of Specification

- It helps a teacher in allotting the questions to different content areas and Bloom‘s learning categories in a systematic manner.
- The blueprint is meant to insure content validity. Content validity is the most important factor in constructing an achievement test. (will be discussed in later unit)
- A unit test or comprehensive exam is based on several lessons and/or chapters in a book supposedly reflecting a balance between content areas and learning levels (objectives).

Two way Table of Specification

A Table of Specifications consists of a two-way chart or grid relating instructional objectives to the instructional content.

Table of specification performs two important functions

1. Ensures the balance and proper emphasis across all content areas covered by teacher
2. It ensures the inclusion of items at each level of the cognitive domain of Bloom's Taxonomy.

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Learning objectives	Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation	Weightage %age
1							
2							
3							
4							
5							

Topic 29: Concept of table of specification (video 29)

It helps a teacher in allotting the questions to different content areas and Bloom's learning categories in a systematic manner.

Learning objectives	Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation	Weightage %age
1							
2							
3							
4							
5							

Topic 30: Elements and Appropriateness in Table of Specification (video 30)

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

In this session student will learn:

1. Elements in table of specification
2. Appropriateness of Table of Specification.

Elements in table of specification

Carey (1988) listed six major elements that should be attended to in developing a Table of Specifications for a comprehensive end of unit exam:

1. Balance among the goals selected for the exam (weighing objectives)
2. Balance among the levels of learning (higher order and lower order
3. The test format (objective and subjective)
4. The total number of items
5. The number of test items for each goal and level of learning
6. The enabling skills to be selected from each goal framework.

A Table of Specifications incorporating these six elements will result in a "comprehensive posttest that represents each unit and is balanced by goals and levels of learning"

Checklist for appropriateness of table of specification

1. Are the Specifications in harmony with the purpose of the test?
2. Do the specifications indicate the nature and limits of the achievement domain?
3. Do the specifications indicate the types of learning outcomes to be measured?
4. Do the specifications indicate the sample of learning outcomes to be measured?
5. Is the number of test items indicated for the total test and for each subdivision?
6. Are the types of items to be used appropriate for the outcomes to be measured?

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

7. Is the difficulty of the items appropriate for the types of interpretation to be made?
8. Is the distribution of items adequate for the types of interpretation to be made?
9. If sample items are included, do they illustrate the desired attributes.
10. Do the specifications, as a whole, indicate a representative sample of instructionally relevant tasks that fits the used to be made of the results?

Topic 31: Balance among Learning Objectives and their Weight in table of specification

In developing a test blueprint first of all it is necessary to select some learning. Objectives and among this list of learning objectives some objectives are more important in sense that more time of instruction is spent on them while some other are less important in terms of time spent on them in classroom so in developing table of specification balance among these learning objectives is important, for this purpose we need to weigh the learning objectives for calculating their relative weightage in test.

Step 1: Instruction Time

To do the calculations for the instruction time for columns of the table of specifications the teacher must use the following formulas for each objective in the table.

Time in class spent on objective (min) / total time for the instruction being examined (min) Percentage of instruction time

$$\text{Percentage of instruction time} = \frac{time\ spent\ onobjective,content,theme\ (min)}{Total\ time\ fo\ rthe\ instruction\ being\ examined\ (min)}$$

$$\text{Percentage of instruction time} = \frac{250}{1000}$$

Percentage of instruction time = 25%

Step 2: Examine value

Then the instructor should look at the number of test items/score to be allocated to objective/content/theme 1.

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Let us assume total marks of there are 100. Then 25 marks should be allocated to questions related to objective/content/theme 1.

Step 3

Percent of instruction time = Percent of examination value (within ±2 percent, if not, redo test)

25 ±2 = 25 ±2

It can be a bit tricky if the total marks of the test are 50. Then 25% of 50 will be 12.5 marks. Point total of questions for objective / total points * on examination = % of examination value

Topic 32: Balance among the Levels of Learning Objectives in Table of Specification

In this session student will learn:

1. Practical example to develop table of specification

We have learnt to give weightage to the content area in a table of specification. Now we look at an example to develop table of specification practically. Following is the table of specification comprised of topics to be cover in test and their weightage that represent percentage of marks for each topic.

Topics/Level	Knowledge	Comprehension	Application	Marks
Pakistan Movement Time: (100/500)*100 = 20%				
Geography of Pakistan Time: (150/500)*100 = 30%				
Climate Change Time: (150/500)*100 = 20%				

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Industries Time: (50/500)*100 = 10%				
Economy Time: (50/500)*100 = 10%				
Total (Time: 500/Marks: 50)				

Let‘s consider that we have to develop a test of 50 marks according to the above discussed table of specification then distribution of marks for each topic is as under.

Topics/Level	Knowledge	Comprehension	Application	Marks
Pakistan Movement Time: (100/500)*100 = 20%				10 (20%)
Geography of Pakistan Time: (150/500)*100 = 30%				15 (30%)
Climate Change Time: (150/500)*100 = 20%				15 (30%)
Industries Time: (50/500)*100 = 10%				5 (10%)
Economy Time: (50/500)*100 = 10%				5 (10%)
Total (Time: 500/Marks: 50)				50 (100%)

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Then we have to consider the importance of each topic for cognitive level of questions according to Bloom‘s Taxonomy.

Topics/Level	Knowledge	Comprehension	Application	Marks
Pakistan Movement Time: (100/500)*100 = 20%	5 (50%)	2 (20%)	3 (30%)	10 (20%)
Geography of Pakistan Time: (150/500)*100 = 30%	2 (10%)	6 (40%)	7 (50%)	15 (30%)
Climate Change Time: (150/500)*100 = 20%		7 (50%)	8 (50%)	15 (30%)
Industries Time: (50/500)*100 = 10%	1 (10%)	1 (20%)	3 (70%)	5 (10%)
Economy Time: (50/500)*100 = 10%	1 (20%)	1 (20%)	3 (60%)	5 (10%)
Total (Time: 500/Marks: 50)	9 (18%)	17 (34%)	24 (48%)	50 (100%)

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Lecture 6: Selection of Test

Topic 33: Selecting pre-designed

In this session student will learn:

1. Selection of pre-designed test

Selecting Pre-designed test

Published test, supplement and complement informal classroom tests, and aid in many instructional decisions.

Published test are designed and conducted in such a manner that each and every characteristic is pre planned and known.

There are many published tests available for school use. The two most value to the instructional program are:

1. Achievement tests
2. Aptitude tests

There are hundreds of tests available for each type. Selecting the most appropriate one is important task. In some cases published tests are used by teachers. But more frequently these are used by provincial or national testing programs.

In classrooms most used published tests are:

1. Achievement tests
2. Reading test

Published tests commonly used by provincial or national testing programs are:

1. Aptitude tests
2. Readiness tests
3. Placement tests

Topic 34: Standards for selecting appropriate test

In this session student will learn:

1. Standards for selecting appropriate test

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Standards for selecting appropriate test

Test users should select tests that meet the purpose for which they are to be used and that are appropriate for intended population.

1. First define the purpose for testing and the population to be tested and select the test accordingly.
2. Investigate the potentially useful sources of information, in addition to the test scores, to validate the information provided by tests.
3. Read the materials provided by test developers and avoid using tests for which unclear or incomplete information is provided.
4. Become familiar with how and when test was developed and tried out.

Topic 35: Standards for selecting appropriate test (Continue)

In this session student will learn:

1. Standards for selecting appropriate test

Standards for selecting appropriate test

Test users should select tests that meet the purpose for which they are to be used and that are appropriate for intended population.

5. Read independent evaluations of a test and of possible alternative measures.
6. Examine specimen sets, disclosed tests or sample questions directions, answer sheets, manuals and score reports before selecting the tests.
7. Select and use only those tests for which the skills needed to administer the test and interpret scores correctly are available.

Topic 36: Fairness in selecting appropriate tests

In this session student will learn:

1. Fairness in selecting appropriate tests

Fairness in selecting appropriate tests

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

1. Evaluate the procedures used by test developers to avoid potentially insensitive content or language
2. Review the performance of test takers of different races, gender, and ethnic groups when sample of sufficient size are available.
3. Evaluate the extent to which the performance differences may have been caused by inappropriate characteristics of test.
4. Use appropriately modified forms of tests or administration procedures for test takers with handicapping conditions.

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Lecture 7: Characteristics of Good Test

Topic 37: Characteristics of Good Test

In this session student will learn:

2. Characteristics of good test

We discussed different types of assessment or how the results are to be used, all assessments should possess certain characteristics. The most essential of these are:

- Validity
- Reliability
- Usability

Validity

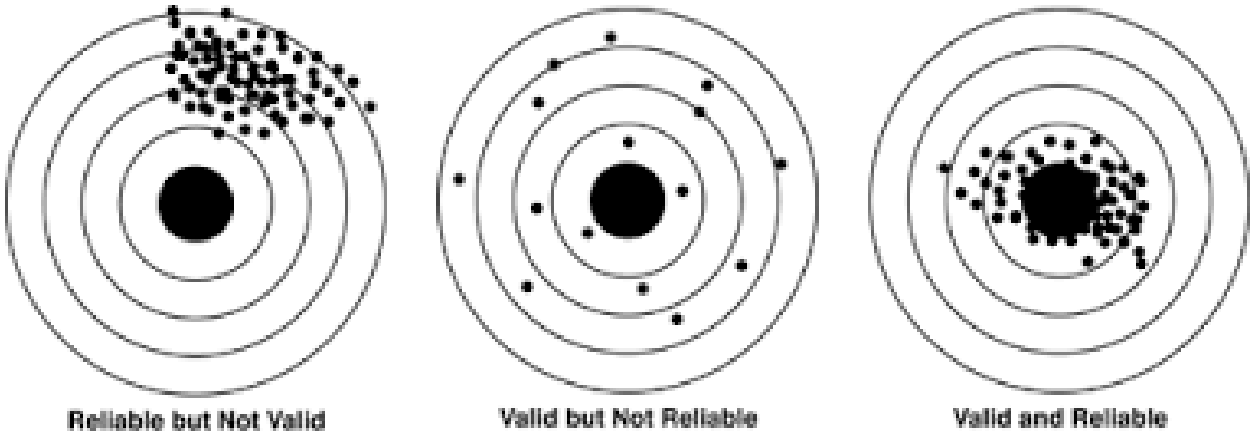
Validity is an evaluation of adequacy and appropriateness of the interpretation and uses of results. It determines if a test is measuring what it intended to measure.

Reliability

Reliability refers to the consistency of assessment results.

Reliability vs Validity

Reliability of measurement is needed to obtain the valid results, but we can have reliability without validity. Reliability is necessity but not sufficient condition for validity.



© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Usability

In addition to validity and reliability, an assessment procedure must meet certain practical requirement which includes feasibility, administration environment and availability of results for decision makers.

Topic 38: Nature of validity

In this session student will learn:

1. Nature of validity.

Following are different points that describe the nature of validity.

1. Appropriateness of the interpretation of the results

Validity is referred as ―validity of test‖ but it is in fact validity of the interpretation and use to be made of the results.

2. Validity is matter of degree

It does not exist on all or none basis. It is best considered in term of categories that specify degree, such as high, moderate or low validity

3. Validity is specific to some particular use or interpretation

No assessment is valid for all purposes. An arithmetic test may have high degree of validity for computational skill and low degree for arithmetical reasoning

4. Validity is a unitary concept

Validity does not have different types. It is viewed as a unitary concept based on different kind of evidences

5. Validity involves a overall evaluative judgment

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

It requires an evaluation in terms of the consequences of interpretations and uses of assessment results

Topic 39: Evidences of validity: Content validity

In this session student will learn:

1. Evidences of validity
2. Concept of content validity
3. Procedure to find content validity
4. Method of ensuring content validity.

Three Evidences of Validity

Content
Construct
Criterion

Meaning

How well the sample of assessment tasks represents the domain of the tasks to be measured.

Procedure

It compares the assessment tasks to the specifications describing the task domain under consideration

Method

1. Classroom instruction determines which intended learning outcomes (objectives) are to be achieved by students
2. Achievement domain specifies and delimits a set of instructionally relevant learning tasks to be measured by an assessment
3. Instructional and assessment priorities specifies the relative importance of learning objectives to be assessed (Table of specification

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Topic 40: Evidences of validity: Construct validity

In this session student will learn:

1. Concept of construct validity
2. Procedure to find construct validity
3. Methods to ensure construct validity

Meaning

How well a test measures up to its claims. A test designed to measure *depression* must only measure that particular construct, not closely related ideals such as *anxiety* or *stress.*

Procedure

Develop a test framework;

1. Defining construct,
2. Identifying sub-constructs,
3. Listing indicators of each sub-construct
4. Writing test items for each indicator

Construct: Essay Writing

Sub-construct	**Meaning/Scope**	**Indicators**
Introduction Paragraph	It introduces the main idea, captures the interest of reader and tells why topic is important.	1. Single sentence called the thesis statement is written 2. Background information about your topic provided 3. Definitions of important terms written

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Supporting Paragraphs	Supporting paragraphs make up the main body of your essay	1. List the points about main idea of essay. 2. Write separate paragraph for each supporting point. 3. Develop each supporting point with facts, details, and examples.
Summary Paragraph	Concluding paragraph comes after you have finished developing your ideas.	1. Restate the strongest points of 2. Restate the main idea 3. Give your personal opinion or suggest a plan for action.

Method

There are two methods to confirm construct validity of a test.

1. Expert judgment
   There are experts of the field. For above example, people who are expert in essay writing will be considered to assess the construct validity of the table and table will be revised under their guidance.

2. Factor analysis
   In this, we group the questions by keeping in view the responses of respondents on them.

Topic 41: Evidences of validity: Criterion validity

In this session student will learn:

1. Concept of criterion validity
2. Procedure to find criterion validity
3. Methods to ensure criterion validity

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Meaning

Demonstrates the degree of accuracy of a test by comparing it with another test, measure or procedure which has been demonstrated to be valid.

Concurrent validity

This approach allows one to show the test is valid by comparing it with an already valid test

Predictive

It involves testing a group of subjects for a certain construct, and then comparing them with results obtained at some point in the future

Procedure

Compare assessment results with another measure of performance obtained at a later date (for prediction) or with another measure of performance obtained concurrently (for estimating present status)

Method

The degree of relationship can be described more precisely by statistically correlating the two sets of scores. The resulting *correlation coefficient* provides numerical summary of relationship

Topic 42: Evidences of validity: Consequence validity

In this session student will learn:

1. Concept of consequence validity
2. Procedure to find consequence validity
3. Methods to ensure consequence validity

Meaning

How well use of assessment results accomplishes intend purposes and avoids unintended effects

Procedure

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Evaluate the effects of the use of assessment results on teachers and students. Both, the intended positive effects (e.g., increased learning) and possible unintended negative effects (e.g.,, dropout of school) need to be evaluated

Considerations

- Does the assessment artificially constrain the focus of student‘s study
- Does the assessment encourage or discourage exploration and creative modes of expression

Factors in Test or Assessment Itself

- Unclear directions
- Reading vocabulary and sentence structure too difficult
- Ambiguity
- Inadequate time limits (construct irrelevant variance)
- Overemphasis of easy to access aspects of domain at the expense of important, but hard to access aspects
- Test items inappropriate for the outcomes being measured
- Poorly constructed test items
- Test too short
- Improper arrangement of items
- Identifiable pattern of answers

Topic 43: Nature of reliability

In this session student will learn:

1. Nature of reliability.

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Reliability refers to the consistency of measurement

1. Reliability refers to the results obtained with an assessment instrument and not to the instrument itself.
2. An estimate of reliability always refers to particular type of consistency (stability, equivalence, internal consistency)
3. Reliability is necessary but not sufficient condition for validity.
4. Reliability is primarily statistical (range +1 and -1).

Topic 44: Method of estimating reliability

In this session student will learn:

1. Method of estimating reliability

Characteristics

1. **Stability:**

Consistency over period of time

2. **Equivalence:**

Over different forms of assessment

3. **Internal consistency**:

Within the assessment itself

Determining Reliability by Correlation Methods

In determining reliability, it would be *desirable* to obtain two sets of measures *under identical conditions* and then to compare the results.

The reliability coefficient resulting from each method must be interpreted according to type of consistency being investigated

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Method to Estimate Reliability

- Test-Retest (stability)
- Equivalent Forms (equivalence)
- Test- Retest with Equivalent Forms (stability and equivalence)
- Split Half (Internal Consistency)
- Kuder- Richardson(Internal Consistency)
- Cronbach Alpha (Internal Consistency)
- Inter-rater Reliability (consistency of rating)

Topic 45: Method of estimating reliability: Test-retest

In this session student will learn:

1. Test-retest method

Test- Retest Method

- *Test-Retest* method is a measure of stability
- It gives the same test twice to the same group with any time interval between tests, Time interval can range from several minutes to the several years

Test- Retest	
September 25	October 15
Form A	Form A
1. Item a yes	2. Item a yes
2. Item b no	2. Item b no
3. Item c yes	3. Item c yes

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Time interval is key point in this type

- Short interval will provide inflated coefficient of reliability
- Very long interval will influence results by instability and actual changes in students over time.

Topic 46: Method of estimating reliability: Equivalent form Method

In this session student will learn:

1. Equivalent Forms method
2. Test- Retest with Equivalent Forms

Equivalent Forms method

- Equivalent Forms method is measure of equivalence
- It gives two forms of the test to the same group in close succession

September 25	September 25
Form A	Form B
1. Item a yes	2. Item d yes
2. Item b yes	2. Item e yes
3. Item c No	3. Item f No

Test- Retest with Equivalent Forms

- Test- Retest with Equivalent Forms is measure of stability and equivalence
- It gives two forms of the test to the same group with increased interval between forms

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Equivalent Forms method	
September 25	September 25
Form A	Form B
1. Item a	2. Item a
2. Item b	2. Item b
3. Item c	3. Item c
Score = 82	Score= 78
Test- Retest with Equivalent Forms	
September 25	October 15
Form A	Form B
1. Item a	2. Item a
2. Item b	2. Item b
3. Item c	3. Item c
Score = 82	Score= 74

Topic 47: Method of estimating reliability: Split Half Method

In this session student will learn:

1. Split Half Method

Split Half Method

- Split Half Method is measure of internal consistency
- It gives test once. Score two equivalent halves of test, correct correlation between halves to fit whole test by spearman –brown formula

Split Half Method		
Sum number of odd items correct	Sum number of even items correct	September 25
Item 1	Item 2	1. Item 1

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Item 3	Item 4	2. Item 2
Item 5	Item 6	3. Item 3
.	.	4. Item 4
.	.	5. Item 5
.	.	6. Item 6
.	.	
Odd	Even	
Score = 40	Score= 42	Total Score=82

Split Half Reliabilities tend to be higher than equivalent form reliabilities because split half method is based on the administration of single assessment

Topic 48: Method of estimating reliability: Kuder-Richardson Method

In this session student will learn:

1. Kuder- Richardson methods and Coefficient Alpha
2. Inter-Rater Method

Kuder- Richardson methods and Coefficient Alpha

- Kuder- Richardson methods and Coefficient Alpha is also measure of internal consistency.
- It gives test once. Score total test and apply Kuder- Richardson

As with the split half method, these formulas provide an index of internal consistency but do not require splitting the assessment in half for scoring purposes

One formula KR20 is applicable only when student responses are scored dichotomously (0 or 1). It is most useful with traditional test items scored correct or incorrect

The generalization of KR20 for assessments that have more than dichotomous, right-wrong scores is called Coefficient Alpha

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Inter-Rater Method

- Inter-Rater Method is measure of consistency of ratings
- It gives a set of students responses requiring judgmental scoring to two or more raters and have them independently score the responses

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Lecture 8: Alternate Assessment Tools

Topic 49: Anecdotal records

In this session student will learn:

1. Types of assessment tools
2. Anecdotal Record

Many outcomes in the cognitive domain, such as those pertaining to knowledge, understanding, and thinking skills, can be measured by paper pencil tests. But there are still many learning outcomes that require informal observation of natural interactions.

Types of assessment tools

Learning outcomes aspects of development can generally be assessed by

1. Observing students as they perform and describing or judging that behaviors (Anecdotal record).
2. Asking their peers about them and assessing social relationships (Peer appraisal).
3. Questioning them directly and assessing expressed interests (Self-appraisal).
4. Measuring progress by recorded work (portfolio).

Anecdotal records

Impressions gained through observation are apt to provide an incomplete and biased picture, however unless we keep an accurate record of our observations. Method to do so is called anecdotal records.

Anecdotal records are factual descriptions of meaning incidents and events that the teacher observes.

Topic 50: Effective use of Anecdotal Records

In this session student will learn:

1. Effective use of Anecdotal Record

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Effective use of Anecdotal Records

One should keep in mind the following points to use anecdotal records effectively.

1. Determine in advance what to observe but be alert of unusual behavior.
2. Analyze observational records for possible sources of bias.
3. Observe and record enough of the situation to make behavior meaningful.
4. Make record of the incident as soon after the observation is possible.
5. Limit each anecdote to a brief description of a single incident.
6. Keep the factual description of the incident and your interpretation of it separate.
7. Record both positive and negative behavioral incidents.
8. Collect a number of anecdotes on a student before drawing inferences concerning typical behavior.
9. Obtain practice in writing anecdotal records.

Topic 51: Advantages and limitations of Anecdotal Records

In this session student will learn:

1. Advantages and limitations of Anecdotal Records

Advantages of Anecdotal Records

Following are the advantages of anecdotal records.

1. It depicts actual behaviors in natural situations.
2. Facilitate gathering evidence on the events that are exceptional but significant.
3. Beneficial for students with less communication skills.

Limitations of Anecdotal Records

Following are the advantages of anecdotal records.

1. It takes long time to maintain.
2. Subjective in nature.
3. Anxiety may lead to wrong observation.

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

Topic 52: Peer appraisal

In this session student will learn:

1. Peer appraisal

Peer appraisal

In this procedure students rate their peers on the same rating device used by their teacher. It depends on greatly simplified procedures.

Techniques of peer appraisal

There are two widely used techniques in this area are:

1. Guess who technique
2. Sociometric technique

Guess who technique

In this technique teacher use a positive or negative behavior of a student as a example. Other students from same group try to guess the statement with that characteristic correctly. Generally behaviors used for example are positive in nature to avoid any adverse effect on the student pointed out in the example.

The guess who technique is based on nomination method of obtaining peer ratings and is scored by simply counting the number of mentions each students receive on each description.

Sociometric technique

Sociometric technique is method for assessing the social acceptance of individual students and the social structure group. It is based on students choice of companion for some group situation.

This form was used to measure student‘s acceptance as seating companions, work companions and play companions.

There are few important principles of sociometric choosing:

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

1. The choices should be real choices that are the natural part of classroom activities.
2. The basis for the choice and restriction on the choosing should be made clear.
3. All students should be equally free to participate in the activity or situation.
4. The choice of each student make must be kept confidential.
5. The choices should actually be used to organize or rearrange the group.

Topic 53: Portfolio

In this session student will learn:

1. Portfolio
2. Weakness and strengths of portfolio

Portfolio

Systematic collection of students work into portfolios can serve a variety of instructional and assessment purposes. The value of portfolios depend heavily on the clarity of purpose the guidelines for the inclusion of materials, and the criteria to be used in evaluating portfolio.

Portfolio of students work

A portfolio is collection of student work selected to serve a particular purpose such as documentation of student growth. It is purposeful; collection of pieces of student‘s work.

Key steps in defining and using portfolios

1. Specify purpose.
2. Provide guidelines for selecting portfolios.
3. Define student‘s role in selection and self-evaluation.
4. Specify evaluation criteria.
5. Use portfolios in instruction and communication.

Strengths of portfolios

1. The can be readily integrated with the instruction.
2. Provide opportunity to student‘s to show what they can do.
3. Encourage to become reflective learner.

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

4. Help in setting goal and self-evaluation
5. Help teacher and student to collaborate and reflect on student‘s progress.
6. Effective way to communicate with parents.
7. Provide mechanism for student centered and student-directed conferences with parents.
8. Provide concrete examples of students development and current skills.

Weaknesses of portfolios

1. Can be time consuming to assemble.
2. Hard to use in summative assessment.
3. Difficult to compare results.
4. Very low reliability.

Topic 54: Purpose of Portfolio

In this session student will learn:

1. Purpose of Portfolio
2. Guidelines for portfolios entries.

Purposes of portfolios

Fundamentally two global purposes for creating portfolios of students work: for student‘s assessment and instruction. It can be used to showcase student‘s accomplishment and document the progress.

Instructional purposes:

When primary purpose is instruction, the portfolio might be used as means of:

1. Helping students develop and refine self-evaluation skills.
2. Providing teacher with more reflecting information regarding students‘ progress.
3. Set criteria of excellence between teacher and student.
4. Student directed conferences with parents.
5. Access to student thought process and awareness of standards.

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

6. Teachers‘ student to communicate with different audience.

Assessment purposes:

When emphasis is on assessment it is important to distinguish between formative and summative roles of assessment.

1. It can be used for formative purposes to measure progress.
2. Basis for certifying accomplishment.
3. For system accountability mechanism.

Current accomplishment and progress

When the focus is on accomplishments, portfolios usually are limited to finished work and may cover only a relatively small period of time.

When focus is on demonstrating growth and development the time frame is longer. It will include multiple version of same work over time to measure progress.

Showcase and documentation portfolios

It contains student selected entries. It demonstrate students ability to choose his best work which demonstrates his ability to do a task.

It intended to provide evidence about breath as well as depth of learning. It needs to be more inclusive and not just limited to special strength of student.

Finished and working portfolios

It implies that work is complete for specific audience. A job application portfolio for example. It is finished product for specific audience.

Guidelines for portfolios entries

Guidelines should specify:

1. The uses that will be made of the portfolio.
2. Who will have to do it?

© Copyright Virtual University of Pakistan

Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3

3. What type of work is appropriate to include.
4. What criteria will be used in evaluating the work?
5. Should define timeline for the portfolios.
6. Minimum and maximum numbers of entries.

© Copyright Virtual University of Pakistan
Join VU Group: https://chat.whatsapp.com/BL4ok4jnM9iJiglhAA1SG3