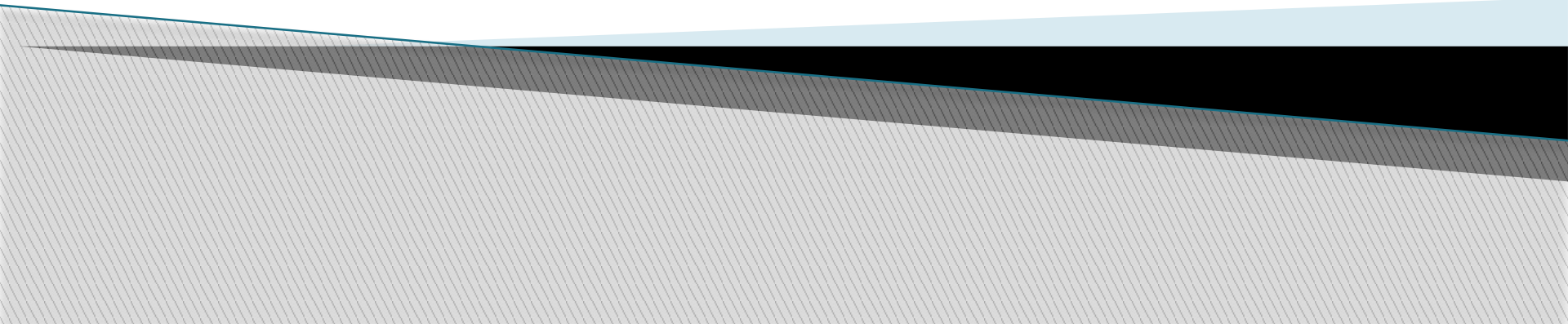


Test construction and its alignment with Instructional objectives

October 30, 2016 E,C,



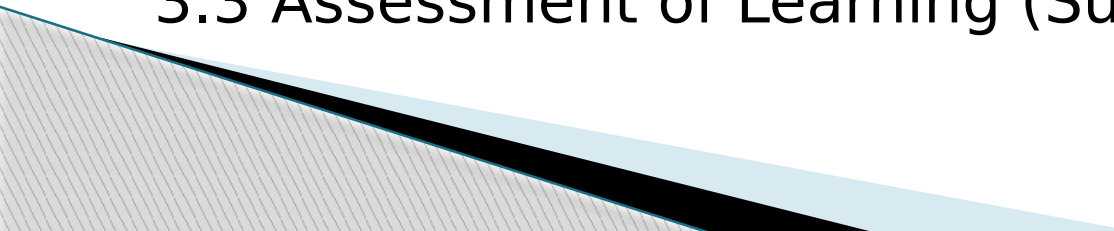
Objectives

- ▶ At the end of this session, the participants are expected to:
- ▶ list the **five basic steps** of teaching
- ▶ show the **place of testing** in teaching
- ▶ Define basic terms in educational objectives and
- ▶ Differentiate the meaning of basic terms in educational testing
- ▶ indicate the alignment of **assessment with learning** objectives
- ▶ Explain the concept of classroom assessment
- ▶ Elaborate the **basic types** of assessment:
- ▶ Explain the basic **Purposes** of assessment
- ▶ Identify the basic **characteristics** of assessment.
- ▶ identify the main strategies of assessment in teaching
- ▶ identify the different **types of test items**
- ▶ prepare **test plan** for the purpose of developing CR test
- ▶ discuss points to be considered in using CATs

Objectives ----

- ▶ Explore the uses, forms, merits, and demerits of the various types of test items
 - ▶ list the specific **rules** to be considered in constructing CAT
 - ▶ construct test items following test preparation rules
 - ▶ **discuss assembling, administering, and scoring of tests**
 - ▶ **Apply the rules in assembling , administering, and scoring test items**
- 

Contents

1. Introduction: The place of testing in teaching
 - 1.1. Basic terms in educational testing : Measurement, Assessment, Evaluation, and tests
 - 1.2. Basic terms related to educational outcomes
 - 1.3. Alignment of assessment to learning objectives
 2. **Assessment: Purposes, importance, Principles, and process**
 3. Types of Assessment
 - 3.1 Assessment for Learning (Formative Assessment)
 - 3.2 Assessment as Learning
 - 3.3 Assessment of Learning (Summative Assessment)
- 

Contents ----

- 4. Assessment strategies
 - 5. Teacher-made Achievement Tests
 - 5.1. Planning tests: uses and steps
 - 5.2. Selecting appropriate types of test items
 - 5.3. constructing various item types basing rules
 - A. General Principles
 - B. Specific rules
 - 6. Assembling test items, Administering and scoring test
 - 6.1. Assembling test items
 - 6.2. Administering test,
 - 6.3. scoring test items
 - 7. Improving quality of test items
- 

Approaches of teaching

- ▶ Brain storming
 - ▶ Group work
 - ▶ Gaped lecturing
 - ▶ Debating
- individual activities
 - discussion
 - Reflection
 - Pyramiding
 - Etc

1. Introduction : The Place of Assessment in the steps Teaching

Activity 1:

1. What is teaching for you? Or Can you list steps for teaching? So, what are they?

2. Where is the place of assessment in teaching?

The 5 basic steps of Teaching

Teaching can be understood as a **five basic steps activity** of teachers

The steps of teaching are:

Step1: Defining **instructional objectives** to be achieved

Step2: Choosing **contents** to be imparted/ preparing the **subject matter** for delivery

Step3: Selecting appropriate **instructional methods or strategies** to be employed

Step4: Transmitting **the main ideas** of the contents. i.e., actual teaching

Step5: Assessing or testing and evaluating the outcome

of teaching or students' learning

Steps - con't

- ▶ The basic steps
 - are **interrelated , overlapping or mutually inclusive** activities of teacher in teaching.
 - can be grouped in to : **before ,during, and after actual teaching** activities.
- ▶ The last step of teaching, i.e., **assessment /evaluation /testing** is the focus of this training session.

1.1. Basic terms in educational testing :

- ▶ There are terms that create confusion in educational testing. The basic terms that create confusion for people including teachers are :
 - Measurement,
 - Assessment,
 - Evaluation, and
 - tests

Activity 2. What do we mean by test, measurement, assessment, and evaluation, and how do the terms differ and relate to one another?

(Answer this question individually. Then discuss your answers with your neighboring friends.)

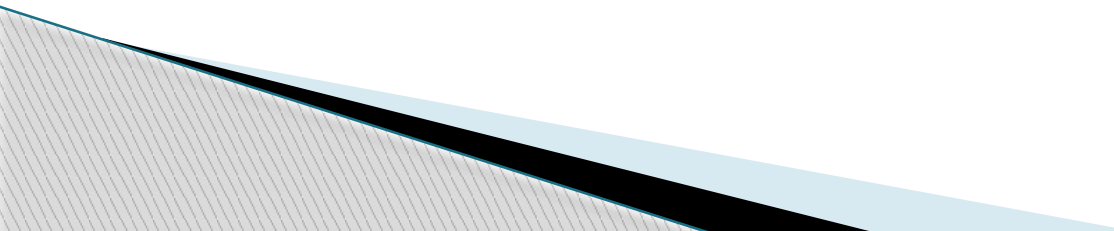
The Meaning of basic terms---

- **Educational Testing**
 - **implies educational assessment**

or measurement or evaluation,
- The terms (*evaluation, assessment, measurement, and test*) are:
 - often used interchangeably as applied to education.
 - different though they are inseparably related one other.
- The distinction and relationship of these terms can be seen as follows.

1. Test

Test :

- systematic **device /tool** useful for measuring **ability, achievement, interest & other traits.**
 - uses, in **education** , to **measure students' achievement of instructional objectives.**
 - uses in the **process of assessing** students' **knowledge** to make judgments or *evaluations*".
- 

Test (Con't)

- ▶ is often used interchangeably with the terms "quiz" and "examination".
But, the terms differ in that:
- ▶ **“Quiz”** - **short and informal questions**
 - is provided in the class hours(as entertainment and test general knowledge).
- ▶ **“Test”** - a **longer and more carefully prepared** series of questions
 - includes **SAT** and **CAT**
- ▶ **“Examination”** - **comprehensive process** of measuring students' performance . e.g. ESLCE.

Quiz

test

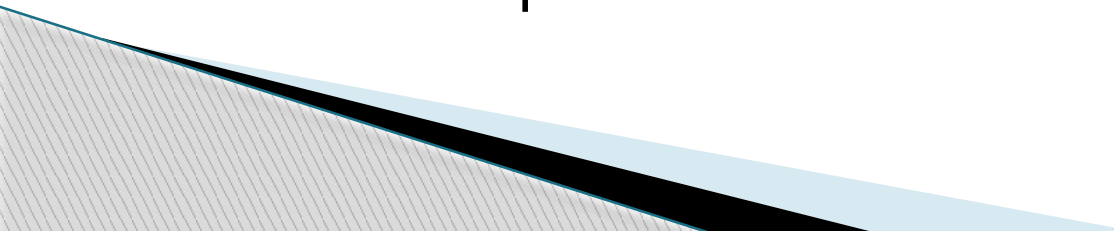
examination

2. Measurement

Measurement

- is the **process of obtaining information** in numbers or **quantitative form** applying **various tools**
- employs **different devices** like **observation, rating scale, tests and any other device.**
- is **the quantitative description** of student's behavior or performance.

E.g. Alemu correctly solved 35 of the 40 problems.



3. Assessment

Assessment

- can be seen in broader and narrow senses.
- is used, in education , to find out **what student's qualities** such as abilities, attitudes, etc. are & **how the qualities are changed** since last assessed to arrive at a judgment of value or worth of the qualities of the person **using not necessarily testing but by asking student's opinion or observing him/her in action** in formal and informal ways.
- ▶ So, **assessment**
- ▶ - refers to the activity of **asking the opinion** of students or **observing them in action** to arrive at a **judgment of the value or the worth of the qualities** of students' work.
- ▶ - is a process of **gathering evidence** of what a student can
- ▶ do to provide **feedback on his/her learning** to encourage
- ▶ their development.

4. Evaluation

Evaluation

- is the broadest of all these terms in meaning
- is **systematic process** of collecting and analyzing data in order to **determine whether and to what degree objectives** have been, or are being, achieved and then **make decisions** about educational programs.
- includes both
 - **quantitative** (measurement) and techniques in the
 - **qualitative** (non-measurement) **descriptions of pupils' behavior**
- plus **Values judgments**.

Examples

- Aster got 50 out of 60 - **quantity** ,
- Aster's work is neat - **quality**
- Aster shows good progress in English - **Value judgment & use**
- is the process of **interpreting evidence** and **making a judgment of performance** to make **informed decisions**, such as assigning a grade or promoting a student to a higher performance level.

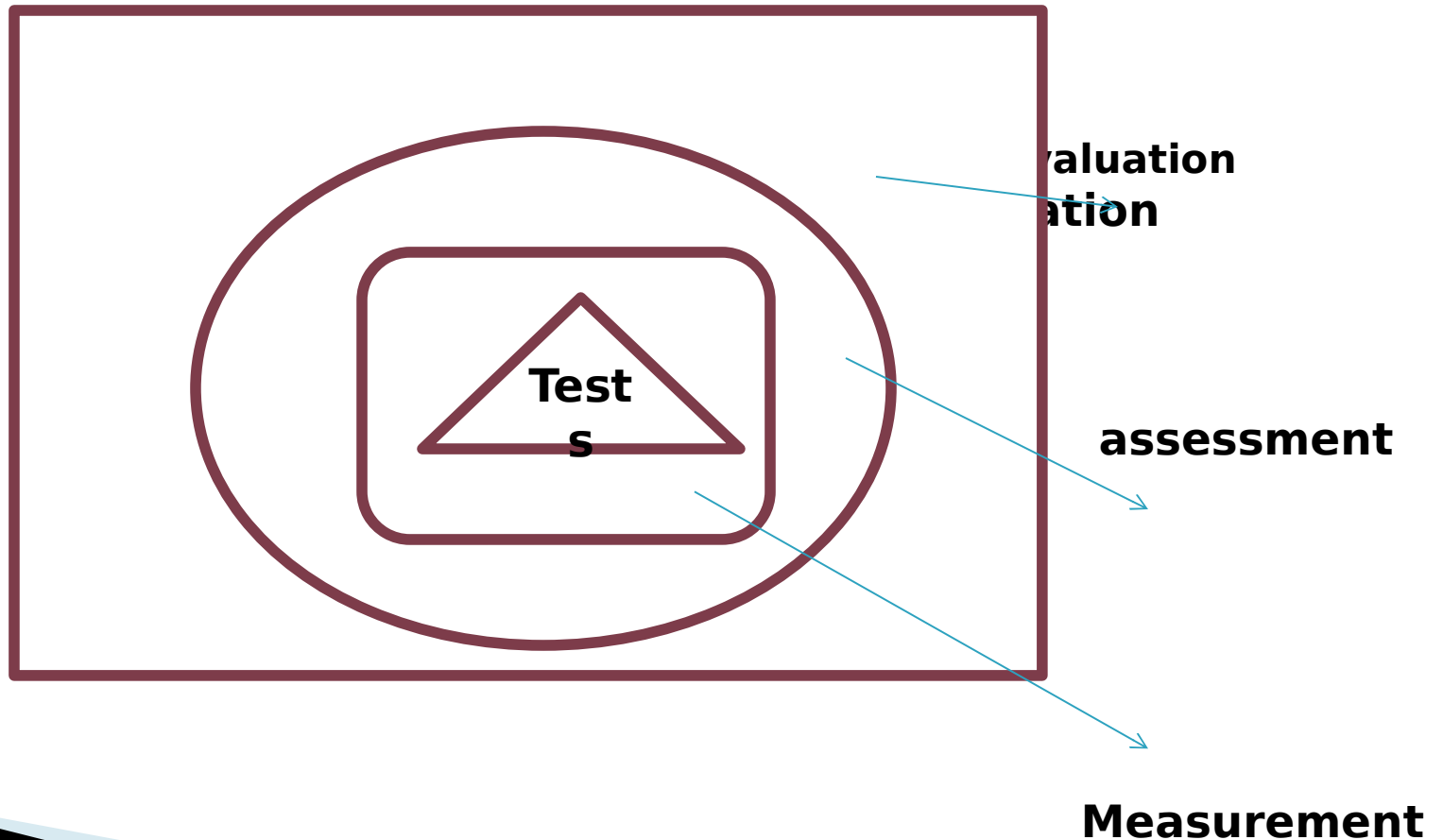
Con't

- These terms (measurement, assessment and evaluation)
 - **show the process of gathering relevant information and then interpreting the information**, for the purpose of **making educational decisions**,
 - are **not instruments** for gathering data or information ,
 - **are not decisions** rather, they provide information that help in making decisions.

Con't

- ▶ To assess/evaluate **behavior/performance** of some body, one has to measure the behavior or performance using a **composite of different types of measuring instruments**/ yardsticks or tools of measurements) such as **tests, assignments, laboratory reports, projects, observation, rating scale**, etc.
- ▶ The **information obtained using such different tools of assessment or measurement or evaluation** (like assignment ,test etc.) can **help us make decisions** with more **accuracy and efficiency**.

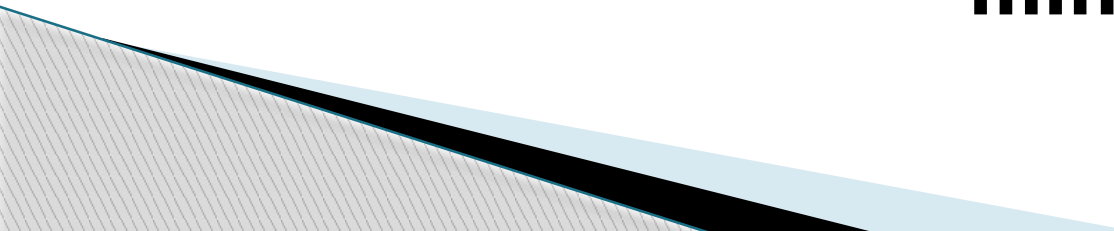
Relationship of tests, measurement, assessment, and evaluation in the educational testing process



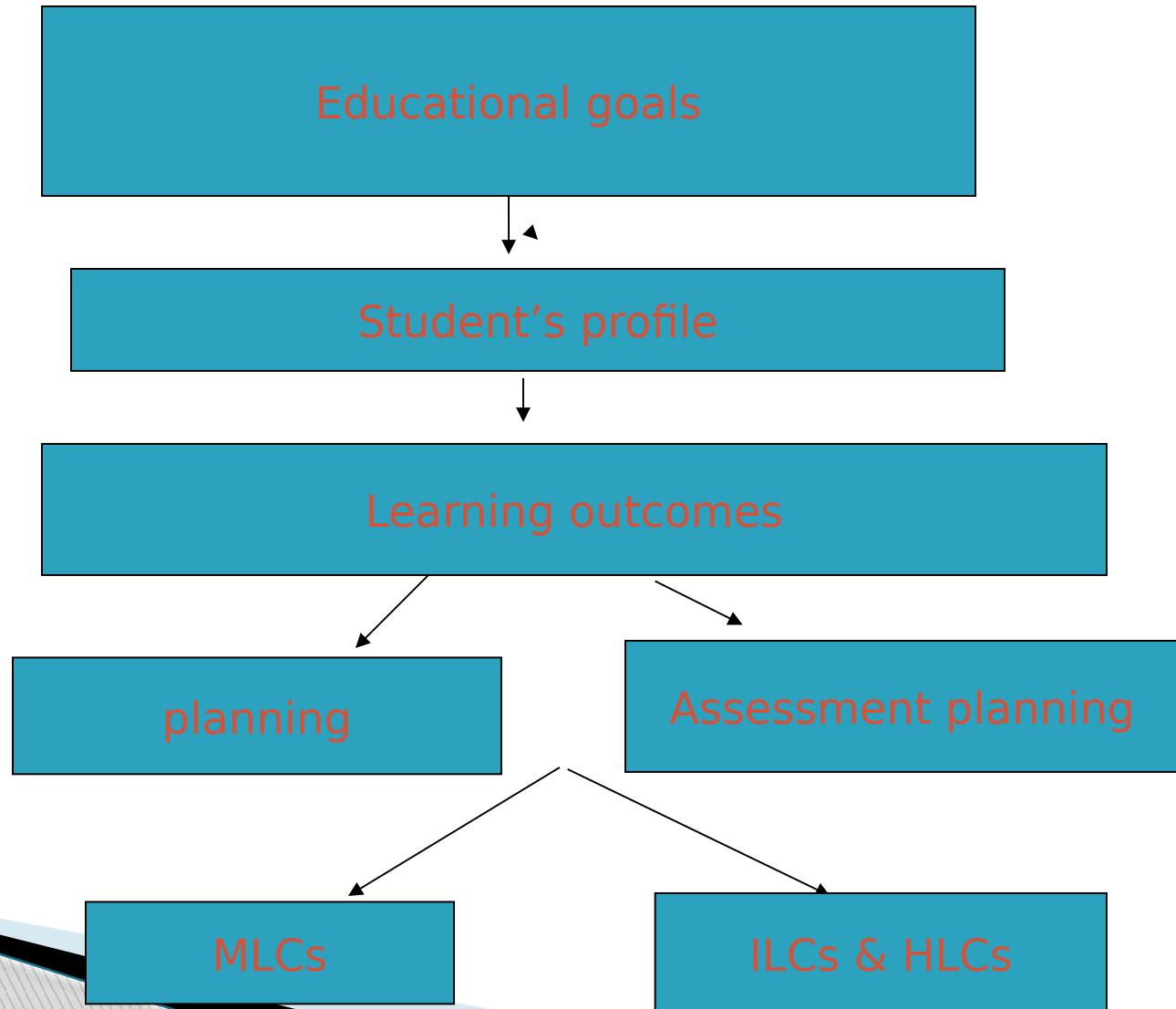
1.2. Basic terms in educational outcomes

- ▶ In assessment, the basic thing to assess in education is students' achievement of educational purpose or outcomes .
- ▶ Activity 3: What do we mean by educational purpose or learning outcome ?

- ▶

- ▶ Educational outcomes or purposes are represented at different levels of complexities/specificities from general up to very specific levels as
 - ▶ **Educational goals,**
 - ▶ **Student's profiles,**
 - ▶ **learning outcomes**
 - ▶ **competences - intermediate and higher**
- minimum**
- 

Educational goals, Student's profiles, learning outcomes



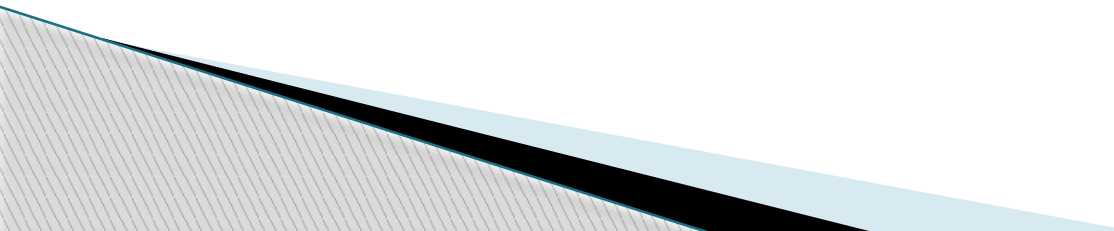
Definitions

- **Educational goals** : broad statements that describe the outcomes and processes of education.
- **student's profile** : is a set of statements that describes what students complete in a particular cycle of schooling. **Useful for planning curriculum**
- **Learning Outcomes**: are statements of what students are expected to achieve at the end of a set of lessons.
- **Intermediate and Higher learning Competencies**: are more advanced LO that most students should achieve beyond the minimum learning competencies.
- **Minimum learning competencies**: are learning outcomes that ALL your students should achieve. They are the least requirements students should achieve.

1.3. Alignment of assessment to instructional objectives

- ▶ Activity 4: What do you understand by alignment of educational components?

- ▶



Alignment

- ▶ Alignment is matching educational components with one another, so that they work together.

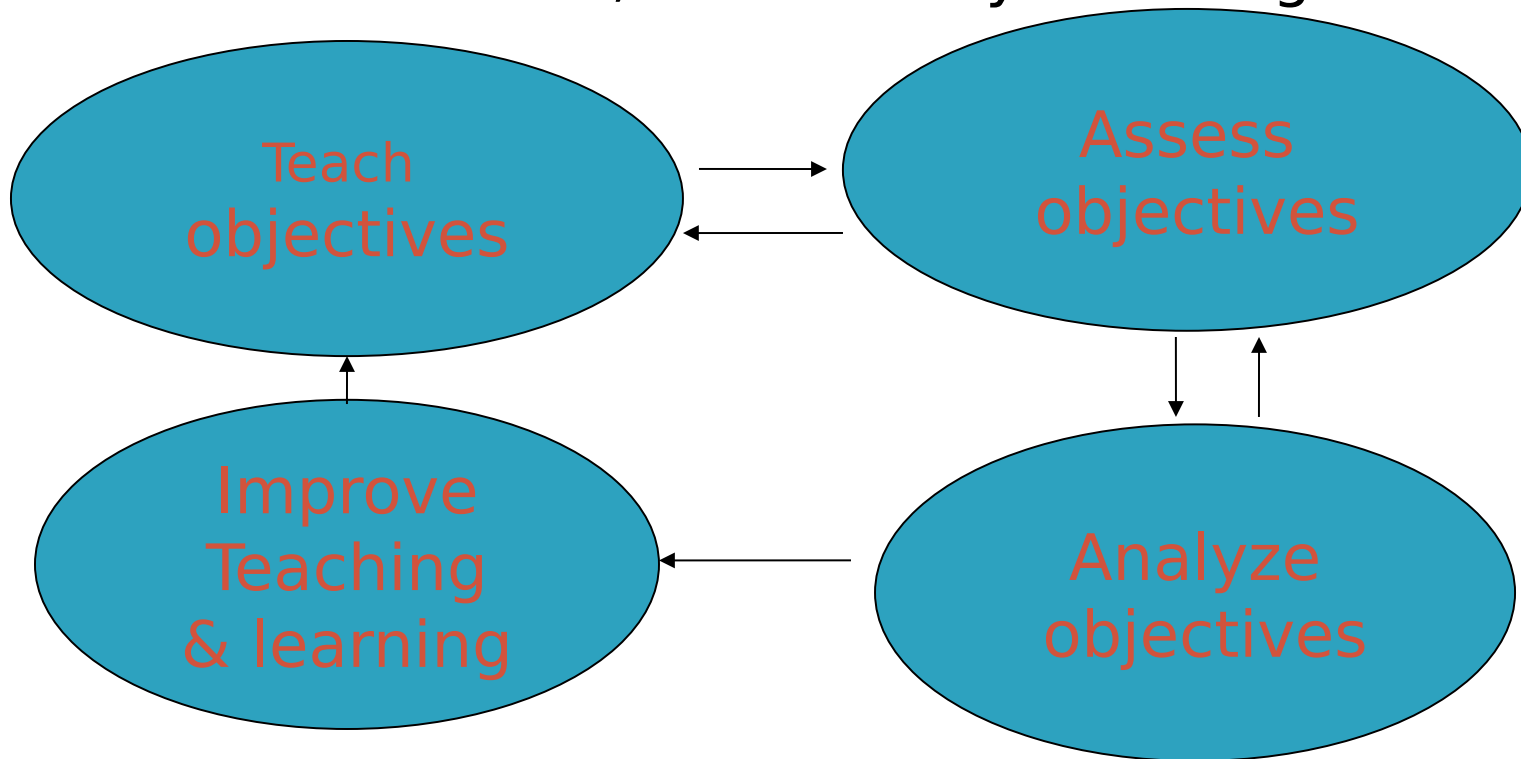
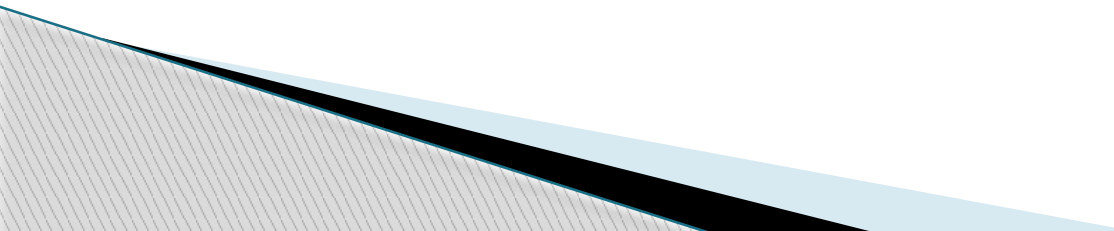


fig. The alignment of educational components

- ▶ Instructional objectives encompass a variety of learning outcomes, and evaluation includes a variety of producers.
 - ▶ The key to effective evaluation of student learning is to relate evaluation procedures as directly as possible to the intended learning outcomes.
 - ▶ This could easily be accomplished if the general instructional objectives and specific learning outcomes have been clearly stated in terms of students' performance.
 - ▶ It is then simply a matter of writing or selecting evaluation instruments that provide the most direct evidence concerning attainment of learning outcomes.
 - ▶ The sequence of steps shown below summarizes the general procedure of relating evaluation techniques to instructional objectives.
 - ▶ The following examples illustrate how the performance indicated in the specific learning
 - ▶ Outcome is related to the evaluation technique.
- 

- ▶ General Instructional Objectives (Intended outcomes that direct our teaching)



- ▶ Specific Learning Outcomes (Types of student performance you are willing to accept as evidence of the attainment of objectives)



- ▶ Evaluation Techniques (Procedures for obtaining samples of student performance like those described in the specific learning outcomes)

- ▶ **1. Specific Learning Outcome:** Identifies the function of a given body structure.
- ▶ **Relevant Test Item:** Which one of the following is the function of kidneys?
 - A) Eliminate body wastes
 - B) Improve the circulation of blood
 - C) Maintain respiration
 - D) Stimulate digestion
- ▶ **2. Specific Learning Outcome:** Identifies common uses of weather instruments.
- ▶ **Relevant Test Item:** Which one of the following instruments is used to determine the amount of moisture in the air?
 - A) Altimeter
 - B) Barometer
 - C) Hygrometer
 - D) Radar
- ▶ **3. Specific Learning Outcome:** Define common terms of educational assessment
- ▶ **Relevant Test Item:** In one or two sentence define the following terms.
 - i) Assessment
 - ii) Test
 - iii) Measurement
 - iv) Evaluation

2. Assessment: its Purpose, importance, principles , nature and process ,

▶ Activity 5.

- ▶ 1.What could be the overall purposes and then importance of assessment?

- ▶

- ▶ 2. What are the principles to be considered in assessment?

- ▶

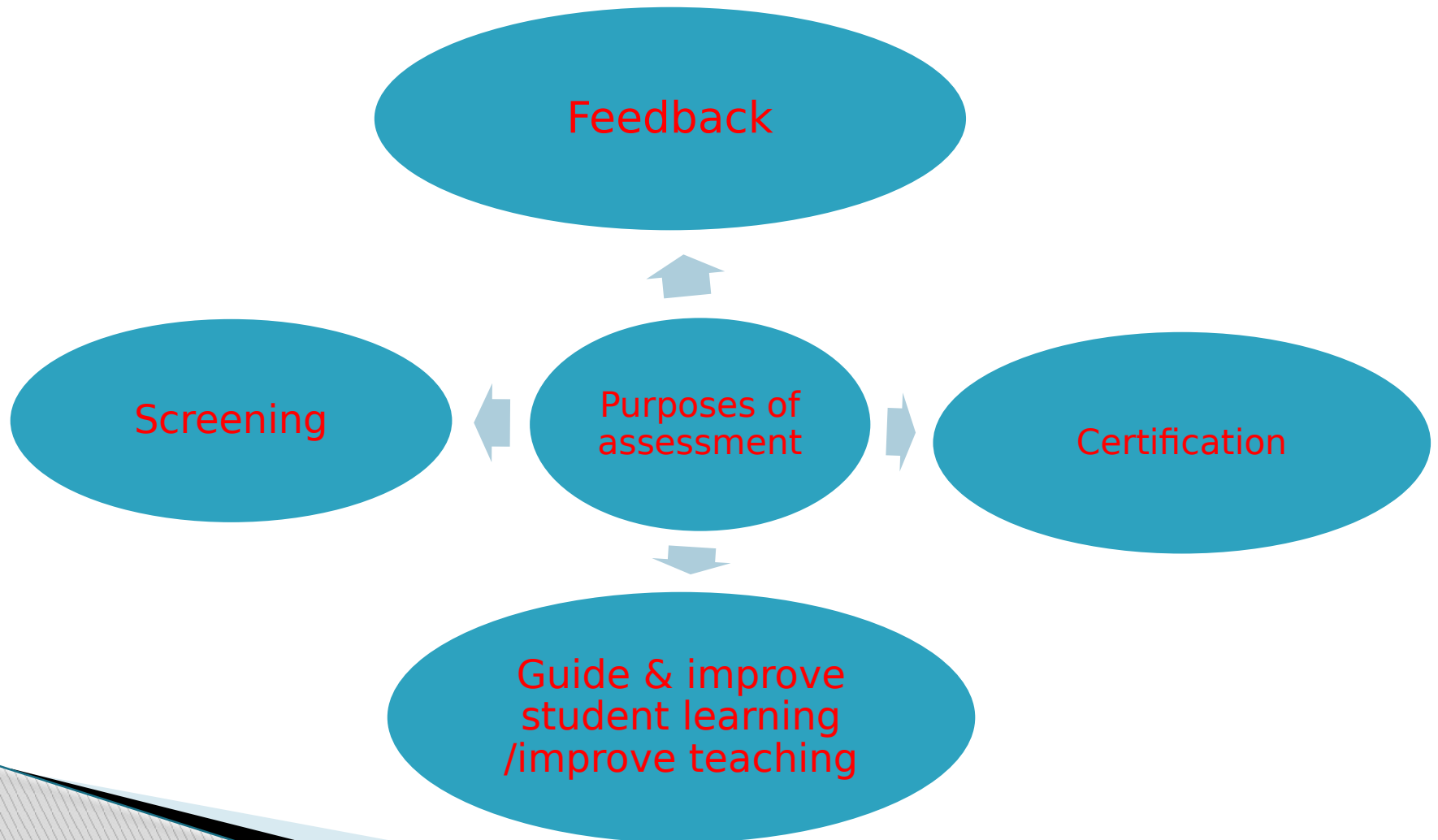
- ▶ 3. what is competency-based assessment?

- ▶

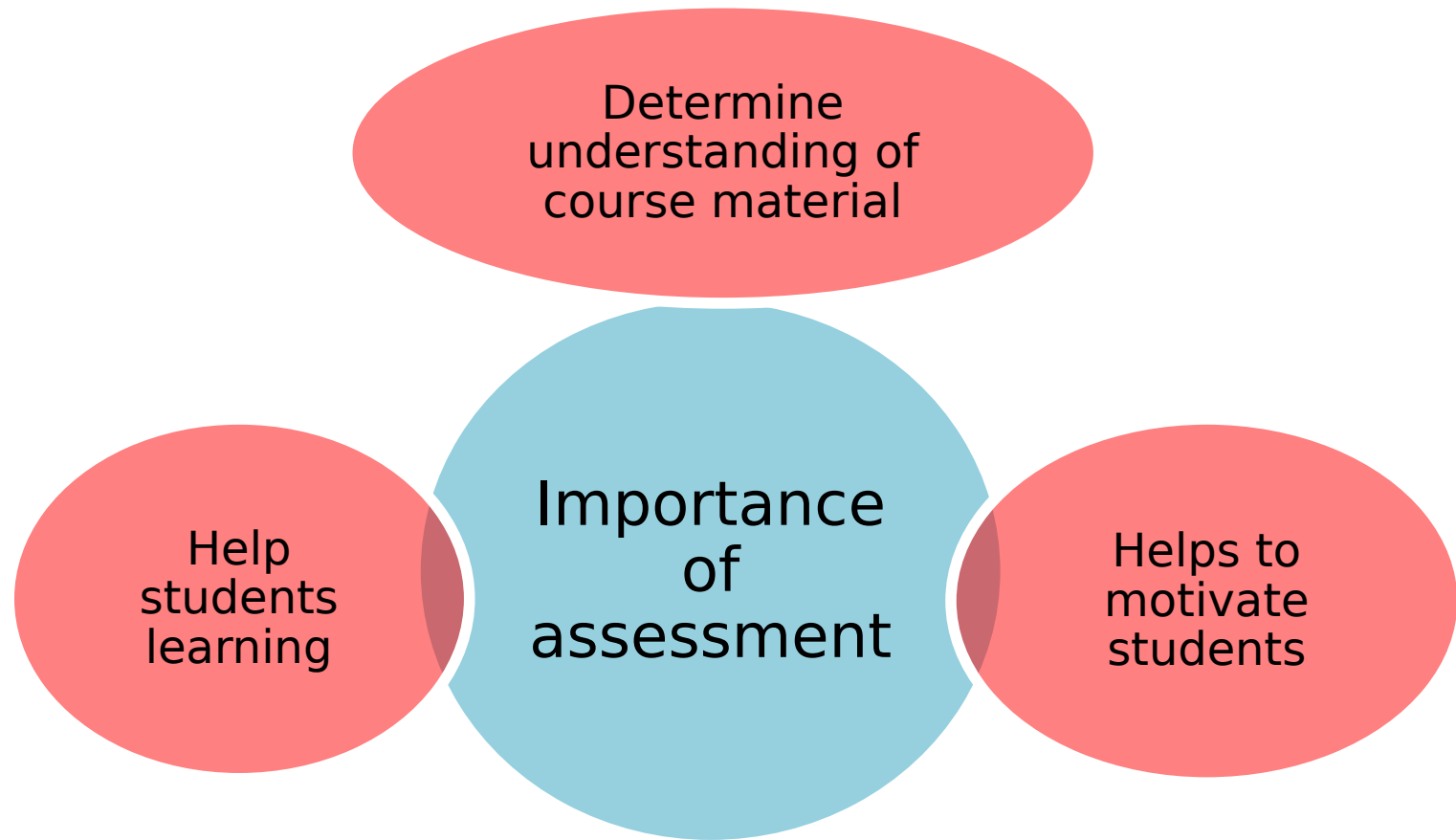
- ▶ 4. what are the main process of assessment?

- ▶

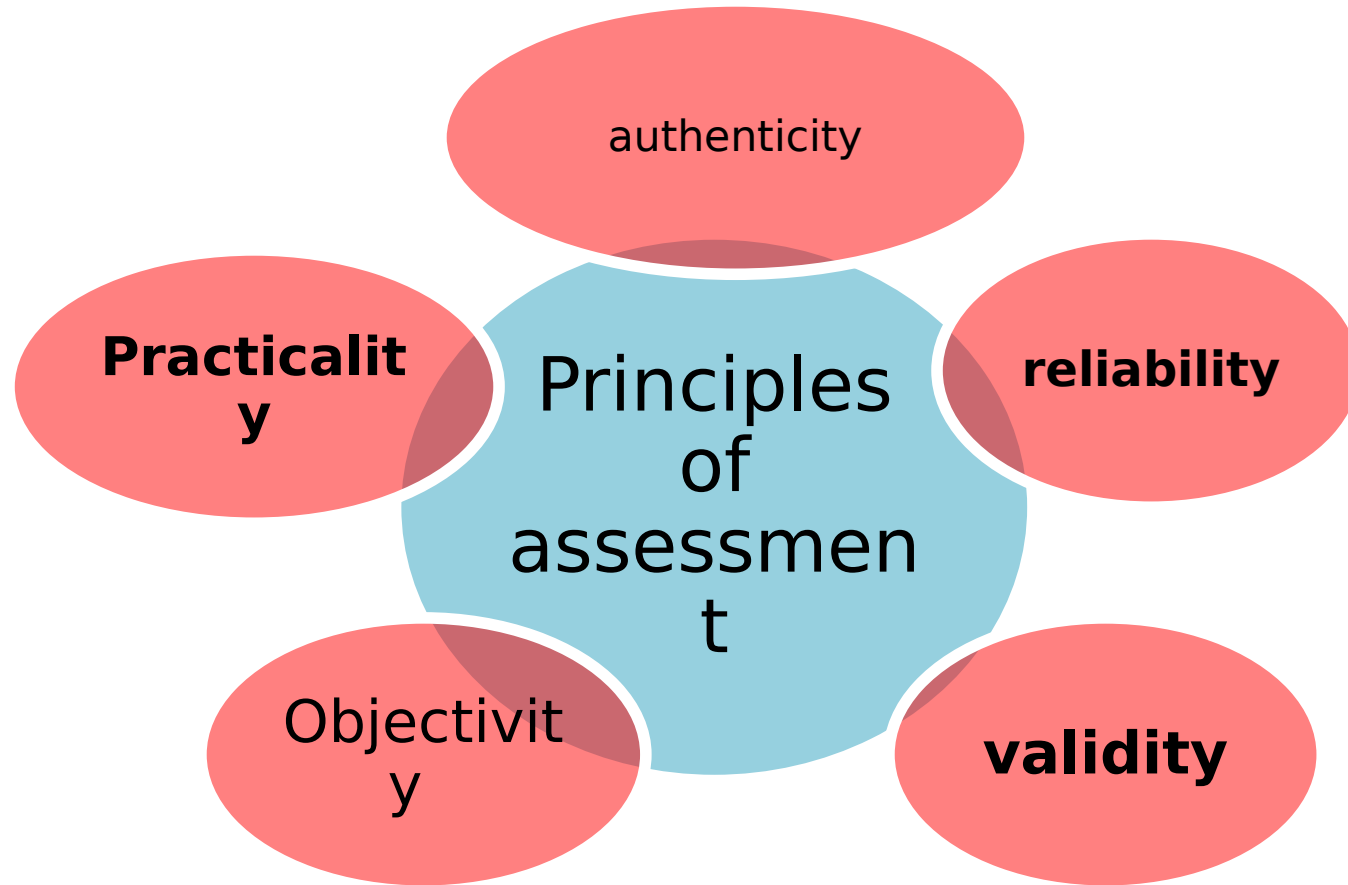
2.1. Purpose of Assessment



2.2. Importance of assessment



2.3. Principles of assessment



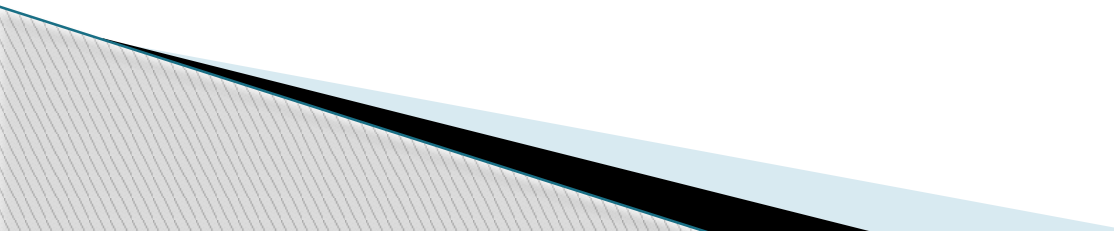
2.4. Competency-Based assessment

- ▶ A competency assessment
- ▶ - is an evaluation of a person's skills and abilities to map existing capabilities to carry out his/her role.
- ▶ -Involves assessing the combination of a person's **strengths, skills and knowledge**.

Competency-based assessment

- ▶ -is the process of collecting evidence and establishing conclusions on the character and scope of the **learner's progress** toward needed **standards or competency**.
- ▶ -seeks to determine whether a person can do a task or group of tasks and how well they can do them.
- ▶ - Uses to assess people usage of their knowledge and skills in an on-the-job situation.
- ▶ - can extract the **successful performance** of the learner through proper demonstration, and clear and tangible actions.

2.5. Main steps in the assessment process

- ▶ The main steps in assessment process are:
 - ▶ Step1: Clearly define and identify the learning outcomes.
 - ▶ Step2: Select appropriate assessment measures and assess the learning outcomes.
 - ▶ Step 3: Analyze the results of the outcomes assessed.
 - ▶ Step 4: Adjust or improve programs following the results of the learning outcomes assessed
- 

• 3. Types of Assessment

- Activity 6:

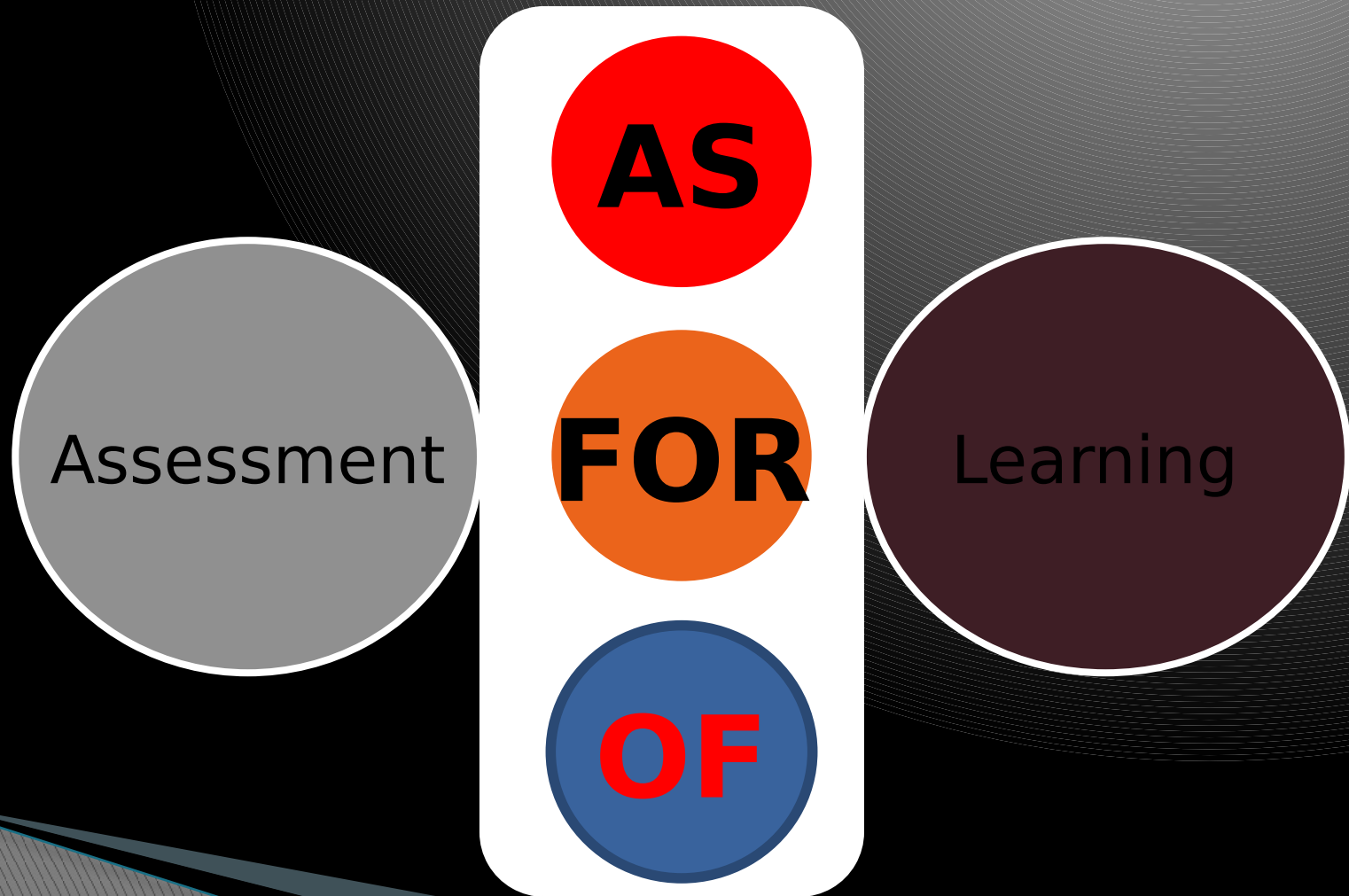
- 1. what are the three basic types of assessment ?

-

- 2. what do we mean by each of the three types of assessment?

-

3.Types of Assessment



• 3.1. Assessment as Learning

- ▶ AAL enables students to:
 - Identify and reflect on **elements of their own** learning
 - Set their **own learning targets**/ setting objectives/Identifying challenges;
 - personally monitor to make adjustment, adaptation, etc.
 - Practice **self and peer assessments**.

3.2. Assessment for Learning (Formative Assessment)

▶ AFL

- uses as an **investigative tool** for teachers to find out as much as they can about
 - what their students know and can do ,
 - what confusions, preconceptions, or gaps they have
 - is an **integral part** of the learning process;
 - Clarifies for students what is to be learned;
 - uses as base to give useful and timely **feedback** ;
 - Helps students to identify the next steps they need to take.
-
- ▶ Example: information collected through worksheets, class observations and ungraded quizzes allows teachers to understand student progress and regulate teaching plans accordingly

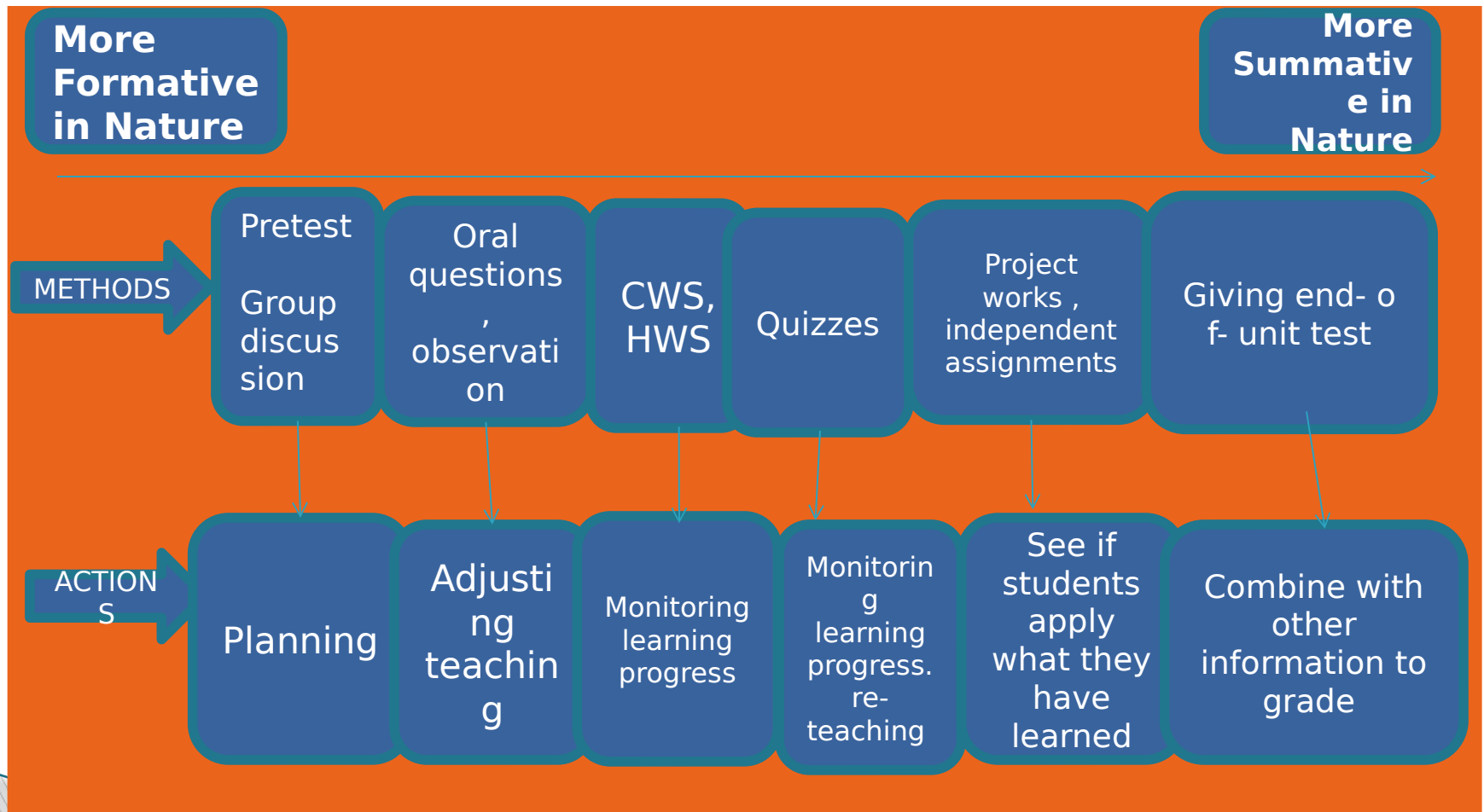
3.3 Assessment of Learning (Summative Assessment)

- ▶ AOL (Summative Assessment)
 - Happens **after** the learning takes place
 - Information is gathered by the teacher and is usually transformed into marks or grades
 - Looks back on past learning
 - uses for Comparison of a child's performance with others' performances
 - Refers to strategies designed to confirm what students know, demonstrate whether or not they have met curriculum outcomes or certify proficiency
- ▶ Example: a final test at the end of the semester constitutes assessment of learning, as it measures how much have the students learnt through the course

4. Strategies of Assessment

Formative strategies	Summative strategies
• Using oral questioning, class & home works, work sheets quizzes, short tests, assignments, etc. • Returning results in time • Explaining how students can improve • Meeting with students to go over their work class	Using unit tests, mid tests, project works, final exams, research works, etc.

FCA and SCA during one's Instructional unit



The formal means /methods/ of CA

- ▶ The basic forms of formal **paper-and-pencil techniques** of assessment/testing/ evaluation that are useful in assessing students' achievement include :
 - **Homework / assignments**
 - **Quizzes,**
 - **Projects,**
 - **Independent investigations,**
 - **End-of-unit/course test and/or examinations.**
- ▶ The **main techniques** are **tests made by teachers** (or classroom achievement tests).

5. Classroom/ teacher-made achievement tests (CATs)

CATs:

- are **basis** for assessing/evaluating students' progress in schools.
 - play **central role** in evaluating students' learning and the curriculum
 - provide **relevant and direct measure** of many important learning outcomes
 - can be **adapted to** and used for **various and special purposes**.
-
- ▶ In utilizing CATs , the tests should give **valid and reliable data or information**.
 - ▶ The validity and reliability of the information gained by the tests depends on:
 - the utilization of **various types of CR tests** ,
 - the **careful planning of tests**, and
 - **careful preparation of the test items**.

5.1. Planning Classroom Achievement Tests

Activity 7:

- a) Why do we need to plan for testing?**
- b) What are the major elements to be considered in planning for a test?**
- c) List the elements needed to be taken into account when preparing /deciding to give a test.**

Planning CATs: Introduction

- ▶ The preparation of good CATs starts from adequate and extensive planning
- ▶ Test plan should relate
 - instructional objectives,
 - the teaching methods,
 - the textual material, and
- the evaluation procedures meaningfully.
- ▶ Planning for test is the base and the primary step in preparing good classroom test.

Why is test planning important?

- ▶ Test plan is important for making the test :
 - appropriate for the intended purpose.
If it is written properly, it helps to ensure the success of the other steps in the process.
 - representative of the contents and skills to be tested.
- ▶ The test plan can :
 - guide in the development of the test.
“To develop a test without a test plan would be like trying to build an office without architectural blue prints.”
 - use as base for selecting items from an item bank in constructing alternative forms of the test

The Steps of Planning and preparation tests

- ▶ From the steps of educational testing, the tasks that are concerned with the planning and the preparation of classroom tests include:
 - Determining the **purpose** of testing
 - Developing the **table of specification**
 - Selecting appropriate **item types**, and
 - Preparing relevant **test items**.

1. The determination of purpose of CR testing

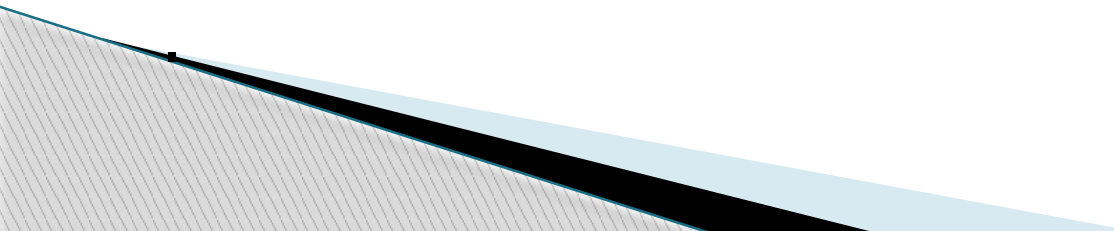
- is the most **crucial task** of test constructor

- is raising question that states “**why am I testing?**”

- indicate the **use of the test** for **various individuals**

- e.g. pupil, teacher, counselor, parents and administrators

- ▶ CATs serve a variety of instructional purposes



Specific purposes of test

The information gathered through CATs can be **used for many purposes.**

Tests use to :

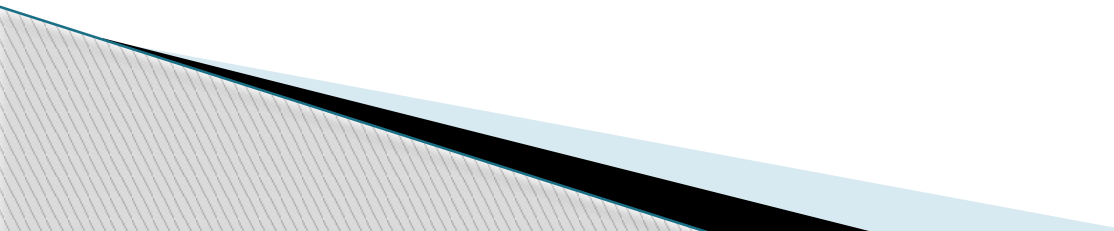
- ▶ **diagnose learner** strengths and needs (so that appropriate teaching can be provided).
- ▶ **provide feedback** to teachers about how well they are teaching (so that they can make needed changes and continue doing things that are effective).
- ▶ **inform and guide instruction** (for example, to help a teacher decide if some learners need remedial teaching or if the whole class needs more teaching on a topic).
- ▶ **inform learners how well they are learning** (so that they continue to do good work and they know when they need to improve).
- ▶ **make clear** to learners what they **are expected to learn.**
- ▶ **motivate** and focus learner attention and effort (to help them learn better).
- ▶ Etc....

Four basic types of CR evaluation

The various purposes/uses of CRTs can be classified into four types of assessment/evaluation:

- 1. Placement evaluation:** evaluation of **prerequisite knowledge / entry behavior.**
- 2. Formative evaluation:** evaluation of students' **learning progress .**
- 3. Diagnostic evaluation:** evaluation of students' **learning difficulties**
- 4. Summative evaluation:** evaluation of students' **end achievement.**

N.B .These purposes provide a good general basis for test planning and preparation.



2. The building of Table of Specification(TS)

- is the second crucial task of test constructor in preparing test .

Table of specification (TS)

- is a **device** that enables to make a test a **valid measure** of the instructional objectives and course content.
- is called **test blue print/ test grid, or content validity chart**.
- is a **two-way chart** useful to obtain **representative sample** of student's **performance in each content area** of the subject .
- ▶ The purposes of the **TS are** to:
 - define the **scope and emphasis** of the test,
 - relate the **objective to the content**, and
 - construct a **balanced test/content validity** tests.

CON'T

- ▶ TS is prepared or developed following 3 steps :

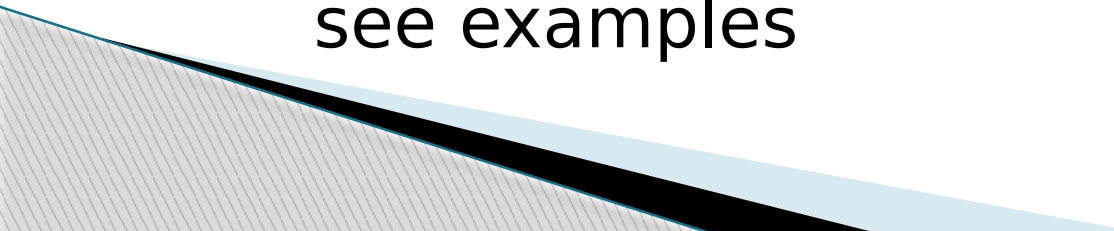
Step 1. Obtaining a list of objectives

Step 2. Outlining the contents

Step 3. Preparing the two-way chart

NB.: TS should be prepared following the above steps.

see examples



Step 1 Obtaining list of instructional objectives

- ▶ Identify the learning outcomes that the student must demonstrate.
- ▶ A learning outcome has three parts.
 - **Condition** - the setting in which the learning will take place
 - **Behavior** - something which can be observed and measured
 - **Performance Level** - measurement criteria
- ▶ Eg. By the end of the semester, the student will summarize the main events of a story in grammatically correct

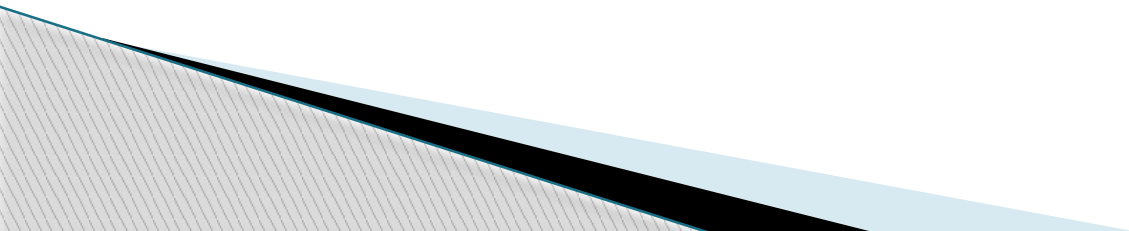
Sample list of objectives

- ▶ Identify the different types of test items (K)
- ▶ *Prepare test plan for the purpose of developing test (A)*
- ▶ *Discuss the uses, forms, merits, and demerits of the various types of test items(U)*
- ▶ *List the rules to be followed in constructing CAT (K)*
- ▶ *construct test items following test preparation rules (A)*
- ▶ Discuss points to be considered in using test (CATs) (U)
- ▶ ***Discuss the assembling, administering , and scoring process of tests(U)***
- ▶ *Follow proper rules in assembling, administering, and scoring tests (A)*
- ▶ Criticize tests on the bases of knowledge of test construction(E)
- ▶ Elaborate the use of the results of tests (S)
- ▶ **Conduct item analysis to improve test items(A)**
- ▶ **Improve test items on the base of the Conduct item analysis (s).**



Step 2 Outlining instructional contents

- ▶ Listing major concepts to be covered by the test
 - should be representative of objectives and material covered
- ▶ Determining contents to be covered by the test
- ▶ Identifying major contents/topics / sub topics



sample Outline of the contents

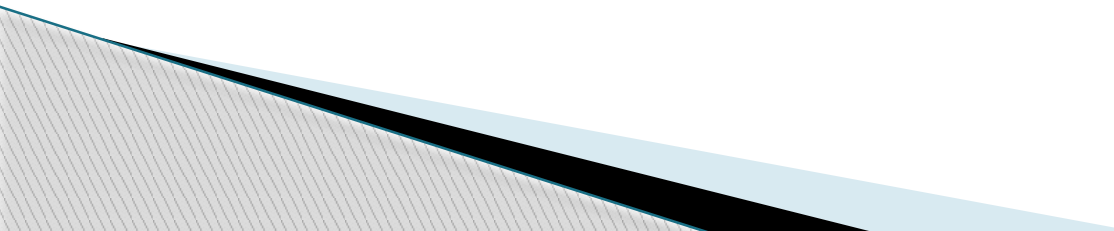
3. Classroom Achievement Test :

3.1. The Types of CAT (2 periods),

3.2. The planning classroom test (6 periods)

3.3. The Construction of CAT (8 periods)

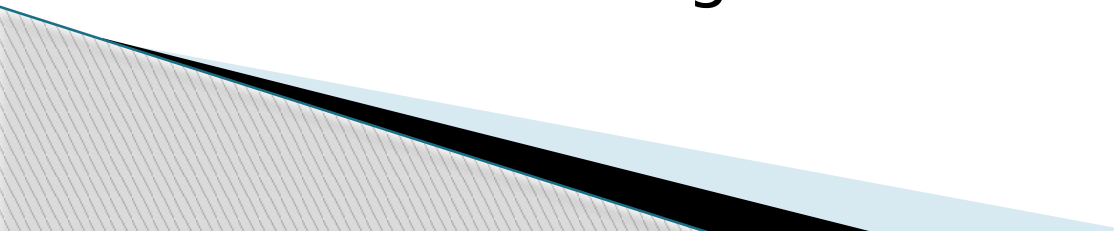
3.4. Assembling items , Administering , and Scoring Tests (4 periods)



Step 3. Preparing the two-way chart

- ▶ The two way chart/table
 - is also known as **Test blue print**
 - is critical document that guides test developers
 - is a framework for the structure of the test
 - is an Engineering design plan to construct a test
 - is usually a grid that contains
 - Objectives/Thinking levels along the horizontal axis
 - contents along the vertical axis

NB: - can be designed across item types





Points to consider in developing test blueprint/ TS

- ▶ Decide the **number of period** devoted for each topic
- ▶ Indicate the **emphasis** given for the category of **objective**
- ▶ Decide the **number of question** to be selected in the **test**.
- ▶ Decide the **number of question** in **each topic** using number of periods devoted for a topic multiplying by total no of question in a test and dividing the result by total period for the unit from which items are prepared
- ▶ Decide the **number of question in each objective category** using percentage of emphasis given to the objective category multiplying by total no of question in a test and dividing the result by total percentage(100) for all objectives in the unit from which items are prepared
- ▶ Decide the **number of question in each topic and objective** in a given cell multiplying row total and column total and dividing by the total question number in the test



Sample of TS : Preparing the two-way chart

R.NO	CONTENTS	No of pds	Knowledge of facts and principles (20%)	Understanding rules and principles (40%)	Application objectives (20%)	analysis , and other higher order objectives (20%)	Total
1	The Types of CAT	2	1 item	2 items	1 item	1 item	5
2	planning classroom test	6	3 items	6 items	3 items	3 items	15
3	Construction of CAT	8	4 items	8 items	4 items	4 items	20
4	Assembling items , Administering , Tests	4	2 items	4 items	2 items	2 items	10
	TOTAL	20pds	10 items	20 items	10 items	10 items	50

• 5.2. selecting item types from the Categories

- ▶ The item types can be grouped in to
- ▶ **1. selected-response items - objective type items/**
- ▶ **2. constructed-response items - subjective or essay types of items)**
- ▶ **Activity 8: What are the types of objective items and subjective items?**

Types of classroom Achievement tests

▶ CATs can be

- **objective type** and **subjective type tests.**

1. The objective tests are :

i. Selection type items:

- **alternative response**(commonly called **true/false**)

- **matching**, and

- **multiple-choice** types.

ii. Supply type items:

- **short answer** and

- **completion/ fill in the blank** items.

Con't

iii. Rank order type items:

- questions answered by **rearranging /ordering** events in sequences basing some **criteria**

iv. Pictorial type items:

- questions answered referring **pictorial events** either by **selecting** from options or **supplying** the correct answer

2. The essay/ subjective tests are :

- restricted response questions** and
- extended/ non-restricted response questions.**
- Orally responded type of questions:**
 - **the** oral examination, less popular today, but can be taken as part of the essay type.

The Selection of the Appropriate item Types

- ▶ **The Selection of the most appropriate item types for a test** is done based on **the nature of the content and the objective** to be measured.
- ▶ This is for the fact that **each type of the test items is efficient for measuring some learning outcomes** and inefficient for measuring others.

e.g., -**Alternate response items** -ability of identifying the correctness of statements of facts, opinions ,e.t.c..
- **Supply type items** -ability of recalling facts
- **Matching items** -ability of relating or classifying similar things E.t.c..

5.3. preparing Good CRT items

- needs **hard work, knowledge of subject matter, and alertness of mind.**
(This is the **only way** to produce good test)
- Besides, to prepare Valid and reliable test, one should construct the test following certain **general principles/ guidelines.**
- To prepare good test, one should know:
 - the unique characteristics
 - the uses,
 - the advantages
 - disadvantages, and
 - the rules for constructing each type of test items

1. Suggestions for constructing items

For preparing test items there are general and specific facts or principles to be taken in to account by test constructors.

Activity 9.List general guidelines to be taken in to account in preparing the different types of tests.

A. **General principles** in the preparation of all types of CAT items

Validity and reliability of CRTs can be achieved only when one constructs the test following certain **general principles and guidelines** .

The general principles to be followed in preparing tests are :

- Principle1.** Test construction procedures must take into account **how the test is going to be used.** (i.e., the purpose of a test)
 - Principle2.** The type of test items should be determined by **the specific learning outcomes** to be measured.
 - Principle3.** Test items construction should be based on a **representative sample** of the course content and specific learning out comes to be measured.
 - Principle4.** Test items should be of the **proper level of difficulty**
- 

Con't -general principles

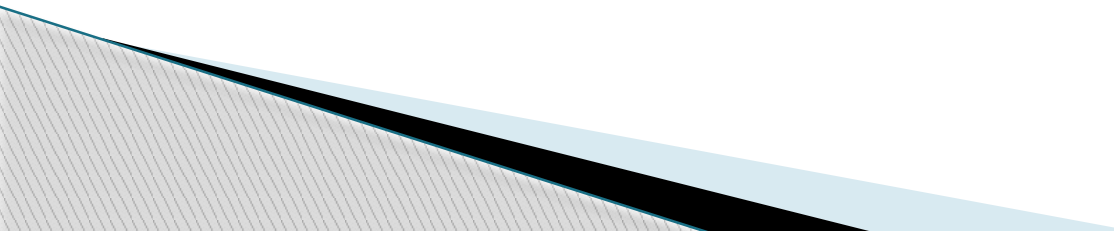
Principle5. Test items should be so constructed that **extraneous factors do not prevent the students** from responding.

Principle6. The test should be so constructed that the students **obtain the correct answer only** if they have attained the desired learning outcomes.

Principle7. The test should be so constructed that it contributes to the **improvement of teaching and learning processes.**

Preparing CR test, having in mind all the above principles helps in constructing good test (i.e., test of good validity and reliability).

General Guidelines in CT Test preparation

- ▶ preliminary draft test should be prepared as early as possible.
 - ▶ A test should include more than one item type
 - ▶ All items of a particular type should be placed together
 - ▶ The content of the test should range from easy to difficult
 - ▶ Items in the test should be arranged in order of difficulty
- 

Cont'd

- ▶ Tests should provide general information
(Date, subject, grade, time , weight etc)
- ▶ The directions should be clear and concise
for each test Items
e.g., do not use Poor directions like
 - Say true or false
 - Match
 - Choose
 - Fill in the blank
 - Explain etc

B. **Specific guidelines** for preparing the different types of test items

Activity 10:

- 1. What are the uses, merits/ advantages and disadvantages of the objective test types(true-false, matching, supply , and multiple choice item types) and subjective (essay) test types ?*
- 2. How can we minimize the limitations and maximize the advantages of each type of the objective items?*

N.B.: This activity needs to form five groups and make each group work on one of the five type of items (T/F, supply, matching, choice, and essay from group 1 up to 5, respectively) and then make them to come up with basic rules/ principles that govern the preparation of each type of items.

1. Writing the Alternate Response (True-false) items

AR(T-F) items

- two response MCIs
- declarative statement answered by saying T or F, R or W, C or INC , Y or N, F or O, A or DisA, an like
- include Common varieties as :
 - right-wrong ,
 - true-false,
 - yes-no ,
 - cluster/ multiple T-F, and
 - correction varieties.

1. The **right-wrong variety** consists of a **sentence, equation, or other**

expressions that are to be marked right or wrong.

2. The **yes-no variety** consists of a **direct question** that is to be answered

saying either **yes or no**.

3. The **true-false variety** consists of a declarative statement that could be

either true or false

Variety(con't)

4. **The cluster/ multiple true-false variety** consists of an **incomplete stem with several suggested completions** each of them is to be judged true or false.

Example1: We can help keep our teeth in good condition by
brushing in the morning and right. T F
visiting the dentist regularly. T F
eating both hard and soft foods. T F

5. **The correction variety** is the variety that **directs examinee to make every false statement true** by suggesting a **substitute for the underlined**.

Examples:

- T F 1. The square root of 64 is 9. _____
T F 2. The use of steam transportation is seen in the 17th c. ____

The uses of AR/T-f/ items

ARIs use to:

- **measure simple learning outcomes** such as the ability to:
 - identify the correctness of definition of terms, facts, principles,
 - distinguish facts from opinions, recognize cause-effect relationships, and
 - Understand some simple aspects of logic and the like.

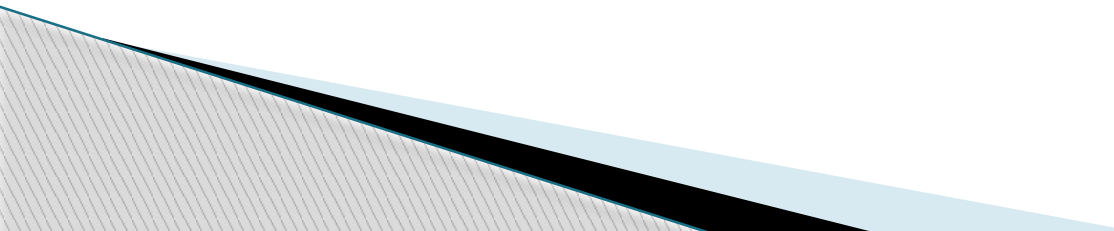
Advantages of AR items

AR/T-F items :

- **easy to construct**
- fit **for young learners or poor readers.**
- **cover a large amount of contents** in a test.
- use to present many items **per units of testing time.**
- **scored quickly, reliably, and objectively** by clerks.
- suitable to **test beliefs in popular misconceptions and superstitions.**
- **adaptable to most content areas**
- may provide a **salvage pool** for discarded multiple choice items
- can measure the higher mental processes of understanding, application, and interpretation.(**if carefully constructed**)
- valid as multiple choice items.

limitations of AR/ T-F items

AR/T-F items :

- Measure **only simple learning outcomes**.
 - **highly susceptible to guessing**.
 - **often copied or picked from the text** directly.
 - **More susceptible to ambiguity and misinterpretation**
 - **Lend** themselves **most easily to cheating**.
 - Tend to be **less discriminating** than multiple-choice tests.
 - **susceptible to response set**; i.e., a pattern of responding.
 - may not have statements **unequivocally true or false**, i.e., difficult to write good true-false items.
 - **expose for specific determiners** (irrelevant clues).
 - **difficult and time consuming** to make them measure the higher mental process.
- 

Specific guidelines *in writing true/false items*

1. Avoid **partly true and partly false** statements, i.e., item should be unequivocally true or false (either **absolutely true or false**) .
2. Avoid the **use of negative statements**, specially double negatives (trick questions)
3. Avoid the **use of clue terms** (specific determiners or qualifiers) that lead to the correct answer
e.g.,- qualifiers like usually, some, generally , often , most, should, as a rule , sometimes , several, may ... make the statement true .
-Like wise, always, never, only completely, none, all make the statements false.
4. Avoid **trivial statements**. Make item deal with important aspects of learning
5. Avoid **including two ideas in one statement** unless cause- effect relationships are being measured

Guidelines in writing true-false items (Con't)

6. Make the true & false statements **approximately equal in length**
7. Include **approximately the same number** of true statements as false ones.
8. **Cite authorities for statements of opinions**
9. Avoid **general truths, highly technical terms, lifting statements verbatim from the text**
10. Word the **item to suggest wrong answer for partial knowledge**
11. Incorporate only **a single idea** in each statement.
12. Write **sharp, clear, and grammatically correct statements** as briefly as possible.
13. Use true statements at first in measuring cause and effect relationship.
14. Avoid ideas and ideologies that reflect an individual's views or attribute to their source.

2. Writing the Matching Exercise

The MEs

- consist of two sets of lists /or parallel columns: **the P &R** lists .
 - premises:** list of problems /questions to be answered.
 - responses:** list of the answers to the questions/solutions.
- demands the examinees to make association b/n each premise and each response on **the basis of criterion of matching of** the responses to the premises that is explained in the direction
- consist of three common varieties :
 - **traditional format,**
 - **"Representation" , and**
 - **classification varieties**

The common varieties of MTIs

- 1. The traditional format variety** consists of **two parallel columns**.
 - 2. The "Representation" variety** consists of **letters or symbols/ numbers that represent locations on maps** and **various parts of an entity** on a diagram or picture/ graph & the like.
 - 3. The classification variety** consists of presentation of **classification scheme** (e.g. parts of speech, types of rocks, etc) for to the **examinee** **classify each** according to the scheme.
- 

Uses of MTIs :

► The MEs

- uses to measure **simple learning outcomes/**

(knowledge of factual information; such as

terms, definitions, dates, and events, involving r/ships) where r/of more or less similar nature are required.

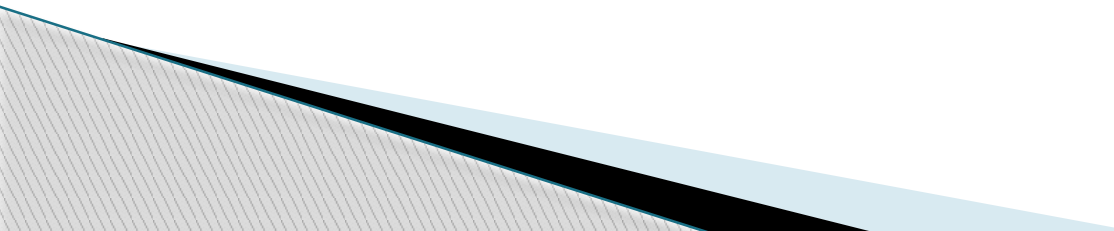
- is well-suited to the "**who**", "**when**", "**what**", "**where**" types of learning,

Advantages of MTIs

The **advantages** of MEs :

- Provide chance to **measure large amount of related materials** in a limited time.
- **Are amenable to machine scoring.**
- Are **easy to construct** but good item need skills.

.



The limitations of MEs

The major **limitations** associated with MEs :

- Encourage serial memorization**

- susceptible to the present for irrelevant clues**

due to

 - difficulty to get homogeneous materials.

- restricted to the measurement of factual materials**

 - based on rote learning.

Some **common faults of teacher-made MEs** :

- vague direction ,

- excessively long sets of matching ,

- lack homogeneity in the lists of responses ,

- the material is set up so that it is not simple for

the

student to respond, and

Guidelines in writing matching items

1. give **clear instructions** for matching the two lists
2. keep the two sets of **items homogeneous** (use only homogeneous material in a single matching exercise)
3. Use a **heading for each column** that accurately describes its content
4. try to **order the responses** logically , alphabetically, and sequence
5. be conscious of the **layout and format**
6. create **more responses** than premises
7. Place **all of the items** for one matching exercise on

Guidelines in writing matching items (CON'T)

8. Maintain **grammatical consistency** within and between columns.
9. Keep **each list reasonably short**, but ensure that each is complete.
10. Provide an **unequal number of premises and responses**
11. Put responses in alphabetical or numerical order.
12. Randomize **position of correct answers**
13. Make the **wording of the premises longer** than the wording of the responses.
14. Identify the items in **first list with numbers** and the

3. Writing the supply items

The STIs

- consist of **the short answer and completion items** that can be answered by numbers, symbols, words, phrases, short-sentences etc.
- are somewhat across between the essay and of the objective items.
- consist of three varieties:
 - **question ,**
 - **completion , and**
 - **association varieties.**

Types of STIs

1. The question variety is presented as **direct question**;

E.g., What is the name of the person who invented the steamboat?

2. The completion variety is written in **an incomplete statement** .

E.g. The name of the person who invented the steamboat is _____.

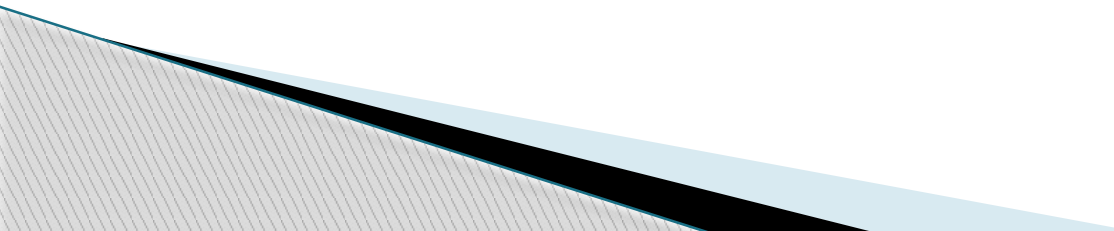
3. The association variety brings **similar ideas under one heading and instruction of one statement**.

E.g. After each city, write the name of the region in which the city is located.

Awasa	(_____)	Bahir Dar	(_____)
Nekemet	(_____)	Jigjiga	(_____)

The uses and advantages of STIs

The STIs use to:

- determine students' ability to **recall of memorized information**
 - **recognize** the appropriate term, concept, or phrase to complete a statement.
 - measure **a wide variety of simple learning outcomes.**
- 

advantages of the STIs

The major **advantages** of STIs :

- the **easiest to construct** and **easier to administer**
- the **best means to measure simple learning out comes**
- **Minimize** the possibility of answering question by **guessing**.
- **Follow the natural and logical method of teaching**-employs the so- called **Socratic question and answer** format.
- **eliminate copying, motivating children,** and providing more **diagnostic information** in the early grades than the recognition type of tests.

Con't

The major **limitations** of STIs :

- **encourage rote memory** and **poor study habits**.
- **Difficult in scoring** because of variety of acceptable answers.
- **unsuitable to measure complex learning outcomes**.
- almost **impossible to construct good items** that
require the student to exhibit their capability

- 1. Word the item** so that the required answer is brief and specific
2. Don't take phrases directly from textbooks.
- 3. Use direct questions** as much as possible instead of incomplete statement
4. Don't include **too many blanks**
5. The parts of the questions should in general be **open ended but requires short answers**
6. The questions should in general be **easy at the start and grow progressively more difficult** towards the end of the set.
7. The questions should be **self contained and not dependent upon a correct response** to an earlier one
- 8. Appropriate space** should be provided on the question paper for the length of the response expected from the student
9. The questions asked should be **consistent as much as possible with the objectives**

4. Writing the Multiple Choice Items

The MCTIs

- are the **most famous, widely used, and applicable items**
- **free from limitations of other types** of objective test items.
- consist of two parts: **the stem and alternatives. i.e., combination of a stem and various alternatives.**
 - **The stem:** problem situations or questions stated in a direct question or an incomplete statement form
 - **Alternatives** : a list of suggested solutions/ or answers/ options or responses
 - include **answer/ key** and **distracters or foils**
or wrong answers.

Types of MCIs (con't)

- **MCIS include different varieties**
- ▶ **The five most frequently used variations are:**
 - One correct answer - Best answer
 - Analogy - Reverse type, and
 - Alternative choice. - **Subset selection technique**
- ▶ **Some other variations are:**
 - Association - incomplete alternative
 - multiple response - combined response,
and
 - substitution

Con't

MCTIs use to :

- measure a **wide variety of learning outcomes** so the most

adaptable/ flexible of the objective item types.

- test

- **Knowledge of Terminology, of specific facts, of principles and generalization, of trends and sequences,**

- of classification and categories**

- **Ability of comprehension/understanding,**

- **ability to apply principles and generalization,**

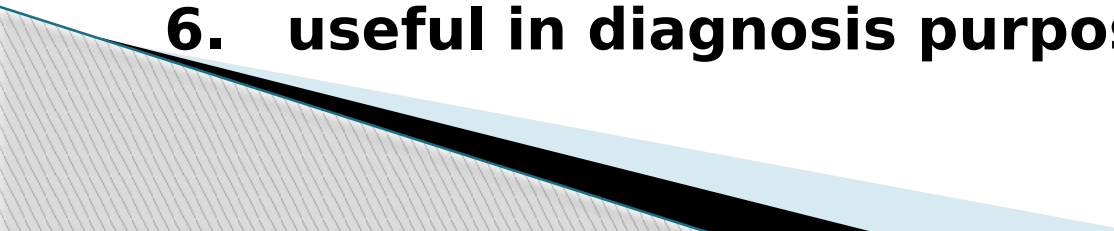
- **ability to interpret causes and effect**

- relationships, and**

- **ability to infer and classify/analysis.**

advantages of MCTIs

The major advantages of MCTIs :

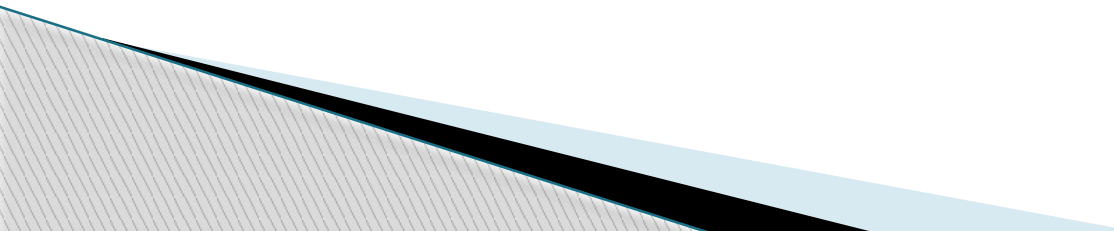
- 1. versatile:** measure a variety of simple & complex learning outcomes.
 - 2. scored very easily:** scoring is not opened for misinterpretation.
 - 3. Reduce the opportunity of guessing the correct answer.**
 - 4. Replace T/F, MIs, and supply items with out undue difficulty.**
 - 5. free from response sets.**
 - 6. useful in diagnosis purpose..**
- 

Advantages of MCIs (Con't)

- 7. Provide
 - a) excellent content sampling leading to more content- valid score
 - b) greater degree to test reliability per item.
- 8. relatively **efficient** (in terms of the number of questions that can be asked in a prescribed time and the space needed to present the questions.
- 9. do not require homogeneous items or qualification to make the answer unequivocally true or false in contrast to matching and true false tests, respectively.
- 10. less prone to ambiguity than the short answer item.
- 11. easier to respond to and better liked by students than true-false items. Students feel that these items type is less ambiguous than true-false items.

Limitation of MCIs

MCIs

- ▶ are limited to verbal level of learning
 - ▶ do not measure ability to organize and present idea
 - ▶ difficulty to find plausible distracters
 - ▶ favor test-wise and risk-taker students
- 

Guidelines in writing multiple choice items

I) Rules about the Stems

1. **The stem of the items** should be meaningful by itself and should present a definite and clear **problem** i.e., **stem** should clearly state **central question or problem** (see pp. 102)
2. The **stem** should be constructed so as to assess the level of knowledge **specified in the objective**
3. The **stem** should be **written at the appropriate language level** for the pupils answering the item
4. The stem should be **kept short** including **only the needed material** that make the problem **clear , specific** and **free from irrelevant** materials
5. Use negatively stated stems/problems only when significant learning outcomes require it.(most Problems should be stated in positive forms).
(avoid **tricky questions or negative wording**)
6. The stem may be a complete or incomplete statement and in the **incomplete statement** **options** should come at the **end of the statement**
7. **Avoid ambiguities** in the stem or the possible responses
8. void type k or key type MCIs

Guidelines in writing multiple choice items

(Con't

II) Rules about alternatives

9. **There should be only one correct or best answer to every question**
10. Position the correct answer at random order and approximately equal number of times , i.e., vary the **position of the best answer**
11. **Avoid the repetition of words in the options**
12. **Make the number of alternatives /choices b/n three and five and should not be greater than five**
13. All the **distracters** should be **feasible** and **plausible**
14. **Avoid choices that are unrelated to the subject of the item**

Guidelines in writing multiple choice items

(Con't)

II) Rules about **options (con't)**

15. Avoid overlapping options
16. Arrange the alternatives as simple as possible
17. Avoid highly technical distracters
18. All the **possible responses** should be of **approximately the same length** and **parallel in structure**
19. Use the alternatives **“none of the above”** and **“all of the above”** **selectively/rarely.**
20. Consider providing an “I don't know options”
21. refrain from **using words** such as always, never, all, none

Guidelines in writing multiple choice items

(Con't)

III) Rules about relation of the stem and answers

22. The responses should be so constructed that **no clues** can be obtained from other alternatives

23. Avoid different types of clues that lead to the correct answer

a) verbal association b/n stem and correct answer

b) length of the alternatives ,

(i.e. all the **possible responses** should be of approximately

the same in length and parallel in structure)

c) grammatical inconsistency or clues

(i.e., the responses should be **grammatically correct)**

5. writing essay test questions

The essay questions are include:

- i. Restricted Response Questions:** limit both the form and the scope of student's answer to the context of focuses.
 - ii. Extended Response Questions:** Provide student complete freedom as to the points he/she discusses and the type of organization he/she will use.
- ▶ Terms such as discuss, explain, apply, determine express, etc, and are commonly used in writing essay test questions.

Con't

- ▶ **The** Essay questions **use to:**

- measure higher order mental processes(application, synthesis, analysis and evaluation) that cannot most probably be measured by objective tests.
- Provide a chance for students to use their own words, style, or organization.

Advantages EQs

The **advantages of** EQs :

- easier to construct
 - measure complex learning outcomes
 - influence study habits and eliminate guessing
 - improve students' writing skills and logical organization of ideas
 - give chance to teacher to comment students' work
- 

Disadvantages of EQs

- ▶ **The limitations of EQs :**

- poor in sampling contents and objectives
- expose for subjective scoring/ low reliability/ because of many extraneous factors
- may not always be understood by students to respond with confidence.
- do not measure factual information effectively.

- ▶ The improvement of EQs requires attention to **two basic problems** .

- ▶ **problem of :**

- constructing essay questions** that call forth the desired behavior, and
- scoring the answer** so that a reliable measure of the achievement is attained.

i. Guidelines in writing essay items

1.Restrict essay questions to those **learning**

outcomes that can not be
satisfactorily

measured by **objective tests.**

2.Formulate questions that will call for the
behavior specified in the **learning out**
come.

3.Phrase the question so that the student
task

is clearly indicated.

4.Indicate an approximate time limit or

ii . Suggestions for effective scoring

In scoring EQs needs to:

1. Prepare an **outline of the expected answer** in advance.

This includes the characteristic of the answers, the organization of ideas, and the expressions of points

2. Use the most **appropriate scoring method** .

- ▶ There are two common methods of scoring EQs:
- ▶ **A. The point method** in which answer is compared to ideal answer .
- ▶ **B. The Rating method** in which each paper is placed in one of a number of piles as the answer is read. If 8 points are allotted to a question, for example, 9 piles might be used ranging in value from 8 points to zero.

Con't

3. Decide on the **provisions for extraneous factors**

that are irrelevant to the learning outcomes being measured.

4. Evaluate **all answers to one question** before going on to the next question.

5. Evaluate the **answers with out looking at** the student's **name**.

6. Score the test when you are **alert enough**.

7. Obtain two or more independent ratings if **important decisions** are to be made based on the results. Independent readings may be

6. Assembling Items, administering and scoring tests

Activity11 :

- 1. What are points to be considered in assembling tests? What do you suggest for teachers in assembling their test items ?**
- 2. what is administering test? What are some of the problems teachers face in administering test? What do you suggest to improve the problem of administering test ?**
- 3. What are some of the problems teachers face in scoring essays? What do you suggest to improve the objectivity of scoring essay tests?**
- 4. What is marking test and list the use of marking and reporting results for different concerned bodies?**

6.1. Assembling the tests

Assembling CR tests involves

- **Recording:** placing each item on separate card for flexible ordering
- **Reviewing:** reading the items again after a few days or asking others to criticize them.
- **Arranging the test item:** listing the types and items of test from simple to complex level.

Items should be arranged by item types in order of ascending difficulty- T/F, MEs, supply, Choice, & Essay

- **Preparing direction for tests:** stating instruction in written or oral form or both.

The directions should indicate the purpose, time, base and procedures of answering and the fate of guessing.

- **Reproducing the test:** arranging items to have a needed quality by spacing, making good viewing, sequencing, numbering, and thinking about clarity of the illustration of material.

6.2. Administering the classroom tests

- ▶ The guiding principle of administering test is:
providing a fair chance to all students to demonstrate their effort of achieving learning outcomes. I.e.,
- ▶ - making both the **physical and psychological conditions conducive or suitable** and
- ▶ - controlling other **factors that interfere** with valid testing or measurement.

Conductive **physical** conditions involve

- adequate work place, light, ventilation,
- optimum temperature
- free from noise or quietness;
- good testing time (e.g., testing time should not be given before or after pleasant things, long vacation/holidays, championship game, and big dance.)

Con't

- ▶ Conductive **psychological** conditions involve :
 - creating/having positive attitude,
 - creating good motivation, and
 - other good internal states of the examinees.

Because pupils can not perform at best if excessive tense or anxiety is there.

Therefore, **anxiety should be avoided** by any means.

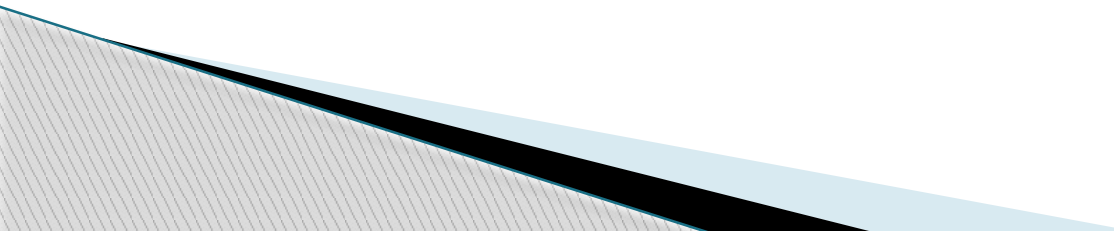
so, during administering test , ones should consider different guidelines.

Guidelines in administering of classroom tests

In administering of tests, one should:

- consider the **suitability of testing place** for all students.
- motivate students/**maintain positive attitudes** of the students
- **Make sure** that the students **understand the direction**
- **make sure** that students **use the answer sheets correctly**.
- **Keep time accurately** and **inform students of the time left** at regular time intervals.

Con't

- Record the **significant events** happened during the test to provide information
 - **Discourage cheating** by special seating arrangement and proper proctoring/ careful supervision.
 - **Avoid distraction of students' work** either by talking or eating or standing behind to read what they are writing. (keep interruption during test at minimum)
 - **Avoid giving hints** to pupils who ask individual items.
 - **Collect the test material properly.**
 - Allow students to **go out after completing the test** but **not at the last few minutes.**
- 

6.3. Scoring of the tests

Scoring tests can be :

- **scoring of objective tests** and
- **scoring of essay tests**

1. Scoring objective tests

Some general considerations for scoring the objective type tests

1. prepare and check Scoring key in advance.
2. factors other than the correctness of the answer (e.g., spelling, grammar, legibility of answer) should be given a separate score.
3. give equal weights to all objective items
4. Give partial credit for partial knowledge

CON'T

2 . Scoring essay questions:

Eqs are criticized for difficulties of scoring- subjective scoring. To score EQs, one should

- 1.Prepare an outline of the expected answer in advance.
- 2.Use the most appropriate scoring method to minimizes biases,

The two methods commonly used are:

- Point (Analytic) Method and
 - global (holistic) method.
- 3.pay attention only to the significant and relevant aspects of the answer,
 - 4.be careful not to let personal idiosyncrasies affect scoring, and
 - 5.apply uniform standards to all papers.

Con't

In scoring EQs, one should

- Use and check key against actual responses
 - Be consistent in scoring/ grading
 - Shuffle the papers before scoring
 - Score all responses to a particular question without interruption.
 - Judge expressions separately from the content
 - Give comments and the correct answer for a question
 - ▶ Set and use a realistic standard of scoring
- 