



EVALUATION

MRS.SUNITHATHERESA
PROFESSOR
ICON





DEFINITION

- Evaluation is a process of making judgements to be used as a basis for planning . it consists of establishing goals, collecting evidence concerning growth towards goals, making judgements about the evidence and revising procedures and goals in the light of judgements. It is for improving the product, the process and even the goals in themselves

Wiles

- Evaluation is the process of determining to what extent the educational objectives are being realized

Ralph Tyler

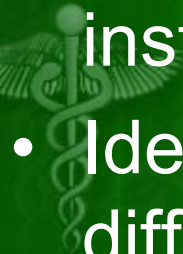



EVALUATION, MEASUREMENT, ASSESSMENT AND TESTING

- **Test**: an instrument or systemic procedure for measuring a sample of behavior
- **Measurement**: the process of obtaining numerical description of the degree to which an individual possesses a particular characteristic
- **Evaluation**: from the stand point of classroom evaluation, it is the systematic process of collecting, analyzing and interpreting information to determine the extent to which pupils are achieving instructional objectives



SCOPE OF EVALUATION

- Value judgment
 - Ascertaining the extent to which the educational objectives have been attained
 - Effectiveness of appraisal or methods of instruction
 - Identifies pupil's strengths and weakness, difficulties and problems, needs and demands
 - Provide baseline for guidance and counseling
- 
- 



- Placements and promotions in jobs
- Development of attitudes, interests, capabilities, creativity, originality, knowledge and skills etc.
- Development of tools and techniques
- Development of curriculum and for its revision
- Interpretation of results
- Helpful for curriculum planners and administrators to improve the curriculum pattern



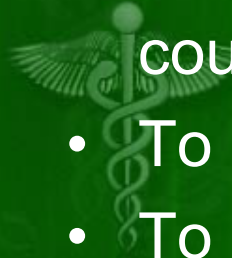


PURPOSE OF EVALUATION

- To discover the extend of competence
- To predict the educational practices
- To certify student's degree, proficiency in a particular educational practice
- To appraise the status of and changes in student's behavior
- To make provision for guiding the growth of the individual student
- To diagnose the individual students educational weakness and strength



- To assess the student's progress from time to time and discloses student's needs and possibilities
- To predict the student's future academic success or otherwise
- To provide basis for modification of curriculum and courses
- To locate areas where remedial measures are needed
- To provide basis for the introduction of experiences to meet the needs of individuals and group of students
- Motivate students towards better attainment and growth





- Test the efficiency of teachers
- Appraise the teachers and supervisors competence
- Improve instructions, measurements and measuring devices
- Bring out the inherent capabilities of a student, such as attitudes, habits, appreciation and understanding, manipulative skills in addition to conventional acquisition of knowledge
- Serves as method of self improvement, improving school learning relations and as a guiding principles for the selection of supervisory techniques

TYPES OF EVALUATION

- **FORMATIVE EVALUATION** (process evaluation)
- **SUMMATIVE EVALUATION** (product evaluation)







FORMATIVE EVALUATION

- It refers to evaluation taking place during the program or learning activity. It is conducted while the event to be evaluated is occurring and focuses on identifying the progress towards purposes, objectives, or outcomes to improve the activities, course, curriculum, program or teaching and student.





Types

- *needs assessment*
 - *evaluability assessment*
 - *structured conceptualization*
 - *Implementation evaluation*
 - *Process evaluation*
- 
- 

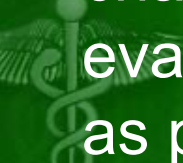



Advantage:

- Emphasize the parts instead the entity
 - The events are recent thus of marking accuracy and preventing distortion by time
 - The results can be used to improve student performance, program of instruction, or learning outcome before the program or course has concluded
- 
- 



Disadvantage

- Making judgment before the activity is completed and not being able to see results before judgments are made
 - intrusive
 - interrupt the flow of outcomes
 - chance for a false sense of security when formative evaluation is positive and the final results are not positive as predicted earlier
- 
- 



SUMMATIVE EVALUATION

- Summative evaluations, examine the effects or outcomes of some object






Types

- *outcome evaluations*
- *impact evaluation*
- *cost-effectiveness and cost-benefit analysis*
- *secondary analysis*
- *meta-analysis*





Advantage:

- All work has been completed and the findings of evaluation is that nothing can be done to alter the results
- 
- 
- 




MAXIMUM PERFORMANCE EVALUATION

- concerned with determining a person's abilities and how well an individual can perform when motivated to obtain as high a score as possible.
- Aptitude and achievement tests are useful in measuring maximum performance



TYPICAL PERFORMANCE EVALUATION

- 
- Determines what individual will do under natural conditions.








CRITERION REFERENCED EVALUATION

- This describes pupils performance according to a specific domain of clearly defined learning tasks.
- 
- 
- 





NORM REFERENCED EVALUATION

- norm-referenced evaluation is based on comparing student achievement to that of others
- 
- 
- 



EVALUATION PROCESS

- identifying the purpose of evaluation
 - identifying a time frame
 - determining when to evaluate
 - selecting the evaluators
 - choosing an evaluation design/ framework of model
 - selecting an evaluation instrument
- 
- 



- collecting data
- interpreting data
- reporting the findings
- using the findings
- considering the costs of evaluation



Identifying the purpose of evaluation

identifying a time frame

determining when to
evaluate

selecting the evaluators




choosing an evaluation
design/ framework of
model

Selection and development of evaluation instrument

- instruments should have the following characteristics
 - Appropriate for what is being evaluated
 - Appropriate for the domain being evaluated
 - Comprehensive
 - Easy to use
 - Cost effective
 - Time efficient
 - Valid and reliable





collecting data

- Consider....
 - Data source
 - Amount of data
 - Timing of data collection
 - Formal versus informal data collection
- 
- 
- 



Interpreting data

- Putting the data in usable form
 - Organizing data for analysis
 - Interpreting data against reestablished criteria
- 
- 



- Consider...
 - Frame of reference
 - Norm referenced interpretation
- Criterion referenced interpretation
- The issues of objectivity and subjectivity
- Legal considerations

Reporting the finding

Using the finding

Considering the costs of evaluation



PRINCIPLES OF EVALUATION

- Determining and clarifying what is to be evaluated always has priority in the evaluation process
- Evaluation techniques should be selected according to the purposes to be served
- Comprehensive evaluation requires a variety of evaluation techniques
- Proper use of evaluation techniques requires an awareness of both their limitations and strengths
- Evaluation is a means to an end, not an end itself

NEED AND IMPORTANCE OF EVALUATION

- a) fulfill classroom objectives
- b) diagnose learning difficulties of students
- C) determine readiness for new learning experiences
- d) form students' classroom groups for special activities
- e) assist students in their problems of adjustment
- f) prepare reports of pupils' progress



CHARACTERISTICS OF EVALUATION

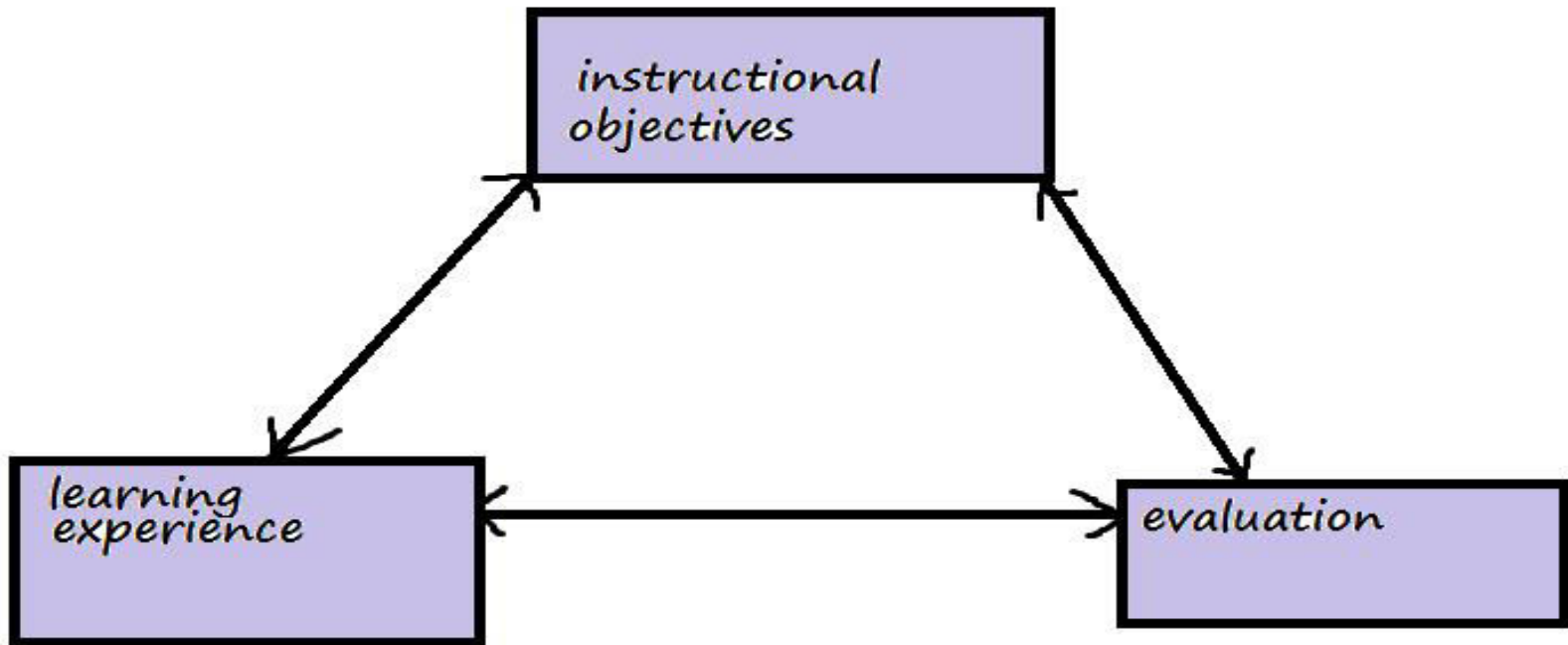
- Evaluation is a continuous process
- Evaluation includes academic and non academic subjects
- Evaluation is a procedure for improving the product
- Discovering the needs of an individual and designing learning experience
- Evaluation is purpose oriented

CHARACTERISTICS OF GOOD EVALUATION

- **Validity**
- A valid evaluation is one which actually tests what is sets out to test
- **Reliability**
- The reliability is a measure of the consistency with which the question, test or examination produces the same result under different but comparable conditions

MODELS OF EDUCATIONAL EVALUATION

- Objective based



- 
- ***Decision-based***
 - ***Value-based***
 - ***Naturalistic approach***

PURPOSES OF EVALUATION IN NURSING EDUCATION

- To determine the level of knowledge and understanding of students
- To determine the level of student's clinical performance
- To become aware of the specific difficulties of individual students or of an entire class, as a basis for further teaching

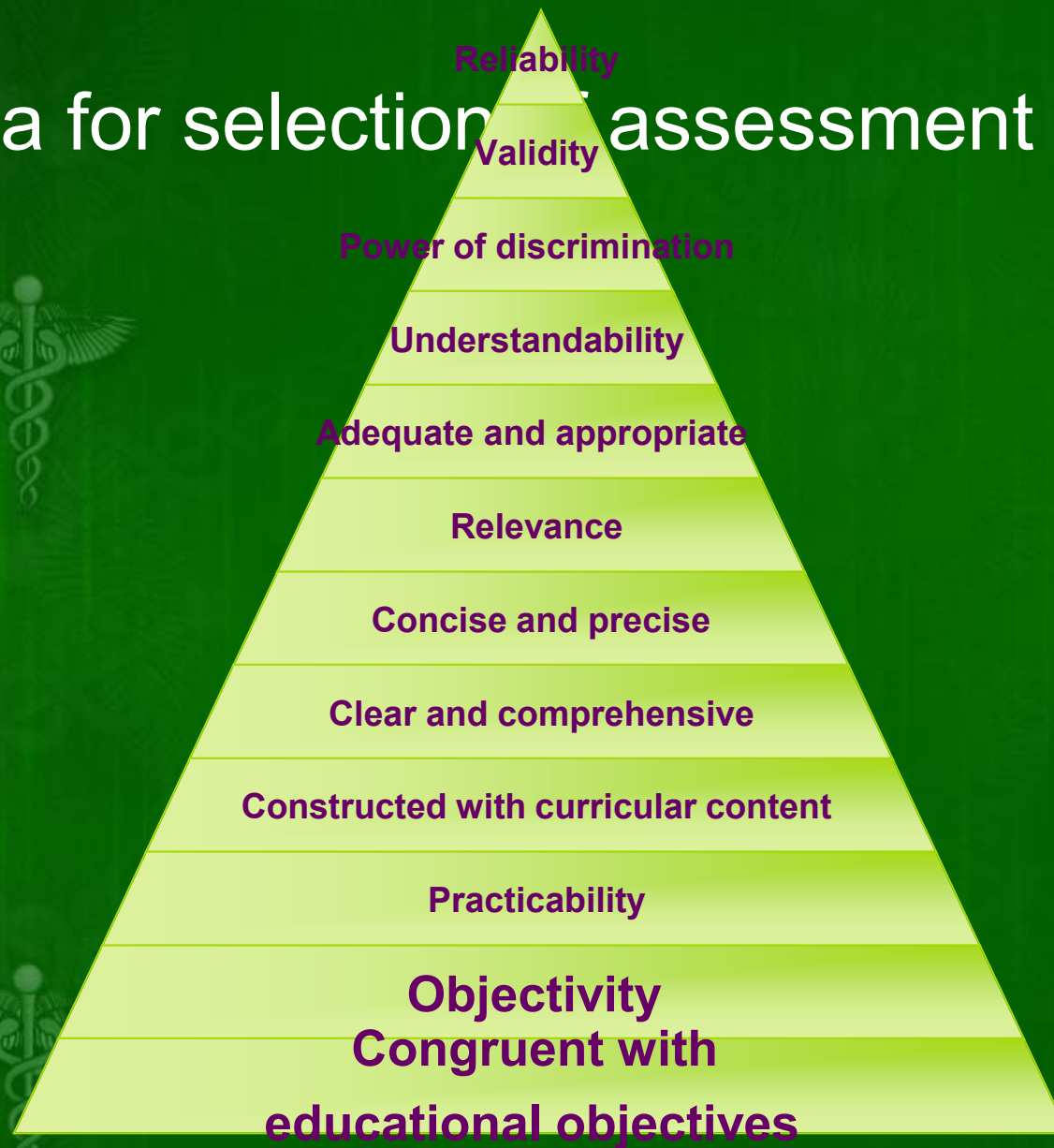


- To diagnose each student's strengths and weakness and to suggest remedial measures which may be needed
- To encourage student's learning by measuring their achievements and informing them of their success
- To help students to become increasingly self directing in their study
- To help students to acquire that attitude of and skills in self evaluation



- To provide the additional motivation of examination that provide opportunity to practice critical thinking, the application of principles, the making of judgments etc.
- To estimate the effectiveness of teaching and learning techniques, of subject content and of instructional media in attaining the goals of the programme
- To gather information needed for administrative purpose

Criteria for selection of assessment methods





Classification of assessment tools and techniques

- Assessment of knowledge
- Assessment of skills
- Assessment of attitude



Assessment of knowledge

- Essay type questions
 - extended response essay
 - Restricted response essay
 - Short answer questions
 - fill in the blanks
 - sentence completion type
 - labeling a diagram
 - short answer
 - Objective type questions
 - Multiple choice questions
 - multiple response question
 - True and false
 - Matching
- 
- 

Assessment of skills



- Observational check list
- Rating scale
- Anecdotal records
- Cumulative records
- Written clinical assignments
- critical incident record
- Practical examination
- Viva voice
- OSCE
- OSPE



Assessment of attitude

- Likert attitude scale
- Semantic differential scale



THE ESSAY TEST

Definition

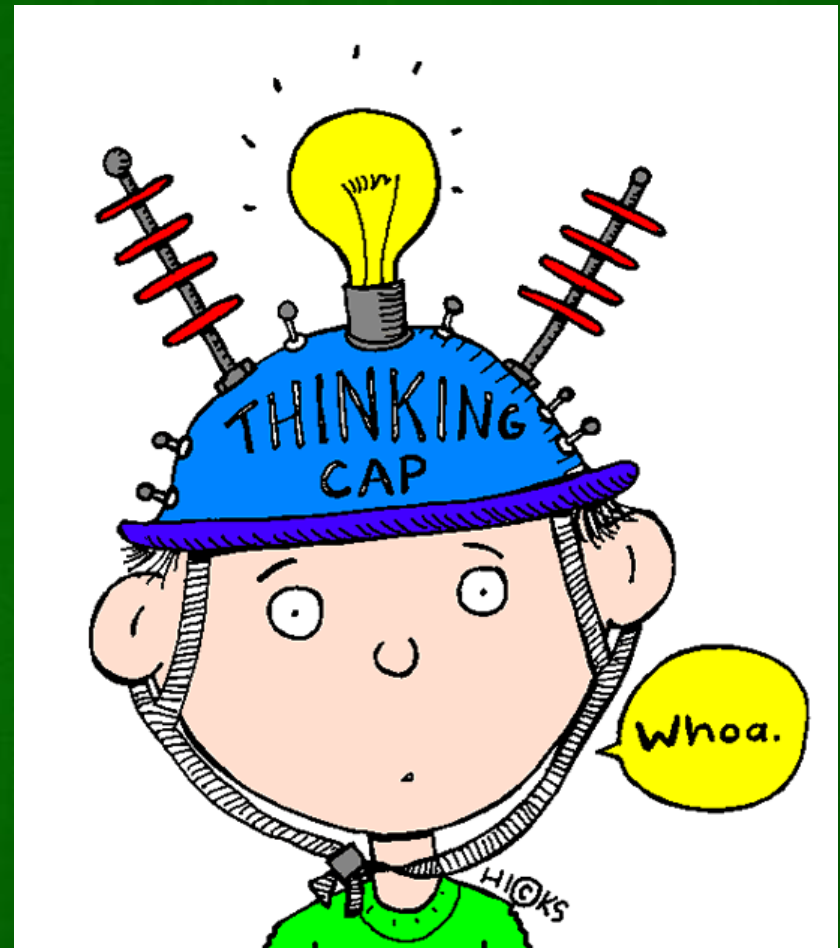
Essay test is a test that requires the student to compose responses, usually lengthy up to several paragraphs.



Essay test measure Higher Level Thinking

Questions that test higher level processes such as

- Analysis
- Synthesis
- Evaluation
- Creativity





Distinctive Feature of Essay Test

The distinctive feature of essay type test is the “***freedom of response***”. Pupils are free to select, relate and present ideas in their own words.





Uses of Essay Test

1. Assess the ability to recall, organize, and integrate ideas.
 2. Assess the ability to express oneself in writing.
 3. Ability to supply information.
 4. Assess student understanding of subject matter.
 5. Measure the knowledge of factual information.
- 
- 

Form of Essay Test

- Restricted Response/ Controlled Response.
- Extended Response/Uncontrolled Response.

Restricted Response

Extended Response





Restricted Response Essay Questions



Restricted response usually limits both the content and the response by restricting the scope of the topic to be discussed.

Useful for measuring learning outcomes requiring interpretation and application of data in a specific area.



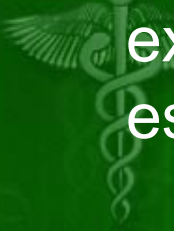



Example of Restricted Response

- Describe two situations that demonstrate the application of the law of supply and demand. Do not use those examples discussed in class.
 - State the main differences between the Vietnam War and previous wars in which the United States has participated.
- 
- 





Advantages of Restricted Response Questions

- Restricted response question is more structured.
 - Measure specific learning outcomes.
 - Restricted response provide for more ease of assessment.
 - Any outcomes measured by an objective interpretive exercise can be measured by a restricted response essay question.
- 
- 



Limitations of Restricted Response Questions

Restricted response question restricts the scope of the topic to be discussed and indicating the nature of the desired response which limits student opportunity to demonstrate these behavior.





Extended Response Essay Questions



Extended response question allows student to select information that they think is pertinent, to organize the answer in accordance with their best judgment, and to integrate and evaluate ideas as they think suitable.

They do not set limits on the length or exact content to be discussed.





Examples of Extended Response Essay Questions



- Compare developments in international relations in the administrations of President William Clinton and President George W. Bush. Cite examples when possible.
 - Imagine that you and a friend found a magic wand. Write a story about an adventure that you and your friend had with the magic wand.
- 
- 

Advantages of Extended Response Questions

- This type of essay item is mostly useful in measuring learning outcomes at the higher cognitive levels of educational objectives such as analysis, synthesis and evaluation levels.
- They expose the individual differences in attitudes, values and creative ability.




Limitations of Extended Response Questions

- They are insufficient for measuring knowledge of factual materials because they call for extensive details in selected content area at a time.
 - Scoring such type of responses is usually difficult and unreliable since the examinees have free will in the array of factual information of varying degree of correctness, coherence and expression.
- 
- 

Advantages Of Essay Questions




- The freedom of response allows the student to express himself in his own words.
 - It measures complex learning outcomes that cannot be measured by other means.
 - Essay tests promotes the development of problem-solving skills.
 - It helps students to improve their writing skills such as writing speed.
 - It encourages creativity by allowing their own unique way.
- 

Advantages Of Essay Questions

- It is easy and economical to administer.
- It encourages good study habits in students.
- Essay item is easy to construct and does not take much time
- It can be used to measure in-depth knowledge especially in a restricted subject matter area.
- It does not encourage guessing and cheating during testing.





Disadvantages of Essay Questions

- Scoring is not reliable because different examiners can grade the score answer differently. In fact, the same examiner can grade the same question differently at different times.
 - Grading of essay tests is time-consuming.
 - Subjective scoring of essay questions.
 - Essay questions do not cover the course content and the objectives as comprehensively as possible.
- 



Disadvantages of Essay Questions

- Evaluating essay questions without adequate attention to the learning outcomes is just like **“three blind men appraising an elephant”** .
One teacher stresses factual content, one organization of ideas, and another writing skill.
- 
- 

Suggestions For Constructing Essay Questions

- Restrict the use of essay questions to those learning outcomes that cannot be satisfactorily measured by objective items.
- *State the question clearly and precisely and make clear what information the answer should contain.*
- Indicate the approximate time limit for each question.
- Avoid the use of optional questions.






Suggestions For Constructing Essay Questions

- Construct question that will call forth the skills specified in the learning standards.

Example:

Write a two page statement defending the importance of conserving our natural resources?

(Your answer will be evaluated in terms of its organization, comprehensiveness, and relevance of the arguments presented.)



Suggestion For Scoring Essay Question

- Chose either the analytical or holistic (global-quality) method.

Analytical Scoring:

This scoring method requires that the instructor develop an ideal response and create a scoring key or guide. The scoring key provides an absolute standard for determining the total points awarded for a response. Student responses are compared to the scoring standard and not to the responses of their classmates.







Suggestion For Scoring Essay Question

Holistic Scoring:

The reader forms an impression of the overall quality of a response and then transforms that impression into a score or grade. The score represents the quality of a response in relation to a relative standard such as other students in the class.





Suggestion For Scoring Essay Question



- Score the responses question-by-question rather than student-by-student.
- Disassociate the identity of students from their responses during the grading process.
- Determine in advance what aspects of the response will or will not be judged in scoring.

Know what
to look for
before
starting to
grade papers



Bluffing-A Special scoring Problem

It is possible for students to obtain higher scores in essay questions than they deserve by the mean of clever bluffing. This is usually a combination of writing skill, general knowledge and use of common tricks like;

- Respond to every question.
 - Stressing importance of topic.
 - Agreeing with teachers opinion.
 - Name dropping.
 - Write on a related topic and “make it fit”.
 - Writing in general terms that fit many situation.
- 
- 

Types of Objective Test Questions

A) RECALL TYPE:

- Sentence completion items
- Fill in the blanks

B) RECOGNITION TYPE:

- Multiple-Choice Items
- True-False Items
- Matching Items

C) OTHERS/PROBLEM SOLVING TYPE:

- Rearrangement
- Analogy
- Context dependent (Pictorial form, Interpretative)

SIMPLE RECALL TYPE

- ❑ Statement is so worded that
 - There is only one response.
 - The response is brief, usually a single word, number, symbol or brief phrase.
- ❑ Blanks should be on right of the question.

Eg: Sphygmanometer is used to measure



SENTENCE COMPLETION TYPE

Certain imp. words or phrase are omitted & students are expected to fill in missing words.

Instructions to be followed:

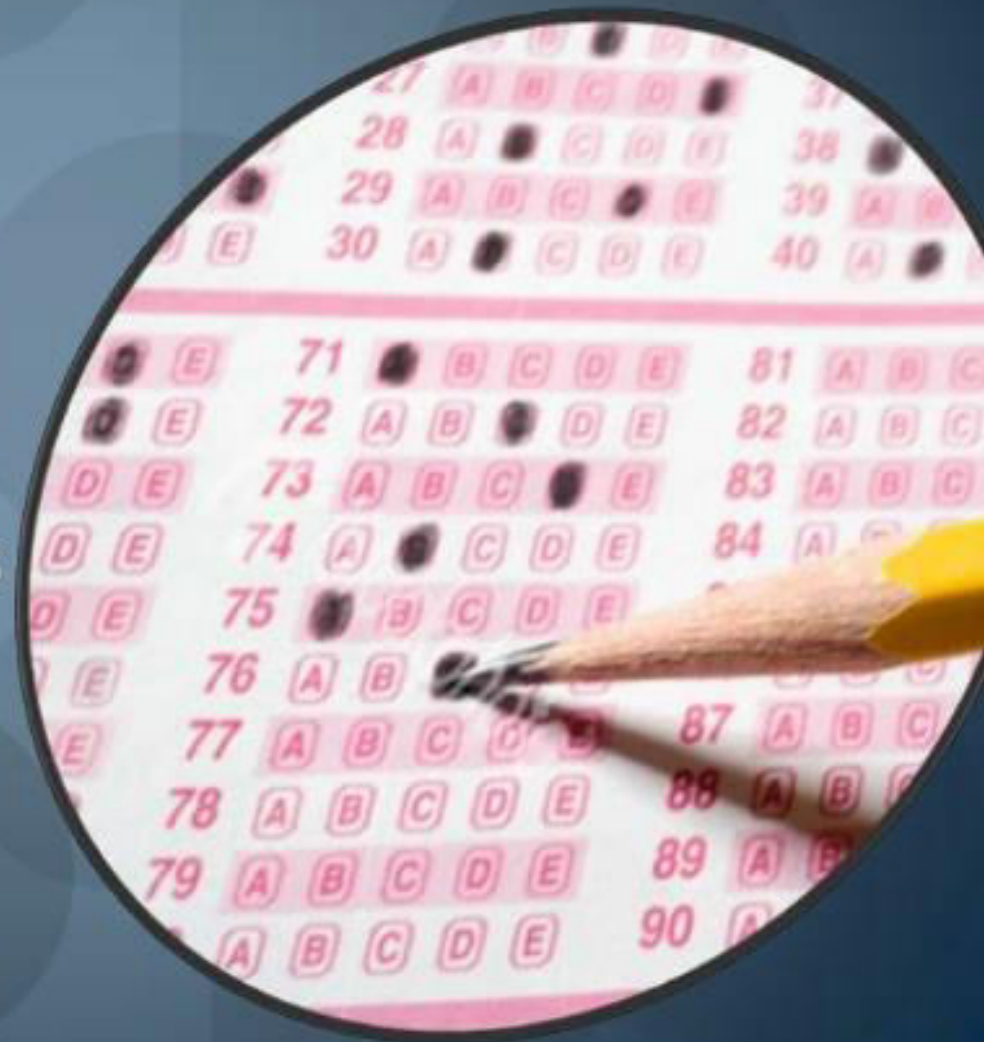
- I. Only key words should be missed.
- II. Don't take sentences directly from the text.
- III. Don't have too many blanks in the statement.
- IV. Don't indicate the expected ans. by putting "a" or "an" before a blank.

Eg: communication is a two way process by which people
..... one another.

MULTIPLE CHOICE ITEMS

Consists of

- I. The stem- direct question /incomplete statement.
- II. The options/responses-
3 to 5 alternative response which includes
 - a) The key
 - b) distractors



MULTIPLE CHOICE ITEMS

- Advantages:
 - Extremely versatile.
 - A compromise between a short answer/essay and T/F item .
 - Can be conveniently used for numerous subject matter areas.
 - Provide reliability , objectivity & validity.
 - Provides student & faculty an opportunity to get the detailed feedback.
 - Reduce the effort of guessing.
 - Cost effective in terms of paper and time.

TYPES OF MULTIPLE CHOICE ITEMS

- Five Basic Types
 - Best answer type
 - Right Answer Type
 - Multiple true false Type
 - Reason-Assertion type
 - Analogy type
- Which Type is Best?
Right Answer
vs.
Best answer



ONE BEST RESPONSE TYPE

- Question:

Active immunization is available against all of the diseases EXCEPT:

- a) Tuberculosis
- b) Small pox
- c) Poliomyelitis
- d) Malaria
- e) Yellow fever

Ans: (d) malaria

ONE BEST RESPONSE TYPE

- Question:

Active immunization is available against all of the diseases EXCEPT:

- a) Tuberculosis
- b) Small pox
- c) Poliomyelitis
- d) Malaria
- e) Yellow fever

Ans: (d) malaria

MULTIPLE RESPONSE/RIGHT ANSWER TYPE

- QUESTION:

A child suffering from an acute exacerbation of rheumatic fever usually have:

1. An elevated sedimentation rate.
2. A decreased P-R interval.
3. An elevated antistreptolysin O titre.
4. Subcutaneous nodule.

- ANS: 1,3,4

MULTIPLE TRUE FALSE TYPE

- This type consists of a stem followed by several true or false statements. The candidate is to determine whether or not each of the four statements which follows is true or false. Then responds according to a code which permits one out of five possible combinations. The directions for this item type are as follows:
 - Ques: Which one is correct:
 - a) If only 1, 2 & 3 are correct.
 - b) If only 1 & 3 are correct.
 - c) If only 2 & 4 are correct.
 - d) If only 4 is correct.
 - e) If all are correct.

ANALOGY TYPE

In this student is requested to deduce the relationship that exists between the first two parts of the item and when apply it to the third and fourth part. Normally ,third part is given & fourth part is missing, which is selected from from the list of options.

Eg: Stem Vitamin A:Night Blindness::Vitamin C:

Response a)Beriberi b) Scurvy c)Cretinism d)Pellagra

DIRECTIONS FOR MCQ'S

- Writing the stem first:
 - A. Be sure the stem asks a clear question.
 - B. Stems phrased as questions are usually easier to write .
 - C. Stems should not contain a lot of irrelevant information.
 - D. Appropriate reading level/terms.
 - E. Be sure the stem is grammatically correct.
- Writing the correct response
 - Use same terms/reading level.
 - Avoid too many qualifiers .
 - Assign a random position in the answer sequence .
 - Avoid negatively stated stems .

DIRECTIONS contd.....

- Read the stem and correct response together .
- Generate the distractors/alternative responses(3-4).
- Be sure that distractors & keys possess homogeneity.
- When dealing with the numerical answers, arrange the answers in order from large to small or vice versa.
- Exclude extraneous words from both stem and alternatives .
- Avoid all of the above (can answer based on partial information).
- Avoid none of the above.

OTHER TIPS FOR MCQ'S

- Alternative responses/distractors should be plausible and as homogeneous as possible.
- Response alternatives should not overlap. eg: Two synonymous terms (arithmetic average/mean)
- Avoid double negatives.
 - None of the following are part of the brain except which one?
- Emphasize **negative** wording.
- Each item should be independent of other items in the test i.e. information in the stem of one item should NOT help answer another item.

TRUE FALSE TYPE QUESTIONS

TEST TIPS:

T
true

The following words usually indicate a true statement:

- Usually
- Many
- Most
- Some
- Often

OR

These words usually make a statement false:

- Always
- Never
- Every
- All
- None

F
false

- These are the statements which the examinee has to think and decide whether the statements are true or false, right or wrong.

Eg: Oslander's sign is the increased pulsation felt in the lateral vaginal fornices.

The Matching Type:

- This consist of 2 columns.Each item of column “A” is matched with the phrase of column “B” on a rationalized basis.
- Good for testing of knowledge of terminologies, relation of facts etc.

Directions:

- ✓ A short title should be given at the top of each column.
- ✓ No. of choices should exceed no of statements.
- ✓ Longer statement should be used in column “A” & shorter in “B”.

THE MATCHING TYPE

Column A

- a) Increased metabolic activity.
- b) Hyperinsulinism.
- c) Lack of storage of glycogen in the liver.
- d) Storage of abnormal glycogen in the liver.
- e) Decreased secretion by pituitary & adrenal glands.

Column B

- 1. Adenoma of islets of langerhans
- 2. Violent exercise
- 3. Hyperthyroidism
- 4. Simmond's disease
- 5. Von Gierke's disease
- 6. Epidemic hepatitis
- 7. Diabetes mellitus

REASON ASSERTION TYPE

Each question consists of 2 paired statements i.e. assertion & reason which are connected.eg:

Cow's milk is preferable to breast milk because cow's milk is richer in calcium.

- a) If both assertion & reason are true & reason is correct explanation of assertion.
- b) If both are true & reason is NOT a correct explanation.
- c) If assertion is true & reason is false.
- d) If assertion is false & reason is true.
- e) Both assertion and reason are false.

ANS: a)

Rearrangement type

Options are given & are to be rearranged according to the priority:

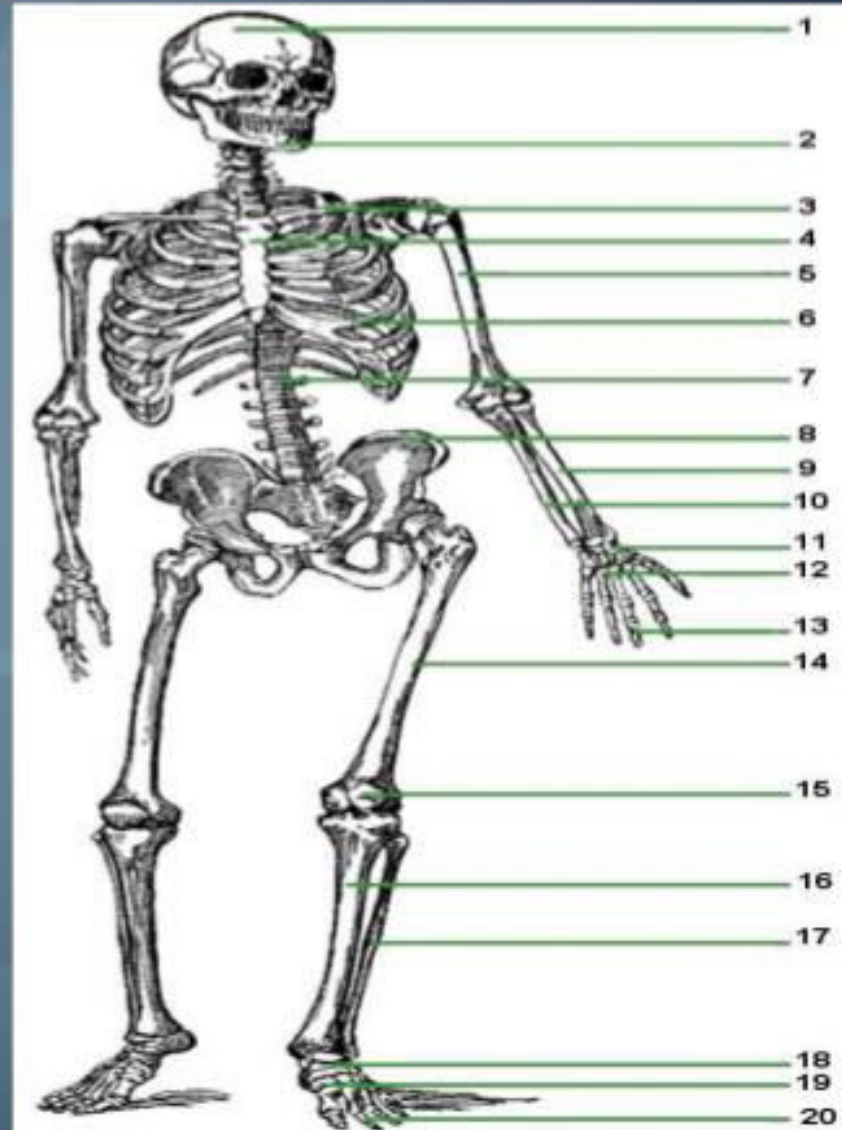
Ques: Rearrange the following conditions according to age priority:

Scurvy, Protein Energy Malnutrition, Neural tube defects, Juvenile Rheumatoid Arthritis.

Ans: Neural Tube Defects, Protein Energy Malnutrition, Scurvy, Juvenile Rheumatoid Arthritis.

THE CONTEXT DEPENDENT TYPE

- Items are based on an external source that may be pictorial or verbal.



Evaluation

- It can be done
 - A) Manually
 - B) computer assisted.
- Negative marking is done.
- Formula used: $(1/n-1)$ where n = number of items
($-1/3$ in 4 item question, $1/4$ in 5 items & -1 in true false type.)
- Setting pass marks (arbitrary).



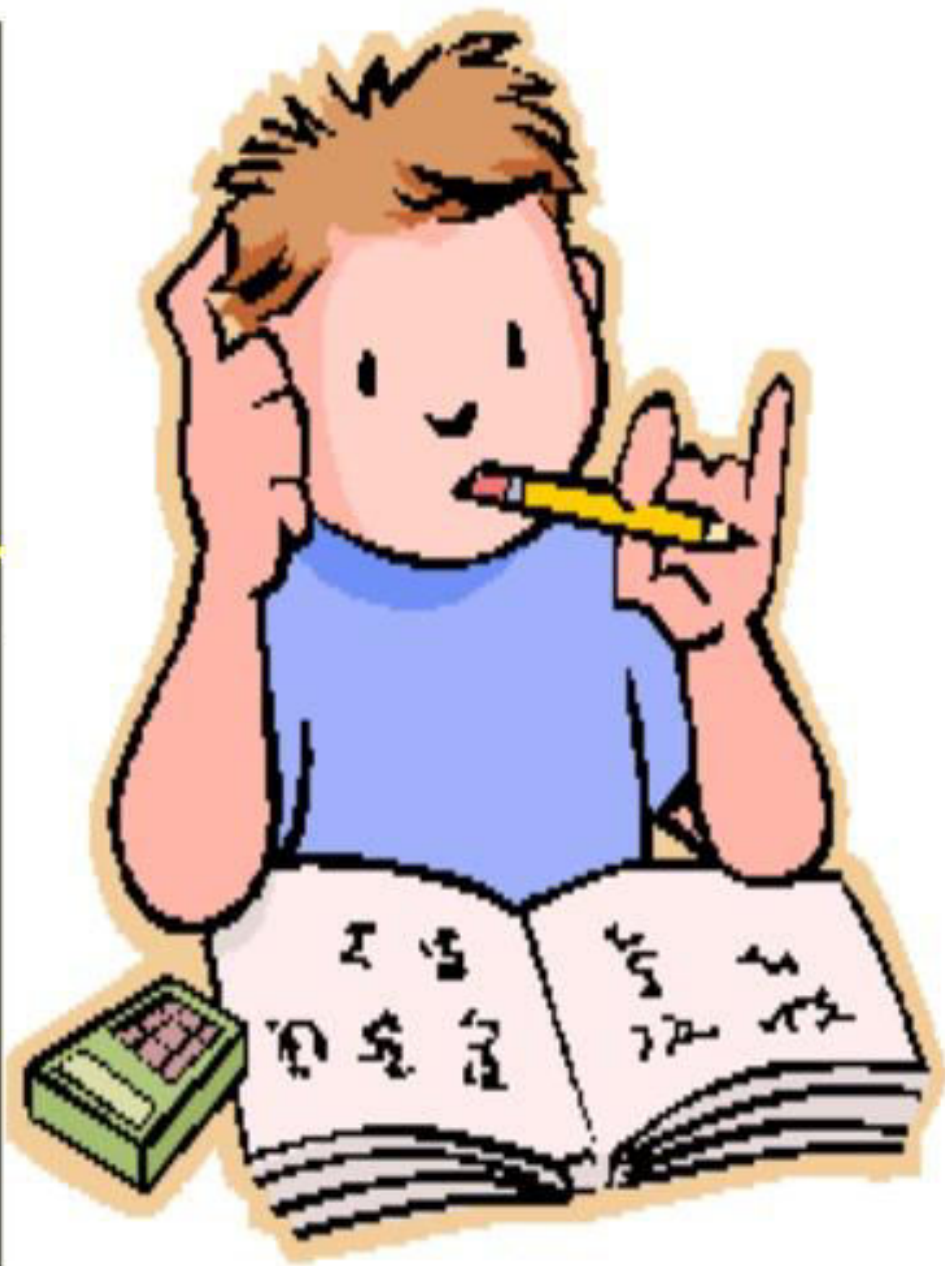
Test Item Check List

- In general:
- It should be wholly realistic & practical
- It should deal with an important & useful aspect of the profession.
- It should be phrased in the working language of the profession.
- It should call for the professional knowledge.
- It should be independent of every other item in the test .
- It should be specific.
- It should be away from the faults of giving the irrelevant details with the best response.

Contd...

- The problem stated :
 - ✓ should be clear.
 - ✓ Should be stated briefly but completely.
 - ✓ It should contain only material relevant to its solution.
- The distractors:
 - Are important ,should be plausible answers rather than obvious distractors.
 - Should deal with similar ideas or data expressed in similar form.

CHECKLIST



OBSERVATIONAL CHECKLIST

- A checklist consists of a listing of steps, activities or behavior which the observer records when an incident occurs.
- A checklist enables the observer to note only whether or not a trait or characteristics is present.

■ SUGGESTION TO FOLLOW WHILE USING CHECKLIST:-

- i. Checklist should be directly related to learning objectives.
- ii. It needs to be confined to performance areas that can be assessed by positive and negative criteria.
- iii. Use checklist when ascertaining a trait or characteristics is present or absent.
- iv. Clearly specify the traits to be observed.
- v. Have a separate checklist for each candidate.
- vi. Multiple observation provide a more accurate assessment.
- vii. Student should be evaluated in natural setting.

BEHAVIOR OBSERVATION CHECKLIST

CHILD'S NAME _____

DATE _____

OBSERVER'S NAME _____

TIME _____

AREA/S OF ROOM OBSERVED _____

Behavior Intervals	1	2	3	4	5	6	7	8	9	10
Attentive										
Competing										
Cooperating										
Demanding										
Dependent										
Disruptive										
Exploring										
Helping Others										
Imitating										
In Transition/Changing										
Inappropriate										
Independent										
Recalls Routine										
Initiates Activity										
Intent										
Leading/Demonstrating										
Making Choices										
On-looker										
Participating										
Requesting Help										
Uncooperative										
Uninvolved/Wandering										
Unsafe Activity										
Using Materials Appropriately										
Using Appropriate Manners										
Waiting										

Comments: (Including changes in behavior, adult interactions, health concerns, etc.)



	yes	some	no
Finishes on time			
Work is organized			
Checks work			
Seeks help when needed			
Works independently			
Works well with others			
Stays on task			

Advantages

- Inter –individual differences
- Simple method to record observations
- Adaptable to subject areas
- Evaluating learning activities expected to be performed
- Evaluating procedure work
- Constrain the direct attention
- Decreases chance of error

Disadvantages

- Do not indicate quality of performance
- Limited component only can be evaluated
- Degree of accuracy cannot be measured
- Limited use in qualitative observation
- Not easy to prepare

Rating Scales



Definition

- 🍎 Rating scales can help to evaluate the quality of the behavior of one student or many students.
- 🍎 It also measure the degree to which a student exhibits a specified behavior.



COMPONENTS OF RATING SCALE

- The rating scale has two components:
- Stimulus variable
- Response options
- The stimulus variable consists of trait names or qualities to be rated.
- The response options consists of numerical or descriptive categories.

- Ratings may be retrospective or concurrent.
- When they are retrospective, they tend to summarize all the impressions gathered by the raters regarding the rates over an extended period of time.
- when they are concurrent, they tend to summarize the impressions that are gathered as it happens in the case of interview.

PRECAUTIONS TO BE TAKEN DURING THE CONSTRUCTION OF RATING SCALE

- The rating scale are very easy to construct and use.
- A problem which is frequently noticed in the use of rating scales is that the traits to be rated are not distinctively defined and may have different meaning for different rates.
- This defeats the purpose of rating scales.

three precautions must be taken while constructing the rating scales

- Each trait to be rated should be clearly defined and explained with specific instances.
- Various intervals or points on the scale should be clearly defined.
- Usually five to seven intervals are used in rating trait, attitude and other sentiments.

According to Guilford, rating scales are divided into six categories:

- Numerical rating scales
- Graphic scales
- Percentage rating
- Standard scales
- Scales of cumulated points
- Forced choice scales

Numerical scales:-

- Numerical scales are the easiest to construct and apply to the objects, persons and events etc to be rated.
- In numerical scale the observer or rater is supplied with a sequence of numbers which is well defined and his task is to rate the objects on the given sequence of numbers on the basis of his impression.

Scales with numerical anchors:-

Extremely disagree

■ Strongly disagree

■ Moderately disagree

■ Mildly disagree

■ Indifferent

■ Mildly agree

■ Moderately agree

■ Strongly agree

■ Extremely agree

Demerits:-

- The rater has to remember the meaning of the numerical anchor while he is writing the the numbers in the box.
- The most common observation is that he forgets the meaning of the number and sometimes write the wrong number in the box.

GRAPHIC RATING SCALE

- The graphic rating scale is the most popular and widely used rating scale.
- the graphic scale may be considered as an improvement over the numerical scale because it tends to overcome some of the difficulties faced with the numerical scale.
- On the graphic scale the scales are presented graphically

Advantages of graphic scales:-

- They have no numerical anchors. As such the rater experience no confusion arising out of the need for numerical discrimination.
- They are simple, easily administered and quickly completed by the raters.

Disadvantage:-

- Graphic scales take time and labour in scoring.

PERCENTAGE SCALE

- Percentage rating is done whenever the investigator wants a quick rating with maximum uniformity from rater to rater.
- The technique requires the rater to place the rates among different specified percentage groups or into different percentiles or quartiles

STANDARD SCALE

- Standard scales are not very popular rating scales for psychological measurement.
- In it the rater is presented with some standards with pre-established scale values.
- These standards usually consist of objects of the same kind e.g: they may be the names of persons.
- The rater is then given only those key names with pre-established five point scale values to rate other persons.

Advantages:-

- Avoids confusion arising out of abstract numerical anchors assigned to the traits to be rated.
- If all the raters use the same key-man in their ratings, their ratings can be comparable both in absolute terms as well as in relative terms.
- Since the scale values of all the key men are pre-established and fixed, raters can't shift over day to day's ratings

Disadvantages:-

- Actual practice has demonstrated that no two raters are alike in rating persons who are well known to them.
- Deliberate overestimation and underestimation of persons by the raters are not controlled by the scale.

FORCED CHOICE SCALE

- In this type of rating scale the rater is given a set of attributes in terms of verbal statements for a single item and he decides which one or ones, represent the individual being rated most appropriately and accurately.
- The items of forced choice scale may have several alternatives-two, three, four or five.
- Of these two alternatives and four alternatives are most common.

- In the two alternative form both the statements regarding the attribute are either favourable or unfavourable.
- However, only one of the statement regarding the attribute in the either case is valid to identify desirable or undesirable attributes though both of them may appear equally favourable or unfavourable to the rater.
- In four alternative forms, two varieties are common- in one variety all the four statements are desirable or undesirable and in second variety two statements are favourable and two statements are unfavorable.

FACTORS AFFECTING THE RATER'S ABILITY

- Thorndike and hagen have mentioned (1977) have mentioned the following factors:-
- **Opportunity to observe the individual being rated**
- **Subjectivity in the traits being rated.**
- **Vagueness in the meaning of the trait rated**

■ **Non uniform standard of reference**

■ **Rater's personal characteristics:- e.g:-** A teacher who is annoyed with the behavior of the student in the class is likely to rate him low.

Practical Examination

- It is concerned with the assessment of practical performance skills and practice competency acquired by the student.
- Most essential part of the overall assessment process in nursing

Purposes

- Assess practical skills and competencies
- Development in affective domain
- Assess student problem solving skills
- Recording and reporting skills of students
- To assess the transformation of theoretical knowledge in to practice
- Assess multiple performance tasks

Guidelines for constructing practical examinations

- Planning phase of practical examination
- Conducting phase of practical examination

Planning phase of practical examination

- Consideration of learning objective and learning experience in advance
- Deciding appropriate place for exam
- Familiar place for students
- Ward in charge should be informed
- Over all evaluation criteria should be planned

Conducting phase of practical examination

- Punctuality
- Rules and regulations
- Random assignment
- Routine nursing procedures/psychologically comfortable
- Stress free atmosphere
- Systematic
- Sufficient time
- Discomfort to patient should be avoided
- Marks to examining body with confidentiality

Advantages

- Opportunity to assess skill and competency
- Assessing the use of compartmentalized knowledge
- Assess communication and interpersonal skills

Disadvantages

- Not a standardized assessment practice
- Higher level of subjectivity and personal bias
- Time consuming process
- Not a feasible method
- Emergencies and complex wards-
disturbance factor

viva voce

- Latin phrase-meaning with living voice-often translated as by words of mouth
- Examiner may ask short questions necessarily related to each other.....

Purposes

- Students ability to communicate
- Supplement the information
- Identify and analyze students presence of mind
- Assess students spontaneity and mannerism
- Diagnose students limitation and weakness

Principles to conduct viva

- Range of different issues
- Short questions
- Allow the student to complete
- Should not exceed 30 minutes
- Candidates should be examined individually
- Adequate notice

Advantages

- Direct contact with individuals
- Facilitates simultaneous assessment by two or more examiners

Disadvantages

- Lacks standardization
- Permits favoritism
- Undue influence of irrelevant factors
- Costly in terms of professional time

ANECDOTAL RECORDS

■ DEFINITION:-



It is a brief description of an observed behavior that appears significant for evaluation purposes.

An objective description by the teacher of a significant occurrence or an episode in the life of the student.

@ CHARECTERISTICS:-

➤ A factual description of an event in which:-

- i. How it occurred,
 - ii. When it occurred,
 - iii. Under what circumstances it occurred,
 - iv. The treatment.....
- Each anecdotal should be of one incident.

● PURPOSES:-

- 1) To stimulate teachers to look for information i.e. to help student in self-adjustment.
- 2) The teacher is able to understand her pupil in realistic manner.
- 3) It provides healthy pupil-teacher relationship.
- 4) Helps the student to improve in their behavior, as it is a direct feedback of an entire observed incident.
- 5) Can be used by students for self appraisal and peer assessment.

❁ PRINCIPLES:-

- Concentrate on only one or two behavior observations should be selective.
- Record should be complete.
- They should be kept by all teachers and not only by the student teacher.
- Record the incident or action as soon as possible after it has happened.
- They should have a compiled and field.
- The teacher should have practice and training in making observations and writing Anecdotal Records.
- Name of the student, Class and School, date of observation
- Setting background of the incident.
- Interoperation of the behavior.
- Recommendation concerning the behavior.

● ADVANTAGES:-

- 1) Provision of insight into total behavioral incidents.
- 2) Use of formative feedback.
- 3) Economical and easy to develop.

● DISADVANTAGES:-

- 1) If careless recorded, the purpose will not be fulfilled.
- 2) Subjectivity.
- 3) Lack of standardization.
- 4) Difficulty in scoring.
- 5) Time consuming.

- For example:-

- ✿ Dated : 09-10-2010 Reena is seen in library, sitting in a corner, preparing nursing care plan of cardiovascular disease.

Dated : 13-10-2010 Reena found to be only girl responding to the teacher's questions on cardiovascular Nursing.

Dated : 16-10-2010 Reena was caught reading a book Medical-Surgical Nursing during "Community Health Nursing" and was asked to leave the room.

- ✿ Recommendation: Reena should be encouraged to do mean in Medical Surgical Nursing and also not to neglect Community Health Nursing and then other subjects

OBJECTIVE STRUCTURED CLINICAL EXAMINATION

- **Objective Structured Clinical Examination (OSCEs)** is a form of performance-based testing used to measure candidates' clinical competence.
- During an **OSCE**, candidates are observed and evaluated as they go through a series of stations in which they interview, examine and treat standardized patients (SP) who present with some type of medical problem.

✿ Features of the Objective Structured Clinical Examination (OSCEs):-

- Stations are short
- Stations are numerous
- A pre-set structured mark scheme is used hence.....reduced examiner input and discretion
- Emphasis on what candidates can do rather than what they know
- The application of knowledge rather than the recall of knowledge
- 5 minutes most common (3-20 minutes)/each patient
- (minimum) 18-20 stations/2 hours for adequate reliability
- Written answer sheets or observer assessed using checklists
- Examination hall is a hospital ward
- Atmosphere active and busy

What is assessed by OSCE?

- Various clinical skills – history taking , physical examination, technical procedure, communication, interpersonal skills.
- Knowledge and understanding.
- Data interpretation.
- Problem solving.
- Attitudes.

How to prepare OSCE ?

Harden 12 tips for organizing an OSCE

- What to be assessed?
- Duration of station
- Number of stations
- Use of examiners.
- Range of approach
- New station

Continued...

- Organization of the examination
- Assigning priority
- Resource requirement
- Plan of examination
- Change signal
- Records

The key to a successful
OSCE is careful planning

Selection of the examination committee

- An examination coordinating committee is made up of members who are committed to the evaluative and educational process
- The number of members who make up this committee is not as important as the intensity of the investment of each member.

Continued...

- The examination committee determines the content of the examination, development and implementation.
- It is important that this committee has the capacity and personnel to address decisions related to reliability and validity

Lists of Skills, Behaviors and Attitudes to be Assessed

- The examination will measure objectively the competencies in specific areas of behavior, techniques, attitudes and decision-making strategies based on the objectives of the course or the requirement of the licensing body.

How to develop case/scenario?

- Define the purpose of the station
- Candidate's instructions
- Scoring checklist
- Standardized patient instructions
- Instruction for station set-up

Define the purpose of the station

- State the skill and domain to be tested

Skill –Physical examination .

Domain –

history collection

physical examination

NG tube feeding

Candidate instructions

- Candidate instruction must be clear and concise.
- Before examination a briefing about whole system is very much effective for a successful OSCE.

Scoring checklist

- The checklist should be complete and include the main components of the skill being assessed.
- Any unnecessary or exaggerated stem must be avoided.

Instruction for station set-up

- List of all equipments required for the station

The Examinees

- The examinee is the student, resident, or fellow in training or at the end of training of a prescribed course

The Examiners

- Most stations will require an examiner, although some stations do not.
- The examiner at the station where clinical skills (history-taking, physical examination, interviewing and communication) are assessed, preferably be either a physician or a standardized patient.

The Examination Site

- The examination site is part of a special teaching facility in some institutions. When such facilities are not available, the examination may be conducted in an outpatient facility .

Examinations Station

- The total number of stations will vary based on a function of the number of skills, behaviors and attitudinal items to be tested.
- For most clerkships or courses, the total will vary from 10-25.

Stations..

- The number of stations in an examination refer the time allocated for each station determines the time required to complete the whole examination.

Duration of station

- Duration of stations has been fixed
- Make sure that the task expected of the student can be accomplished within the time
- If necessary some stations which are allocated double the standard time. Such double stations will require to be duplicated in the examination.

Duration of station

- Times ranging from 4 to 15 minutes have been reported in different examinations and a five minute station probably most frequently chosen.
- This times depend to some extent on the competencies to be assessed in the examination.

Couplet Station

- Some competencies may best be assessed by coupled or linked stations.
 - The use of linked stations extends the time available to complete a task.



Observer Assessment Method

- Checklist
- Rating scale

Check list for assessment of physical finding

- Mr.C. presents with a sore swollen ankle for 6 weeks

	Don't	Do
1-introduces self to patient		
2-Explain to the patient what will be do		
3-Demonstrate concern for patient.i.e.is not excessive rough		
4-Inspectin for any of swelling , erythema ,deformity		
5-Inspection: Standing From anterior Posterior		
6- Inspection pt Gait		
7- palpation		

Communication skills checklist (rating scale)

	Poor 1	Fair 2	Good 3	V Good 4	Excellence 5
1- Interpersonal skill: Listen carefully					
2-Interviewing skill: Uses words patient can understand organized					

Running The Exam

- Space requirements
- Signaling station change
- Collecting result
- Budget

Advantages of OSCE

- Provides an opportunity to test a student's ability to integrate knowledge, clinical skills, and communication with the patient
- Less complexity.
- Valid examination.
- Summative and well formative.
- Time limit for each station.

- Can be used with large number of students.
- Reproducible. **Advantages**
- Provides unique programmatic evaluation.
- More objective.
- Test not only skills and knowledge but attitudes also.

Disadvantage

- Development and administration are time consuming and costly.
- Provides assessment of case-specific skills, knowledge, and/or attitudes
- Knowledge and skills are tested in compartments

Disadvantage

- Rely on task specific checklists and scoring.
- Needs more observation skills of staff.
- Standardization of simulated patients and examiners.
- Repetitive and boring.

STEPS: -

- Registration
- Orientation
- Escorting to exam position
- Station Instruction Time
- The Encounter
- Post Encounter Period
- Repeat Steps 4 to 6
- Exam ended / Escorting to dismissal area

● ATTITUDE SCALE:-

- Used for measuring social attitudes.
- Questionnaires is prepared, an object or system will be allotted for each item.
- Some relevant or direct statements will also be used to reveal the attitude.

◆ TYPES:-

- 1) Point scale
- 2) Differential scale
- 3) Summated or likert scale
- 4) Scalogram
- 5) Semantic scale

POINT SCALE

What is your overall satisfaction with our product?

**Not at all
satisfied**

**Extremely
satisfied**

What is your overall satisfaction with our product?

**Not at all
satisfied**

1 2 3 4 5

**Extremely
satisfied**

What is your overall satisfaction with our product?

1 2 3 4 5

What is your overall satisfaction with our product?

**Not at all
satisfied**

**Slightly
satisfied**

**Moderately
satisfied**

**Very
satisfied**

**Extremely
satisfied**

Likert scale

- Measures attitude , feeling, value of people.
- It is developed with 5 point likert scale
- Mixture of positive and negative declarative statements

Uses

- Measures attitude, feeling and values
- Quantified measurement of quantified attribute
- Assess the opinion of the people

Characteristics

- Psychological measurement tool
- Illustrative in nature
- Neutral statements
- Bipolar scaling method
- Measurement of specific number of scaling objectives

■ Scoring

Guidelines

- Five point or seven point can be constructed
- Statement should be brief, concise
- each statement should deliver one thought
- Simple language

Advantages

- Easy
- More reliable tool
- East to administer
- Less time consuming

Disadvantages

- preplanned statement and categories

Likert Scales

Please circle the number that represents how you feel about the computer software you have been using

I am satisfied with it

Strongly Disagree ---1---2---3---4---5---6---7--- Strongly Agree

It is simple to use

Strongly Disagree ---1---2---3---4---5---6---7--- Strongly Agree

It is fun to use

Strongly Disagree ---1---2---3---4---5---6---7--- Strongly Agree

It does everything I would expect it to do

Strongly Disagree ---1---2---3---4---5---6---7--- Strongly Agree

I don't notice any inconsistencies as I use it

Strongly Disagree ---1---2---3---4---5---6---7--- Strongly Agree

It is very user friendly

Strongly Disagree ---1---2---3---4---5---6---7--- Strongly Agree

Semantic differential scale

- Most effective and most widely used technique
- It measures attitude towards stimulus, words, objects and concepts
- It consists of series of contrasting adjective pairs listed on opposite ends of a bipolar scale
- Scale is popular and easy to construct.

Semantic Differential Scale

For each pair of adjectives place a cross at the point between them which reflects the extent to which you believe the adjectives describe policemen

clean	:	:	:	:	:	:	:	:	dirty
honest	:	:	:	:	:	:	:	:	dishonest
kind	:	:	:	:	:	:	:	:	cruel
helpful	:	:	:	:	:	:	:	:	unhelpful
fair	:	:	:	:	:	:	:	:	biased
strong	:	:	:	:	:	:	:	:	weak
foolish	:	:	:	:	:	:	:	:	wise
energetic	:	:	:	:	:	:	:	:	lazy
unreliable	:	:	:	:	:	:	:	:	reliable

(Robson, 1993)

Definition

- Semantic differential scale is a type of rating scale designed to measure the connotative meaning of object, events and concepts.

Advantages

- Convenient method to assess attitude, belief and values
- Easy to administer
- Provides reasonable data

Disadvantages

- Difficult to select the relevant concepts
- Time consuming



Thank you...