



CONSTRUCTING EFFECTIVE MULTIPLE CHOICE QUESTIONS (MCQ's)



Alfred Tenore

*Prof. and Chairman of Paediatrics, University of Udine, Italy
Vice President, European Academy of Paediatrics (CESP)
President, European Board of Paediatrics (EAP/CESP)*

INTRODUCTION

Outline

1. Assessment - Principles that guide assessment
2. Purpose of "Testing"
3. What should be "tested"
4. Issues of "Sampling"
5. Types of test item formats
6. Writing effective questions

ASSESSMENT - AN IMPORTANT COMPONENT OF INSTRUCTION

1. Assessment is a critical component of Instruction
2. Tests are a powerful motivator
3. Students will learn what they believe you value.
4. Because tests have such a powerful influence on student learning, it is important to develop tests that will further the educational goals
 - a. Hands-on clinical skills test drives students out of the library into the clinic for learning
 - b. Tests assessing only recall of isolated facts, drives them to "cram" course review books

PURPOSE OF TESTING

FORMATIVE assessment

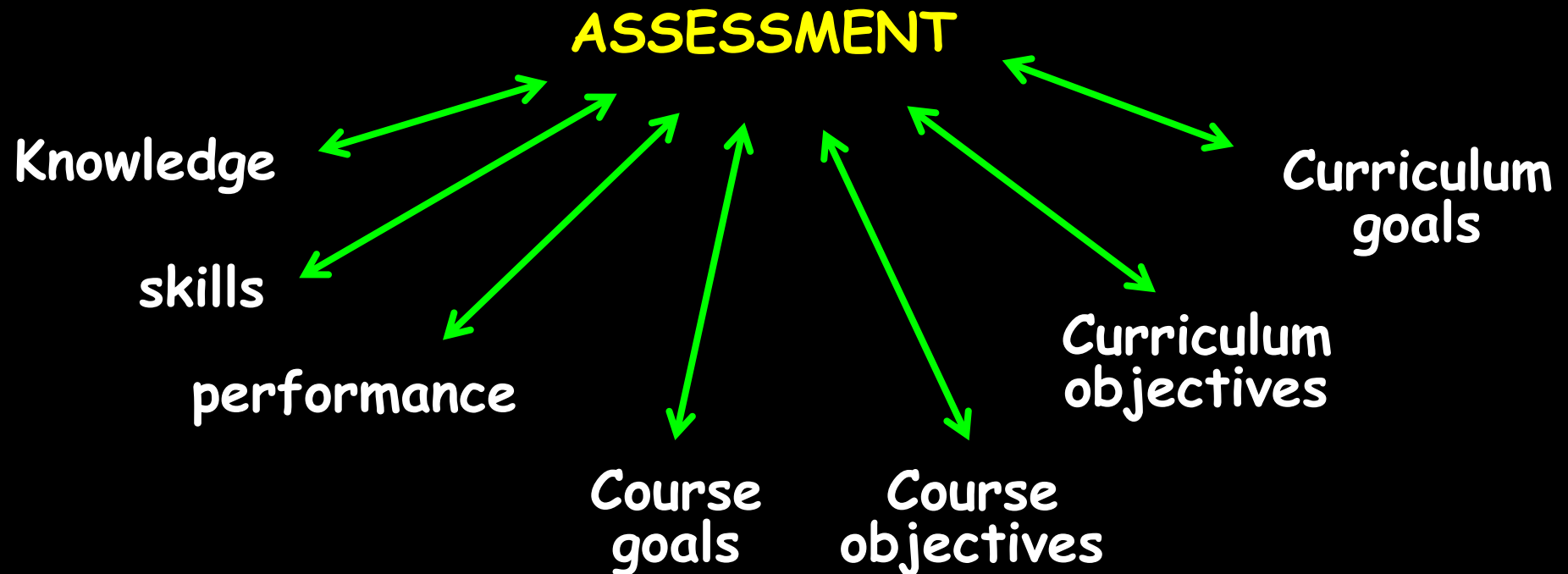
1. Communicate to students what material is important
2. Motivate Students to study (while in school)
Promote life-long learning (in practice)
3. Identify areas of deficiency in need of remediation or further learning
4. Identify areas where the course curriculum is weak

SUMMATIVE assessment

1. Provide basis for grading, passing, graduation

IMPLICATIONS OF ASSESSMENT

Assessment is an essential tool in meeting educational goals



Exam content should match course/clerkship objectives

What Should Be Assessed ?

“Goals of the Educational Process”

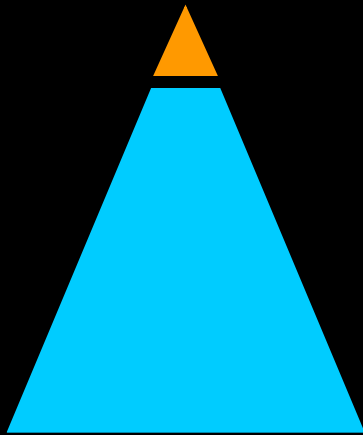
Cognitive Domain

Affective Domain

Psychomotor Domain

What Should Be Assessed ?

2. Bloom's* Taxonomy of Cognitive Levels of "Knowledge"



- a. Memory
- b. Comprehension
- c. Application
- d. Analysis
- e. Synthesis
- f. Evaluation

Intellectual abilities

(* Bloom, BB: "Taxonomy of Educational Objectives: The Classification of Educational Goals, by a committee of college and university examiners". 1st Ed. New York: Longmans, Green, 1956)

ISSUES OF SAMPLING

The purpose of any assessment is to permit inferences to be drawn concerning the **skills of the examinees**.

1. With MCQ you first need to decide what you want to include on the test.
2. The amount of attention given to evaluating something should reflect its relative importance
3. You need to sample **topics** and also sample **skills**:
 - a. Determine the diagnosis
 - b. Deciding on the next step in management

NB. You **CANNOT** ask everything !

ISSUES OF SAMPLING

4. The nature of the samples determines the extent to which the estimate of true ability is:
 - a. reproducible (**reliable**)
 - b. accurate (**valid**)

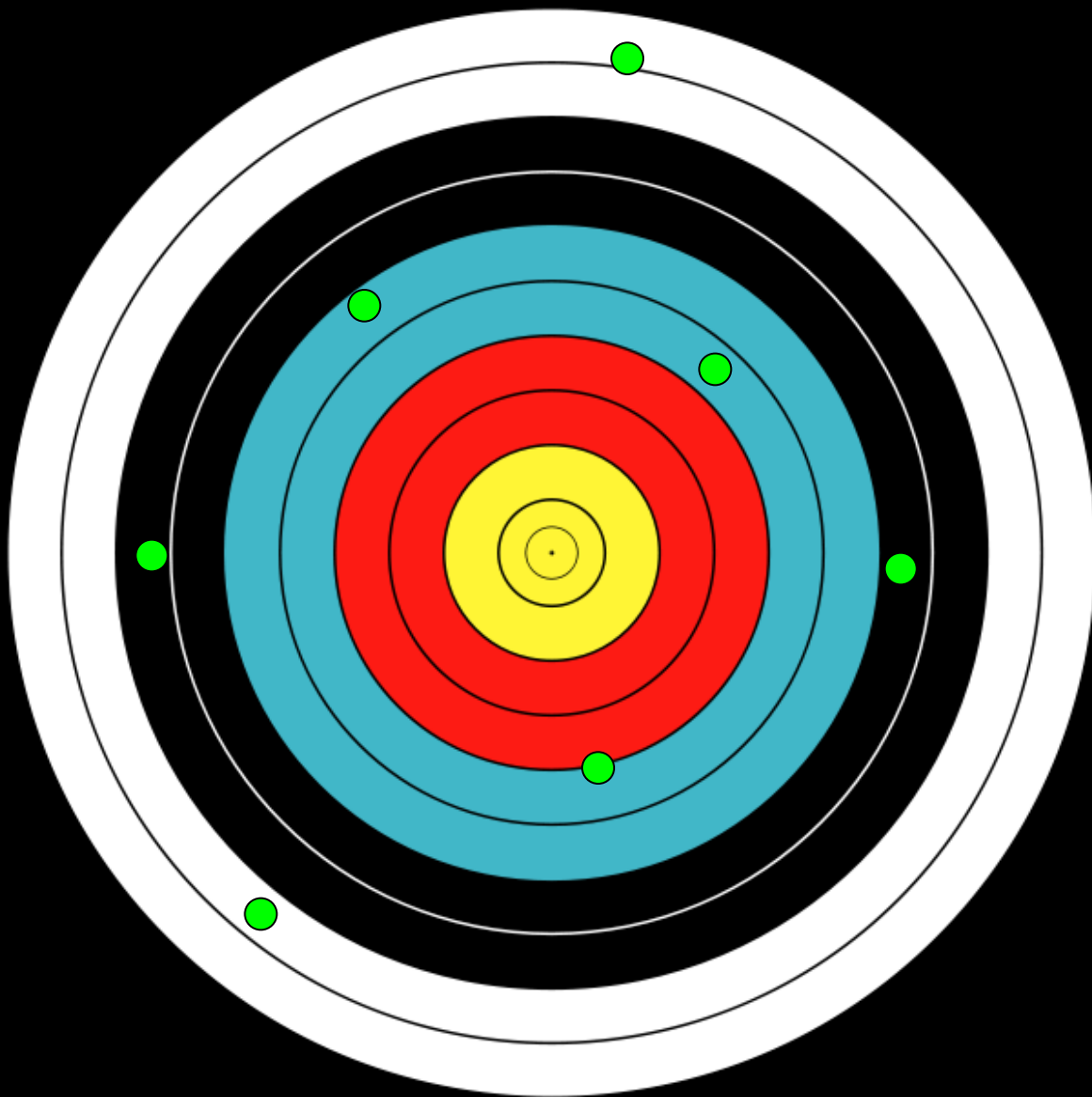
Reliability Score accuracy/stability

- a. Would the scores be reproduced if tested again ?
- b. Would the scores be reproduced by different raters?

Validity Score meaning

- a. Does the score measure what you intend to measure ?

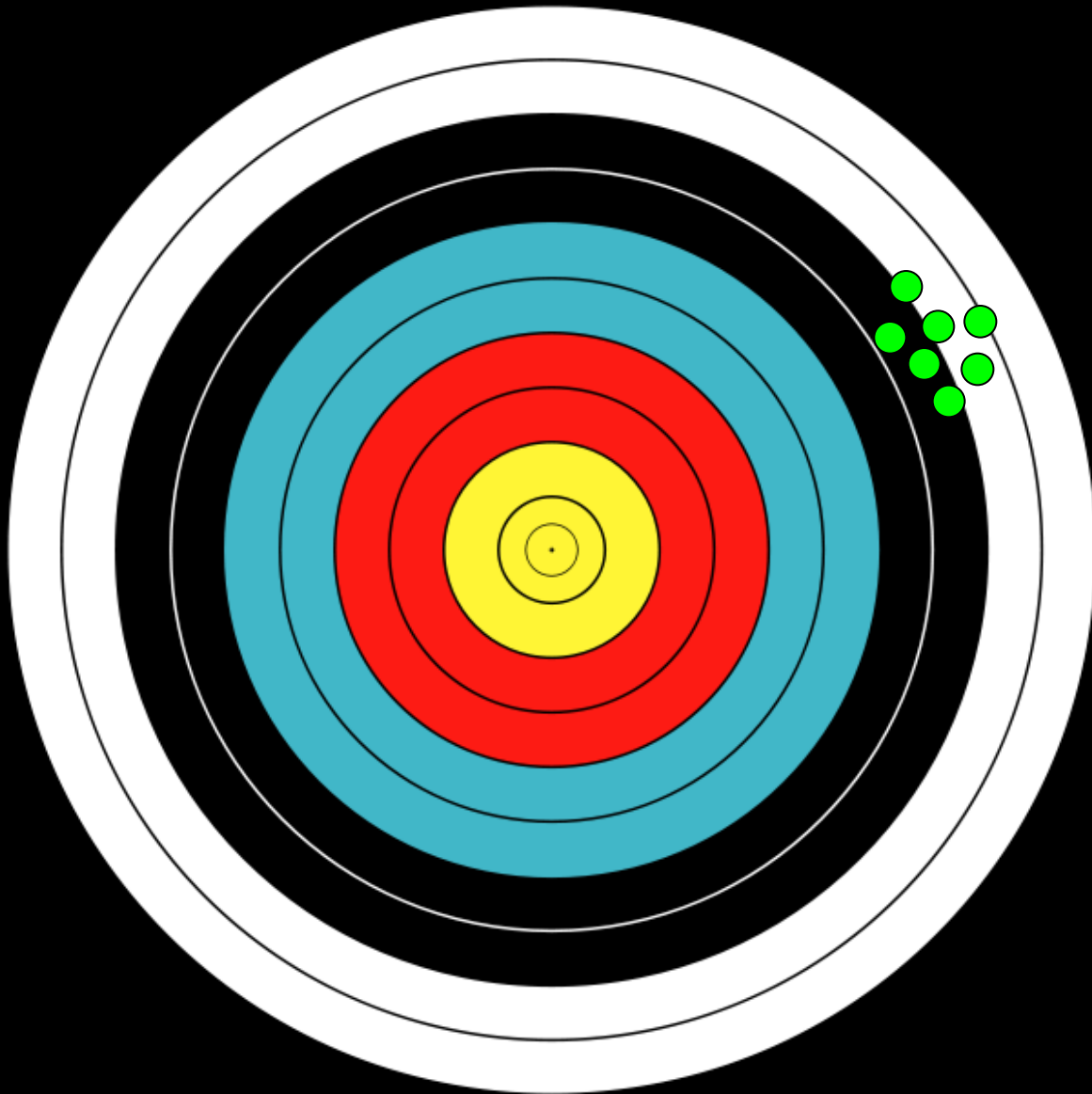
Is the Question **Reliable** & **Valid** ?



Reliability : **Poor**

Validity : **Poor**

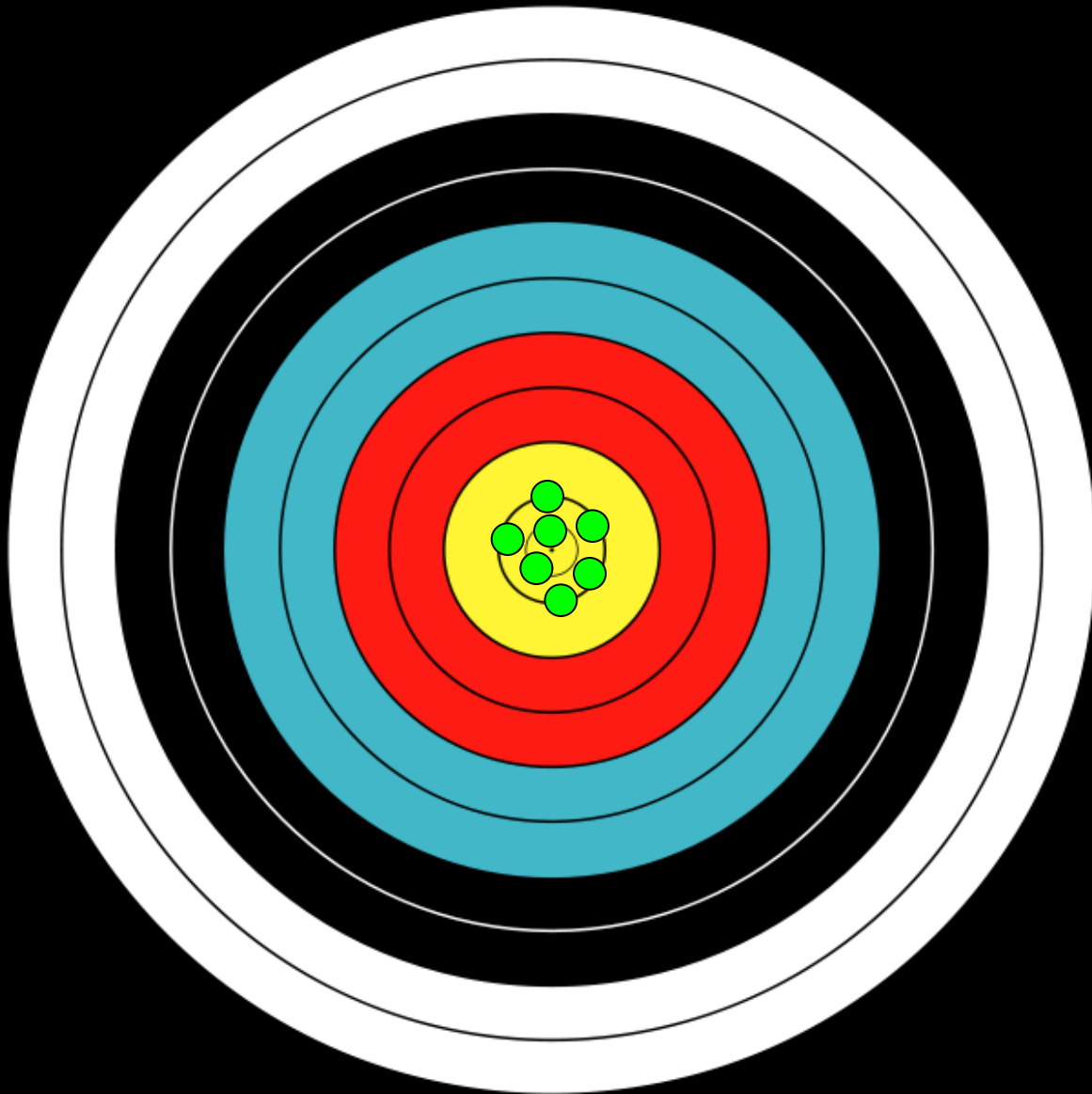
Is the Question **Reliable** & **Valid** ?



Reliability : **Good**

Validity : **Poor**

Is the Question **Reliable** & **Valid** ?



Reliability : **Good**

Validity : **Good**

ISSUES OF SAMPLING

5. If the sample is too small, exam results may not be sufficiently precise (**reproducible, reliable**) to ensure that they reflect true proficiency:
 - a. In order to generate a reproducible score, you need to sample content broadly:
 - ① 12 or more **cases**
 - ② 100 or more **MCQs**

TESTING FORMATS

1. No single testing method is likely to assess all the Cognitive levels of "knowledge"
2. Basic Test Item formats
 - a. Multiple-choice
 - b. True-False
 - c. Matching
 - d. Completion - short answer
 - e. Essay
 - f. Demonstration
3. What we will focus on is the writing of high-quality multiple choice questions that assess skill in interpreting data and making decisions

The four Rules to Follow in Writing Questions

1. Choose an important topic
2. Write your stem as if it were a question on a short answer written test
3. Make your choices homogeneous
4. Use straightforward prose and avoid technical flaws

The four Rules to Follow in Writing Questions

1. Choose an important topic
2. Write your stem as if it were a question on a short answer written test
 - a. You should be able to cover up the choices, read the stem, and write down a short answer
 - b. In general, the stem should be long, the choices short.
3. Make your choices homogeneous
 - a. The choices should all be similar in length, grammar and syntax and theme
 - b. You should be able to order your choices from least to most true along a single thematic dimension, with the correct answer being that which you would have written down as a short answer
4. Use straightforward prose and avoid technical flaws
 - a. don't make the stem unnecessary long or complicated
 - b. Keep the choices simple and in a logical order (where possible)
 - c. Don't use imprecise terms (*frequently/rarely*) & avoid absolutes

MULTIPLE-CHOICE-ITEM FORMATS

2. All of the multiple-choice questions can be divided into two families of items :

Those that require the examinee to indicate

- a. all responses that are appropriate (true/false)
- b. a single response (one best answer)

NB. Many variants of these 2 subgroups exist

The TRUE / FALSE FAMILY

1. T/F items require an examinee to select all the options that are "true"
 - a. The examinee must decide where to make the cut-off (to what **extent** must a response be "true" to be "correct")
 - ① While this task requires additional judgment (with respect to the one-best-answer), this additional judgment may be unrelated to clinical expertise or knowledge
 - ② Too often examinees have to guess what the item writer had in mind because the options are not either **completely true** or **completely false**

Example of **FLAWED TRUE / FALSE** ITEM

True statements about cystic fibrosis (CF) include ?

1. The incidence of CF is 1:2000
2. Children with CF usually die in their teens
3. Males with CF are sterile
4. CF is an autosomal recessive disease

Options 1, 2, 3 cannot be judged as absolutely true or false
(a group of experts would not agree on the answers)

In thinking about Option 1, note that the incidence is not exactly 1:2000
(experts would want more information: Is this in Italy?
Is this among all ethnic groups?)

Modifying the language to "approximately 1:2000" does not help
because of lack of specification

Option "4" is the only clear option

Example of **FLAWED TRUE / FALSE** ITEM

The way to a man's heart is through his:

1. aorta
2. pulmonary arteries
3. pulmonary veins
4. stomach

Illustrates a common problem: items for which the stem is unclear

Depending on your perspective: Options 1, 2 and 3 may be true

Alternatively: Options 1, 2 and 3 may be false and 4 is true

The "Family" of the SINGLE-BEST - ANSWER

1. Which of the following statements concerning pseudogout is true?
 - A. It occurs frequently in women
 - B. It seldom associated with acute pain in a joint
 - C. It may be associated with a finding of chondrocalcinosis
 - D. It is clearly hereditary in most cases
 - E. It responds well to treatment with allopurinol

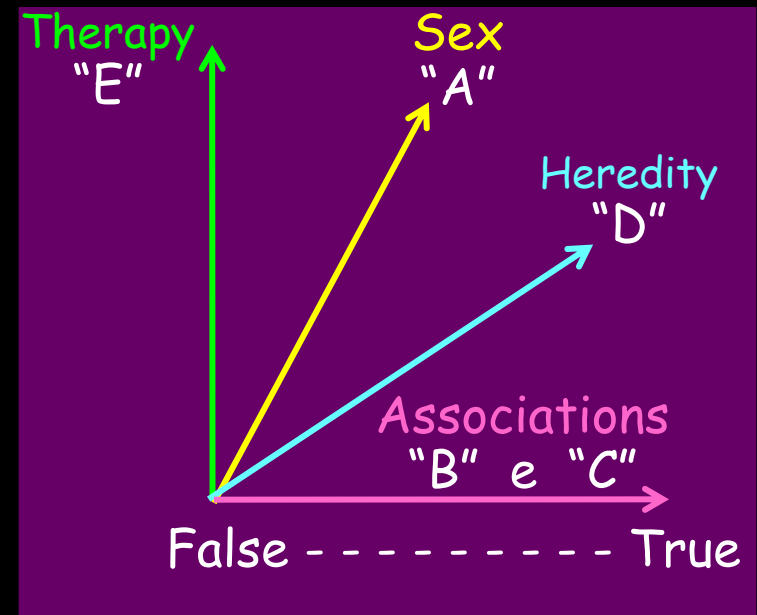
In the composition of this item there are problems !

1. In attempting to determine the "best" answer, the test taker has to decide whether its occurrence "frequently in women" is more or less true with respect to the fact that "it is seldom associated with acute pain in a joint"
2. This is a comparison between "apples" and "potatoes"
3. In order to be able to evaluate the progressive order of the relative correctness of the options, these must differ in one single dimension, otherwise all of the options must be either 100% true or false

The "Family" of the **SINGLE-BEST - ANSWER**

1. Which of the following statements concerning pseudogout is true?
 - A. It occurs frequently in women
 - B. It seldom associated with acute pain in a joint
 - C. It may be associated with a finding of chondrocalcinosis
 - D. It is clearly hereditary in most cases
 - E. It responds well to treatment with allopurinol

1. A schematic representation of these **options** could be the following →
2. The options are **heterogeneous** and reflect different factors that cannot be evaluated in a progressive order from "false" to "true" in a single dimension.
3. Although this question may appear to assess the knowledge of a variety of different things, the question is not clear!



The question **CANNOT** be answered without looking at the options !

What "Rules" are being Violated ?

1. Which of the following is true about pseudogout ?
 - A. It occurs frequently in women
 - B. It seldom associated with acute pain in a joint
 - C. It may be associated with a finding of chondrocalcinosis
 - D. It is clearly hereditary in most cases
 - E. It responds well to treatment with allopurinol

It violates rules 2, 3 and 4

2. There isn't a clear question in the stem and you are forced to look at the choices to make any sense of it
3. The choices don't fall along a single dimension (you are asked to compare apples and oranges, or in this case: gender, location, pathology, genetics and pharmacology)
4. to repair this question you would have to rewrite the stem and focus the choices on one theme, you would also have to be careful about terms like "frequently" and "clearly" as they are frequently unclear

Characteristics of **FLAWED TRUE / FALSE** ITEMS

In each of the previous "flawed" samples:

- the stem is unclear

- the options contain vague terms

- the options are "partially" correct

In each instance, a group of experts would have difficulty reaching a consensus on the correct answer

Example of **ACCEPTABLE TRUE / FALSE** ITEM

Which of the following is/are X-linked recessive conditions ?

1. Hemophilia A (classic hemophilia)
2. Cystic fibrosis
3. Duchenne's muscular dystrophy
4. Tay-Sachs disease

The options can be diagrammed as follows

2
4

1
3

Totally
wrong
options

Totally
correct
options

The TRUE / FALSE FAMILY

2. Because T/F items require an examinee to select all the options that are "true", these types of questions must satisfy the following rules:
 - a. Stems must be clear and unambiguous.
 - ① AVOID using **imprecise** phrases as:
 - is associated with
 - is useful for
 - is important
 - ② AVOID words that provide "**cueing**", such as
 - may
 - could be
 - ③ AVOID using **vague** terms such as
 - usually
 - frequently
 - b. Options must be **absolutely** true or false (no shades of gray are permissible)

It is Recommended **NOT TO USE** "true/false" type questions

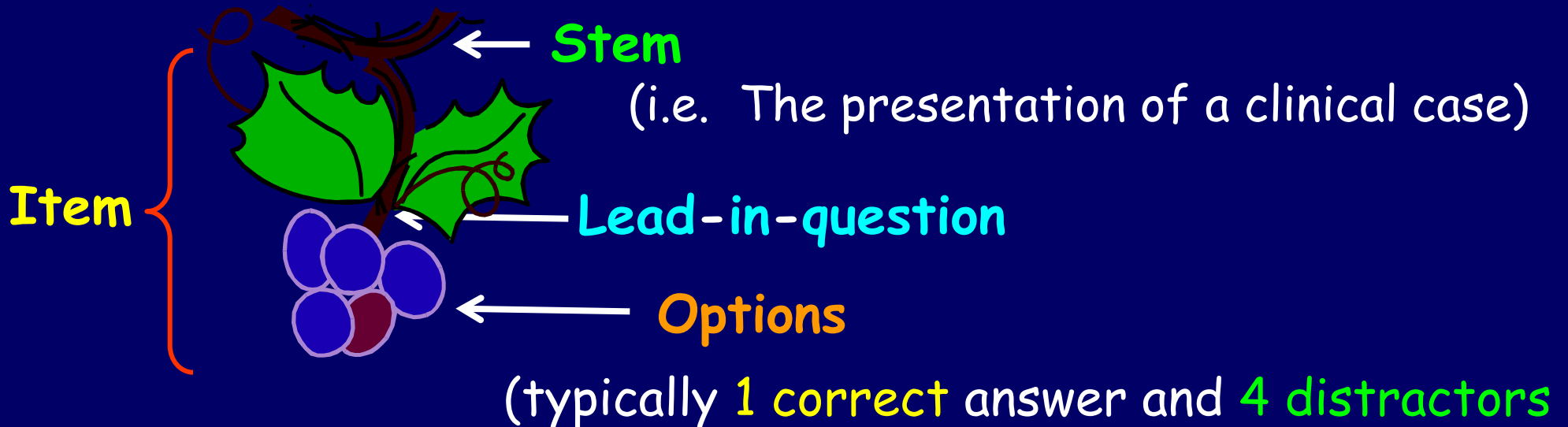
1. Even though many believe that they are easier to write, they are the most problematic
2. Often the distinction between "true" and "false" is not clear, and the reviewers of the questions often change the key of the answers
3. Tend to emphasize "recall" of isolated information.
4. Most options do not fall at the extremes of the "truth" continuum

It is Recommended **NOT TO USE** "true/false" type questions

5. The person who writes the question has something specific in mind but a careful revision usually uncovers subtle difficulties which were not evident to the writer of the question
6. Probably, the most important reason is that :
 - a. To avoid ambiguity, we are pushed towards evaluating the recall of a **single piece of information** (*that which should be specifically avoided*)
 - b. The use of questions which assess "**knowledge**", "**integration**", "**synthesis**" and "**judgment**" may best be evaluated with the "**one-best-answer**" type question

TERMINOLOGY USED IN MCQs

ITEM	The complete question with all its answers
STEM	Everything that comes before the question
LEAD-IN-QUESTION	The specific question
OPTIONS	All of the answers, both correct and wrong
DISTRACTORS	The wrong answers
KEYED ANSWER	The single best answers indicated as correct



The "Family" of the **BEST - SINGLE - ANSWER**

Stem:

A 32 yr old man complains of a progressive muscular weakness of the extremities over the last 4 days. He has always been in good health. 10 days ago he had an upper respiratory tract infection.

On physical exam, his temp was 37.8°C, BP: 130/80 mmHg, HR 94/min and RR 42/min with superficial respirations. He is noted to have symmetrical weakness on both sides of the face and of the proximal and distal muscles of the extremities. The sensibility is intact. The deep tendon reflexes are absent; the plantar reflex is flexion.

Lead-in :

Which of the following is the most likely diagnosis?

Options :

- A. Acute disseminated encephalomyelitis
- B. Guillain-Barré Syndrome
- C. Myasthenia gravis
- D. Poliomyelitis
- E. Polymyositis

The "Family" of the BEST-SINGLE-ANSWER

Stem:

1. Note that the "wrong" options are not totally wrong.
(the options could be schematized as follows:)

D

C

A

E

B

the LEAST
correct

the MOST
correct

Opzioni :

- A. Acute disseminated encephalomyelitis
- B. Guillain-Barré Syndrome
- C. Myasthenia gravis
- D. Poliomyelitis
- E. Polymyositis

The "Family" of the **BEST - SINGLE - ANSWER**

- A. Acute disseminated encephalomyelitis
- B. Guillain-Barré Syndrome
- C. Myasthenia gravis
- D. Poliomyelitis
- E. Polymyositis

1. Even if the wrong answers are not completely wrong, they are less correct than the established answer
2. The examinee is asked to choose, "**the most probable diagnosis**"
3. Experts would all agree that the most probable diagnosis is "**B**"; however, they would also agree that the other diagnoses could be considered
4. As long as the options can be placed along a single continuity line (in this case from the least common to the most common) they cannot be totally wrong.

Advantages and Disadvantages between T/F and MCQ

MCQ

T/F

Advantages

Can measure all levels of student ability
Enables wide sampling of subject content
Quick and easy to score
Enables objective scoring
Can be analyzed for effectiveness

Efficient for testing large sample of information
Enables efficient and objective scoring

Disadvantages

Difficult to construct good items
Tendency to measure simple recall

Permits high guess factor
Difficult to construct effective items

The five Rules to Follow in Writing Questions

1. Choose an important topic
2. Write your stem as if it were a question on a short answer written test
3. Make your choices homogeneous
4. Use straightforward prose and avoid technical flaws
5. Make your questions test the **application of knowledge** rather than recalling a single fact !
 - a. Having the stem first describe an experimental/clinical situation and then asking for the result/diagnosis can accomplish this

MULTIPLE-CHOICE-ITEM FORMATS

1. In order for a test question to be a good one and to generate valid scores it must satisfy two basic criteria:
 - a. Address **important content**
 - ① Item content is of critical importance and therefore an essential condition
 - ② However, alone, it is not sufficient to guarantee that the test question is a good one
 - ③ Items that attempt to assess critically important topics cannot do so unless they are **well-structured**
 - b. Avoid "flaws" (*that benefit the **testwise** examinee*) and irrelevant difficulty

SUMMARY of TECHNICAL ITEM FLAWS

Issues Related to "Testwiseness"

1. Grammatical clues

One or more distractors does not follow grammatically the structure of the "stem"

2. Logical clues

One of the subgroups of the options is exhaustively complete

3. Absolute terms

The finding of terms such as "always" or "never" in some options

4. Long correct answer

The correct answer is longer, more specific and more complete than the others

5. Repetition of words

A word or phrase is included in the "stem" and in the right answer

6. Convergence strategy

The correct answer includes the majority of the elements which are found in the other options

Guiding Principles in Assessment

Principles

1. Learning

“Assessment drives learning”

2. Content

■ Focus on important and relevant information

3. Integration

■ Promote integration of knowledge and skills across disciplines

4. Validity

■ Promote valid and reliable data on high stakes performance

“Examinations are formidable even to the best prepared,
for the greatest fool may ask more
than the wisest man can answer”

Charles Caleb Colton (1780?-1832)

Thank you for your Attention

