

Successfully Writing Multiple Choice Questions (MCQ's)

Alfred Tenore

Vice Chair, CESMA-UEMS

Past President, European Academy of Paediatrics (EAP-UEMS)

Chair, Global Pediatric Education Consortium (GPEC)

President, National Medical Licensure Examination Board , Ministry of Education, Italy

*Prof. of Pediatrics, Senior Associate Dean for Medical Education
and Chair of the Department of Medical Education
California University of Science and Medicine, CA, USA*



Outline

I Introduction

- A. Assessment
- B. MCQ Format

II Multiple Choice Questions

- A. "Auto-analysis"
- B. Terminology Used in MCQ's
- C. Criteria for writing "good" test items

- 1. Important
- 2. Clear
- 3. No Tricks
- 4. No Clues

} "Take Home Messages"

III. Conclusions

- A. Checklist for writing single-best answer MCQ's

Introduction

Assessment

Allows inferences to be drawn concerning the **knowledge and skills** of the examinees.

NB. You CANNOT ask everything !

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

Introduction

Assessment

2. The nature of the MCQ's determines the extent to which the estimate of true ability of the examinee is:
 - a. Reproducible (reliable - accuracy of the score)
 - ① Would the scores be reproduced if tested again and by different raters ?
 - b. Accurate (valid - meaning of the score)
 - ① Does the score measure what is intended to be measured ?

Introduction

Assessment

3. Size of sample

- a. If too small, exam results may not be sufficiently precise (**reproducible reliable**) to ensure that they reflect true proficiency (**validity of the score**)
- b. In order to generate a reproducible score, the content of the discipline needs to be sampled broadly.
 - ① NEVER less than 100 MCQ's

Multiple Choice Items

Formats

All MCQ's can be divided into items that require the examinee to indicate:

- a. All responses that are appropriate (True / False)
- b. A single response (one best answer)

Multiple Choice Questions (MCQ's)

MCQ's : **effective & efficient** way to assess learning outcomes (LO)

(statements that identify what the learner will **know** and be able to **do** by the end of a program)

■ **Versatile**

Applicable to assess various **levels** of learning outcomes

Recall

→ Application

→ Analysis

→ Evaluation

Shortcoming :

Not effective way to test ability to articulate **explanations** or **creative ideas**

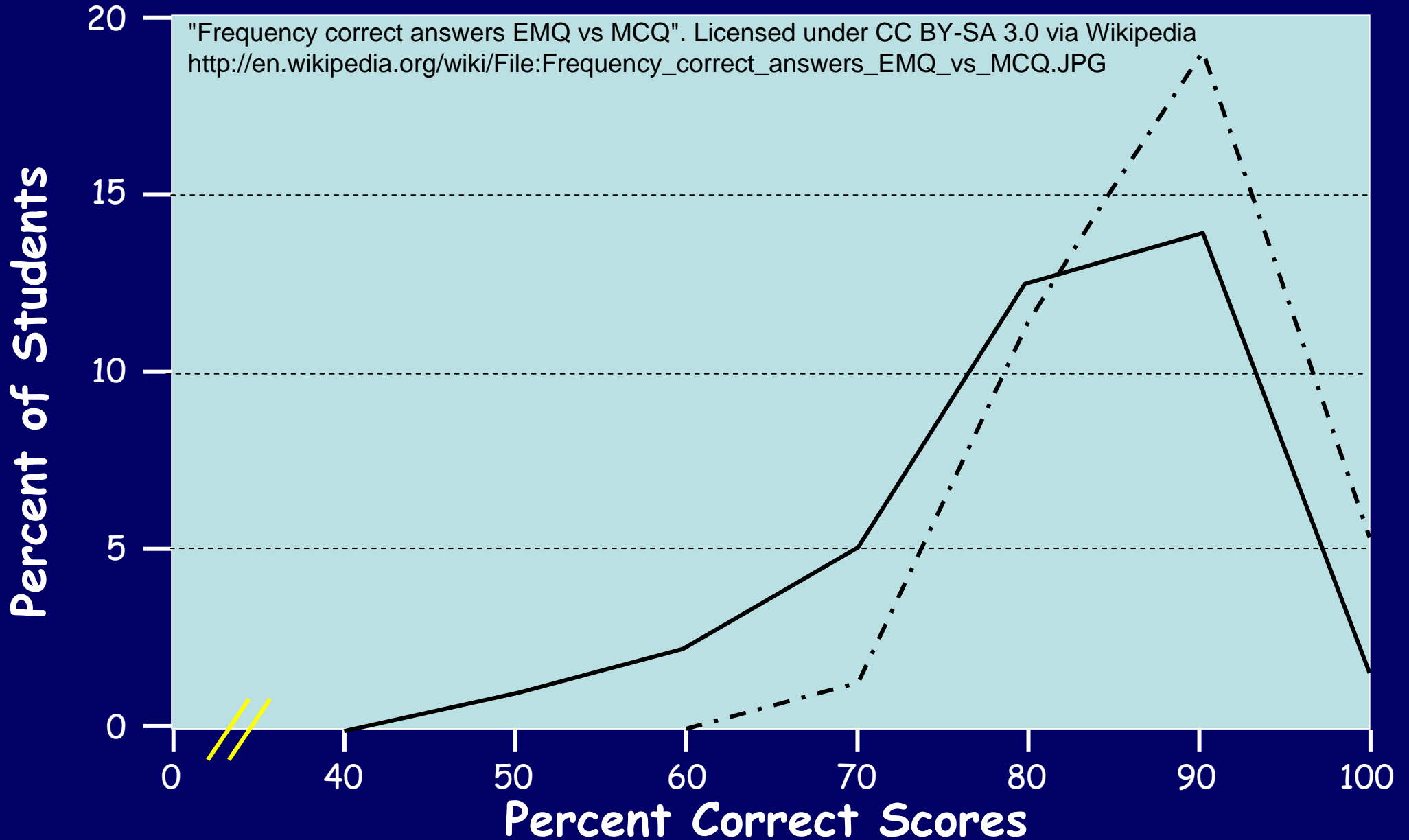
■ **Reliable**

The ability of a test to **consistently measure** learning outcomes
Less susceptible to guessing → More reliable means of assessment
(related to number of options)

■ **Valid**

Frequency Distribution of Scores

Comparing sets with 5 (---) or 15 (—) options



Multiple Choice Questions (MCQ's)

MCQ's : **effective & efficient** way to assess learning outcomes (LO)

(statements that identify what the learner will know and be able to do by the end of a program)

■ **Versatile**

Applicable to assess various levels of learning outcomes

Recall

→ Application

→ Analysis

→ Evaluation

Shortcoming :

Not effective way to test ability to articulate explanations or creative ideas

■ **Reliable**

The ability of a test to consistently measure learning outcomes

Less susceptible to guessing → More reliable means of assessment
(related to number of options)

■ **Valid**

Ability to measure learning outcomes that it claims to measure.

Typically focuses on a relatively broad representation of material



increasing the validity of the assessment

Successfully Writing Multiple Choice Questions (MCQ's)

What we **are** saying

vs.

What we **want** to mean



Which of the two is important
in writing MCQ's

?

Successfully Writing Multiple Choice Questions (MCQ's)

Definition:

"**Success**" "Achieving a desired **aim**"

└── Writing MCQ's

In writing MCQ's, the "**aim**" is to make sure we write an item that has the **characteristic components** of a "MCQ"

Components of a MCQ

a **question** with a number of answer options of which 1 is correct

Successfully Writing Multiple Choice Questions (MCQ's)

What we **are** saying



make sure we write
an item
that **has the components**
of a "MCQ"

vs.

What we **want** to mean



make sure we write
an item
that **is written well**

"Successful"

has no bearing on whether the MCQ written is **"good"** or **"bad"**!

Writing **Good** Multiple Choice Questions (MCQ's)

Take Home Message

We must make sure that the question is **clear** and **unambiguous**

Writing **GOOD** Multiple Choice Questions (MCQ's)

In order to write "good" MCQ's

we need to :

- ■ identify the **various components** that make up a MCQ

The complete question
with all its answers

← **Cordon** (Stem of the MCQ)

← **Stem** (Lead-in-question)

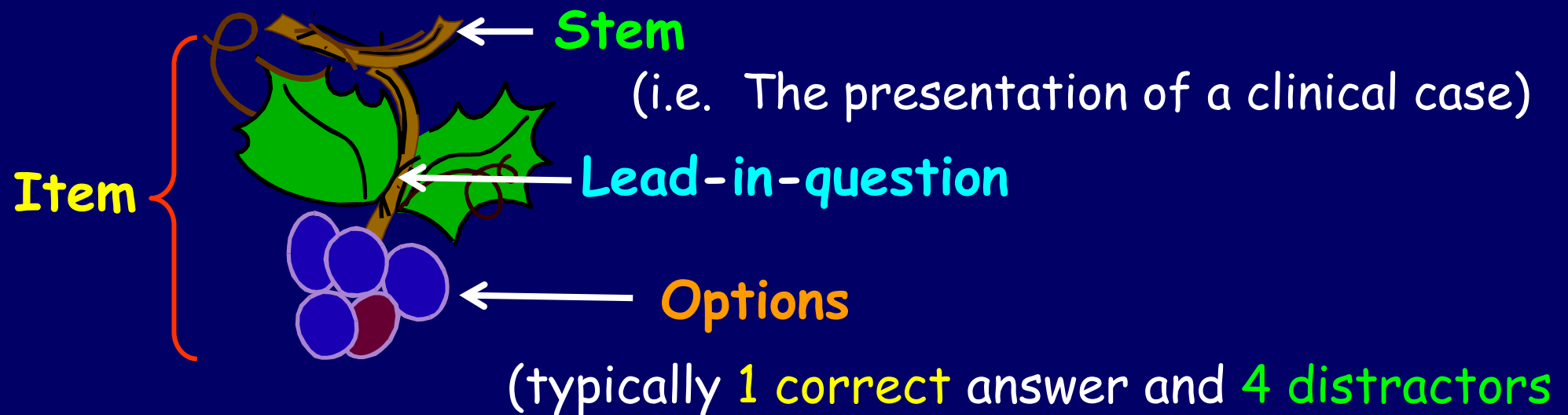
Item

← **Grapes**
(Options: correct and incorrect answers)

**Paradigm
of a MCQ**

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 incorrect
DISTRACTORS	The incorrect answers
KEYED ANSWER	The single best answers indicated as correct



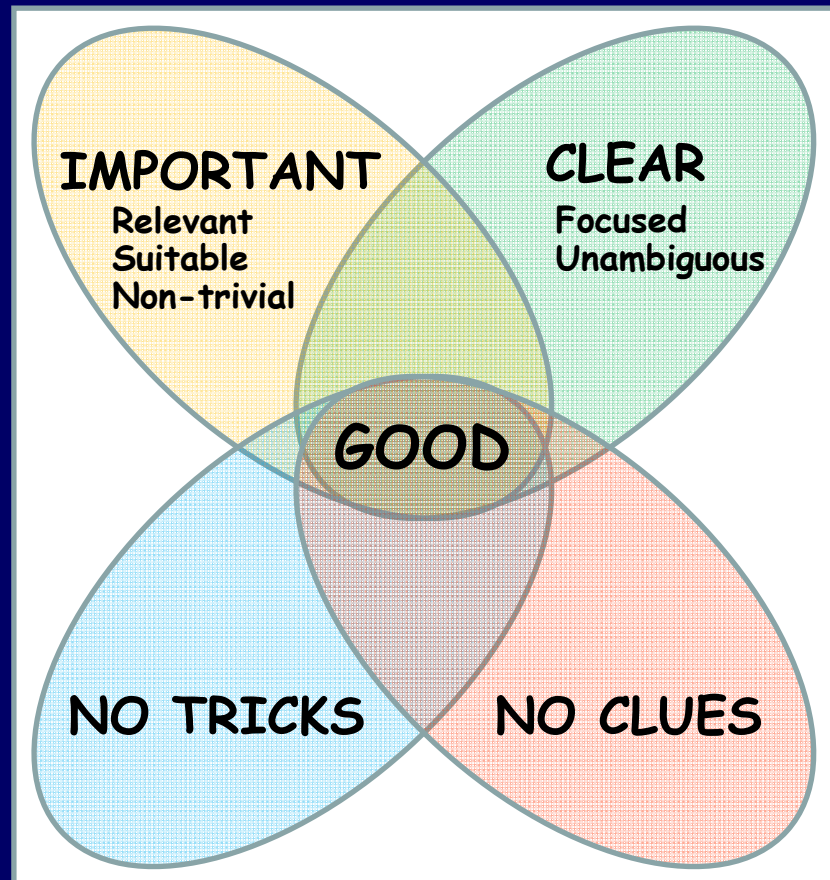
4 Criteria for Good Test Items (MCQ's)

IMPORTANT

Provides examinee with a realistic (clinical) situation in which knowledge and/or judgment must be applied

NO TRICKS

Avoid tricks that may cause examinees to select **incorrect** responses



CLEAR

Easy to understand and interpret in the same manner by all examinees

NO CLUES

Avoid responses that may help unknowledgeable, but test-wise examinees to select the **correct** answer

4 Criteria for Good Test Items (MCQ's)

IMPORTANT

Relevant
Suitable
Non-trivial

Criteria for Good Test Items (MCQ's)

Important

Application of Knowledge (Thinking, Understanding, Reasoning)

- Using pathophysiology to **predict** findings from specific conditions
- **Explaining** why something has occurred
- Using formulas to **calculate** a result
- **Interpreting** data, graphs, tables
- **Ordering** and interpreting the results of diagnostic studies
- **Formulating** a differential diagnosis
- **Summarizing** a patients' finding into a diagnosis
- Initiating **treatment**
- Determining **prognosis**

Criteria for Good Test Items (MCQ's)

Important

Testing Application of Knowledge (Commonly used phrases in the lead-in question)

■ Which of the following

- is the most likely cause of the patient's condition ?
- is the most likely diagnosis ?
- is the most appropriate next step ?
- is the most appropriate next step in the patient's management ?
- additional findings is most likely to develop in this patient ?
- physical findings is most likely ?
- results of diagnostic studies is most likely ?
- findings is ABNORMAL ?

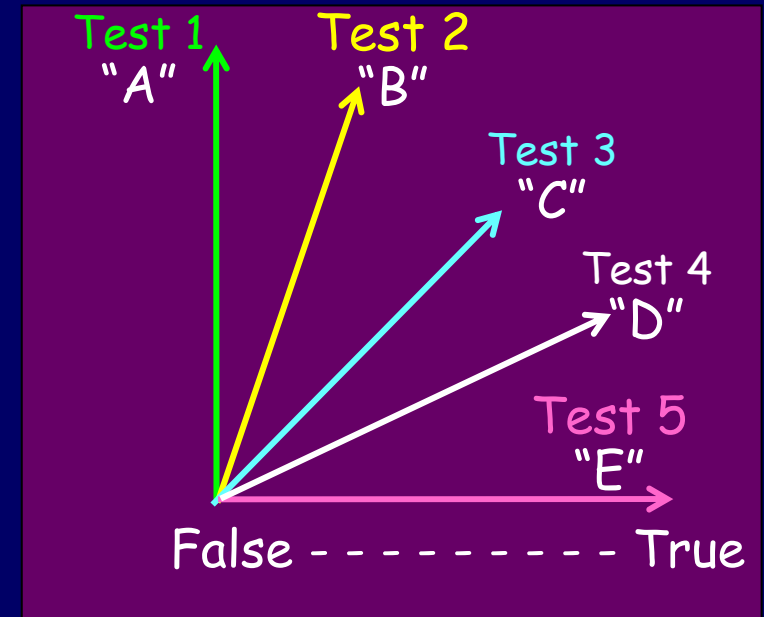
■ Results of which of the following tests will most likely be diagnostic ?

■ For which of the following conditions is the patient most at risk ?

Criteria for Good Test Items (MCQ's)

- Results of which of the following tests will most likely be diagnostic ?

All options are either (100%) True or False



- For which of the following conditions is the patient most at risk ?

All options differ along one single dimension

D C A E B

the LEAST correct the MOST correct

Criteria for Good Test Items (MCQ's)

Important

Utilization of Clinical Vignettes

- **MCQ's that assess application of clinical and basic science knowledge** (*require examinees to apply their knowledge to interpret, synthesize and make some type of judgment from information presented in a clinical scenario*)
- **Consistent use of clinical vignettes when assessing basic science knowledge emphasizes the relevance of basic science for clinical medicine?**
- **Use of vignette worksheet, structures the writing and success of producing effective MCQ's that assess the examinee's ability to apply knowledge from the basic and clinical sciences**

Worksheet for the Construction of a VIGNETTE

(modified from NBME-U)

1	a) Patient's age and sex :
	b) Site of care :
	c) Presenting (Chief) complaint (<i>including nature of onset</i>) :
	d) Duration of symptoms (<i>include changing nature of symptoms</i>) :
	e) Pertinent history (<i>HPI; personal; medication; family</i>) :
	f) Physical Examination findings :
	g) Results of diagnostic studies :
	h) Response to initial treatment :
2	Specific Question

Stem

Lead-in question

3	Option-set
	A.
	B.
	C.
	D.
	E.

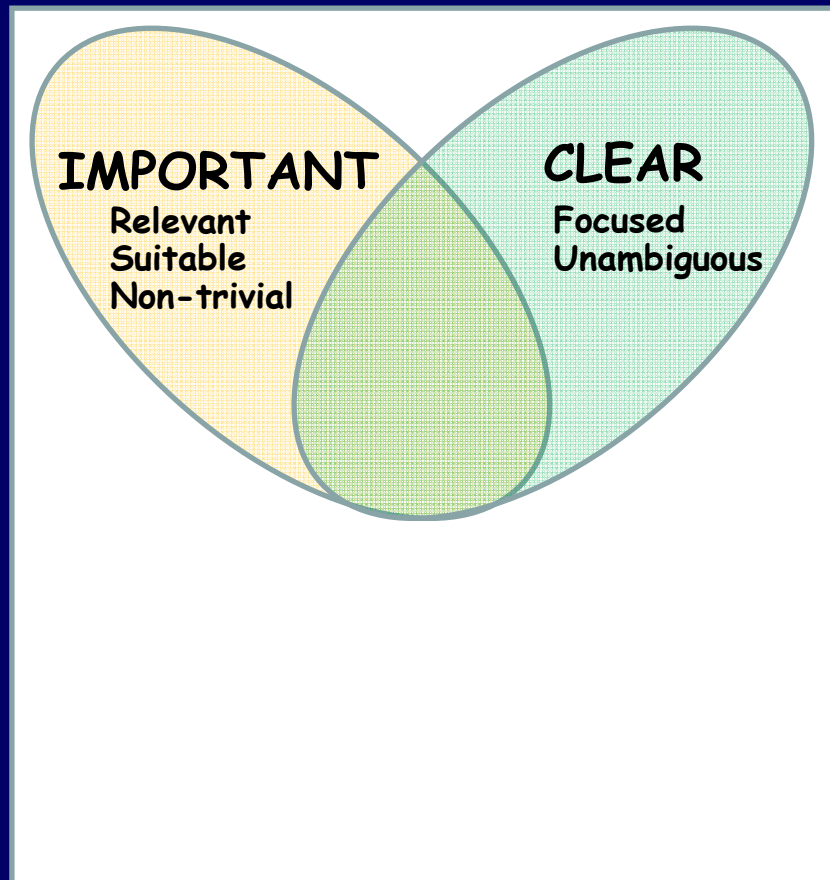
Options

Take Home Messages

Important

- 1 The three components of a single - best answer multiple choice question are:
 - Stem
 - Lead - in
 - Option set
- 2 Write items only on important concepts
(*avoid trivia - avoid superfluous information*)
- 3 Consistent use of task templates promotes consistent phrasing of stems and lead-in questions

4 Criteria for Good Test Items (MCQ's)



Criteria for Good Test Items (MCQ's)

Clear

Stem

- Focused, unambiguous { Allows examinee to give an answer before reading the responses ("cover test")
- Sufficient information to answer question
- NO excess verbiage; NO "red herrings"; NO lectures

Responses

- Homogeneous, parallel in content { All responses similar (drugs; tests)
Allow examinee to compare responses (rank responses on single dimension)

Ranking Responses on a Single Dimension

Stem:

A 16 yr old girl complains of muscular weakness of the extremities which has been progressive over the last 4-5 days. She has always been in good health. A little over a week ago she had an infection of the upper respiratory tract. On physical exam, her temp is 37.9°C, BP: 112/78 mmHg, HR 90/min and RR 42/min with superficial respirations. She 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 plantar reflex shows flexion; and the deep tendon reflexes are absent.

Lead-in :

Which of the following is the most likely diagnosis ?

Responses:

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

Ranking Responses on a Single Dimension

- 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.

Criteria for Good Test Items (MCQ's)

Clear

Stem

- Focused, unambiguous { Allows examinee to give an answer before reading the responses ("cover test")
- Sufficient information to answer question
- NO excess verbiage; NO "red herrings"; NO lectures

Responses

- Homogeneous, parallel in content { All responses similar (drugs; tests)
Allow examinee to compare responses (rank responses on single dimension)
- Only one correct answer
- Non-controversial
- Responses similar in length and grammatical construction
- Plausible
- Represent common misconceptions, outmoded ideas

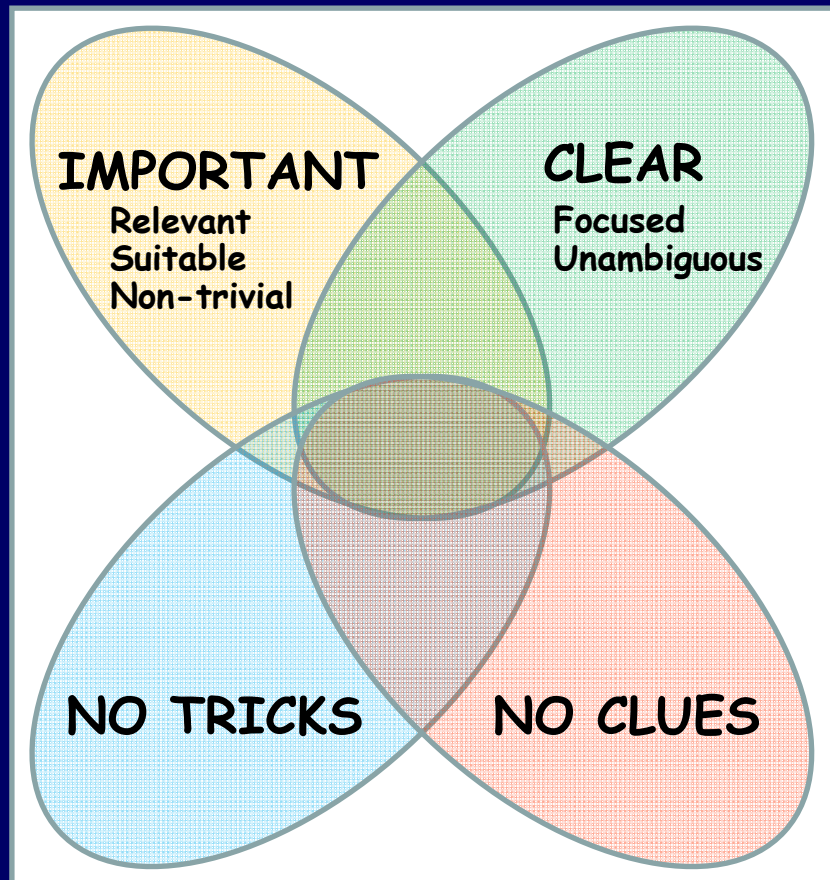
Take Home Messages

Clear

Stem & Responses

- 1 The “**stem**” presents the stimulus and should be written in the form of a vignette that describes a specific clinical scenario
- 2 The “**lead-in**” should be focused to ask a specific question that the examinee can answer without seeing the options
- 3 Include most of the information in the stem, **NOT** in lengthy distractors
(responses should rarely exceed one line)
- 4 The “**option set**” should include homogeneous answers (**correct and incorrect**) that can be rank-ordered (**from most to least correct**) on a single dimension

4 Criteria for Good Test Items (MCQ's)



Criteria for Good Test Items (MCQ's)

There are two major types of **technical flaws** that are commonly introduced by question writers:

Flaws related to **irrelevant difficulty**

Avoid things that may cause examinees to select an **incorrect** response

No Tricks

Flaws related to examinees' **testwiseness**

Avoid responses that may help examinees select the **correct** answer

No Clues

Criteria for Good Test Items (MCQ's)

No Tricks

Avoid tricks that may cause examinees to select incorrect response

- Vague terms { frequently
usually
commonly
rarely
- Negative terms { Especially when negative terms are combined
(*The physician should **not** avoid*)
- Reverse truths { A known fact which is turned it into a falsehood
("*Thyroxine is not produced by the thyroid*")
- Double options { Contain more than one element (lab studies; treatments)
(*Confusion if one correct and other not*)
Some option acceptable if often they belong together
(*height and weight; CBC and diff; nausea and vomiting*)

Criteria for Good Test Items (MCQ's)

No Tricks

Avoid tricks that may cause examinees to select incorrect response

- **Medical jargon**
 - **Popular slang**
- } Usually country specific therefore to be avoided
- **Eponyms**
- { To be avoided unless they are very well known
(*"Turner syndrome"*)
- **Abbreviations**
- { May be unfamiliar or misinterpreted by examinees
(*Allowed if abbreviation is explained previously in question*)

Tips for Avoiding MCQ Flaws

No Tricks

No Clues

- Write the lead-in as a complete sentence that ends in a question mark.

This will prevent writers from providing **grammatical cues**

This will help to ensure that the lead - ins are :

Focused

Satisfy the "cover-the-options" rule

- Only include **logical options**
- Place the options in a **logical order**
- Avoid **absolute** terms ("always", "never")

Tips for Avoiding MCQ Flaws

No Tricks

No Clues

- Avoid **vague** terms (“frequently”, “usually” “rarely”)
- Avoid long options.
Check that the length and specificity for correct answer is similar to that of distractors
- Watch for “**clangs**”
Avoid repeating words in the responses that have been used in the stem
- State numerical data in a consistent form
(all as whole values; with decimal; as ranges; *etc*)
- Do not use “**None**” or “**All**” **of the above** as an option

Take Home Messages

No Tricks

Avoid tricks that may cause examinees to select incorrect response

- 1 Write the **lead-in** as a complete sentence ending with a question mark
- 2 The **lead-in** should be focused to ask a specific question that the examinee can answer **without seeing the options**
- 3 Include most of the information in the stem, **NOT** in lengthy distractors
(*responses should rarely exceed one line*)

Take Home Messages

No Clues

Avoid responses that may help examinees to select the **correct** answer

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

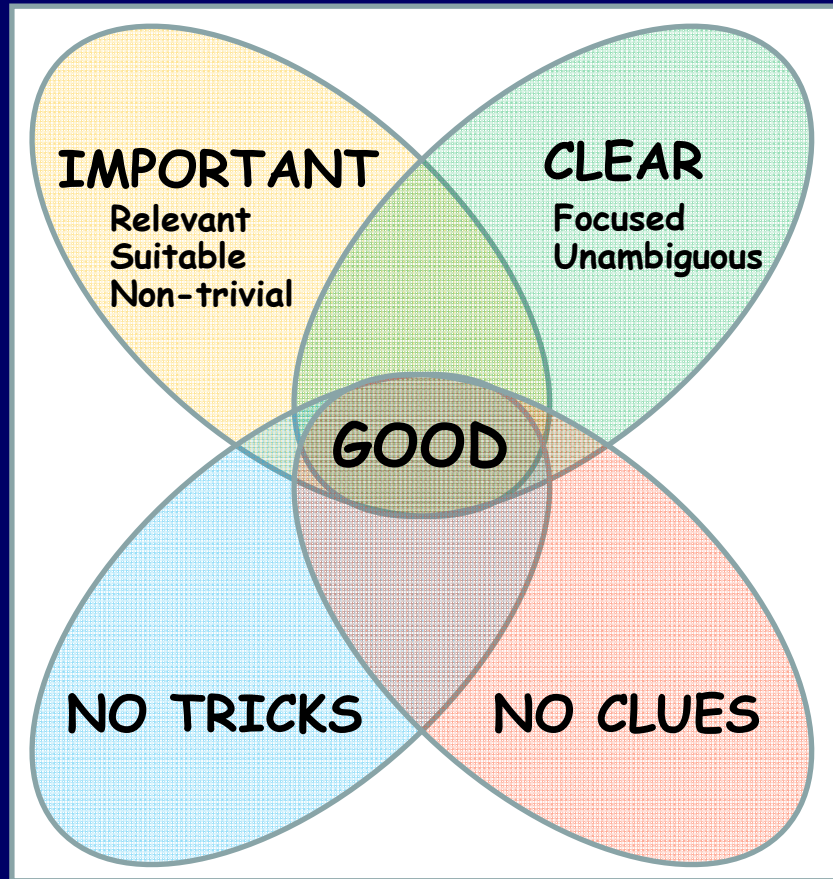
5 Repetition of words

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

Conclusion



Checklist for Writing Single-Best Answer MCQ's

(modified from NBME-U)

- Item**
- Has a single-best answer format been used ?
 - Does it satisfy the "cover test" rule ?
 - Does it test application of knowledge rather than recall of isolated facts ?
 - Is it appropriately "balanced" (most of reading in stem & relatively short options) ?
- Stem**
- Is the stem structured as a vignette, describing a specific situation ?
- Lead-in**
- Is it phrased positively rather than negatively ?
 - Does the phrasing avoid making the vignette irrelevant ?
 - Is it focused so that it poses a clear question to be solved ?
 - Is it structured as a complete sentence ending with a question mark ?
- Options**
- Are they homogeneous in content and phrasing ?
 - Are they similar in length and parallel in structure ?
 - Does each follow the lead-in both grammatically and logically ?
 - Can they be rank-ordered on a single dimension (from most to least correct) ?
 - Does the correct answer avoid repeating words used in the stem ("clang" clue) ?
 - Are distractors phrased to avoid repetition that clues correct answer (convergence) ?
 - Has the option-set been constructed to avoid "none of the above" "all of the above" ?
 - Does each avoid the use of absolute terms (e.g. "always", "never") ?
 - Does each avoid the use of vague terms (e.g. "rarely", "usually") ?

“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

