How to write a multiple choice question (MCQ)- Analysis of the exam results

Amira Taman, PhD, CPHQ

Professor of Medical Parasitology, Member of Medical Education Department

Agenda

Define assessment and its purpose.

Recognize the anatomy of MCQ.

Evaluate the quality of “some” MCQs.

Recognize the analysis of the exam results.

Assessment

An ongoing process of gathering and interpreting information about learner's knowledge, skills and attitude.

The purpose of assessment

-Pass/fail

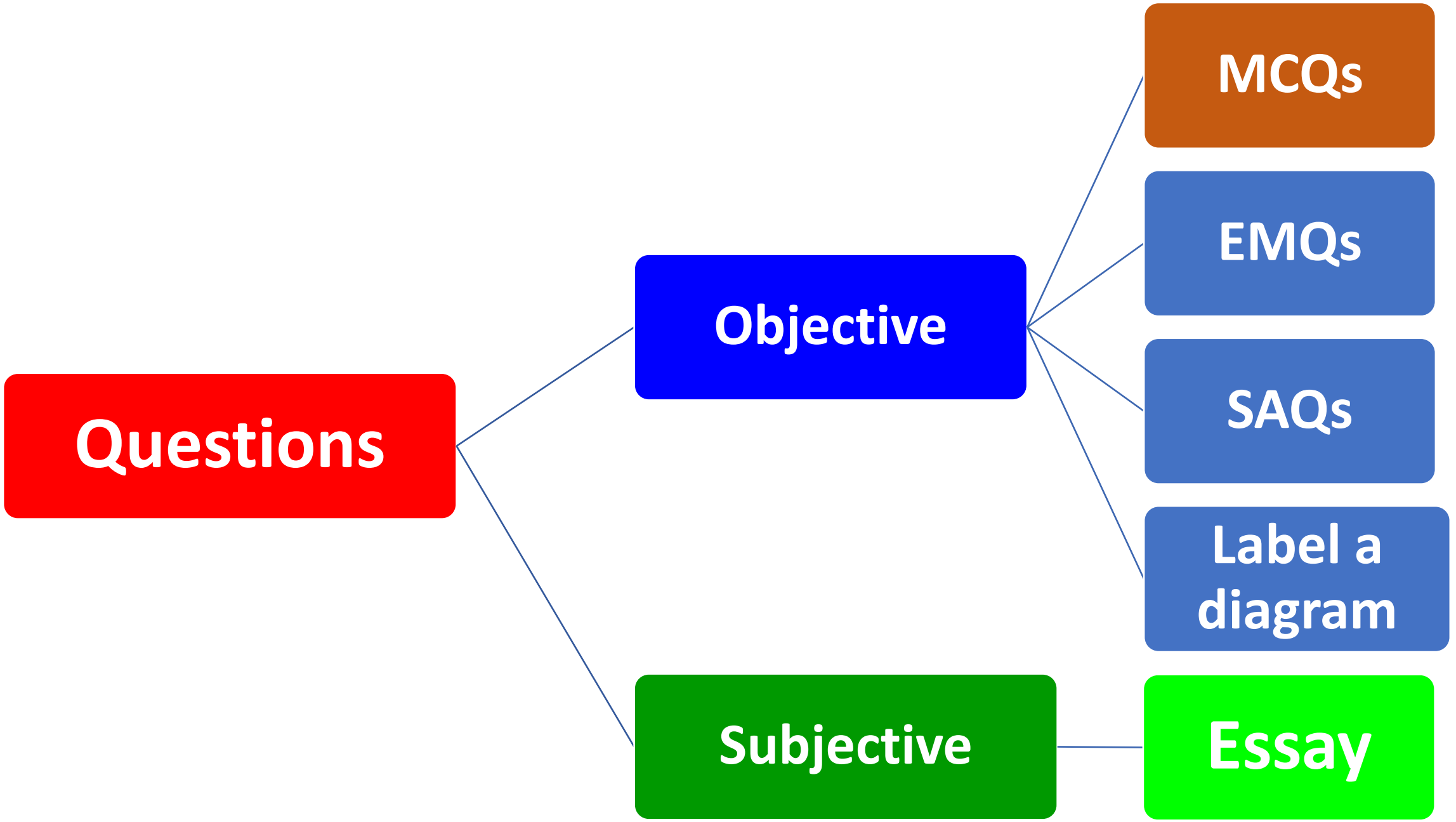
- Certify

-Rank

-Feedback

Purpose of Student Assessment

- ❑ Evaluating the effectiveness of the course
- ❑Showing the effectiveness of the curriculum
- ❑Introducing curriculum change
- ❑Identifying effective teaching
- ❑ Motivating teachers and learners
- ❑Testing students' performance
- ❑ Measuring improvement over time


Questions
Objective
MCQs
EMQs
SAQs
Label a diagram
Subjective
Essay

Types of MCQs

Single Best Answer (SBA)

True/False (TF)

Multiple choice questions (MCQs)

Advantage

A

- Increases objectivity by removing inter/ intra-examiner variation.

B

- Asses broad range of knowledge, for a big number of learners, in a short period of time.

C

- Can test higher order thinking such as "application and evaluation".

Multiple choice questions (MCQs)

Disadvantage

- Time consuming
- Need training and experience
- Guessing
- Cheating
- Influenced by the reading ability of the examinee

Anatomy of MCQ

Q. A 1-year-old infant is known to have heart disease and is noted to be cyanosed

← **Stem**

Which of the following is the most likely diagnosis?

← **Lead-In**

A. Atrial septal defect
B. Patent ductus arteriosus
C. Ventral septal defect
D. Tricuspid atresia

Distractors

→ **Key**

Stem

A

- **Direct question or statement.**
- **Preferred to be case scenario (clinical vignette, authenticity)**

B

- **Comprehensive and contain all facts needed.**
- **(Raw data?)**
- **Clear words, avoid negative words (NOT or EXCEPT)**

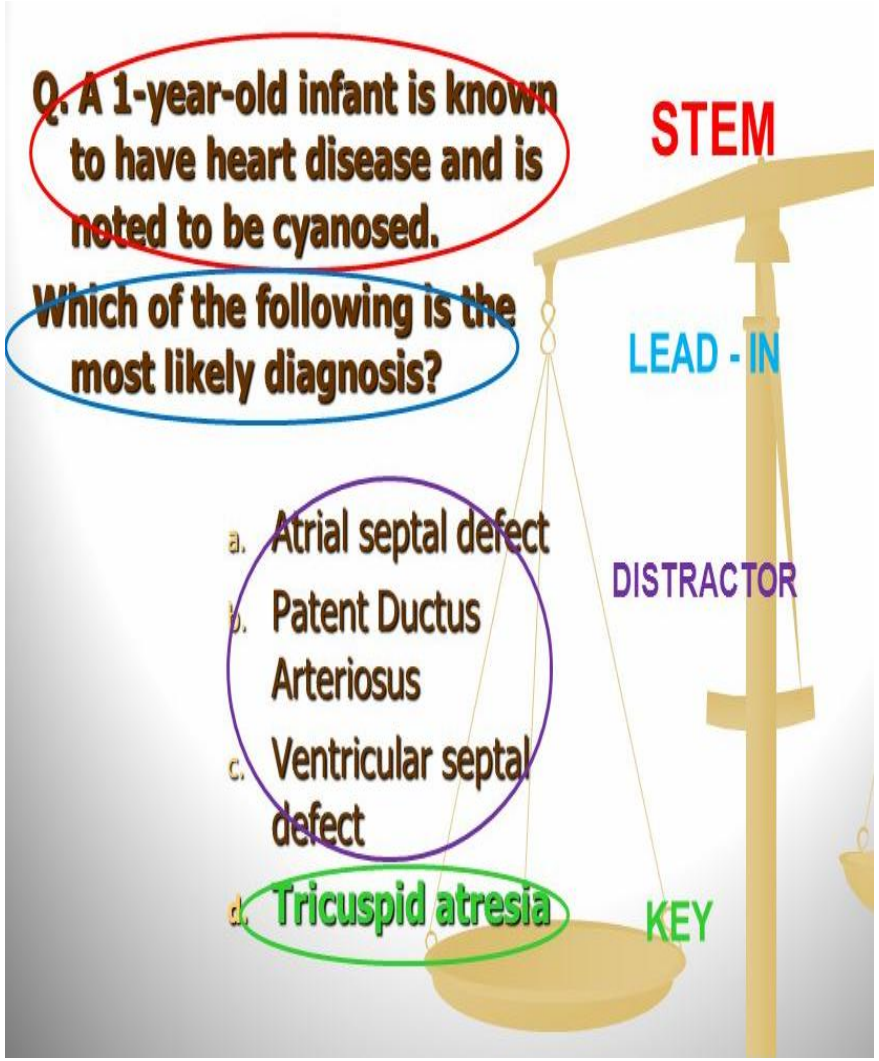
C

- **Long stem compared to distractors**
- **Avoid vague and absolute terms (sometimes, always, rarely)**

Lead in question???

1- Should be related to the stem

2- Cover test (answer question without looking to the options)



Distractors

Number (3-5)

Same category (homogenous), reasonable

Independent (not overlapping)

Avoid “all /non of the above”

Keep length almost equal

Avoid Grammatical inconsistencies that cue to the correct choice

Avoid clang associations (choices identical to or resembling words in the stem)

Avoid specific determiners including always, never, completely, and absolutely

Clinical vignette with single best answer

A 12-year-old boy is known to have **neurofibromatosis type 1.** His blood pressure is **148/110 mmHg** (>95th centile for height and sex).

- What is the most important **pathophysiological mechanism** for his **high** blood pressure?

A. Increased intravascular volume

B. Increased release of anti-diuretic hormone

C. Activation of renin, angiotensin system

D. Increased catecholamine activity

E. Increased sympathetic activity

Recall/Application of knowledge

Recall of Isolated Fact: Which of the following vitamins or minerals is involved in clotting factor synthesis?

Application of Knowledge: A 70-year-old widower has ecchymoses, perifollicular petechiae, and swelling of the gingiva. His diet consists primarily of cola and hot dogs. Which of the following vitamins or minerals is most likely to be deficient?

A. Vitamin B
B. Vitamin K
C. Zinc
D. Iodine
E. Biotin

Q 1.

In children, ventricular septal defects are associated with which of the following?

A. Systolic murmur
B. Pulmonary hypertension
C. Tetralogy of Fallot
D. Cyanosis
E. Dyspnea

Q 2.

- **A 50-year-old alcoholic man is confused, agitated and describes the world as being unreal. What is the name of this symptom?**

A. Depersonalization

B. Derailment

C. Derealization

D. Focal memory deficit

E. Signal anxiety

Q 3.

What is the incidence of infertility following the second attack of Salpingitis?

A. 1 to 20%

B. 2 to 13%

C. Greater than 50%

D. Less than 20%

E. 20-50%

Q 4.

Arrange the parents of the following children with Down's syndrome in order of the highest to the lowest risk of recurrence . Assume the maternal age to be 22 in all cases and that a subsequent pregnancy occurs within 5 years. The karyotypes of the daughters are:

I- 46 xx (14q21q),pat

II-46xx,(14q21q)de novo

III-46xx,(14q21q),mat

IV-46xx, (14q21q)pat

V- 47xx,(21q21q),parents not karyotyped

A- III, IV, I, V,II

B- IV, III, V, I, II

C- III, I, IV, V, II

D-IV, III, I, V, II

E- III, IV, I, II, V

Q 5.

Which of the following is not seen in a case of malaria?

A. Cough
B. Nusea
C. Vomiting
D. A and B
E. B and C

Q 6.

Which of the following is an autosomal recessive disorder?

A. Neurofibromatosis.

B. Thalassemia.

C. Marfan syndrome.

D. Polycystic kidney.

E. Non of the above

Q 7.

- **A 60-year-old alcoholic man is brought to the emergency hospital with status epilepticus. After ensuring a patent airway what is the first step in the treatment of this patient?**

A. CSF examination

B. CT scan of head

C. Phenytoin

D. Diazepam

E. Glucose with vitamin B

Q 8.

Which of the following is true about DM?

A. Not occur in children

B. Treated with glucagon

C. Treatment not involve diet regulation

D. Primarily a disturbance of CHO metabolism due to relative or absolute insulin deficiency

E. No oral therapy is available

Q 9.

Local anesthetics are most effective in the :

A. anionic form, acting from inside the nerve membrane.
B. cationic form, acting from inside the nerve membrane.
C. cationic form, acting from outside the nerve membrane.
D. uncharged form, acting from inside the nerve membrane.
E. uncharged form, acting from outside the nerve membrane.

Q 10.

Which of the following best describes the commonest organism causing bacterial pneumonia in childhood?

A. Gram positive coccus, occurring in chains

B. Gram positive coccus, occurring in pairs

C. Gram negative coccus, occurring in pairs

D. Gram negative coccus, occurring in chains

E. Gram positive bacillus, occurring in pairs

Q 11.

In medico legal law suits, associated requirements to assure the physician receives due process include:

A. Notice, an impartial forum, council, a chance to hear and confront evidence against him/her
B. Proper notice, a tribunal empowered to make the decision, a chance to confront witnesses against him/her, and a chance to present evidence in defense
C. Reasonable and timely notice, impartial panel empowered to make a decision, a chance to hear evidence against himself/herself, and the ability to confront witnesses and present evidence in defense

Q 12.

A 12- year- old girl presents with an acute onset of breathlessness and a 2 week history of malar rash. Her temperature is 37.4°C and heart rate is 120/min. Respiratory rate is 35/min with widespread inspiratory crackles over the bases. Chest X ray shows cardiomegaly and ESR is 63 mm/hr.

Which of the following is the most important investigation to aid diagnosis?

A. Perform an electrocardiogram

B. Measure the blood pressure

C. Anti-double stranded DNA antibody titre

D. Consider lung biopsy

E. Serum calcium and ACE levels

Q 13.

Which of the following statements is the most accurate about croup in children?

A. May be treated adequately with oral dexamethasone

B. Could be a complication of hypercalcaemia

C. Is never seen in patients under 3 months

D. Is rarely fatal

E. Possibly relates to immaturity of cartilaginous development

Q 14.

A 22-year old male is diagnosed as meningitis. CSF examination showed low glucose level, elevated protein and increased Lymphocytes.

This is suggestive of:

A. *Neisseria meningitides.*

B. *Trypanosoma brucei*

C. *Herpes Simplex Virus.*

D. *Naegleria fowleri.*

E. *Salmonella typhi*

In general .....

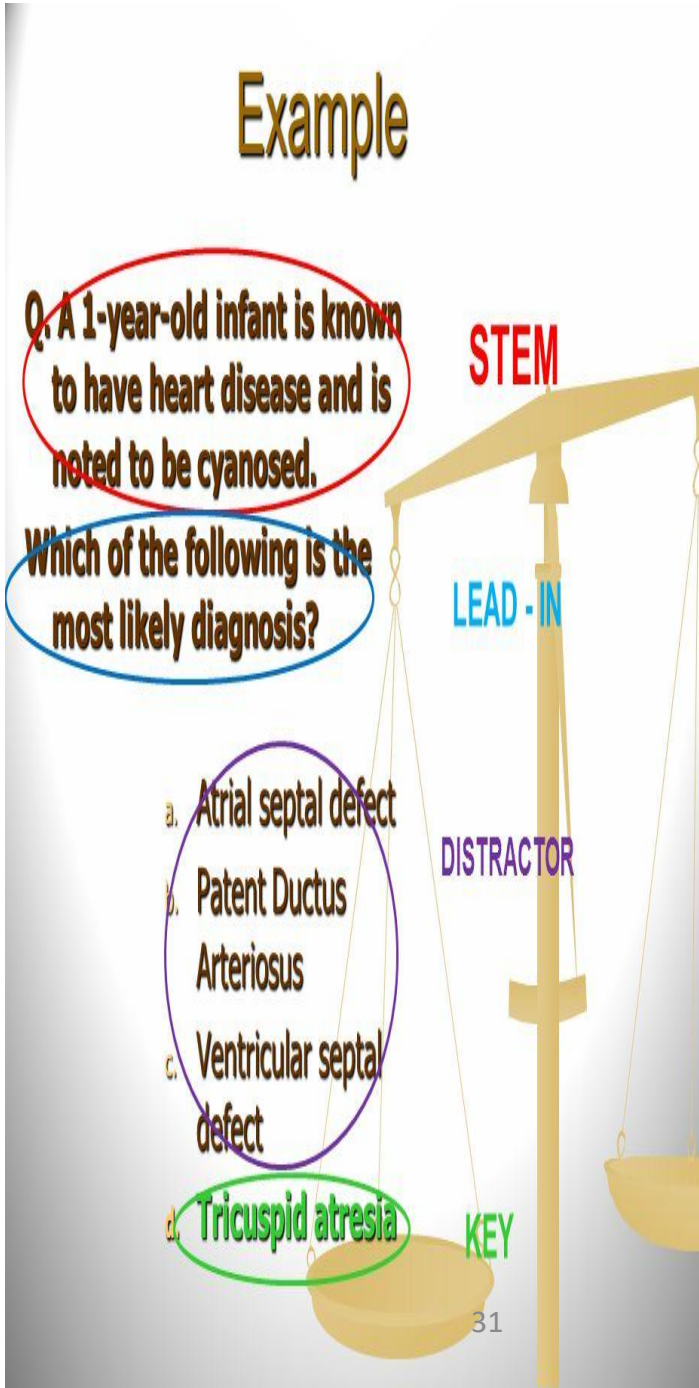
- Better to use "The single Best Answer" format.

- Make the question simple and straight forward.

- Each question should have focus on a single problem.

- Avoid over specific and over general content.

Question can be answered without looking at options.



In general .....

-Avoid trivial content.

- Every item should reflect specific content area as defined by blueprint.

- Try to test higher level learning such as application not simple recall only.

- Ensure that content of each item does not cue answers other items on test.

- Avoid the use of acronyms except if nationally accepted terms

- Avoid tricky items.

In general .....

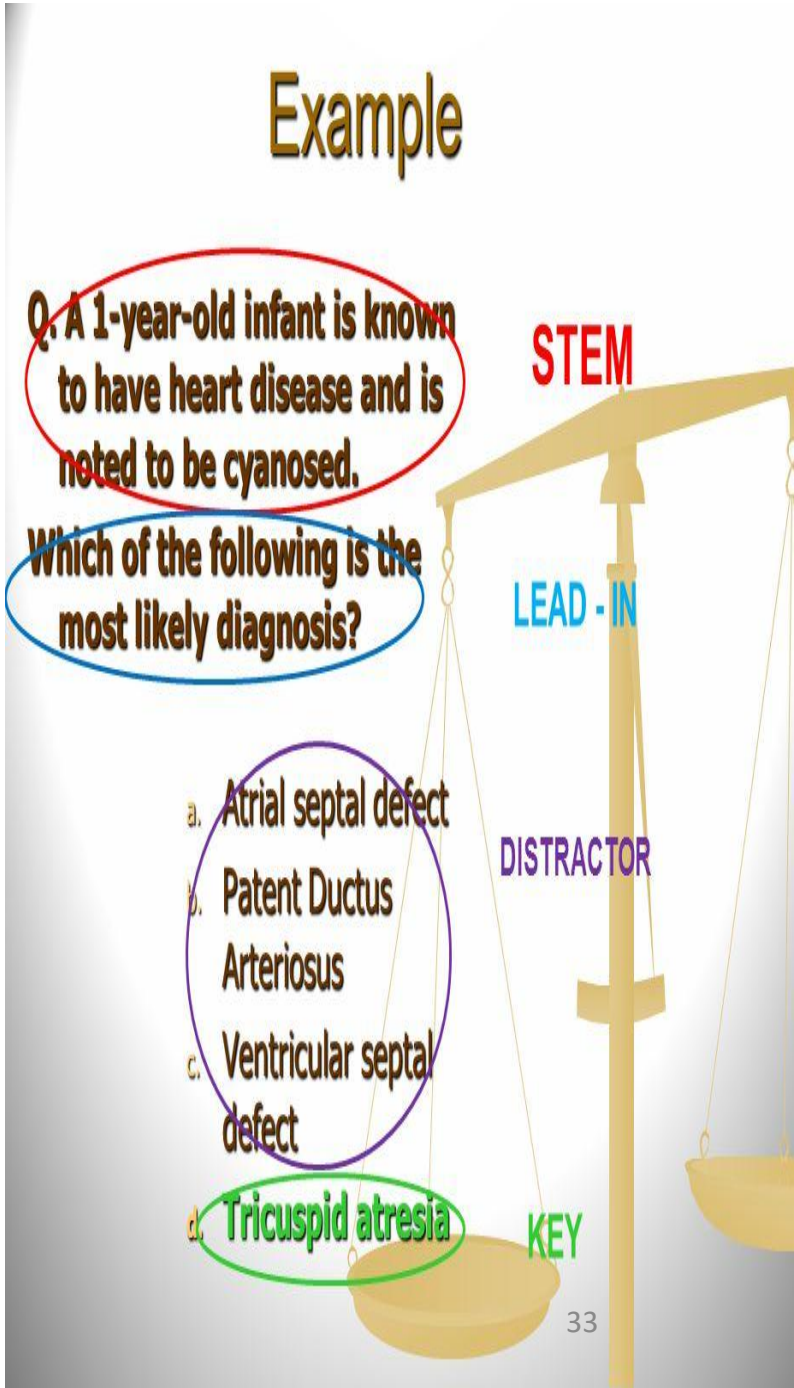
- Keep the entire question on one page.

- Vary the location of the right answer.

- Use correct and consistent grammar,

  punctuation, capitalization, and spelling.

- Check and edit.



MCQ Flaws that aid the “test-wise” examinee

Grammatical clues

Logical inconsistencies

Absolute or vague terms

More detailed correct answers

Word repeats (clanging)

Convergence

All of the above and

Non of the above

Analysis of the exam results

Item analysis

A process of analyzing the performance of MCQ after it has appeared in a question paper.

Item analysis provides:

1. Difficulty Index (P):

The question difficulty is the **percentage of students** who selected **the correct response**

2. Discrimination Index (item effectiveness-D):

Indicates how well the question **separates the students who know the material well** from those who don't

3. Distracter Effectiveness (DE): Effectiveness of the (alternatives) given

The essential steps of item analysis are:

1. Score the whole test for all the students.

2. Rank the students based on their test scores.

3. Take the top third as high achievers and bottom third as low achievers.

Prepare a table for each item

Options	No. selecting the option among high achievers (H)	No. selecting the option among low achievers (L)	Total response N(%)
A	5	10	15 (15%)
B	5	10	15 (15%)
C	30	10	40 (40%)
D	10	10	20 (20%)
E	Nil	2	2 (2%)
Not responded	Nil	8	--
Total (N)	50	50	100

Difficulty index: Indicated by the symbol '**P**' and is calculated by the formula

$$P = \frac{H + L}{N} \times 100$$

Where,

H = No. of students answering the item **correctly in the high achieving group**.

L = No. of students answering the item **correctly in the low achieving group.**

N = **Total number** of students in two groups including non-responders.

Thus, for the given example the value of **'P'** $= \frac{30 + 10}{100} \times 100 = 40\ \%$

Interpretation of the difficulty index (P)

Cut off values	Grading	Interpretation/ recommendation
0.9 - 1.0	Very easy Conditional acceptance	**Educationally, we want some of the items to be answered correctly by a majority of the class.**
0.80 – 0.9	Easy	**Some questions should have this range**
0.30 to 0.80	Excellent	**the mid range difficulty demonstrating adequate discrimination.**
0.15 – 0.29	difficult	**Some questions are desirable to be in this range**
Below 0.15	Hard	**This value indicates serious problem with question or teaching. Judgment analysis may be used to retain the question.**

Discrimination index: indicated by the symbol ‘d’ and is calculated by the formula

$$d = \frac{H - L}{N} \times 2$$

Where,

H = No. of students answering the item correctly in the high achieving group.

L = No. of students answering the item c**orrectly** in the low achieving group.

N = Total number of students in two groups including non-responders.

Thus for the given example the value of ‘d’ = $\frac{30 - 10}{100} \times 2 = 0.4$

Simply,

$$d = \frac{30}{50} - \frac{10}{50} = \frac{60}{100} - \frac{20}{100} = 0.6 - 0.2 = 0.4$$

Prepare a table for each item

Options	No. selecting the option among high achievers (H)	No. selecting the option among low achievers (L)	Total response N(%)
A	5	10	15 (15%)
B	5	10	15 (15%)
C	30	10	40 (40%)
D	10	10	20 (20%)
E	Nil	2	2 (2%)
Not responded	Nil	8	--
Total (N)	50	50	100

Interpretation of the discrimination index (D)

Cut off values	Grading	Interpretation/ recommendation
Above 0.35	Very Good	Reflects good construction and teaching
Between 0.35 – 0.10	Good	Reflects effective teaching
Below 0.10	Poor	The item has some room for improvement.
A negative value	Negative discrimination	Low scorer is doing well in these items, which is unacceptable. The item should be revisited or rejected. One should check the key answer.

Correlation between difficulty and discrimination indices

Positive correlation between difficulty and discrimination indices	
High *P value* items are Good discriminator	
Moderately easy/difficult (30-70%) items	Maximal discriminative ability
Very difficult/ Very easy item	Poor discrimination
Very easy item with Negative discrimination index	Indicating a faulty item, or incorrect keys.

3. Distracter effectiveness (DE)

- Any of the distracters in the item, which has not **attracted even 5%** of the total responded item, is said to be a "**non-functional**" distracter.

- In the given example, option number '**E**' is nonfunctional having attracted only **2%** of students. So, "alternative E" should be changed

Prepare a table for each item

Options	No. selecting the option among high achievers (H)	No. selecting the option among low achievers (L)	Total response N(%)
A	5	10	15 (15%)
B	5	10	15 (15%)
C	30	10	40 (40%)
D	10	10	20 (20%)
E	Nil	2	2 (2%)
Not responded	Nil	8	--
Total (N)	50	50	100

Take home message……

Five steps for MCQ writing

- Focus on one concept per item
- Choose a level of cognition
- **Avoid common errors**
- Write the correct answer
- Validate the items (pre and post)

An ideal item (MCQ) will be the one which has:

- An average difficulty index (31% to 70%)
- High discrimination index (>0.35)
- Maximum distractor efficacy with (3-5) functioning distractors

KNAPP
Our experience. Your success.

Checklist for the Review of Multiple-Choice Questions

Item Content

- Does the concept directly impact professional/occupational competence? ("Need to know" vs. "Nice to know")
- Is it reasonable to assume that in general, competent individuals should know the answer to the problem posed in the stem?
- For scenario- or case-based questions, is the situation presented a realistic and reasonable one (i.e., not trivial or unusual)?
- Is the item of appropriate difficulty for the level of competence targeted by the certification?
- Does the item deal with a noncontroversial topic – one on which all experts would agree?

Stem

- Can you understand what is being asked without reading the options? (Cover the options during your initial review to determine whether the stem is clear and unambiguous.)
- Does the stem include the least amount of information necessary for understanding the question and selecting the correct answer (i.e., no extraneous information and not overly detailed)?

Options

- Do the options fit logically and grammatically with the stem?
- Are the options parallel and structurally similar to each other?

Distractors

- Are the distractors plausible, but clearly wrong? (What is the rationale for examinees to choose each distractor)?

Key

- Is there only one correct option listed?
- Is the key one on which there is consensus within the field?
- Is the key the same length and level of detail as the distractors?

References

- **Rao C, Kishan Prasad H, Sajitha K, Permi H, Shetty J (2016)** Item analysis of multiple choice questions: Assessing an assessment tool in medical students. International Journal of Educational and Psychological Researches 2: 201-204.
- **Coughlin, P. A., Featherstone, C. R. (2017).** How to write a high quality multiple choice question (MCQ): A guide for clinicians. European Journal of Vascular and Endovascular Surgery, 54, 654–658.
- **Gupta P, Meena P, Khan AM, Malhotra RK, Singh T (2020).** Effect of Faculty Training on Quality of Multiple-Choice Questions. Int J Appl Basic Med Res., 10(3):210-214.
- **Przymuszała P, Piotrowska K, Lipski D, Marciniak R, Cerbin-Koczorowska M (2020).** Guidelines on Writing Multiple Choice Questions: A Well-Received and Effective Faculty Development Intervention. SAGE Open.
- **Owolabi LF, Adamu B, Taura MG, Isa AI, Jibo AM, Abdul-Razek R, Alharthi MM, Alghamdi M (2021).** Impact of a longitudinal faculty development program on the quality of multiple-choice question item writing in medical education. Ann Afr Med.; 20(1):46-51.

e. mail: amirataman@mans.edu.eg