

**ISTANBUL MEDİPOL UNIVERSITY**  
**INTERNATIONAL SCHOOL OF MEDICINE**



**MULTIPLE CHOICE QUESTION**  
**EXAM PREPARATION GUIDE**

**MEASUREMENT AND EVALUATION COMMISSION**

**2023**

## **MULTIPLE CHOICE QUESTION EXAM PREPARATION GUIDE**

### **PREPARED BY**

Prof. Aylin REZVANI

Prof. Dilek UNAL

Prof. Ilknur TURKMEN

**(2018)**

### **TRANSLATED BY**

Asst.Prof. Ayse Guliz DIRIMEN ARIKAN

Sumeyye CAKIR

Sinem KUTLU

**(2023)**

### **Advisor:**

**Marmara University School of Medicine**

**Department of Medical Education**

**Prof. Mehmet Ali GULPINAR**

## Contents

<b>INTRODUCTION</b> .....	4
<b>Exam Types</b> .....	4
<b>Exam Qualifications</b> .....	4
<b>Detailed Exam Types</b> .....	5
<b>Measurement and Evaluation Method and Tools</b> .....	5
<b>Table 1:</b> Measurement - Evaluation Method and Tools - 1 .....	5
<b>Table 2:</b> Measurement - Evaluation Method and Tools - 2 .....	5
<b>Measurement and Evaluation Areas and Methods</b> .....	6
<b>Assessment Areas and Methods</b> .....	7
<b>Table 3:</b> Assessment Areas and Methods .....	7
<b>Evaluation of Multiple-Choice Questions</b> .....	9
<b>QUESTION PREPARATION CHECKLIST</b> .....	10
<b>SOME QUESTION EXAMPLES</b> .....	14
<b>REFERENCES</b> .....	25

## INTRODUCTION

Measurement and evaluation have special importance in medical education. The Measurement and evaluation process is necessary for a planned education and should be well planned in itself.

Medical education; requires acquiring knowledge, applying it, being able to do it, and developing a behavior pattern for things all learned.

Goals, which are the last step of education, can only be achieved with a well-planned measurement and evaluation system;

- Assessing means "classifying with symbols" according to the degree to which objects or people have certain characteristics.

- Evaluation, on the other hand, means "reaching a verdict/judging, making a decision" about the results of assessments or observations in line with the determined criteria.

While determining the measurement and evaluation methods, care should be taken to consider the educational outcomes and to use appropriate measurement and evaluation methods and tools for each output.

## Exam Types

**A well-planned Measurement and Evaluation system should have at least 2 main objectives.**

**1- Descriptive exams (formative, for the benefit of the student):** These are the evaluations made to determine the status of the student in the education process and to give feedback on their further development. (such as developmental exams, quizzes, etc.)

**2- Decision-making exams (Summative, for the benefit of the society):** These are the exams (final exam, midterm exam, etc.) held to determine the success status of the student (successful - unsuccessful, pass/fail) related to a program.

## Exam Qualifications

For the effective assessment, the examination must be appropriate and objective. Valid, reliable, and distinctive features of the exam are indispensable. These features are defined below:

**Validity:** Having appropriate criteria to assess the learning goal (knowledge-skill)

**Reliability:** Getting the same result at repeated assessments

**Applicability:** Cost, practicality

**Effect on Learning/Effect on Development:** Directing the student to higher-level learning such as problem-solving, decision-making, planning, execution, evaluation, and development.

## Detailed Exam Types

### 1. Written Exams

a- **Open-Ended Exams;** Fill-in-the-blank, Case-based Modified written exam

b- **Exams with Choices;** Best answer, match type MCQ, true-false

### 2. Structured Oral Examination

### 3. Practical Exam

### 4. Clinical Oriented Reasoning Exam (CORE)

### 5. Objective Structured Clinical Exam (OSCE)

### 6. Hands-on Evaluation

## Measurement and Evaluation Method and Tools

**Table 1:** Measurement - Evaluation Method and Tools - 1

	Reliability	Validity	Applicability	Effect on learning
<b>Multiple choice question (MCQ)</b>	High	High for first-level knowledge	High	Prompts the student to memorize. Fragmented/Partial and inactive/inert information.
<b>Case-based MCQ (Best answer, match type)</b>	High	High for second-level knowledge	High	It directs the student to higher-level learning such as reasoning, decision-making, and evaluation.

**Table 2:** Measurement - Evaluation Method and Tools – 2

	Reliability & Validity	Applicability	Effect on learning
<b>Case-based modified written question (Case-MWE)</b>	Multiple cases and Evaluator, Acceptable by Evaluation criterion – High	High	It directs the student to higher-level learning such as problem-solving, decision-making, planning, execution, evaluation, and development. Contextual thinking, problem-solving, decision-making, and evaluation. Evaluating the rationale and perspective behind the decision. Evaluation of professional attitudes and behaviors.
<b>Case-based modified oral question (Case-SOE)</b>	High for higher-level thinking competencies	Medium	

## Measurement and Evaluation Areas and Methods

In each period of medical school education, the measurement of professional approach, ethics, and behavioral styles changes from the plain knowledge level to the decision-making level, from the simulated patient to the real patient. At the bottom, which is likened to a pyramid, first acquiring knowledge, comprehending knowledge, and integrating knowledge are placed. The best method for measuring this level is by asking '**multiple-choice questions**'. This method has high validity, reliability, and applicability for first-level information.

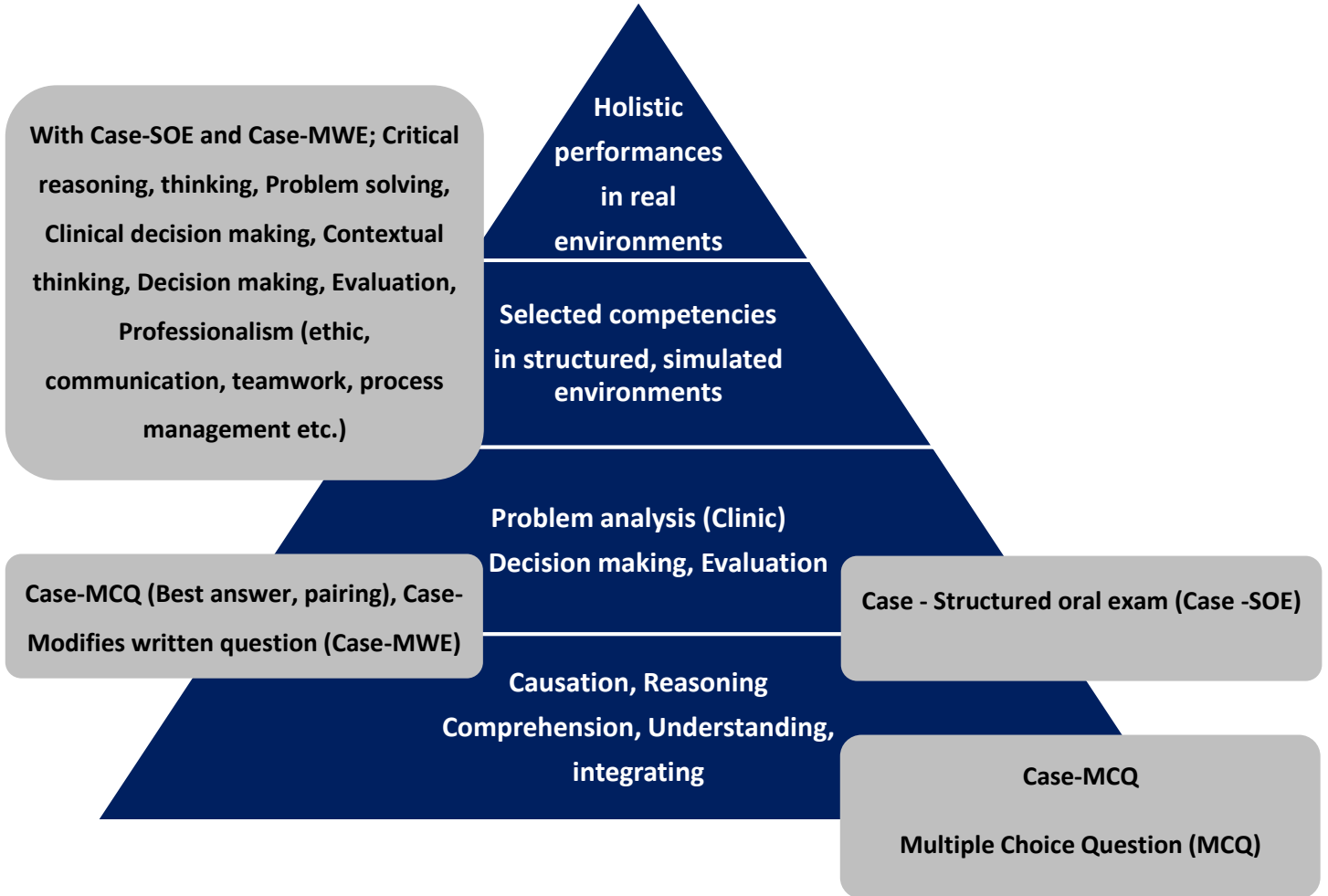
Measuring the use of knowledge is best understood through cases. The best method for the causation of problems and assessment of problem-related reasoning skills is asking '**multiple-choice questions accompanied by case**'. This method has high validity, reliability, and applicability for second-level information. It directs the student to higher-level learning such as reasoning, decision-making, and evaluation.

The ability to solve case problems and make clinical decisions can be measured with a 'written and oral structured exam'.

As we move towards the top of the pyramid, the unity of knowledge, skills, and attitudes can be measured by 'structured oral exams and modified written exams'. It directs the student to higher-level learning such as problem-solving, decision-making, planning, execution, evaluation, and improvement. It is a useful method for evaluating the rationale and perspective behind the decision and for evaluating professional attitudes and behaviors (Table 1).

## Assessment Areas and Methods

**Table 3:** Assessment Areas and Methods



In this guide, basic information, and examples about methods of preparing multiple-choice questions will be provided. The objective is to ensure that the training is completed with measurement and evaluation methods in accordance with the determined targets.

**While preparing Multiple Choice Questions (MCQ), it is necessary to act according to the following rules;**

1. While preparing multiple choice questions, it is necessary to use the “Learning Outcome Table”. Thus, the “Content validity” of the exam is increased.
2. It should be stated which of the learning outcomes in the Learning Outcome Table, the question is asked.
3. It should be stated according to which level of knowledge the question is asked.
4. Questions should be revised with the checklist before sending to Measurement and Evaluation Commission.
5. Questions should evaluate by other lecturers of the branch in terms of detecting possible unnoticed errors before sending them to the Measurement and Evaluation Commission.
6. To assure the quality of the questions, item analysis (discriminability and difficulty level of the questions) should be performed.
7. To prevent the negative effects of the questions on students, “true/false” type questions should be avoided.

**While preparing multiple choice questions, the following important points should be considered;**

1. The question stem should consist of a single sentence. Single words or phrases without verbs should be avoided.
2. The question must be simple and clear.
3. The answer to another question should not be hidden inside the question stem.
4. Negative statements should NOT be included in the question stem. If included, the negative word should be underlined.
5. Reasonable or logical distractors should be used.
6. Distractors should appear to be relevant to the question in phrasing, content, and nature.
7. Distractors that are irrelevant to the topic should NOT be used because they immediately reveal themselves by disrupting the harmony between question-and-answer options.
8. Hints that point to the correct answer should NOT be used.
9. The distractors and the correct answer should be similar and homogeneous in terms of content and total word count.
10. The words 'ALL' and/or 'NONE' should NOT be used in the answers.
11. If the answer to the question consists of numbers, it should be written in order from highest to lowest or from lowest to highest.
12. The correct options should not be in the same order often.



## Evaluation of Multiple-Choice Questions

After an exam attended by at least 40 students, the difficulty and discrimination level of the question can be determined for each of the multiple-choice questions. Questions that do not have the desired characteristics should be prevented from being used in the next exam. For this purpose, students are ranked from the highest to the lowest according to the grades they get from the exam. The answers of the students in the upper 27% group and the students in the lower 27% group who received high grades are included in the calculation:

### Difficulty index:

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

H = Number of students who answered the question correctly in the upper group (in the first 27%)  
L = Number of students who answered the question correctly in the lower group (in the last 27%)  
N = Total number of students in upper and lower groups.

### The difficulty index is calculated for each question and evaluated according to the following criteria:

50 – 60% is the recommended difficulty level.

30 – 70% acceptable

Over 70% response rate very easy

Answering less than 30% indicates that the question is very difficult.

### Discrimination index:

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

H = Number of students who answered the question correctly in the upper group (in the first 27%)  
L = Number of students who answered the question correctly in the lower group (in the last 27%)  
N = Total number of students in upper and lower groups.

### The discrimination index is calculated for each question and evaluated according to the following criteria:

0.35 and above is perfect

0.34 – 0.25 is good

0.24 – 0.15 should be revised

0.15 and below indicate that it is not sufficient in terms of discrimination.

## QUESTION PREPARATION CHECKLIST

Eligibility of the question	Yes	No
Compatible with learning objectives.	<input type="checkbox"/>	<input type="checkbox"/>
Appropriate for the year/grade level of the students.	<input type="checkbox"/>	<input type="checkbox"/>
There are NO apparent distractors that make the correct option obvious or ineffective.	<input type="checkbox"/>	<input type="checkbox"/>
Case (For Case-Based Questions):	Yes	No
The content of the case was written considering the students' level of knowledge. The case is NOT too complex or simple.	<input type="checkbox"/>	<input type="checkbox"/>
The phrases used in the case are simple and easy to understand.	<input type="checkbox"/>	<input type="checkbox"/>
The case is realistic.	<input type="checkbox"/>	<input type="checkbox"/>
Focused on a specific problem in the case and question.	<input type="checkbox"/>	<input type="checkbox"/>
<b>Adequate content of basic, clinical, and social-behavioral information about the condition/problem:</b> <ul style="list-style-type: none"> <li>● Inside the case, there are adequate positive clues [1] that focus the student on the correct answer. [1] Condition-specific information which is important to be seen in that condition (descriptive, diagnostic) and that directs the student to the problem (clinical/normal condition).</li> <li>● Inside the case, there are adequate negative clues [2] that focus the student to the correct answer / distract them from another relevant condition/problem. [2] Information that is important not to appear in the clinical condition/normal condition, and when it is not seen, that focuses the student to the clinical condition/normal condition described in the scenario.</li> </ul>	<input type="checkbox"/>	<input type="checkbox"/>
Inside the case, there is contextual information (age, gender, family status, etc.) that will affect the student's decision.	<input type="checkbox"/>	<input type="checkbox"/>
Inside the case, even if it is about the question, there are NO unnecessary information and irrelevant, confusing phrases used.	<input type="checkbox"/>	<input type="checkbox"/>
Question Stem	Yes	No
The question stem is NOT too long for the answer. The way of requesting the answer is NOT complicated or indirect.	<input type="checkbox"/>	<input type="checkbox"/>
Items related to the information to be measured are specified directly, it should not contain information pertaining to more than one question.	<input type="checkbox"/>	<input type="checkbox"/>
Phrases containing personal opinions and judgments such as "In your opinion...", "...what do you think" did NOT used.	<input type="checkbox"/>	<input type="checkbox"/>
There are NO negative statements and double negative statements in the question stem.	<input type="checkbox"/>	<input type="checkbox"/>
In case-based question stems, terms such as "Most probable diagnosis/Most appropriate treatment/Most probable cause" SHOULD be used.	<input type="checkbox"/>	<input type="checkbox"/>

<b>Options</b>	<b>Yes</b>	<b>No</b>
There are <b>NO</b> options such as “all above true” or “all above false”.	<input type="checkbox"/>	<input type="checkbox"/>
There are <b>NO</b> ambiguous phrases such as “usually”, “often”, and precise expressions such as “always”, “never”.	<input type="checkbox"/>	<input type="checkbox"/>
The negative suffix is written <b>ATTENTION-GRABBING</b> (bold, underlined). Negative statements are avoided, both at the question stem and in the options.	<input type="checkbox"/>	<input type="checkbox"/>
Options are expressed in short and clear words that everyone can understand.	<input type="checkbox"/>	<input type="checkbox"/>
Each option is meaningful on its own and can be read as a continuation of the question.	<input type="checkbox"/>	<input type="checkbox"/>
All options are of similar length (Is the correct option more specific, longer, and more comprehensive?).	<input type="checkbox"/>	<input type="checkbox"/>
Keywords and phrases are <b>NOT</b> repeated both in the question stem and in the correct option.	<input type="checkbox"/>	<input type="checkbox"/>
There are <b>NO</b> two opposing options among the options.	<input type="checkbox"/>	<input type="checkbox"/>
Correct options are in a different order for each question (if more than one question is written).	<input type="checkbox"/>	<input type="checkbox"/>
If abbreviations are used; They <b>SHOULD</b> be proper names that are well-known and accepted worldwide and internationally adopted abbreviations.	<input type="checkbox"/>	<input type="checkbox"/>
The unit is <b>NOT</b> repeated in the options, the unit is given in the question stem.	<input type="checkbox"/>	<input type="checkbox"/>
<b>Using Number/Value</b>	<b>Yes</b>	<b>No</b>
Except for special cases, it should be given as lower and upper bounds instead of just a number (value).	<input type="checkbox"/>	<input type="checkbox"/>
The number ranges specified in the options <b>SHOULD NOT</b> overlap with each other.	<input type="checkbox"/>	<input type="checkbox"/>
Unit measurements are specified once in the question stem and are <b>NOT</b> written repeatedly in each option.	<input type="checkbox"/>	<input type="checkbox"/>
Numbers are given in ascending or descending order.	<input type="checkbox"/>	<input type="checkbox"/>

## Check-List Created by the Measurement and Evaluation Commission for Checking the Multiple-Choice Questions;

### In Evaluating the 'Question' Unity:

- 1- Is the question compatible with the aim and learning objectives of the course?
- 2- Is the question asked for a single learning objective of the course?
- 3- Is the question appropriate for the student's education year/grade level?
- 4- Between the correct answer and the question stem, an analogy of phrase or terminology that distinguishes the correct answer from the distractors should NOT be allowed. If such a situation is unavoidable, it must be ensured that the same similarity is found in at least one other distractor.

### In the Evaluation of "Case" Unity:

- 1- Is the case realistic?
- 2- Is there contextual (age, gender, family status, etc.) information in the case that will affect the student's decision-making?
- 3- Is the content of the case written (neither too complex nor too simple) considering the level of knowledge of the students?
- 4- Are the phrases used in the case simple and easy to understand?
- 5- Are there enough positive clues (descriptive, diagnostic, useful information) to focus the student on the correct answer in the case?
- 6- Is there a negative clue in the case to focus the student on the correct answer and distract him/her from another situation in the differential diagnosis list?
- 7- Is there any unnecessary or confusing information in the question? (Even if related to the case unnecessary information should NOT be provided.)

### In Evaluating the Unity of the "Question Stem":

- 1- Is the question written in plain language and in accordance with spelling rules?
- 2- Is the question stem presented in a short, clear, and direct way? (The answer should NOT be asked in an indirect way, words that make it difficult to understand should NOT be used.)
- 3- Is there any information sharing about the topic at the question stem? (There should NOT be an information sharing about the topic as if reminding a topic in the question.)
- 4- At the question stem, are there statements that include personal opinions and judgments such as "In your opinion...", and "...what do you think"?

**5-** Are there negative statements at the question stem? [Negative verbs should be avoided as much as possible. If they should be used, they should be highlighted (bold, underlined)].

**6-** Is there an inexplicit statement(s) such as “usually” or “often” or an absolute/ambiguous statement such as “always” or “never” in the question stem?

**7-** A statement that can be put in the question stem should be placed at the stem and should NOT be repeated in every option.

#### **In Evaluating the Unity of “Answer Options”:**

**1-** The options should have unity. (If treatment is asked, all options should be related to treatment if a diagnosis is questioned, all options should be related to diagnosis.)

**2-** Does the correct option stand out clearly among other options?

**3-** In the options, is the correct answer striking (apparent distractor) or is there any option unrelated to the question (ineffective distractor)? (All options should contain part of the information, and this information should be necessary to answer the question. The options other than the correct answer should be similar and as close as possible to the correct answer.)

**4-** Does the elimination of distractors require that the knowledge or behavior being tested is learned?

**5-** Are the distractors such a nature that they do not hint at the correct answer?

**6-** Are there statements as “All above is true” or “All above is false” in the options?

**7-** Are the options briefly and clearly expressed in words that everyone can understand?

**8-** Are the phrases in the options arranged in such a way that they do not hint at each other?

**9-** Are the opposite phrases used in the options?

**10-** If a number/value is required in the options

**a.** Are the values given in ascending and descending order?

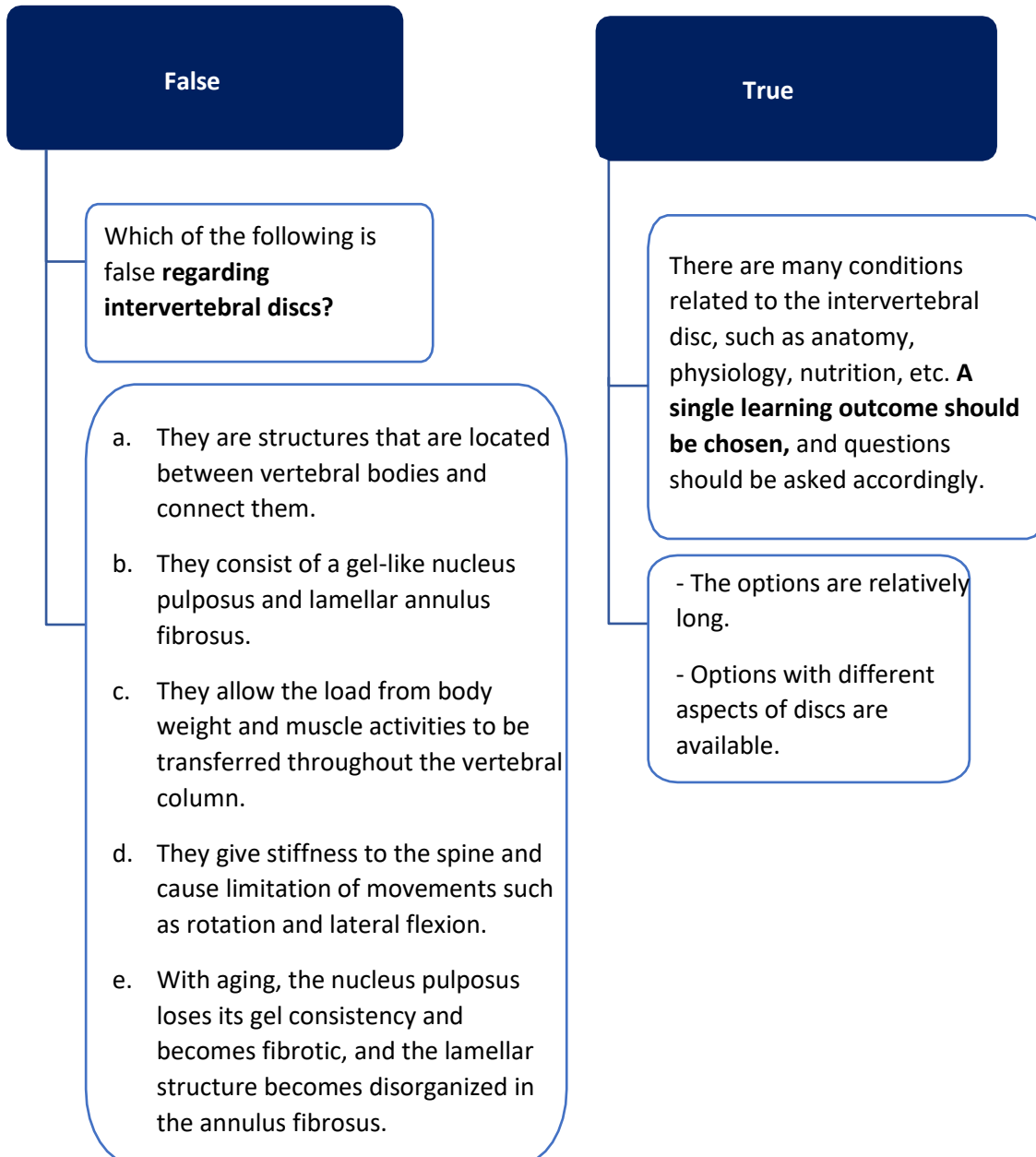
**b.** Are the values expressed inclusive of each other? (The values should NOT be arranged to cover each other.)

**c.** Are the values given as ranges rather than just a number? (Odd numbers should be avoided, except in special cases where it is absolutely necessary.)

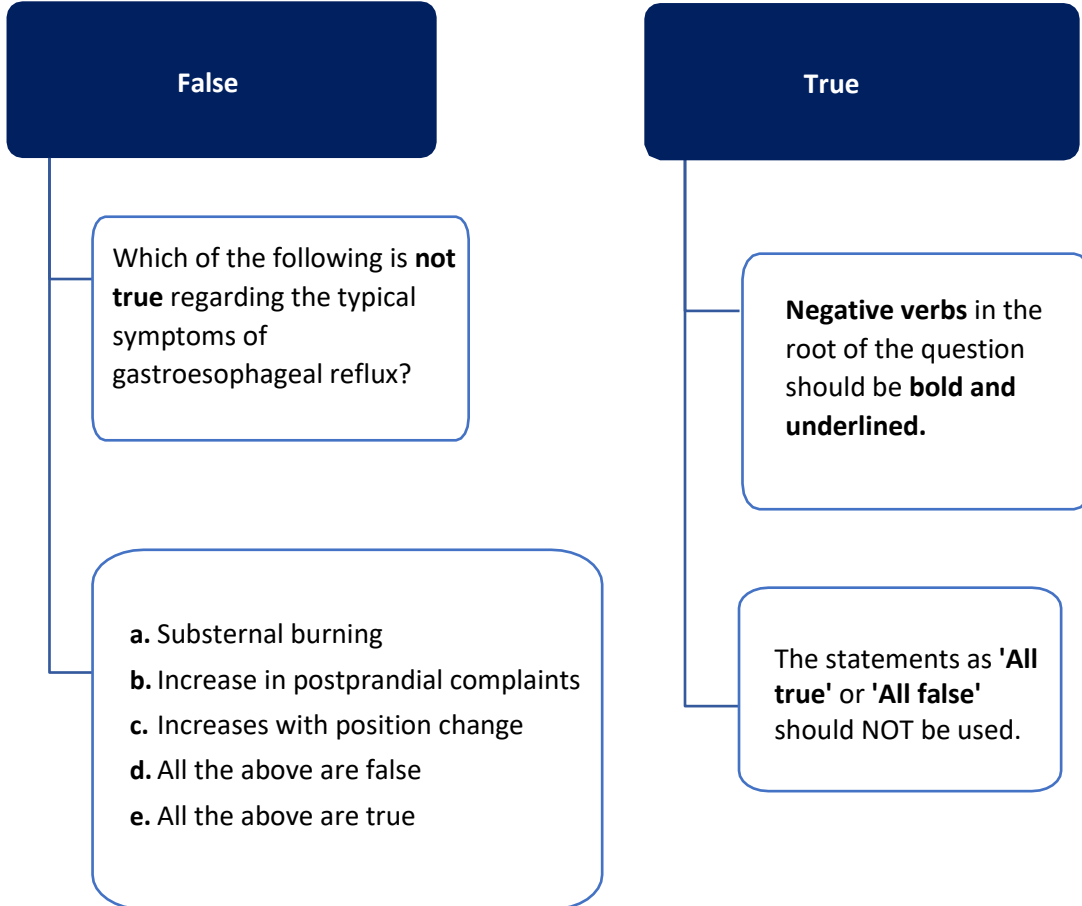
**d.** The unit should not be repeated in the options, the unit should be given at the question stem.

## SOME QUESTION EXAMPLES

1- Does the question addresses a single learning objective of the course? Topics should address one of the objectives specified in the Learning Outcome Table. For example; anatomy, biomechanics, diagnostic methods, etiopathogenesis, treatment, etc.



2- Do the options include "all above true" or "all above false"? Such statements are undesirable.



**3- The options should have unity. (If treatment is questioned, all options should be related to treatment, if a diagnosis is questioned, all options should be related to diagnosis.)**

**False**

1. Which of the following is **true?**

- a. During the pronephrosis and mesonephrosis stages of kidney development, the kidney is temporarily functional.
- b. The metanephros originate from the ureteric bud and the metanephrogenic blastema.
- c. The kidneys get their adult position in the abdomen just before birth.
- d. The bladder mainly develops from the pelvic part of the urogenital sinus.
- e. The male urethral epithelium and the female urethral epithelium all originate from the urogenital sinus ectoderm.

**True**

**Regarding the question stem:**

- question stem alone is **meaningless**.
- in the question, '**it is not stated which subject it is related to**'.

- The question is not asked for a **single learning outcome of the course**. There is a generalization.

- **Negative verbs** should be bold and underlined, not positive verbs.

**Regarding the answers:**

- the answers are **relatively long**.

- Since the question was **not asked for a single outcome, there is no** homogeneity and unity in the options. A kidney and a bladder are mentioned.



4- In the question stem, is there any information sharing about the topic? (There should NOT be information sharing about the topic, as if reminding a topic in the question.)

The unit should not be repeated in the options, the unit should be given in the root of the question.

**False**

1. The structures that make the subunits of the ribosome have various precipitation (S) coefficients.

According to this, which of the following represents the size of ribosome subunits in eukaryotes?

- a. 60S – 40S
- b. 50S – 30S
- c. 60S – 30S
- d. 50S – 40S
- e. 40S – 30S

**True**

**Information is given at the question stem.**

**There is no homogeneity in the answers.**

Some options are separated by “-“ and some are separated by “and”.

The unit is repeated in the options.

5- Does the question stem include personal opinions and judgments such as “In your opinion...”, “...what do you think”?

**False**

In your opinion, which of the following is the first disease to be considered in a **patient presenting with monoarthritis?**

- a. Acute Rheumatic Fever
- b. Psoriatic Arthritis
- c. Septic Arthritis
- d. Rheumatoid Arthritis
- e. Crystal Arthritis

**True**

- The words asking for opinion such as “**in your opinion, according to you, etc.**” should NOT be used in the question.

- If “**the first thing to consider**” is asked in this question, it should be prepared as a **case and contextual information (age, gender, weight, etc.) should be given.**

If there is no **contextual information**, the answer is unsound because the choices “a” and “e” appear at different ages.

6. A statement that can be put in the question stem should be placed at the stem and should NOT be repeated in every option.

False

1. Which two of the followings provide the formation of tractus iliotibialis?

- a. Musculus gluteus medius-  
Musculus gluteus maximus
- b. Musculus gluteus medius-  
Musculus gluteus minimus
- c. Musculus gluteus maximus-  
Musculus tensor fascia lata
- d. Musculus tensor fascia lata-  
Musculus gluteus medius
- e. Musculus tensor fascia lata-  
Musculus gluteus minimus

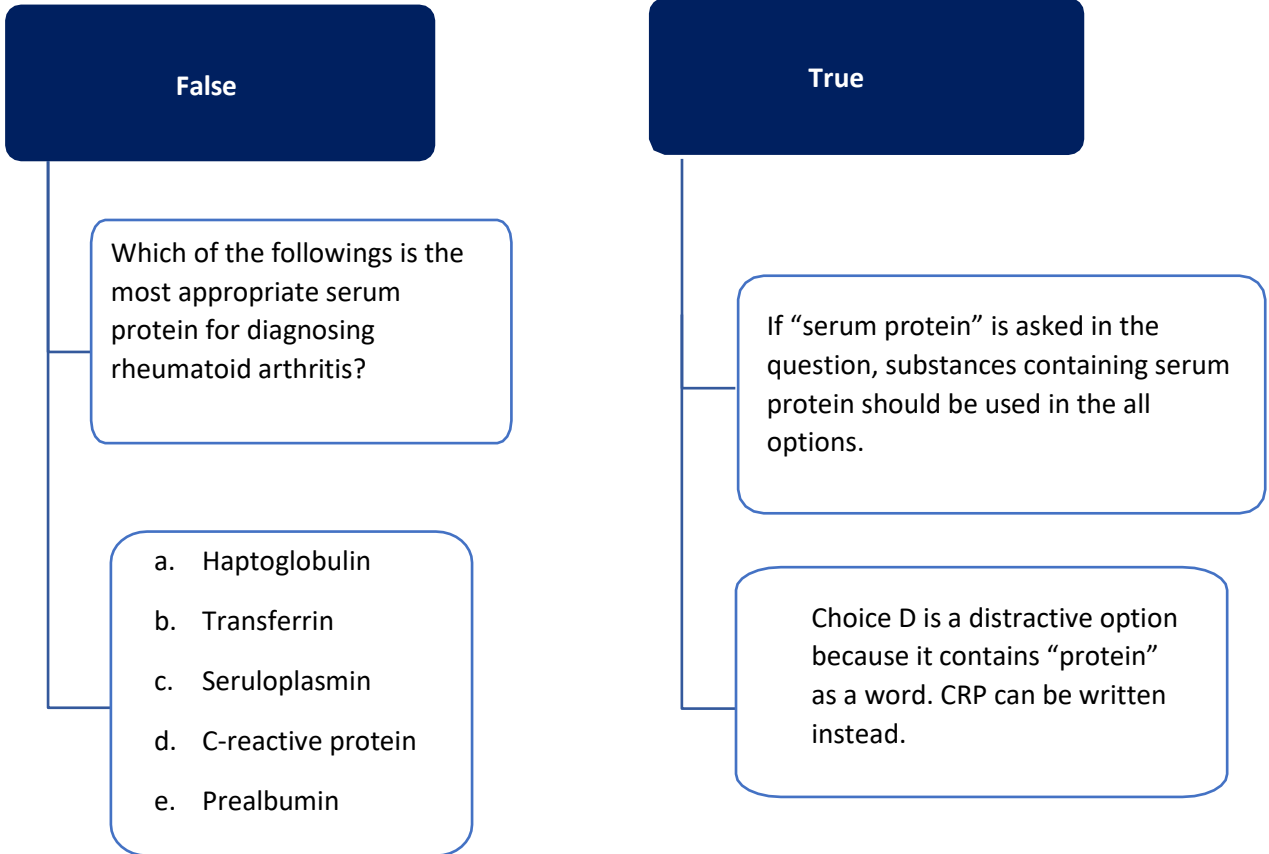
True

The phrase '**which two muscles**' is more accurate rather than "which two".

In all options, the word "musculus" was repeated twice. In this case, **unnecessary repetitions should be avoided** since all answer options are already muscle.

**Example: A- gluteus medius gluteus maxim**

7. Between the correct answer and the question stem, an analogy of statement or terminology that distinguishes the correct answer from the distractors should NOT be allowed. If such a situation is unavoidable, it should be provided that the same similarity is found in at least one other distractor.



**8. Are opposite phrases used in the options? (Opposite statements are undesirable.)**

**False**

For HIV, please mark the wrong one.

- a. It is a member of the Retroviridae family.
- b. It is an enveloped virus
- c. It is a RNA virus
- d. pH 2 is inactivated by acids
- e. It is a non-enveloped virus.

**True**

Negative questions should NOT be asked as much as possible. If unavoidable, it should be **bold and underlined.**

Choices “b” and “e” are opposite answers of each other.

## 9. If a number/value is required in the options;

- Are the values in ascending and descending order?
- Are the values expressed inclusive of each other? (Values should NOT be arranged to cover each other.)
- Are the values as ranges rather than just a number? (Using only a number should be avoided, except in special cases where it is absolutely necessary.)
- The unit should NOT be repeated in the options, the unit should be given in the question stem.

False

In a district with a population of 1000, 7 newly diagnosed diabetes mellitus (DM) were detected in 1 year, a total of 30 people with the previous cases were followed up with the diagnosis of DM, 3 of the cases died due to various reasons. What is the annual incidence of DM in this county?

- 0,30%
- 0,4%
- 0,23%
- 0,7%
- 0,37%

True

a. What is the annual incidence of DM in this province?

b. In the answer options, **'The values are mixed in order, they are not given in ascending and descending order'**

c. The options should be as follows;

- 4
- 7
- 23
- 30
- 37

**False**

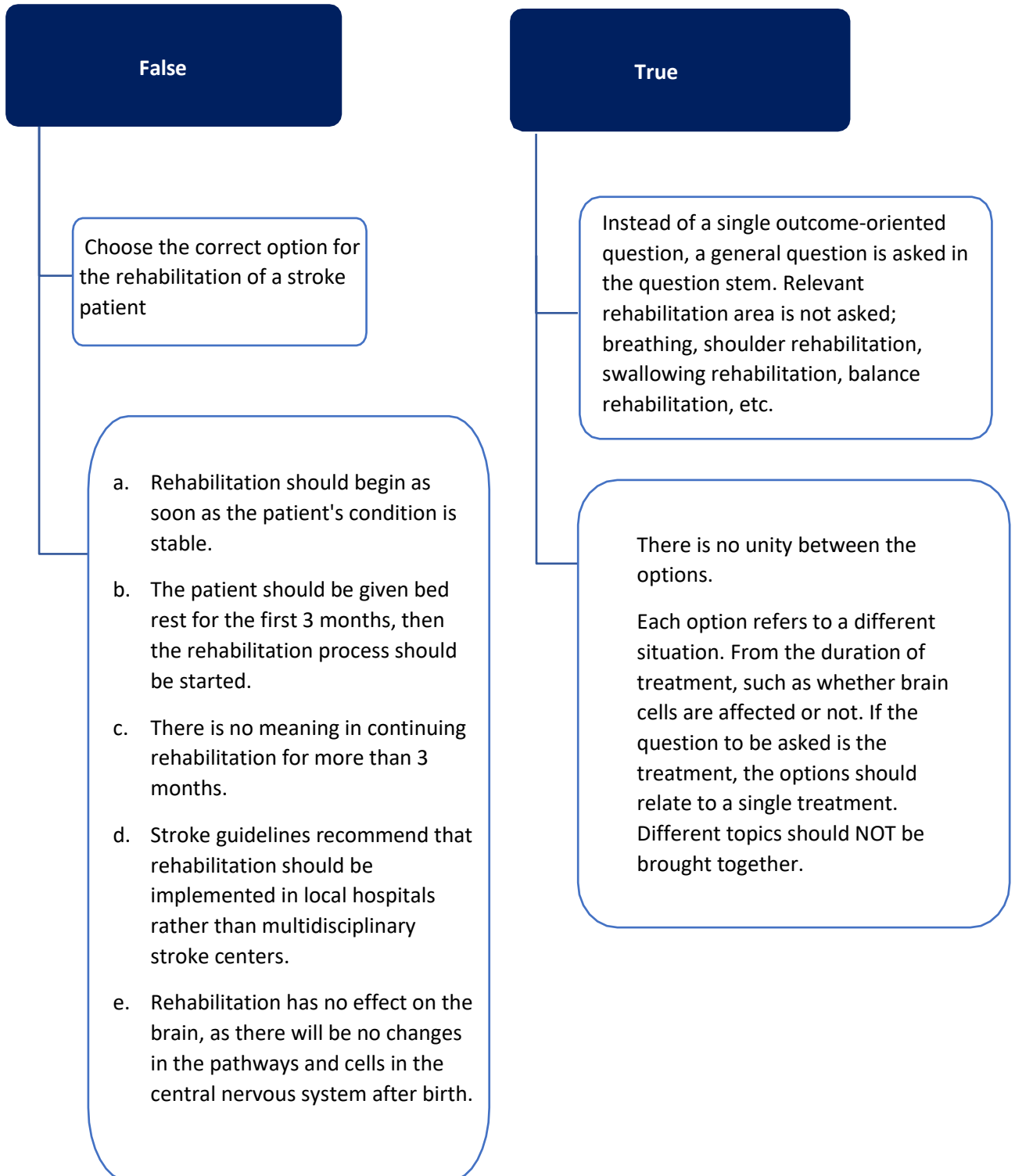
What are the feeding intervals for a term newborn?

- a. Hourly
- b. With 2-3 hours intervals
- c. With 2-5 hours intervals
- d. With 1-3 hour intervals
- e. Every time it cries

**True**

- In these options, the time intervals are intertwined.
- The “hourly intervals” as phrase are repeated unnecessarily.

**10. The options should have unity. (If treatment is questioned, all options should be treatment, if diagnosis is questioned, all options should be related to diagnosis.)**





## REFERENCES

- Guilbert JJ. (Akalin A, Solakoğlu Z (Çeviri Editörleri)). Sağlık Çalışanları İçin Eğitim Kitabı, Dünya Sağlık Örgütü. İstanbul Üniversitesi Tıp Fakültesi, İstanbul, 2000.
- İstanbul Üniversitesi Tıp Fakültesi, Eğitici Eğitimi, Kurs Kitabı.
- Amin Z, Seng CY, Eng KH. (Kara CO, Sarioğlu Büke A. (Çeviri Editörleri)). Tıp Eğitiminde Ölçme ve Değerlendirme için Pratik Rehber. İstanbul Tıp Kitabevi, 2011.
- Susan M. Case, David B. Swanson, (Contributing Authors), 2003, Constructing Written Test Questions For the Basic and Clinical Sciences, National Board of Medical Examiners (NBME®), Philadelphia, USA.
- Steven M. Downing & Rachel Yudkowsky (Ed). Assessment in Health Professions Education. New York: Routledge, Taylor & Francis Group, 2009.
- Mike Davis, Judy McKimm & Kirsty Forrest. How to Assess Doctors and Health Professionals. UK, Oxford: Blackwell Publishing Ltd, John Wiley & Sons, 2013.
- Margery H Davis, Gominda Ponnampereuma, Sean McAleer & David Rowley. Joint Committee on Intercollegiate Examinations Examiners' Manual. UK, University of Dundee, Centre for Medical Education,
- Susan M. Case & David B. Swanson. Constructing Written Test Questions For the Basic and Clinical Sciences. 3rd Edition, Philadelphia, National Board of Medical Examiners®, 2001