ROYAL COLLEGE
OF PHYSICIANS AND SURGEONS OF CANADA

Developing Multiple-Choice Questions for the Royal College Certification Examinations

Editorial Revisions
September 2013

Educational Research Unit

TABLE OF CONTENTS

Preface

This booklet consists of two sections. The first section provides a description of how to write well-constructed multiple-choice questions (MCQs) for the Royal College and includes information about the characteristics of well-constructed questions and the common problems that occur in their formulation. The second section contains two worksheets. The first worksheet consists of an exercise designed to help you recognize well-constructed and poorly constructed questions. The second worksheet consists of item templates that are designed to facilitate the construction of MCQs.

Part I

I. Introduction

MCQ examinations are, arguably, the most reliable, valid, and cost-effective method of assessing the clinical competence of candidates, especially for measuring their medical knowledge. From the candidates' perspective, MCQ examinations often consist of questions that are trivial, irrelevant, and ambiguous. Why would this discrepancy in the perceived value of MCQ examinations exist? The main reason is related to how the questions are typically constructed. Many MCQs are constructed to test the simple recall of textbook knowledge or, in an attempt to make the questions difficult for candidates, test the knowledge of relatively uncommon medical conditions. In addition, questions are often constructed in a language or format that is clear to the author but is ambiguous when read by the candidate.

The purpose of this booklet is to describe how to develop MCQs that are constructed in a format that candidates will find clear and relevant.

II. Characteristics of a Well-Constructed MCQ

A well-constructed MCQ consists of a stem, a lead-in question, and a series of response options. As an example, consider the following question:

> A 60-year-old man presents with progressive weakness in his arms and legs. He reports difficulty climbing stairs and combing his hair. He also has difficulty swallowing, but has no visual complaints. On physical examination, you note a maculopapular eruption on the eyelids, nose, cheeks, and knuckles. Joint examination is normal. What is the MOST likely diagnosis?
>
> **a) dermatomyositis**
> b) myasthenia gravis
> c) polymyalgia rheumatica
> d) rheumatoid arthritis

The first component of the question is called the stem. The stem is actually a clinical case presentation and usually consists of a presenting problem along with relevant signs, symptoms, laboratory data, etc. The second component of the question is called the lead-in question. This is the question that the candidate is being asked to answer. The last component of the question contains the response options. One of the options is chosen to be the correct answer, and the remaining options are called distractors.

This example of a well-constructed MCQ has three characteristics that should be noted. First, the question has four response options, one of which is correct. Although many types of MCQs exist, the Royal College recommends the use of the "one best answer" type of question, which has one clearly correct answer and three distractors. Second, note that the question contains a clinical scenario, and the candidate is asked to use that information to answer the question. This approach emphasizes the application of medical knowledge and makes the question appear to be more clinically relevant and valid to the candidate. The third characteristic deals with the shape of the MCQ. The stem of a well-constructed MCQ includes a clinical scenario that should contain all of the information that is necessary to answer the question. For these reasons, the stem will tend to be long but the response options should be relatively short. The figure below shows the shape of a well-constructed MCQ.

Shape of a well-constructed MCQ[1]

Long Stem (includes a clinical scenario and all relevant clinical and laboratory data)

a.
b.
c. Response Options (short)
d.

III. Steps in Developing a Well-Constructed MCQ.

Choose a topic for the question

The topic is the theme for a specific question; that is, the topic is the specific medical knowledge that a question is designed to test. When choosing a topic for a question, focus on one important concept, typically a common clinical problem encountered in your specialty.

In most cases, the topics will be given to you by your examination board chair and will be chosen from the examination blueprint. The blueprint is a guide that is used for creating an examination, and consists of a list of the competencies and topics that should be assessed on an examination.

B. Choose the appropriate context for the question

The context specifies the clinical scenario that will be used to test the topic. Context is important because it determines what type of information should be included in the stem and the response options. Consider the following two examples.

Example 1.
Topic: Turner syndrome
Context: Physical Examination
A 16-year-old adolescent girl presents with primary amenorrhea. Which of the following signs BEST supports the diagnosis of Turner syndrome?

a) hypertension
b) hirsutism
c) short stature
d) epicanthal folds

Example 2.
Topic: Turner syndrome
Context: Diagnosis
A 16-year-old adolescent girl presents with primary amenorrhea. On examination, you note that she is 148 cm tall. In addition, you note that her external genitals are immature and there is no breast development. What is the MOST likely diagnosis?

a) Turner syndrome
b) mixed gonadal dysgenesis
c) pure gonadal dysgenesis
d) Noonan syndrome

[1] Case, S.M. & Swanson, D.R. (1998). Constructing written test questions for the basic and clinical sciences. (pp. 42). National Board of Medical Examiners: Philadelphia

Notice that both examples are testing the topic of Turner syndrome. The context of the questions differs, however, and this difference influences the type of information that is presented in the question stem. In the first example, the context is a physical examination, so the stem and response options contain information likely to be found during a physical examination. In the second example, the context is diagnosis, so the stem contains relevant signs and symptoms, and the response options consist of potential diagnoses.

Common clinical contexts that could be used for constructing an MCQ include the following: interpreting data, eliciting data (physical examination, history-taking), further investigations, diagnosis, initial management, long-term care, risk factors, side effects and contraindications, counselling, and ethical issues.

C. Create a stem

1. Use clinical scenarios

Clinical scenarios provide a good basis for a question stem. The clinical scenario should describe the presenting complaint first and be followed by a listing of relevant signs, symptoms, results of diagnostic studies, initial treatment, subsequent findings, etc. In essence, the stem should contain all the information that is necessary for a competent candidate to answer the question.

2. Use a clear lead-in question

The lead-in question should give clear directions as to what the candidate should be doing to answer the question. The lead-in question should always be a complete sentence. Consider the following examples of lead-in questions:

Example 1: With respect to myocardial infarction:
Example 2: What is the MOST likely diagnosis?

In Example 1, the candidate is not presented with a task. This type of lead-in statement will often result in an ambiguous or unfocused question. In Example 2, the task is clear and will result in a more focused question. If the lead-in question is well constructed, the candidate should be able to answer the question without looking at the response options. As a check, cover the response options and try to answer the question.

3. Ensure content is at an appropriate level of difficulty

Well-constructed MCQs should reflect an appropriate level of difficulty to test the candidate's knowledge. For Royal College examinations, the questions should be designed to test the knowledge of residents who are ready to practice their profession competently. In other words, would specialists on their first day of practice know how to answer the question?

Testing residents' knowledge does not mean that a question must be extremely difficult. If the question is testing knowledge that is essential to the practice of the specialty, then the question may in fact be quite easy.

4. Make the question clinically relevant

Try to focus on problems that would be encountered in clinical practice rather than assessing the candidate's knowledge of trivial facts or obscure problems that are seldom encountered. The types of problems that you frequently encounter in your own practice can provide good examples for developing questions.

5. Test the application of medical knowledge
Well-constructed MCQs should test the application of medical knowledge rather than just the recall of information. The benefits of testing the application of knowledge include the following: the question will be focused on clinically important information rather than trivia; the question will identify those candidates who have memorized factual information but are unable to use that information effectively, and; from the candidate's perspective, the validity of the question will be improved. The use of a clinical scenario as the basis for a question will help ensure that a question tests the application of medical knowledge.

D. Create the correct answer

The correct answer should be clearly correct. If the "best answer" is sought, then this should be stated in the lead-in question.

When creating the correct answer, try to avoid clues that would reveal an option as being the correct answer. Some common problems to avoid include the following:

1. The correct answer is longer than the other distractors.
2. Textbook wording is used for the correct answer but not for the distractors.
3. Specific qualifiers (e.g. always, never) are used in the correct answer but not in the distractors.

E. Create the distractors

A good distractor should be an inferior choice to the correct answer but should also be plausible to an incompetent candidate. When creating a distractor, it may help to consider how an inexperienced resident would manage the clinical scenario described in the stem. In addition, try to avoid clues that would reveal a response option as a distractor. Some common problems to avoid include the following:

1. The distractors and the correct answer are not homogenous in content (i.e. the correct answer is a treatment, the distractors are tests).
2. The grammar of the distractors does not match the grammar of the stem.
3. The distractors are not the same length as the correct answer.
4. The grammar of the distractors and the correct answer is not the same (i.e. options and answer should have the same grammatical construction).

IV. Other Guidelines to Consider When Constructing an MCQ

A. Avoid the use of "all of the above" as a response
A candidate has to identify only two response options as correct to know that "all of the above" is the correct response. This reduces the value of the question. In addition, "all of the above" implies that there is more than one correct answer. The Royal College recommends that MCQs be constructed so that only one option is considered the best response.

B. Avoid mutually exclusive options
For questions that require a single best answer, options that contradict one another cannot both be correct, and therefore mutually exclusive options reduce the number of plausible responses.

C. Avoid overlapping content in the response options

The response options should be independent of one another in terms of their content. For example, imagine that one had a written question about pain management in which the correct answer was to "prescribe an analgesic" and one of the distractors was "prescribe acetaminophen". There is an overlap in the content of these two response options and therefore they are not independent of one another.

D. Avoid imprecise terms such as sometimes, frequently, often, etc.

The definition of these terms is ambiguous and will cause confusion if used in an examination question.

E. Avoid the use of negative terms in the lead-in question (i.e. all of the following except)

Negative terms tend to overly complicate a question. In addition, you are primarily interested in whether the candidates know the best response, not necessarily the poorest response.

F. Avoid the use of "None of the above" as the correct option.

"None of the above" should not be used as a correct option because it does not demonstrate that an examinee knows the best answer; it only shows that the examinee knows what the answer is not.

Part II

Worksheet 1: Examples of MCQs

The following is an exercise to help you recognize well-constructed and poorly constructed MCQs. Some of the questions that follow are well constructed and others are not. Read the questions and, if you think one is poorly constructed, then list the problems associated with this question. The last part of this worksheet displays which items we feel were well constructed, which items we feel were poorly constructed, and a list of problems that have been identified.

1. A 32-year-old man who is unemployed and suffers from alcoholism underwent a mastoidectomy in his youth. He now presents with headaches, nausea, vomiting, drowsiness, and confusion. He does not have a fever. You are unable to visualize his right eardrum but there appears to be some discharge from his ear. He also experiences slight neck stiffness. What is the MOST appropriate investigation at this time?

    a) lumbar puncture
    b) ECG
    c) x-ray of the skull
    d) CT scan of the head

    Problems:

2. When a tendon is cut and repaired, what would be the strength of the repaired tissue after one year?

    a) almost always less than normal
    b) usually greater than normal
    c) almost the same as normal
    d) more or less than normal, depending on the age of the patient

    Problems:

3. According to the American Heart Association guidelines, in what way should a patient with a prosthetic heart valve be given prophylactic antibiotic treatment before a surgical procedure?

    a) in routine fashion to everyone
    b) according to the magnitude of the procedure
    c) according to the type of microbial flora most likely to cause endocarditis
    d) only for gastrointestinal procedures

    Problems:

4. A 32-year-old woman presents with a two-week history of diarrhea associated with heat intolerance, sweating, and restlessness. Physical examination reveals a blood pressure of 150/60 mm Hg and a pulse of 106/bpm. She has a fine tremor of her outstretched arms. Her thyroid is diffusely enlarged, firm, and tender. Which of the following tests will help to establish the etiology of her thyrotoxicosis?

    a) antithyroid antibodies
    b) sensitive thyroid-stimulating hormone assay
    c) free triiodothyronine (T_3)
    d) radioactive iodine uptake

    Problems:

5. With respect to the management of foot ulcers in a patient with diabetes, which of the following statements about assessment for arterial revascularization of the lower limbs is TRUE?

   a) It is not beneficial because nonvascular factors, such as neuropathy and infection, minimize the benefits of revascularization.
   b) It is not beneficial because artherosclerosis is too widespread for surgical correction to be beneficial.
   c) It is not beneficial because arterial revascularization is too limited for surgical correction to be beneficial.
   d) It is advisable because artherosclerosis is sometimes segmental and amenable to surgical correction.

   Problems:

6. Pulmonary embolism:

   a) always associated with a fever.
   b) never seen in non-smokers.
   c) always confused with pneumonia.
   d) treated by administering heparin.

   Problems:

7. A 55-year-old man presents with shortness of breath and purulent sputum. There is no history of hemoptysis or chest pain. On several occasions in the past few days, he has experienced episodes of feeling hot or cold, but there have been no rigors. Chest examination shows hyperinflation and decreased breath sounds without dullness or crackles, but with scattered wheezes. Chest radiograph is normal. Spirometry shows the following results: FEV_1 1.68 (58% predicted), FVC 2.12 (75% predicted). In managing this patient, you would suggest

   a) intravenous antibiotic.
   b) oral theophylline.
   c) inhaled bronchodilator therapy.
   d) smoking cessation.

   Problems:

8. Which of the following drugs given in the setting of acute myocardial infarction has NOT been shown to reduce mortality?

   a) intravenous r-tissue plasminogen activator
   b) intravenous streptokinase
   c) acetylsalicylic acid
   d) nifedipine

   Problems:

Worksheet 1 (continued)

Potential problems with the questions

Question 1: - this is a well-constructed MCQ.

Question 2: - uses vague descriptors like “more or less” and “usually”.

Question 3: - the correct answer is longer than the distractors. One of the distractors is cued as a wrong answer. The question is asking for methods of administering a treatment and distractor D is not a method.

Question 4: - Either all distractors should mention the word “test” or “assay”, or none of them should as the idea of a test is in the question.

Question 5: - the stem is vague and can’t be interpreted without reading the options
- the correct answer is cued because it is the only positive option.
- the shape of the item is incorrect. The response options are almost as long as the stem. A well-constructed MCQ has a long stem and short response options.

Question 6: - the lead-in question is unclear
- uses absolute terms (e.g. always, never)
- response options are not homogenous (signs, diagnosis, risk factors, treatments)
- the distractors are not grammatically correct when combined with the stem
- the shape of the item is incorrect. The response options are longer than the stem. A well-constructed MCQ has a long stem and short response options.

Question 7: - the lead-in question is unclear. Is the question testing the first step in managing the patient, long-term care for the patient, or the most effective treatment?

Question 8: - it is not clear what is being measured because of the negative lead-in statement. Is the question testing whether the candidate knows that nifedipine does not reduce mortality or that the other drugs do reduce mortality?

Worksheet 2: Item Templates

Constructing good MCQs can be difficult and some question writers find an item template to be a useful tool. Item templates are designed to outline the structure of a well-constructed MCQ but do not contain the content of the question. As a question writer, one would choose a particular template and fill in the blanks with the appropriate information. The following item template is an excerpt from a book by Case and Swanson[2] and should prove to be useful for creating questions.

An item template can depict the overall structure of an item. You can typically generate many items using the same template. For example, the following template could be used to generate a series of questions related to gross anatomy:

> A *(patient description)* is unable to *(functional disability*). Which of the following structures is MOST likely to have been injured?

This is a question that could be constructed using the above template:

> A 65-year-old man has difficulty rising from a seated position and straightening his trunk, but has no difficulty flexing his leg. Which of the following muscles is MOST likely to have been injured?
>
> **a) gluteus maximus**
> b) gluteus minimus
> c) hamstring
> d) iliopsoas

Many basic science questions can be presented within the context of a clinical scenario or patient vignette. The patient vignettes may include some or all of the following components:

> Age, Gender (e.g. a 45-year-old man)
> Site of Care (e.g. comes to the emergency department)
> Presenting Complaint (e.g. because of a headache)
> Duration (e.g. that has continued for two days)
> Patient History (e.g. with a family history of)
> Physical Findings
>
> +/– results of diagnostic studies
> +/– initial treatment, subsequent findings, etc.

[2] Case, S.M. & Swanson, D.R. (1998). Constructing written test questions for the basic and clinical sciences. (pp. 42). National Board of Medical Examiners: Philadelphia

Additional Templates

A *(patient description)* has a *(type of injury and location)*. Which of the following structures is MOST likely to be affected?

A *(patient description)* has *(history findings)* and is taking *(medications)*. Which of the following medications is the MOST likely cause of his condition *(on history, physical examination or laboratory finding)*?

A *(patient description*) has *(abnormal findings)*. Which [additional] finding would suggest/suggests a diagnosis of *(disease 1)* rather than *(disease 2)*?

A *(patient description)* has *(symptoms and signs)*. These observations suggest that the disease is a result of the *(absence or presence)* of which of the following *(enzymes, mechanisms)*?

A *(patient description)* follows a *(specific dietary regime)*. Which of the following conditions is MOST likely to occur?

A *(patient description)* has *(symptoms, signs, or specific disease)* and is being treated with *(drug or drug class)*. The drug acts by inhibiting which of the following *(functions, processes)*?

A *(patient description)* has *(abnormal findings)*. Which of the following *(positive laboratory results)* would be expected?

(time period) after a *(event, such as trip or meal with certain foods)*, a *(patient or group description)* became ill with *(symptoms and signs)*. Which of the following *(organisms, agents)* is MOST likely to be found on analysis of *(food)*?

Following *(procedure)*, a *(patient description)* develops *(symptoms and signs)*. Laboratory test results show *(findings)*. What is the MOST likely cause of this presentation (or result)?

A *(patient description)* dies of *(disease)*. What would be the MOST likely finding on autopsy?

A patient has *(symptoms and signs).* What would be the MOST likely explanation for the *(findings)*?

A *(patient description)* has *(symptoms and signs)*. Exposure to which of the *(toxic agents)* is the MOST likely cause of this presentation?

What is the MOST likely mechanism of the therapeutic effect of *(drug class)* in patients with *(disease)*?

A patient has *(abnormal findings)*, but *(normal findings)*. What is the MOST likely diagnosis?