ROYAL COLLEGE
OF PHYSICIANS AND SURGEONS OF CANADA

Developing Multiple-Choice Questions for the Royal College Certification Examinations

**Editorial Revisions
November 2011**

Educational Evaluation and Analysis Unit

TABLE OF CONTENTS

Preface

As a member of an examination board, test committee, or a specialty committee, you may be asked to develop multiple-choice questions (MCQs) for an examination. This document is intended to assist you in developing MCQs that are well constructed, reliable, and valid.

The booklet consists of two sections. The first section provides a description of how to write well constructed MCQs for the Royal College and includes information about the characteristics of well constructed questions and common problems that occur when writing them. The second section contains four 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. The third worksheet consists of a checklist that can also be used to facilitate the construction of questions. The fourth worksheet consists of a form that is used for submitting questions to the Royal College. These worksheets were designed to augment the information found in the first section of this booklet but can also be used as tear away worksheets to help you work on specific tasks.

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 candidates.

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. For example consider the following question:

> A 60-year-old man presents with progressive weakness of arms and legs. He reports difficulty climbing stairs or combing his hair. He also has difficulty swallowing, but he 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, lab tests, etc. The second component of the question is called the lead-in question. This is the actual question that the candidate is 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 question has three characteristics that should be noted. First, the question has four response options, one of which is correct. Although many types of multiple-choice questions 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 consists of a clinical situation and asks the candidate 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 candidates. The third characteristic deals with the shape of the item. The stem of a well constructed question consists of a clinical case and 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 options should be relatively short. The figure below shows the structure of a well constructed MCQ.

Shape of a well constructed question[1]

> **Long Stem: consists of a clinical case and all relevant facts.**

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

III. Steps in Constructing a Well Constructed MCQ

A. Choose a topic for the question

The topic is the theme for a specific question; that is, it 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 from your specialty.

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

B. Choose the appropriate context for the question

Context defines the clinical situation that will 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's Syndrome
   Context: Physical Examination
   An 18-year-old woman presents with primary amenorrhea. Which of the following signs best supports the diagnosis of Turner's syndrome.
         a) Hypertension
         b) Hirsutism
         c) Short Stature
         d) Epicanthal folds

Example 2.
   Topic: Turner's Syndrome
   Context: Diagnosis
   An 18-year-old woman presents with primary amenorrhea. On exam you notice that she is 148 cm tall. In addition, you note that her external genitals are immature and there is no breast development. What is your most likely diagnosis?
         a) Turner's Syndrome
         b) Mixed Gonadal Dysgenisis
         c) Pure Gonadal Dysgenisis
         d) Noonan's 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 same topic, which is Turner's Syndrome. The context of the questions differs, however, and this difference influences the type of information that is presented in the question. For the first example, the context is a physical examination so the stem and response options contain information likely to be found during a physical exam. For the second question 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 exam, history taking), further investigations, diagnosis, initial management, long term care, risk factors, side effects and contraindications, counseling, and ethical issues.

C. Create a stem

1. Use clinical cases
Clinical cases provide a good basis for a stem. The clinical case should begin with presenting a problem and be followed by relevant signs, symptoms, results of diagnostic studies, initial treatment, subsequent findings, etc. In essence, all the information that is necessary for a competent candidate to answer the question should be provided in the stem.

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. For example, consider the following examples of lead-in questions.
>   Example 1: Regarding myocardial infarction:
>   Example 2: What is the most likely diagnosis?

Note that for the Example 1, no task is presented to the candidate. This type of lead-in statement will often lead to an ambiguous or unfocused question. In the second example, the task is clear and will lead to a more focused question. To ensure that the lead-in question is well constructed, the question should be answerable without the candidate having to look at the response options. As a check, cover the response options and try to answer the question.

3. The content should be at an appropriate level of difficulty
Well constructed MCQs should be written at an appropriate level of difficulty to test the knowledge level of the candidates. For Royal College examinations, the questions should be designed to test the knowledge of a resident who is ready to practice their profession competently. In other words, would a specialist on his/her first day of practice know how to answer the question?

Note that testing the appropriate knowledge level of a resident 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 actually 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 on obscure problems that are seldom encountered. The types of problems that you commonly 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. Benefits to 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 candidates perspective, the validity of the question will be improved. The use of a clinical case 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 determiners (always, never) are used in the correct answer but not in the distractors.

E. Create the distractors

A good distractor should be inferior to the correct answer but should also be plausible to a noncompetent candidate. When creating a distractor, it may help to think how an inexperienced resident would respond to the clinical case 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 (e.g. 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.

IV. Other Guidelines to Consider When Constructing an MCQ Item

A. Avoid the use of "all of the above" as a response
A candidate only has to identify 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 correct.

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 information in the response options should be independent of one another. 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 Tylenol". 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 like sometimes, frequently, often, etc
The definition of these terms is ambiguous and will cause confusion if used on 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. Use "none of the above" sparingly
Sometimes it is difficult to create four plausible options and therefore the response "none of the above" can be used. If "none of the above" is used, however, it should be the correct answer on at least 1/5 of the questions used on the examination and it should be clearly correct or clearly incorrect.

Worksheet 1: Examples of MCQ items

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

1. A 32-year-old unemployed alcoholic who underwent a mastoidectomy as a youth presents with headaches, nausea, vomiting, drowsiness, and confusion. He does not have a fever, but his right eardrum is not visualized and there appears to be some discharge. There is slight neck stiffness as well. What is the most appropriate investigation at this time?
    a) Lumbar puncture
    b) ECG
    c) X-ray the skull
    d) CT scan the head
    Problems:

2. When a tendon is cut and repaired, what is the strength of repaired tissue after one year?
    a) almost always less then 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 guidelines of the American Heart Association, 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 2-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/minute. She has a fine tremor of her outstretched arms. Her thyroid is diffusely enlarged, firm and tender. Which one 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 (T3)
    d) Radioactive iodine uptake
    Problems:

5.   In the management of foot ulcers in diabetics, 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) is never seen in non smokers

      c) is always confusable with Pneumonia

      d) is 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 exam shows hyperinflation and decreased breath sounds without dullness or crackles but with scattered wheezes. Chest radiograph is normal. Spirometry shows the following: FEV1: 1.68 (58% predicted),FVC 2.12 (75% predicted). In managing this patient you would suggest:

      a) intravenous antibiotic

      b) oral theopylline

      c) inhaled bronchodilator therapy

      d) smoking cessation

Problems:

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

      a) intravenous r-tissue plasminogen activator

      b) intravenous streptokinase

      c) acetylsalicylic acid

      d) nifedipine

Problems:

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:   - this is a well constructed MCQ

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 imprecise terms (e.g. frequently)
                - options are not homogenous (signs, diagnosis, risk factors, treatments)
                - the first distractor is not grammatical 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 people find an item template to be a useful tool. Item templates are designed to have the structure of a well constructed MCQ but are missing 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.

The overall structure of an item can be depicted by an item template. 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 is most likely to have been injured?

This is a question that could be constructed using this template:

> A 65-year-old man has difficulty rising from a seated position and straightening his trunk, but he has no difficulty flexing his leg. Which of the following muscles is mos t likely to have been injured?
>
> **A. Gluteus maximus**
> B. Gluteus minimus
> C. Hamstrings
> D. Iliopsoas

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

> Age, Gender (eg, A 45-year-old man)
> Site of Care (eg, comes to the emergency department)
> Presenting Complaint (eg, because of a headache)
> Duration (eg, that has continued for 2 days)
> Patient History (with Family History ?)
>
> 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 *(one history, PE or lab 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 findings show *(findings)*. Which of the following is the most likely cause?

A *(patient description)* dies of *(disease)*. Which of the following is the most likely finding on autopsy?

A patient has *(symptoms and signs).* Which of the following is 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?

Which of the following is the most likely mechanism of the therapeutic effect of this *(drug class)* in patients with *(disease)*?

A patient has *(abnormal findings)*, but *(normal findings)*. Which of the following is the most likely diagnosis?"

Worksheet 3: Checklists for the Development of MCQs

A   Follow these five steps when developing a question.

1. Choose a topic for the question   ●
   Topics are the specific knowledge that the question is designed to test. They are related to the competencies that the examination should be testing.

2. Decide on the context for testing the objective   ●
   Context is the clinical situation (interpreting data, diagnosis, management) that will determine what information should be provided in the question.

3. Write the stem of the question   ●
   The stem should use a clinical case as the basis of the question. It should also contain all relevant information necessary to answer the question and should end with a clear question.

4. Create four response options   ●
   Choose one option to be the correct answer and this option should be clearly correct. The remaining three options are called distractors and should be clearly incorrect but plausible to a weaker candidate.

5. Try the question on a colleague   ●
   Other people often notice problems that the author may have missed or not considered when writing the question.

B   Use this checklist when creating a stem.

1. Is the question related to a topic from the blueprint?   Yes
2. Is the stem relevant to clinical practice?   Yes
3. Was a clinical case used as the basis for the question?   Yes
4. Does the stem consist of all information that a competent candidate will require to answer the question?   Yes
5. Does the lead-in question clearly indicate how to answer the question?   Yes
6. Is the stem written at an appropriate level of difficulty?   Yes
7. Can the question be answered without looking at the options?   Yes
8. Does the question test the application of medical knowledge rather than recall?   Yes

C.   Use this checklist when creating the response options.

1.	Is there one clearly correct answer?	Yes
2.	Are all the distractors plausible to a weak candidate?	Yes
3.	Are there any obvious clues to the correct answer (all options are homogenous, grammatical, same length, same degree of technical language)?	No
4.	Were terms like "all of the above", and "all of the following except" used?	No
5.	Were terms like "none of the above" used sparingly and if so were the items occasionally correct?	Yes
6.	Were any imprecise terms (frequently, sometimes, often) used?	No

D.   Does the item have an appropriate shape?

The stem of a well constructed MCQ usually consists of a clinical case and should contain all of the information necessary to answer the question. For these reasons, the stem will tend to be relatively long but the options should be relatively short. The figure below presents a diagram of how a well constructed MCQ should look.

Shape of a well constructed question.[3]

Long Stem: consists of a clinical case and all relevant facts.

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

---

[3] 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

Worksheet 4: Item Submission Form

NAME: _____

SPECIALTY: _____

TOPIC: _____

In the space below please **type** your question**

CORRECT ANSWER:

KEYWORDS:

REFERENCE: Journal   - Author(s), volume, page(s) and year
                         Book   - Title, edition, page(s)

CLASSIFICATION:

** ITEMS MUST BE CLASSIFIED AND HAVE A REFERENCE

Multiple-Choice (MC) Item Writing Guidelines for Royal College Exams

General Approach
1. Ensure that the "One Best Answer" (or Type A) format is used.
   a) Do not use the True/False, matching, multiple-correct answer or complex Type K format (that tests logic and reading skills rather than content knowledge).
2. Although research suggests three options may be adequate, Royal College examinations like other high stakes exams require 4 options including only one correct answer.
3. Make the question simple and straightforward by minimizing the amount of reading in each item and focusing on what you are assessing. Can you assess what you want with simpler wording?
4. The purpose of Royal College examinations is to assess competence of residents preparing for independent practice. Some questions should be designed to discriminate between competent and non-competent candidates whereas others may be mastery-level questions (questions that test knowledge that all competent candidates should know

Content
5. Avoid trick items.
6. Each question should have one main focus or focus on a single problem.
7. Base each item on important content to assess:
   a) Avoid trivial content.
   b) Every item should reflect specific content area as defined by an exam blueprint.
8. As appropriate, include a clinical setting to increase the face validity of the examination.
9. Use relevant material to test higher level learning such as application of content; minimize testing for simple recall only.
10. Ensure that the content of each item does not cue answers other items on the test.
11. Avoid over specific and over general content.
12. Do not create opinion-based items.
13. Keep vocabulary appropriate for the group being tested.
    a) Avoid the use of acronyms.
    b) Use nationally accepted terms common to the specialty.

Writing the stem
14. Include a lead in question and phrase the item as a question when possible
15. Ensure that the directions in the stem are very clear.
16. Include the central idea in the stem instead of the choices.
17. Avoid window dressing such as excessive verbiage or unnecessary "red herrings".
18. If using a "fill-in-the-blank" format, avoid using multiple blanks and do not put blanks at the beginning of a statement.
19. Word the stem positively, avoid negatives such as NOT or EXCEPT. If negative words are used, use the word cautiously and always ensure that the word appears CAPITALIZED and **boldface**

Writing the choices
20. Make all three distractors plausible yet definitively incorrect.
21. Keep choices independent; choices should not be overlapping.
22. Keep choices homogeneous in content and grammatical structure.
23. Keep the length of choices roughly equal.
24. None of the above should be used carefully.
25. Do not use All of the above.
26. To optimize your choices:
    a) Use typical errors of candidates to write your distractors;
    b) Use true statements and familiar phrases that do not correctly answer the question;
    c) Use item analysis to optimize your distractors and item discrimination.
27. Avoid giving clues to the right answer, such as
    a) Specific determiners including always, never, completely, and absolutely.
    b) Clang associations (choices identical to or resembling words in the stem).
    c) Grammatical inconsistencies that cue the test taker to the correct choice.
    d) Conspicuous correct choice.
    e) Pairs or triplets of options that clue the test taker to the correct choice.
    f) Blatantly absurd, ridiculous options.

Formatting concerns
28. Keep the entire question on one page.
29. Format the item vertically and not horizontally.
30. Vary the location of the right answer according to the number of choices.
31. Place choices in logical or numerical order.

Style concerns
32. Use correct and consistent grammar, punctuation, capitalization, and spelling.
33. Given the high-stakes nature of Royal College examinations, do not use humour.
34. Edit and proof items. Have others edit and proof your items as well.