MD Program
UNIVERSITY OF TORONTO

Office of Faculty Development

The Essentials of Item-writing for Multiple-choice Examinations

Participant Package

Included in this package:

1) Writing Effective Multiple-Choice Questions - Guidelines for faculty who are creating MCQs for the assessment of medical students

2) Writing Effective Multiple-Choice Questions (MCQ) for the Non-Medical Expert Competencies

3) Steps to Guide the Review and Editing of Existing Multiple-Choice Questions for Use on a Mastery Exercise

4) Worksheet - Creating and Editing Multiple-Choice Questions (MCQ) for use on Mastery Exercises

MD Program
UNIVERSITY OF TORONTO

Writing Effective Multiple-Choice Questions (MCQ)

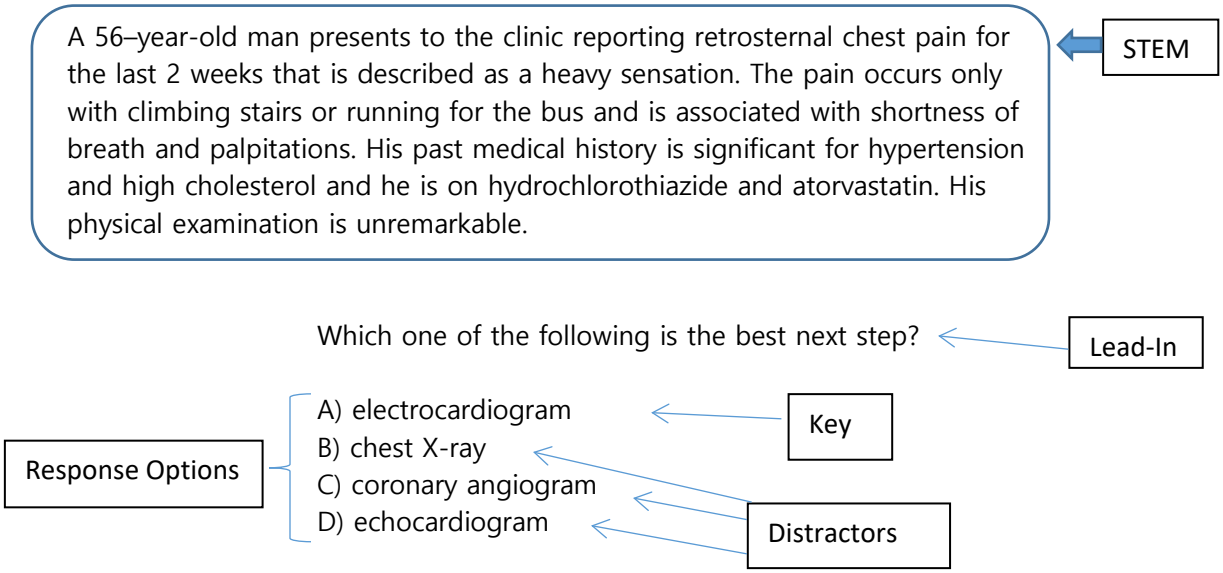
Guidelines for faculty who are creating MCQs for the assessment of medical students

BACKGROUND AND FORMAT

This document is meant to guide you to create well-constructed MCQs for the assessment of medical students in the Foundations and Clerkship Curriculum. It takes you through a series of steps that help you to conceptualize a question, create it in a well-constructed fashion, and confirm that the pitfalls common to item generation have been avoided.

Terminology

The main body of the question is the **Stem**. The task is called the **Lead-In** question. The response options are the **Distractors** (wrong answers) and the **Key** (the right answer). See example below:



Levels of learning targeted by MCQs

Bloom's Revised Taxonomy (Anderson, 2001) is used as a framework to understand different levels of cognitive behaviour or levels of learning and MCQs can be written to specifically test student learning at a particular level. The three levels that are most applicable to our assessments are: 1)Remembering 2) Understanding and 3) Applying. These represent a continuum of increasing cognitive complexity.

©2016. Faculty of Medicine, University of Toronto. These guidelines were developed and approved by the MD Program Test Committee, December 2016.

DESCRIPTION OF DIFFERENT TYPES OF QUESTIONS

Remembering Questions

What is it? Questions that target this level of knowledge require simply, that the student be able to recall information. Answering "Remembering" questions correctly can be accomplished with memorization of facts, and does not require an understanding of the material. These questions typically do not have a clinical scenario as a stem.

Example. In considering the topic of acute renal failure, a "Remembering" MCQ could be the following:

> Which one of the following best estimates the prevalence of acute renal failure requiring hospitalization in Canada from 2000-2015 (unit: Per million)?
> (a) 15
> (b) 800
> (c) 2000
> (d) 15,000

Examples. The following are additional examples of "Remembering" types of MCQs:

> Which one of the following symptom is likely to occur in coronary artery disease?
> What is the median survival of stage 4 small cell lung cancer?
> Which one of the following formulas is correct for the calculation of GFR?
> Which one of the following best describes the lesion in the attached picture?
> Which one of the following is a potential side effect of opioid medications?

Understanding Questions

What is it? Rather than a simple recall of facts, "Understanding" questions require a deeper comprehension of the concepts but do not ask students to apply their knowledge to a specific clinical case example. As such, there is no clinical decision making task. Similar to "Remembering" questions, and unlike "Applying" questions, "Understanding" questions typically do not have a clinical scenario as a stem. "Understanding" questions may ask students to identify examples of a given term or concept or interpret the meaning of an idea or concept.

©2016. Faculty of Medicine, University of Toronto. These guidelines were developed and approved by the MD Program Test Committee, December 2016.

Example. The following is an example of an "Understanding" question related to the example of acute renal failure:

> Which one of the following is the best estimate of the glomerular filtration rate (GFR) given the following parameters?

Age	86 years
Gender	Female
Weight	50 mg
Creatinine	700 mmol/L

a) 15
b) 30
c) 50
d) 90

Comments. In this example, the student is expected to know the formula for GFR and know how to use it. They are not, however, presented with a concrete clinical scenario and asked to apply their knowledge to a patient case.

Example. The following are additional examples of "Understanding" questions:

> Which one of the following arterial blood gas results would be consistent with a metabolic acidosis and compensatory respiratory alkalosis?

> An elementary school has 200 students. A group of 100 students is exposed to a toxic substance. Over half (N=56) of the students who are exposed develop a skin rash 3 days later. Another 12 students who are not exposed also develop the rash. What is the relative risk of developing the rash with exposure to the toxic substance?

Cognitive Integration: An "Understanding" question may ask a student to make a connection between the presentation or management of illness and the underlying mechanism of disease or treatment, and if so, it is assessing the student's ability in the area of "cognitive integration". In fact, many questions at the "Understanding" level also test "Cognitive Integration." These types of questions require the learner to understand the mechanistic explanation or the cause and effect story that links clinical concepts and underlying basic science concepts. These questions are trying to get at the "why" or "how" explanations. They allow an assessment of those basic science concepts that are most relevant to a clinician and are thus most likely to be perceived by medical students to be a fair assessment of what they should know.

©2016. Faculty of Medicine, University of Toronto. These guidelines were developed and approved by the MD Program Test Committee, December 2016.

Example. Re-visiting the example of acute renal failure, the following would be a "Cognitive Integration" question at the "Understanding" level:

Which one of the following are you most likely to find on kidney biopsy in a patient with acute renal failure secondary to severe dehydration?

a) Acute tubular necrosis
b) Acute interstitial nephritis
c) Mesangioproliferative glomerulonephritis
d) Necrotizing granulomatous vasculitis

Examples. The following are additional MCQ examples of these "Cognitive Integration" questions at the "Understanding" level:

A lesion in which one of the following sites is most likely to result in a lower facial nerve palsy?

Which one of the following conditions is most likely to result in p wave enlargement on a patient's ECG?

Applying Questions

What is it? These questions demand that a student understands the information and can *apply it* to a *clinical context*. A typical question would present a patient scenario and ask the student to elicit factors on history that would be important, predict physical examination or test results, determine the most likely diagnosis, suggest management (drug and non-drug), or recommend a monitoring plan. The question presented at the beginning of this primer is an example of an MCQ that would fall under this category.

Example. Using the example of acute renal failure, the following would also be an example of an "Applying" question.

An 86-year old woman presents to the emergency department with a 3-day history of severe lethargy. On further questioning she also reports multiple daily episodes of non-bloody, watery diarrhea, and vomiting, for the past 7 days. She resides in a nursing home and other clients are also suffering from a similar illness. Her blood work shows a creatinine of 700 mmol/L. Her creatinine one month ago, was normal. Which one of the following additional abnormalities are you most likely to find?

a) Hyperkalemia
b) Hypercalcemia
c) Hypernatremia
d) Hypoglycemia

Comments: This question requires that the student:

1) *understand* that a viral diarrheal illness can cause dehydration, which can lead to acute renal failure, that in turn can cause hyperkalemia **AND**
2) *apply* all that knowledge to the clinical setting to recognize, diagnose, and begin to manage

©2016. Faculty of Medicine, University of Toronto. These guidelines were developed and approved by the MD Program Test Committee, December 2016.

Because the question is presenting a patient scenario and asking them about anticipated bloodwork abnormalities, it is assessing whether a student might be able to manage such a patient at the bedside.

Example. The following are additional types of lead in questions that you can use when constructing "Applying" questions:

> **W**hich **o**ne **o**f the **f**ollowing (**WOOF**)
> ... is the best next step? (used to ask about investigations or management)
> ... is the most likely diagnosis?
> ... would you ask about on history?
> ... would you expect to find on physical examination?
> ... would you counsel the patient about in order to ....?

Cognitive Integration: "Cognitive Integration" questions can also be written at the "Applying" level by asking the student to make a connection between the patient's presentation or management plan and the underlying mechanism of disease or treatment. Like cognitive integration questions at the "Understanding" level, these questions ask about the underlying science of the patient's symptoms, physical exam, investigations, or treatment. The main difference between "Cognitive Integration" questions at the "Appling" level and at the "Understanding" level, is that those at the "Applying" level will present the student with a specific clinical case and ask them to relate their knowledge to that specific patient example.

Example. An example of a MCQ that is related to the one above on acute renal failure and that tests cognitive integration at the "Applying" level could be:

> An 86–year-old woman presents to the emergency department with a 3-day history of a severe lethargy. On further questioning she also reports multiple daily episodes of non-bloody, watery diarrhea, and vomiting, for the past 7 days. She resides in a nursing home and other clients are also suffering from a similar illness. Her blood work shows a creatinine of 700 mmol/L. Which one of the following are you most likely to find if you were to biopsy her kidneys?
> a) Acute tubular necrosis
> b) Acute interstitial nephritis
> c) Mesangioproliferative glomerulonephritis
> d) Necrotizing granulomatous vasculitis

Comments. Here the student is required to *understand* that a viral diarrheal illness can cause dehydration which can lead to acute renal failure, that in turn can cause hyperkalemia, AND apply that knowledge to the clinical setting to recognize acute renal failure on the basis of dehydration AND be able to *link the underlying pathology* to that diagnosis.

©2016. Faculty of Medicine, University of Toronto. These guidelines were developed and approved by the MD Program Test Committee, December 2016.

Examples: Additional examples of "Applying" questions that also assess cognitive integration include:

> A 67-year-old man is seen in the family practice clinic for symmetrical swelling of his ankles for the last 3 weeks. His past medical history is significant for severe COPD and sleep apnea. On examination, his jugular venous pressure is 7 cm above the sternal angle and there is evidence of a C-V wave. Which one of the following abnormalities do you expect to find on his echocardiogram?
>
> A 23-year-old man is seen in the emergency room for fever and weakness for the last 5 days. He has a known history of injection drug abuse of heroin. On examination, he is febrile at 38.9C. His JVP is 8 cm above his sternal angle and he has a pan-systolic murmur at his left sternal border that increases with inspiration. On examination of the pads of his figures you note painful, erythematous papules. Which one of the following pathology results would you expect on a biopsy specimen from his hand lesions?

TEMPLATE FOR APPLICATION MCQS

All questions:

> A [*age*]-year-old [*man or woman*] presents to [*clinic, emergency room, family practice office*] with [*symptom*].

As relevant to the question:

> The patient reports [*more HPI details*]. His/her past medical history is significant for [*relevant PMH here*]. His/her medications are [*relevant meds here, doses only if needed*]. Family history is significant for [*relevant Fhx here*]. On examination [*relevant physical examination here*]. Investigations showed [*relevant investigations here*].

All questions:

> Which one of the following [question here].
>
> -one correct answer and 3 plausible distractors:

©2016. Faculty of Medicine, University of Toronto. These guidelines were developed and approved by the MD Program Test Committee, December 2016.

STEPS FOR WRITING A MCQ

Step 1: Choose the learning objectives for which you want to write an MCQ

The goal is for you to write 1-3 MCQs per learning objective. The objective should guide the *CONTENT* area and the *LEVEL OF LEARNING that should be targeted with the MCQ*. If writing an MCQ is not possible for a given objective because it does not define the content area of the level of learning well, you should omit writing an MCQ for that objective and note this in order to provide feedback to the course director for the purpose of improving the objectives.

You need to be clear about what level of learning you are targeting. You will need to decide whether you want to test the student's ability to remember, understand, or apply their knowledge. The level which you choose to target will depend on a few factors. As mentioned, *the objectives should naturally lend themselves to questions that test a particular level of learning*. For example, if the objective is "to be able to list the side effects of medication X", your question can only check whether a student remembers the information, not whether they can understand it or apply the information to a clinical context. A question could be, "Which one of the following is a potential side effect of medical X?" Alternatively, if the objective is "to be able to recognize the clinical presentation of typical angina," then one could choose to write an application question where the stem describes a patient with typical angina chest pain and the question asks, "Which one of the following is the most likely diagnosis?"

Hint. The blueprint for your assessment exercises has been set by your course. You should be familiar with this blueprint and write the correct proportion of remembering, understanding, and applying questions as guided by your blueprint. The proportion of MCQs should also reflect the time allotted to the teaching of the different topics. Resist the temptation to write many MCQs that test one very restricted piece of the teaching or that primarily test a student's ability to recall a list on a teaching slide. This will be perceived as an unfair assessment of the student's learning.

Hint. Focus each MCQ on a single piece of knowledge. Ask yourself, "What specific knowledge is it that I want to test?"

Hint. We suggest that you start your writing with questions pertaining to the objectives that you feel are the most critical for students to understand, given the stage of their training. It may be that you choose those objectives that are foundational and upon which other learning will build, or are critical (for example, they test knowledge that is absolutely required to ensure the safe care of patients). Alternatively, you may choose to focus a question on those areas of common misconceptions or that commonly pose challenge to students who do not grasp the material in order to identify areas of need. After these more critical areas are covered you may move on to write items covering other objectives.

©2016. Faculty of Medicine, University of Toronto. These guidelines were developed and approved by the MD Program Test Committee, December 2016.

Step 2: Write your stem

For MCQs that target learning at the "Remembering" or "Understanding" levels, the stem is very brief and often is simply amalgamated with the lead-in.

For MCQs that target the "Applying" level, you will need to write a clinical scenario. In the scenario always include the following information: Age, gender, location, and reason for assessment. These typically should all appear in the first line. Next include the relevant additional factors on history, physical examination, or investigations that are necessary to answer the question and *nothing more*.

Hint. Straightforward and easy to understand scenarios work best.

Hint.. Avoid the temptation to make the question "trickier" by excluding any information that is unnecessary to get to the answer or adding in extraneous information. If pieces of information are not included in the stem the presumption is that they are negative.

Step 3: Write your lead-in

This is the question that you are asking the student to answer. If possible, formulate your lead in using the WOOF template: "Which one of the following..."

Step 4: Write your correct answer

The correct answer should be undeniably correct. There may be additional correct answers but they cannot appear in the list of options. Essentially, you are asking students to choose the most correct answer from the options that they have been given.

Step 5: Write your distractors

These are the additional options the student has to choose from. They should be plausible but clearly wrong. Do not make up any distractors that don't actually exist.

Hint. To think of distractors, it is often helpful to try to think of what answers the student who does not know the material sufficiently would give. Ask yourself, "How would my weaker student answer this question?" It is ideal if the distractors are all of one category (e.g. all investigations, all management). This is not always possible, however, and if they represent different categories, ensure that there is a balance and that the lead in question does not allow for any distractors to be obviously wrong (e.g. lead asks for investigations and distractor is a management option).

Step 6: Check for pitfalls

Check to ensure that you have followed the best practices for writing effective MCQs by evaluating whether your MCQ has fallen victim to any of the pitfalls listed in Table 1 - and if so, edit it.

Hint. See Table 1 – Common Pitfalls in developing MCQs

Step 7: Edit with a partner

Every MCQ should be edited with a partner who is familiar with the content and who can look for both errors in format and also in content.

©2016. Faculty of Medicine, University of Toronto. These guidelines were developed and approved by the MD Program Test Committee, December 2016.

SUMMARY

WHAT MAKES A MCQ A GOOD MCQ

- ✓ A good MCQ assesses one clearly identified concept or learning objective.
- ✓ A stem is included if the question is at the "Applying" level.
- ✓ The stem contains all information that is necessary to answer the question and nothing more.
- ✓ The writing is clear and the task is defined and, where possible, follows the WOOF format.
- ✓ There is only one correct answer.
- ✓ The distractors are plausible but clearly wrong.
- ✓ The question is positively worded (i.e. does **not** use EXCEPT, NOT, etc.).
- ✓ The correct answer and distractors contain no overlapping options.

Table 1. COMMON PITFALLS IN DEVELOPING MCQS

	Check each MCQ for the following:	**Rationale**
✓	1. Have I avoided negative wording (such as Except for, Not true, or Least likely) in the **lead-in (question)?**	Makes the question confusing and causes the students who know the information to answer incorrectly
	2. Have I excluded all irrelevant or useless information in the **stem**?	Increases time to complete the questions without affecting accuracy
	3. Have I avoided turning the MCQ into a True False question? Or used **response options** such as None of the above, or All of the above?	Makes the question too easy for the students
	4. Have I avoided using obviously wrong **distractors**, (i.e. any response options that are highly unlikely or outside the realm of possibility)?	Allows for easy guess
	5. Have I avoided using "never" or "always" in the **response options**?	Easy to eliminate and therefore cues to the answer
	6. Have I avoided wordy **response options**?	Increases time and could be a give away as longer options are frequently correct
	7. Have I avoided making any **responses options** visibly longer or shorter than the rest?	Longer ones could be easily identified as right while short ones would be easily eliminated. Provides cuing to the students.
	8. Have I avoided using grammar **for response options** inconsistent with the **stem**? (e.g. tense, singular, plural)	Allows for process elimination and is confusing
	9. Have I avoided complex **response options** (A and B, A and C but not B)	Allows student to use process of elimination to come up with answer. Tests logic, not content knowledge.

©2016. Faculty of Medicine, University of Toronto. These guidelines were developed and approved by the MD Program Test Committee, December 2016.

References

Anderson, L. W. and Krathwohl, D. R., et al (Eds.) (2001) *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives.* Allyn & Bacon. Boston, MA (Pearson Education Group)

Haladyna, Thomas M. and Downing, S. M.. Validity of a taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education*, 2(1), 51-78, 1989.

Morrison, Susan and Free, Kathleen. Writing multiple-choice test items that promote and measure critical thinking. *Journal of Nursing Education* 40: 17-24, 2001.

Case, S. M. and D. B. Swanson. (1998). "Writing Written Test Questions for the Basic and Clinical Sciences." 3rd Edition. from http://www.nbme.org/pdf/itemwriting_2003/2003iwgwhole.pdf.

MD Program
UNIVERSITY OF TORONTO | Office of Evaluations

MD Program
UNIVERSITY OF TORONTO | Office of Faculty Development

©2016. Faculty of Medicine, University of Toronto. These guidelines were developed and approved by the MD Program Test Committee, December 2016.