Test Question Formats for M1 and M2 Quizzes and Exams
Creighton University School of Medicine

While we welcome essay-type questions (from lecturers who are willing to grade them), most of our exams consist solely of objective questions. Writing items that are discriminating, reliable, and efficient is difficult and time-consuming. Fortunately, the NBME has done a lot of research on the types of objective questions that are most effective. After fifty years of experience, they have chosen only **two types** to use on the Medical Licensing Exams. Both of these require the student to choose the **single best answer** from the options given. The two subtypes are the multiple choice questions (which they call *Type A*) and Extended Matching (*Type R*) items. These are the two types used on quizzes and exams at Creighton University School of Medicine. Examples of each are printed at the end of this document. The NBME's excellent guide to item construction, a 181 page Adobe Acrobat file, is free for download at
http://www.nbme.org/PDF/ItemWriting_2003/2003IWGwhole.pdf

Guidelines for One-Best-Answer Items
- Create a well-written stem (e.g., clinical case presentation)
- Ask a clear, focused question in the stem (avoid "Concerning topic X,")
- Include language in the stem instead of repeating it in each option
- Use parallel grammatical structure in the answer options
- Use plausible distractors – use words familiar to students
- Arrange answer choices alphabetically or numerically
- Supply measurement units for laboratory values – use standard units
- Use generic names for drugs
- Please do not use
  - negatively phrased items
  - vague (e.g., many) or definitive (e.g., always) language
  - correct answers obviously longer than the distractors
  - distractors with double options ("A & C") unless all the options are double
  - "none of the above" as a distractor
  - "of the following" in the stem of each question

These question formats are no longer used by the NBME :

- Please do NOT use true/false items. The NBME has completely stopped using true/false formats on its exams. True/false questions may seem more challenging, but attempting to reduce ambiguity leads to testing recall only.

- Please do NOT use a stem followed by options where one or more may be correct. Although these Type K questions appear to test in-depth knowledge, research has shown they are less discriminating, less reliable, and less efficient -- and sometimes they demonstrate a "cueing" effect.

- Please do NOT use negative Type A questions, e.g., "Each of the following is correct EXCEPT…"

Items should focus on important concepts and assess the *application* of knowledge – NOT recognition of a term or recall of an isolated fact.

Please turn the page to see examples of test item formats.

Examples of Effective Test Item Formats

Type A (self-contained multiple choice question)

A 26-year-old man who is HIV positive has a CD4+ T-lymphocyte count of 250/mm3 (N>500). After 5 weeks of therapy with a nucleoside polymerase inhibitor and a protease inhibitor, he feels weak and is easily fatigued. His hemoglobin concentration has decreased from 12.8 g/dL to 8.2 g/dL. Which is the most likely cause of the anemia in this patient?
(A) Decreased formation of erythrocytes
(B) Folic acid deficiency
(C) Increased formation of erythrocyte antibodies
(D) Increased fragility of erythrocytes
(E) Iron deficiency

R Type (One set of choices used for several questions).

A. Acute leukemia
B. Anemia of chronic disease
C. Congestive heart failure
D. Depression
E. Epstein-Barr virus infection
F. Folate deficiency
G. Glucose 6-phosphate dehydrogenase Deficiency
H. Hereditary spherocytosis
I. Hypothyroidism
J. Iron deficiency
K. Lyme disease
L. Microangiopathic hemolytic anemia
M. Miliary tuberculosis
N. Vitamin B12 (cyanocobalamin) deficiency

For each patient with fatigue, select the most likely diagnosis.

1. A 19-year-old woman has had fatigue, fever, and sore throat for the past week. She has a temperature of 38.3 °C (101 °F), cervical lymphadenopathy, and splenomegaly. Initial laboratory studies show a leukocyte count of 5000/mm^3 (80% lymphocytes, with many lymphocytes exhibiting atypical features). Serum aspartate aminotransferase (AST, GOT) activity is 200 U/L. Serum bilirubin concentration and serum alkaline phosphatase activity are within normal limits. **Ans: E**

2. A 15-year-old girl has a two-week history of fatigue and back pain. She has widespread bruising, pallor, and tenderness over the vertebrae and both femurs. Complete blood count shows hemoglobin concentration of 7.0 g/dL, leukocyte count of 2000/mm^3, and platelet count of 15,000/mm^3. **Ans: A**

What Not to Ask: Examples of Unacceptable Test Item Formats

Most of the types of questions that we ask that you avoid are variations on the True-False question format. What's wrong with True-False questions? In brief, it's hard to write an unequivocally true statement without making it a simple recall of an isolated fact.

1. True-False question.

The sun rises in the East.   T   F

To answer a True-False question, the student has to decide how much "truth" the instructor wants.   In this seemingly straightforward example, the student may know that the sun rises directly in the East only on the two equinoxes and otherwise rises in the northeast or southeast, except at the North Pole (rises from the south) and South Pole (rises from the north). So the statement is both true and false.

2. "None of the Above"

From what direction does the sun rise?
> A. East
> B. North
> C. South
> D. West
> E. None of the above

Without option E, it would be clear that A is the best answer. Adding "None of the Above" as an option transforms a single best answer multiple choice into a True-False question. Now E represents any better answer, including a precise answer that mentions the equinoxes, North and South Poles etc. Unless one of options (A-D) is a metaphysical certainty, "none of the above" will always be the best answer.

3. Variable number of correct answers.

From what direction does the sun rise?
> A. East
> B. North
> C. South
> D. A & B
> E. A & C
> F. A, B, & C

This format also converts a multiple choice into a True-False question. The student knows that A is a correct answer, but it's not clear whether B and C should also be considered correct (to account for the South and North Poles).

4. All of the Above.

From what direction does the sun rise?
> A. East
> B. North
> C. South
> D. All of the above.

This type of question is a variation on the "variable number of correct answers" and presents the same challenge to the student.

Single best answer questions require the student to choose the BEST answer from the alternatives given. Even if the student believes that none of the answers is perfect, he or she can still pick the best of the alternatives. On the other hand, True-False questions require the student who knows the subject well to guess the level of precision the instructor is seeking.

We also ask you to avoid types of questions that tend to end up confusing students as they try to answer them. These include:

5. Choose the WRONG answer.

This type of question usually contains the words EXCEPT or NOT in the stem:

> All of the following appeared in my lecture notes EXCEPT

> Anti-thrombin III does NOT have the following characteristic:

Picking out the wrong answer is counter-intuitive; as they move through the distractors, students often forget that they are looking for the worst answer and reflexively choose the best answer, as they are (correctly) used to doing. An extreme example of confusion occurs when a negative stem is paired with a negative statement in the choice (a double negative) or with "None of the above".

6. Unfocused Question.

> Which of the following is TRUE about the sun?
> >   A. it is 150 million kilometers from Earth
> >   B. it is brighter as it rises than as it sets.
> >   C. it rises in the East
> >   D. it rises at 6 AM

To choose the best answer, students must be able to compare the truth of one choice with each of the others. If all of the choices address the same idea, this can be done. However, when the choices are heterogeneous, it is difficult to make this comparison, especially when a student finds the statements to be partially correct but not absolutely so.

7. Grammatically Faulty Questions

Grammatical problems crop up when the question writer forgets the stem after running out of ideas for the distractors. This type of question often starts off ok, then falls apart:

> In a well written multiple choice question, the stem contains
> >   A. most of the information necessary to answer the question.
> >   B. a negative, such as NOT or EXCEPT, in bold uppercase letters.
> >   C. there should be no more than two incorrect distractors.
> >   D. students should be reminded to choose the one best answer.

Choices C and D do not grammatically complete the sentence in the stem. The student has to try to rewrite the question before answering it. (By the way, the correct answer is A).

Developed by the Office of Medical Education at the Creighton University School of Medicine