

Writing Multiple Choice Tests

Examinations that require no after-the-test judgment about correct answers and that can be machine-scored are called objective tests. The most common format for an objective test is the multiple-choice item.

A test used in a medical school course should be both **reliable** and **valid**. When you scan and score an examination in the Office of Medical Education, the Remark software will report an index of reliability called the Kuder-Richardson 20. See the companion Helpful Teaching Hints entitled *Test and Item Analysis* for an explanation of the KR-20 and internal consistency reliability.

The type of test validity of most interest to medical school faculty is **content** validity – i.e., does the test measure the objectives and what students were expected to learn? There is no statistical index for content validity. However, when course faculty collaborate in the planning of test material and in the reviewing of test items, the prospects for content validity are increased.

What types of multiple-choice test items should I use?

Extensive measurement research has demonstrated that two types of multiple-choice test items are the most psychometrically sound: the **single-best-answer multiple-choice** item with four or more options and the **extended matching** item (a variant of a multiple-choice item). Happily, the USMLE Step exams, for the most part, only use these two formats. When we write well-constructed objective test items using these two formats, we are preparing our students for their licensure examinations.

Single-best-answer multiple-choice

A 30-year-old woman being treated for hypertension has the sudden onset of fever and malaise. Temperature is 38.3 C (101 F) orally and blood pressure is normal. She has a malar rash, swelling and tenderness of the wrists and knees, and a friction rub at the left lower sternal border. Which of the following drugs is the most likely cause of these findings?

- A. Captopril
- B. Hydralazine**
- C. Minoxidil
- D. Nitroprusside
- E. Propranolol

Extended matching

- | | |
|-------------------------|-------------------|
| A. Acetaminophen | J. Nalidixic acid |
| B. Amiodarone | K. Nitrofurantoin |
| C. ACE inhibitors | L. Penicillin |
| D. Aspirin | M. Prednisone |
| E. Atenolol | N. Procainamide |
| F. Bleomycin | O. Propranolol |
| G. Cytosine arabinoside | P. Sulfasalazine |
| H. Furosemide | Q. Tetracycline |
| I. Metronidazole | R. Verapamil |

For each patient, select the drug most likely to have caused the adverse effect.

1. A 62-year-old man with chronic obstructive pulmonary disease begins therapy with an antihypertensive drug. Two weeks later, he has marked worsening of dyspnea and clearly audible wheezing.

Ans: O. Propranolol

2. A 56-year-old man with recurrent ventricular arrhythmias began taking an antiarrhythmic drug 5 months ago. He now has progressive dyspnea, cough, and a low-grade fever. Erythrocyte sedimentation rate is increased. X-ray film of the chest shows a diffuse interstitial pneumonia. Pulmonary function tests show that diffusing capacity for carbon monoxide is decreased.

Ans: B. Amiodarone

What are some principles for writing a reliable and valid multiple-choice test?

- Match the test blueprint to the course objectives and content.
- Write test items on important concepts and application of knowledge.
- Balance short-stem items with longer stems using patient vignettes.
- Construct the test before the course starts; revise the test shortly before the exam is scheduled.
- If a multi-instructor course, have a test construction group.
- Review the test and item analysis after the test has been administered (see Helpful Teaching Hints *Test and Item Analysis*).

What are the four principles for writing a sound multiple-choice test item?

- Test an important concept or application of knowledge.
- Write item stems that are relatively long and ask a question.
- Make item choices:
 1. as brief as possible.
 2. alphabetically or logically ordered.
 3. homogeneous.
- Write test items that can be answered with the choices covered.

Practice: Does the test item below satisfy these four principles?

What research design is used for both diagnostic test investigations and patient survey research?

- a. case control study
- b. case series
- c. cohort study
- d. cross sectional study
- e. randomized clinical trial

Where can I learn more?

www.usmle.org - Click on Step 1 more for content descriptions by discipline and sample test questions.

www.nbme.org - You can download a complete and excellent Item Writing Manual at www.nbme.org/about/itemwriting.asp