

# **ITEM WRITING STRATEGIES FOR MEDICAL EDUCATORS**



## **WORKBOOK**

**National Board of Osteopathic Medical Examiners  
Chicago, Il.  
Conshohocken, Pa.  
2007**

## **Welcome**

The NBOME presents this Item Writing Workshop as an educational service to the faculty of Colleges of Osteopathic Medicine nationwide. The objective of the sessions is to familiarize item writers with the techniques and approaches that reflect the content of a domain, and have produced psychometrically sound test items.

During the program you will receive information on the proper guidelines for item-writing, and then be asked to apply them to a series of test items so that the necessary skill sets that are learned are reinforced. Prior to beginning the course, you will be asked to take a pre-test. After completing the course, you should review what you believed to be the best practices in item-writing, and compare them to what you learned during the program.

It is also pointed out that the introduction of the computer-based test will permit the introduction of new and varied test item material. Familiarity with multi-media test item writing is on the cusp of a new and challenging future in test item construction. The NBOME encourages all faculty members to consider writing test items for the COMLEX-USA examination and COMAT, the subject examination, offered by the NBOME. Faculty and students both will benefit from this endeavor.

We hope you enjoy the course, and consider the opportunity to join the ranks of other professional item-writers producing quality test items for all Levels of COMLEX-USA.

## **Making the Test Writing Assignment**

- Fits the table of specifications (blueprint) to cover content
- The test items are consistent with course or rotation objectives
- Properly weighted for the subject areas
- Specifications are made clear to the item writer
- Designed with the end in mind
  - Formative
  - Summative
    - » Punitive

## **Why Are Test Items Flawed?**

- Choice of inappropriate content
- Disparity between what is tested for and what has been taught, stressed or deemed important
- Technical considerations
- Failure to heed stem and distractor construction guidelines
- Irrelevant difficulty in the test items
- Failure to consider the psychometric performance of the test items
- Failure to pre-test items or measure item unfit

## **Consider the Seven Core Competencies of the Osteopathic Profession**

1. Osteopathic Philosophy and Manipulative Medicine
2. Medical Knowledge
3. Patient Care
4. Interpersonal and Communication Skills
5. Professionalism
6. Practice-Based Learning and Improvement
7. Systems-Based Practice

What is to be measured?

What can and cannot be measured?

How is the competency to be measured?

What is the best instrument by which the competency can be tested?

### **Item Content**

- First-things-first
  - Use high priority, high frequency, high impact clinical or basic science situations
  - Begin with the end in mind
    - What knowledge do you want to test for in the question you are writing
    - The interrogatory portion of the test item is essential to good item performance

## **Ideal Item Appearance**

*(Long Stem with clear interrogatory)*

A 40 year old male presents with xxxxxx xxxxxxxxxxxx xxxxxx xxx xxxxx xxxxxx  
xxxx xxxxxx xxxx xxxxxxxxxxxx xxxx. What is the most likely cause of this  
presentation?

*(Short Distractors of approximately equal length)*

- A. aaaa aaa aaaaaaaaaa
- B. bbbb bbbbb bbbbb bbb
- C. ccc cccc ccccccccc
- D. dddd dd dddd dd
- E. eeeeeee eee eeeee

## **Less Than Ideal Item Shapes**

- Short Stems and Long Distractors
- Long Stems and Long Distractors
- Short Stems with Unequal Distractors
- Long Stems with Unequal Distractors
- Matching Items with One Word Distractors

## **Using Item Patterning**

**A** (*describe the patient with age and gender, if pertinent*) **presents with** (*give the signs, symptoms and other findings*) **to the** (*list site of presentation*) **in** (*describe the condition of the patient*). **Exposure to which of the following** (*list of five drugs, i.e. the appropriate and related distractors*) **most likely caused this presentation?** (i.e. the interrogatory)

or

**A** (*patient description*) **presents complaining of** (*somatic/visceral symptoms*). **A structural examination reveals** (*structural and palpatory findings*). **The most likely diagnosis is** (*list five diagnoses as distractors*) **or the most appropriate treatment for this presentation is** (*list five treatments as distractors*).

## **Guidelines That Make For Better Test Items**

- Simple and concise presentation
- The interrogatory portion of the item must be clear
- Responses should parallel the stem
- All distracters should be plausible
- Distracters should be of roughly equal length
- Avoid often, never, always (“most likely” is acceptable in most instances)
- Avoid double negatives
- Avoid abbreviations
- Use proper grammar in stems and distracters (do not give hints with improper tense usage)

## **Guidelines for Better Items**

- Avoid regional influences (for national exams, may be fine for local exams)
- When numerical responses are used arrange in increasing order
- Avoid ranges of values that overlap
- Do not give cues, particularly with grammar
- Do not teach in the stem
- Arrange distractors alphabetically (convention)
- Use positively written items rather than negative ones
- Capitalize only proper nouns and for grammar purposes
- Distractors should not be opposites
- Distractors should not vary greatly in relationship to the question
- Distractors should be independent
- Be sure items follow the blueprint or item assignments
- Test items should be referenced
- Item writers should attempt to require higher thought processes of test takers
- Avoid “syndrome” names where possible
- Avoid one word options when matching sets to trailers
- Have a general sense of how much time it will take the examinee to complete the item and test

# **Taxonomy Levels**

or

“Thinking about raising the level of the thought process, but not necessarily increasing the difficulty of a test item”

## Level 1 - Knowledge & Memory

- Defining
- Recalling
- Identifying

## Level 2 - Comprehension

- Determining
- Explaining
- Inferring

## Level 3 - Application

- Applying
- Choosing
- Organizing

## Level 4 - Analysis

- Comparing
- Contrasting
- Distinguishing

## Level 5 - Synthesis

- Combining
- Deriving
- Specifying

## Level 6 - Pure Reasoning

- Assessing
- Arguing
- Considering

## **Levels of Evidence**

- 1a: Systematic reviews of homogenous randomized controlled trials
- 1a: Systematic review of randomized trials with worrisome heterogeneity
- 1b: Individual randomized controlled studies
- 1c: All or none randomized controlled trials
  
- 2a: Systematic reviews with homogeneity of cohort studies
- 2b: Systematic review of cohort studies displaying worrisome heterogeneity
- 2b: Individual cohort study or low quality randomized trial
- 2c: Outcomes research; ecological studies
  
- 3a: Systematic review with homogeneity of case-controlled studies
- 3a: Systematic review of case-controlled studies with worrisome heterogeneity
- 3b: Individual case-controlled study
  
- 4: Case-series (and poor quality cohort and case-controlled studies)
  
- 5: Expert opinion without explicit critical appraisal, or based on physiology, bench research or “first principles”

# The Quick Item Review Guide

- ✓ Does it follow the specifications of the blueprint?
- ✓ Is content appropriate for the level of the candidate?
- ✓ Is the desired knowledge or skill objective being elicited?
- ✓ Is the format of the item best for asking about the desired information?
- ✓ Is the shape of the item correct?
- ✓ What level of thought process is required to answer the item?
- ✓ Has a desired item pattern (boiler plating) been employed?
- ✓ Is the item logical and grammatically correct?
- ✓ Have the construction rules for the stem (trailers) been followed?
- ✓ Have the constructions rules for the distractors (set) been followed?
- ✓ Is the interrogatory well worded and clear?
- ✓ Is there really only one correct answer?
- ✓ Does the average candidate have sufficient time to read and answer the item?
- ✓ Have you reviewed the psychometric data on the item?

## **References**

All test items used in moderate to high-stakes testing should be referenced.

References should be:

- Accepted as an authority in the domain being tested
  - Current in concept, classification, criteria, and/or requirement
  - Universally available to the population being tested
  - Use the highest level of evidence practicable
  - Quoted accurately as to title, author, edition and year of publication
- Can include reputable internet and journal article quotes

## **WRITING A NEW TEST ITEM (Assignment)**

Objective:

Descriptors:

Competency addressed:

Evidence Level:

Item type:

Psychometric value:

Taxonomy Level:

Reference:

*New Test Item:*

*Patterning*

*Cloning*

*Banking and coding*

## **TRIAL QUESTIONS**

Using the guidelines provided from the course and in this workbook, review each test item and decide whether or not the question is written in the best manner possible, and that is suitable for use in an examination. Attempt to re-write the test item, if necessary, to improve its quality. Space is provided below each question for use in this exercise.

Keep in mind that most any question can be re-written better, after through analysis.



**Test Item # 2**

A middle-aged woman is admitted from the ED to a telemetry unit after having had a prolonged episode of SVT. She has had two episodes in the past resulting in emergency treatment and discharge. This time she had Hemodynamic compromise with sustained rapid rate requiring observation and treatment. Upon admission to the unit, she is stabilized with a BP of 106/60, pulse rate of 110, ST rhythm and is alert but appears fatigued. The woman is quiet, tearful, withdrawn. Old records show she has had a GI work-up for abdominal pain which showed irritable bowel and she has a history of migraine headaches requiring occasional trips to the hospital for Demerol. Her family has gone home. In this instance, what is the most appropriate action?

- A. During the admission process, conduct the standard abuse screening process
- B. During the admission process, ask the woman if she has any idea what is causing her rapid pulse problems
- C. The woman should not be asked about domestic violence in her history at this time
- D. Return to the patient as the end of the shift and do an abuse screening

From: Behrman, *Nelson Textbook of Pediatrics*, 17th Ed., W. B. Saunders

---

---

---

---

---

---

---

---























***Note:***

***Other faults may exist in some of the test item samples used in this workbook, and there may be some errors of content. The items presented are for instructional purposes only, and should not be used in any written, oral or CBT examination.***