Test Item Writing Workshop
Creating the Most Effective and Reliable Examination Questions

Sarah Jacobs, M.Ed.
Faculty Development Specialist
OHSU Teaching and Learning Center
meltonsa@ohsu.edu

Pat Kenney-Moore EdD, PA-C
OHSU School of Medicine
Division of Physician Assistant Education
kenneymp@ohsu.edu

Education grand rounds – February 7th, 2018

Disclosure Information

• Planning committee – Constance Tucker, Ph.D., Patty Carney, Ph.D., Yi Cao, Ph.D., Lisa Hatfield, Ed.D., Judith Bowen, M.D., Sarah Jacobs, M.Ed., Sarah Drummond Hays, Ed.D., Janet Wheeler, M.Ed. have nothing to disclose.

• Speakers – Pat Kenney-Moore, Ed.D., PA-C, Sarah Jacobs, M.Ed. have nothing to disclose.

Objectives

By the end of this session, participants will be able to:
• Identify flaws to avoid in writing MCQs
• Describe how to write a clinical vignette
• Analyze poorly-written MCQs

Terms

• Multiple-choice question (MCQ)
• Objective Test = Multiple Choice Test

Bloom’s Taxonomy

Objective Test/Written assessment
Bloom’s Taxonomy

Performance observation, SIM, portfolio, case studies

Purpose of Testing / Assessment

• For the purpose of determining knowledge, comprehension and application, paper-pencil / objective tests remain the testing format of choice
• Test item construction is best undertaken by faculty who share an understanding of the curricular goals and the students’ needs

Testing / Assessment best practices

• For the purpose of determining knowledge, comprehension and application, paper-pencil / objective tests remain the testing format of choice
• Test item construction is best undertaken by faculty who share an understanding of the curricular goals and the students’ needs

What level are your assessments?

What Makes a Good Test?

• Complete the pre-test that will be passed out
• Do the best you can with picking the best answer despite the fact that the questions may not make sense!
• We will talk about testing flaws that may benefit / harm examinees
Pre-Test

Directions: For each of the following items, circle the one best answer.

1. The primary purpose of the stam is to remove the
   a) Carm
   b) Denton
   c) Menice
   d) Stam bar

2. Which of the following pairs has won the greatest number of Abby awards?
   a) Jones & Smith
   b) Smith & Taylor
   c) Smith & White
   d) White & Allen

3. How many pounds of pressure are exerted by a callam?
   a) 2.6
   b) 150
   c) 260
   d) 2600
Pre-Test

4. The stanon is aided by a
   a) Anstel
   b) Immon
   c) Octal
   d) Port

5. The stanon frequently overheats because
   a) All greetsels are bilious
   b) No immons are directly fectitious
   c) Octals are usually casable
   d) Ports are always actial

6. Stamation normally occurs when the
   a) Anstels rupture
   b) Immon falls and the denton is in place
   c) Octal rotates easily
   d) Ports pass over the caren
Bottom Line.....

• Beware “test-wise” examinees (see handout) – don’t give them clues!

Netwiseness flaws
• Grammatical Cues
• Logical Cues
• Using absolute terms
• Long correct answer
• Word repeat
• Converging elements
• Item-writing flaws
• Long/complicated stem
• Inconsistently stated numeric data
• Frequency terms (e.g. “rarely”, “often”, “usually”)
• Non-parallel language
• Trivia / fascinomas / tricky / superfluous items
• Style questions that focus on local customs
• “all of the above”
• “none of the above”
• Multiple - multiples (K-type)
• Negatively phrased questions
• Red Herrings
• Eponyms (e.g. Aschoff, Dressler, Homan, Marfan, Raynaud)
• “Which of the following is true about XYZ?”
• “all of the above”
• “none of the above”
• Multiple - multiples (K-type)
• Negatively phrased questions
• Red Herrings
• Eponyms (e.g. Aschoff, Dressler, Homan, Marfan, Raynaud)
• “Which of the following is true about XYZ?”

What to avoid (see handout)

Goal of Item Writing

Determine:
• If they know it and how difficult (p-value)
• To distinguish those who know from those who don’t
  • Discrimination
  • Comparing the top quartile to the bottom quartile
• To determine validity
  • The extent to which an exam measures what it is supposed to measure
• To identify reliability
  • The consistency of test scores from one administration to the next

Development of the Question:

• Decide whether to test the content by recall or by application. A direct question will assess recall, whereas a vignette requires the examinee to apply information they know to a clinical problem
• Direct questions should pose a clear question that can be answered without looking at the distractors. Test for this using the cover test. Unfocused questions cannot be answered without looking at the choices
• Vignettes should include enough information to paint a clinical picture that mirrors reality. Include patient age, gender, clinical setting, medical history, physical examination findings, results of diagnostic studies and initial treatment – as warranted

General Guidelines for Item Construction:

• Select a topic from the content outline and define the objective to be tested
• Determine the type of question to be asked
  • Single-Best Answer Item
  • Vignette Items
• Focus each item on an important concept, a common problem or a catastrophic event
• Emphasize the application of information, not recall of isolated facts
General Guidelines for Item Construction:

- Write a stem (test question) that poses a clear question. It should be possible to arrive at the answer without seeing the options.
- Write 3-5 plausible distractors (answers or foils) and identify the correct answer. Distractors should be about the same relative length as the correct answer.
- Include (or obtain) a complete reference with date of publication – should be within the last 3 to 5 years or so.

Sample Lead-Ins

- These findings are most likely associated with a history of:
- Which of the following is the most likely site of a lesion?
- Which of the following arteries is most likely to be involved?
- Which of the following is the most likely cause of these findings?
- Which of the following additional findings is most likely?
- Which of the following is the most likely explanation for these findings?
- Which of the following is the most likely finding on laboratory evaluation?
- The interaction between these drugs is most likely to result in
- The most likely site of additional abnormalities is
- Which of the following is the most likely outcome of (specific treatment)

Templates for Designing Vignettes

- A (patient description) has (history findings) and is taking (medications). Which of the following medications is the most likely cause of his (history, PE, or lab findings)?
- A (patient description) has (abnormal findings). Which of the following additional findings suggests a diagnosis of (disease 1) rather than (disease 2)?
- A (patient description) has a (type of injury and location). Which of the following structures is most likely affected?
- A (patient description) follows a (specific dietary regimen). Which of the following conditions is most likely to occur?

A note on hinged questions

- A set of questions with the same clinical vignette, that each ask something different.
- Be sure each question stands alone and can be answered independently of the others!

Examples of Well Written Questions

Nonvignette:

Which of the following is the most likely abnormality in a child with nephrotic syndrome and normal renal function:

A. Acute poststreptococcal glomerulonephritis
B. Hemolytic-uremic syndrome
C. Minimal change disease
D. Focal and segmental glomerulonephritis
E. Henoch-Schoenlein purpura with nephritis
Examples of Well Written Questions

Vignette
A 2-year-old boy has had swelling of the face and extremities for the past week. His blood pressure is 100/60 mmHg. Physical examination shows generalized edema and ascites. Laboratory studies show:

- Serum Creatinine: 0.4 mg/dL
- Serum Albumen: 1.4 g/dL
- Serum Cholesterol: 569 mg/dL
- Urinalysis: 4+ protein, no blood

Which of the following is the most likely diagnosis?
A. Acute poststreptococcal glomerulonephritis
B. Hemolytic-uremic syndrome
C. Minimal change disease
D. Focal and segmental glomerulonephritis
E. Henoch-Schoenlein purpura with nephritis

Your turn – Critique the following questions and identify their flaws!

Each of the following statements is true about transposition of the great vessels EXCEPT:
A. It commonly presents shortly after birth
B. It is a cyanotic heart disease
C. Cardiac enlargement is frequently present
D. The aorta lies anterior and to the right of the pulmonary artery
E. A cardiac murmur is commonly present

Your turn – Critique the following questions and identify their flaws!

At how many months of age do children normally begin to walk, talk and pick up objects?

<table>
<thead>
<tr>
<th>Pick up objects</th>
<th>Walk</th>
<th>Talk</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. 10</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>B. 10</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>C. 10</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>D. 12</td>
<td>18</td>
<td>10</td>
</tr>
</tbody>
</table>

Your turn – Critique the following questions and identify their flaws!

Which of the following is true about pseudogout?
A. It occurs frequently in women
B. It is seldom associated with acute pain or swelling in a joint
C. It may be associated with the finding of chondrocalcinosis on a radiograph of the affected joint
D. It is clearly hereditary in most cases
E. It is not usually associated with increased serum calcium concentration

Your turn – Critique the following questions and identify their flaws!

Three weeks after discharge from the hospital after acute MI, a patient develops pericarditis. Which of the following is the name of this syndrome?
A. Aschoff
B. Dressler
C. Homan
D. Marfan
E. Raynaud

Your turn – Critique the following questions and identify their flaws!

Which of the following cytokines is most likely to be associated with sepsis?
A. Interleukin-2 (IL-2) and interferon-γ
B. IL-3 and IL-7
C. IL-4 and IL-5
D. IL-10 and IL-12
E. Tumor necrosis factor-α and IL-1
Your turn – Critique the following questions and identify their flaws!

Which of the following is the best definition of the specificity of a laboratory test?

A. True positives/(true positives + false positives)
B. True positives/(true positives + false negatives)
C. True negatives/(true negatives + false positives)
D. True negatives/(true negatives + false negatives)

Resources

National Board of Medical Examiners
- Constructing Written Test Questions for the Basic and Clinical Sciences

National Board of Osteopathic Medical Examiners
- http://www.nbome.org/item-writing.asp

National Association of Boards of Pharmacy
- www.nabp.net