Quality Examination Items

Edward Barnes, MD, FACP
Assistant Dean of Longitudinal & Experiential Learning
Also in part by..........
Ensuring a Quality Exam

1. Examination Preview Process
   1. Training on quality item-writing practices
   2. Team Pre Exam Editing Process

2. Examination Review Process
   1. Post-Exam Team Item Item Analysis
   2. Answering Exam Question Challenges
Examination Preparation Process

- 3 Weeks Prior: Questions Due
- 2 Weeks Prior: Preview Meeting
- 1 Week Prior: CD Final Review
- Exam
- Afternoon Review Meeting
- 24 hr Result Release
- 3 day Challenge Period
- 2 day Faculty Response Period

Calendar Invite  Calendar Invite  Calendar Invite
Examination Preview Process
How to Construct an Effective Exam Item

1. Define the Question’s Objective

2. Craft Quality Exam Items

3. Item analysis-to improve item performance
1. Define the Question Objective

Question Objective  Lecture Objective  Course Objective

The Item Writer should ALWAYS be able to trace back the item to Lecture/Learning Objective
2. Crafting Quality Items

<table>
<thead>
<tr>
<th>ITEM STEM</th>
<th>CLINICAL SCENARIO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A 48-year-old female presents to the office with a three-week history of head, back, and neck pain. She has not been sleeping well and is tired all the time. She states that she “just aches all over,” but she cannot be more specific about the pain. She has been healthy and is on no medications. Physical examination reveals bilateral tenderness over the occiput, sternum, medial scapulae, PSIS, low back, and pes anserine bursae.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LEAD-IN</th>
<th>The most appropriate recommendation to help relieve her symptoms is</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(A) stretching and exercise*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANSWER OPTIONS</th>
<th>DISTRACTORS</th>
</tr>
</thead>
<tbody>
<tr>
<td>(B) long-acting opioid analgesics</td>
<td></td>
</tr>
<tr>
<td>(C) myofascial release to affected areas</td>
<td></td>
</tr>
<tr>
<td>(D) nonsteroidal antiinflammatory drugs</td>
<td></td>
</tr>
<tr>
<td>(E) oral corticosteroids</td>
<td></td>
</tr>
</tbody>
</table>
Item Stem

A [patient description including gender & age] presents to the [site of care] with [signs/symptoms] for [duration]. He/she states that [relevant info about condition]. [Vital signs]. Physical examination reveals [findings]. Diagnostic studies reveal [imaging, electrophysiologic or laboratory results].
Lead-In Check-List

- An ideal lead-in is focused, specific, and clear.

- A strong lead-in comes from a strong concept.

- A lead-in should be positive and ask for the single BEST, MOST APPROPRIATE or MOST LIKELY option.

- A lead-in should depend on the Stem (Avoid Pseudo-Stems/Cases)

- An examinee should be able to determine the answer, or type of answer, from the lead-in without having to read the answer options.
Effective Answer Options

- Items must have one—and only—one best answer.
- Items should have at least 5 answer options—1 answer and 4 distractors.
- Distractors shouldn’t necessarily be completely wrong—but they must be less desirable.
- Each option should be chosen by examinees.
Effective Answer Options

- Similar in length and style
- Unique in content
- Plausible and attractive to the uninformed
- Independent
- Only one single-best answer
Common Exam Question Pitfalls and No-nos

• Withholding Information
• Poor Lead-In
• All of the Above/None of the Above
• Too few distractors
• Both A & B are correct
Examination Review Process

Meeting to occur on the afternoon of the examination
How to increase Examination Validity

1. Set **specific** learning outcomes

2. Exam items **match** learning outcomes

3. Exam items are a **representative sample** of learning outcomes

4. Exam items are the **appropriate level of difficulty**
3. Item Analysis-to improve item’s performance

- Difficulty Index (p-value)

- Discrimination Index (DI) = Point Biserial correlations (PBS)
Difficulty Index (p-value)

- Calculated as the percentage of students that answered correctly.
- Range 0-100% or 0.0-1.00
- The higher the value the easier the item

Difficulty Level:
- $d > 75\% = \text{very easy}$
- $d > 70\% = \text{easy}$
- $d > 30\%-70\% = \text{mod difficult-mod easy}$
- $d \leq 30\% = \text{difficult}$
- $d \leq 25\% = \text{very difficult}$
Difficulty Index (p-value)

General Guidance:

P-Value > 90% are very easy items and might be considered for adjustment of lead-in, distractors or consider not being used on subsequent exams.

P-Value < 20% should be reviewed for possible confusing language, area for instruction adjustment, or removed from subsequent tests.
Discrimination Index= Point Biserial correlation (PBS)

• Ability of item to distinguish between high & low scorers (upper & lower 27%)

• The range is from 0.0-1.00

• Higher the value the more discriminating
  – Top 27% got the item correct
  – Bottom 27% got the item incorrect

• Negative or items near 0.0 should be removed; the bottom 27% performed better
Discrimination Index/Point Biserial Guidance

• 0.40 or higher = very good discrimination

• 0.30-0.39 = reasonably good discrimination

• 0.20-.29 = marginal/acceptable discrimination

• 0.00-0.19 = poor discrimination (increased scrutiny)

• Negative = low performing students selected correct answer more often than high scorers (recommend rejection)
General Guidance

1. If a Question’s Difficulty Index is <0.50 look at the question’s Point Biserial
2. If the Point Biserial is >0.30 then likely good discrimination
3. If the Point Biserial is <0.30 then look at the question and discuss if there were any issues with the question
4. Decide as a team what to do with the question