



# Writing Exam Items

August 6, 2013

Annual Meeting of the American Association  
of Physicists in Medicine  
Indianapolis, Indiana

Donna Breckenridge, MA  
Director of Communications and Editorial Services  
American Board of Radiology



# Writing Exam Items

Nothing to disclose,  
but I do have a disclaimer.

**ABR**  
THE  
AMERICAN  
BOARD OF RADIOLOGY  
EXCELLENCE • PROFESSIONALISM • PUBLIC TRUST • EST. 1948

## Writing Exam Items Can Be Challenging!



## Main Topics

- ✓ **Item and test validity**
- ✓ What is an exam item?
- ✓ What makes a good stem?
- ✓ What makes a good key?
- ✓ What makes a good distractor?
- ✓ Summary



## Item and Test Validity

- Effective item writing is important to ensure that the materials used to evaluate candidates are valid assessments of their knowledge.
- Downing et al. (1995) evaluated the validity of a classroom achievement test in medical education that contained flawed test items (e.g., unfocused item stems, use of *none of the above* and *all of the above*, and negatively worded stems).
- He found that flawed items failed nearly 25% more students than non-flawed items.



## Main Topics

- ✓ Item and Test Validity
- ✓ **What is an exam item?**
- ✓ What makes a good stem?
- ✓ What makes a good key?
- ✓ What makes a good distractor?
- ✓ Summary



## What is an exam item?

- **ITEM** = Test question
- **STEM** = Background and situational information, followed by request for answer.
- **OPTIONS** = The **KEY** (correct answer) and 3 or 4 **DISTRACTORS** (incorrect answers).



# What is an exam item?

All items should be:

- **FOCUSED** – Each item should focus on one concept.
- **CLEAR** – Use language that is straightforward, easy to interpret, and unambiguous, and that will **not** confuse or trick the candidate.
- **CONCISE** – Include only the information that is necessary to answer the question.



## Main Topics

- A. Item and Test Validity
- B. What is an exam item?
- C. What makes a good stem?**
- D. What makes a good key?
- E. What makes a good distractor?
- F. Summary



## What makes a good stem?

1. Linear delivery of information
2. A complete, clear question
3. Focused on a single concept
4. Positively worded format
5. Relevant to practice
6. Noncontroversial



### WHAT MAKES A GOOD STEM?

#### 1. Linear Delivery of Information

Background + situational info + request for answer

A 32-year-old woman is treated for . . .

+ One year later, she presents with. . . .

+ To diagnose her condition, what examination is most appropriate?



## WHAT MAKES A GOOD STEM?

### 1. Linear Delivery of Information

Background + situational info + request for answer

#### **Nonlinear:**

What is the probability that a child will have a defect due to induced genetic damage to the ovum, if the mother gets pregnant 30 days after receiving an ovarian dose of 110 mGy from two CT scans of the pelvis?

#### **Linear:**

A woman has two CT scans of the pelvis and receives an ovarian dose of 110 mGy. Thirty days later, she becomes pregnant. What is the probability that the child will have a defect due to induced genetic damage to the ovum?

## WHAT MAKES A GOOD STEM?

### 1. Linear Delivery of Information

#### **Better example (without extensive background)**

Situational info (includes image, if any) + request for answer:

A lamp is placed 3 m in front of a concave mirror with a 1-m focal length. How far in front of the mirror will reflected light converge?

### WHAT MAKES A GOOD STEM?

## 2. A Complete, Clear Question

### Incomplete Questions

1. Kerma:
2. The dose-rate constant:

Do you have any idea what these questions are asking?  
They don't pass the "cover test."



### WHAT MAKES A GOOD STEM?

## 2. A Complete, Clear Question

### Better Questions:

1. What does kerma measure?
2. The dose-rate constant indicates the rate at which energy is absorbed by what?

You know exactly what these questions are asking.  
These questions pass the "cover test."





### WHAT MAKES A GOOD STEM?

## 2. A Complete, Clear Question

### Confusing Question:

Exposure values are measured with corresponding added aluminum filters from a mammography x-ray tube with a molybdenum target and filter. The values are as follows: 0.0 mm Al: 671.1 mR, 0.2 mm Al: 458.8 mR, 0.3 mm Al: 385.7 mR, 0.4 mm Al: 327.5 mR, and 0.0 mm Al: 673.9 mR. What is the calculated half-value?



### WHAT MAKES A GOOD STEM?

## 2. A Complete, Clear Question

### Better Question:

Exposure values are measured (see below) with corresponding added aluminum filters from a mammography x-ray tube with a molybdenum target and filter. What is the calculated half-value?

mm Al	mR
0.0	671.1
0.2	458.8
0.3	385.7
0.4	327.5
0.0	673.9



### WHAT MAKES A GOOD STEM?

## 3. Focused on a Single Concept

### Unfocused example:

Which of the following statements about electron capture radioactive decay is true?

- A. It occurs when an atomic electron ventures inside the nuclear volume, is captured by a proton, and triggers a proton-neutron transformation.\*
- B. It occurs when an atomic electron ventures inside the nuclear volume, is captured by a neutron, and triggers a neutron-proton and  $\beta^-$  transformation.
- C. It occurs in a neutron-rich unstable parent nucleus.
- D. It occurs when the "L" shell electron has the greatest probability of venturing into the nuclear volume because of its binding energy.

### WHAT MAKES A GOOD STEM?

## 3. Focused on a Single Concept

### Better example:

In electron capture radioactive decay, an atomic electron ventures inside the nuclear volume, is captured by a proton, and triggers what type of transformation?

- A. Proton-neutron\*
- B. Neutron-proton
- C.  $\beta^-$
- D.  $\beta^+$

## WHAT MAKES A GOOD STEM?

### 4. Positively Worded Format

- Items should be structured to ask for the **correct** answer and not a wrong answer. **WHY?**
- Negatively worded items, if used, should be limited to situations that require a negative emphasis, e.g., practices to be avoided.



## WHAT MAKES A GOOD STEM?

### 4. Positively Worded Format

#### Negatively worded example:

According to the NRC, for each occupationally overexposed individual involved in a reportable event, which of the following does **NOT** have to be reported?

- A. Name
- B. Social security number
- C. Date of birth
- D. Occupation\*



### WHAT MAKES A GOOD STEM?

## 4. Positively Worded Format

### Negatively worded AND unfocused example:

Which of the following statements about the design of shielded facilities for x-ray sources used in radiation therapy is **FALSE**?

- A. For units operating above 500 kVp, the workload (W) is described as a weekly dose (Gy/week) delivered at 10 m from the source.\*
- B. Room surfaces that are **not** exposed to the primary beam are called “secondary barriers.”
- C. The use factor (U) is the fraction of the operating time during which the primary radiation beam is directed toward a particular protective barrier.
- D. The rule that radiation intensity decreases with  $(1/\text{distance})^2$  applies exactly for only a point radiation source under conditions of good (narrow beam) geometry.

### WHAT MAKES A GOOD STEM?

## 4. Positively Worded Format

### Better example:

In the design of shielded facilities for x-ray sources used in radiation therapy, the workload (W) for units operating above 500 kVp is described as a weekly dose (Gy/week) delivered at what distance from the source?

- A. 0.5 m
- B. 1.0 m\*
- C. 5.0 m
- D. 10.0 m

## WHAT MAKES A GOOD STEM?

### 4. Positively Worded Format

A double negative is the most confusing of all!



**ABR**  
THE AMERICAN BOARD OF RADIOLOGY  
EXCELLENCE • PROFESSIONALISM • PUBLIC TRUST • EDUCATION

## WHAT MAKES A GOOD STEM?

### 4. Positively Worded Format

#### Double Negative Item

A radiologic dispersal device ("dirty bomb") explodes in a crowded public area. For affected individuals, the NRC recommends all of the following **EXCEPT**:

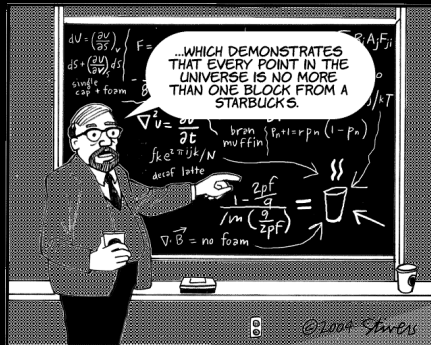
- A. **Not** staying in the immediate area.
- B. Turning on local radio or TV channels for advisories from emergency response and health authorities.
- C. Taking off the clothes they are wearing.
- D. **Not** taking a shower to wash off dust and dirt.\*

**ABR**  
THE AMERICAN BOARD OF RADIOLOGY  
EXCELLENCE • PROFESSIONALISM • PUBLIC TRUST • EDUCATION

## WHAT MAKES A GOOD STEM?

### 5. Relevant to Practice

- Even a well-constructed item is of little value if it does not relate to use in practice.



## WHAT MAKES A GOOD STEM?

### 6. Noncontroversial

Avoid topics that are:

- Trendy, but not yet proven
- Not yet accepted in the mainstream
- Dependent on or specific to the work or studies of a particular person/group/institution

## Main Topics

- ✓ Item and Test Validity
- ✓ What is an exam item?
- ✓ What makes a good stem?
- ✓ **What makes a good key?**
- ✓ What makes a good distractor?
- ✓ Summary



## What makes a good key?

1. Clearly the right answer
2. Properly completes the stem
3. Similar to distractors in length and structure
4. Not controversial or dependent on regional or institutional practice patterns
5. Supported by research



## Main Topics

- Item and Test Validity
- What is an exam item?
- What makes a good stem?
- What makes a good key?
- **What makes a good distractor?**
- Summary



## What makes a good distractor?

1. Focuses on same concept as the key
2. Properly completes stem
3. Similar to key in length and structure
4. Plausible to some degree
5. Contains no tricky language or clues





## WHAT MAKES A GOOD DISTRACTOR?

### 1. Focuses on same concept as the key

- All options—including distractors—should be conceptually related to each other.
- Stem, key, and distractors need to test knowledge of a specific area.



## WHAT MAKES A GOOD DISTRACTOR?

### 1. Focuses on same concept as the key

#### Example: Mixed options

Which of the following statements about receiving a package containing radioactive materials is true?

- A. It does **not** have to be monitored for radioactive contamination. **(requirement)**
- B. If received after normal working hours, the package must be monitored no later than 3 hours from the beginning of the next working day.\* **(requirement)**
- C. An external survey is conducted using an ionization chamber. **(technique)**
- D. A wipe test for removable contamination is carried out by smearing a 100-cm<sup>2</sup> area of the package surface using absorbent paper and counting the swabs in a Csl well counter. **(technique)**

## WHAT MAKES A GOOD DISTRACTOR?

### 1. Focuses on same concept as the key

#### Better example:

If received after normal working hours, a package containing radioactive materials must be monitored how soon after the beginning of the next work day?

- A. Within 1 hour
- B. Within 3 hours\*
- C. Within 6 hours
- D. Within 8 hours



## WHAT MAKES A GOOD DISTRACTOR?

### 1. Focuses on same concept as the key

#### Example: Mixed options

The risk of cancer from irradiation of \_\_\_\_\_ is 1 death per 10,000 individuals exposed.

- A. 0.01 Sv\*
- B. 1.0 mSv
- C. 5.0 rem
- D. 10.0 rem
- E. 50.0 rem



### WHAT MAKES A GOOD DISTRACTOR?

## 1. Focuses on same concept as the key

Better example:

What radiation dose causes a cancer risk of 1 death per 10,000 individuals exposed?

- A. 0.001 Sv
- B. 0.01 Sv\*
- C. 0.05 Sv
- D. 0.1 Sv
- E. 0.5 Sv



### WHAT MAKES A GOOD DISTRACTOR?

## 2. Properly Completes the Stem

Example:

A single, whole-body dose of 1 Gy of low LET radiation causes:

- A. sterility in women for about 3 to 4 months.
- B. increases cataracts.
- C. erythema in about 10 days.\*
- D. diarrhea in > 50% of exposed individuals.
- E. decreases the lymphocyte count.



### WHAT MAKES A GOOD DISTRACTOR?

## 2. Properly Completes the Stem

### Better example:

A single, whole-body dose of 1 Gy of low LET radiation will cause what adverse effect after about 10 days?

- A. Sterility
- B. Cataracts
- C. Erythema\*
- D. Diarrhea
- E. Decreased lymphocyte count



### WHAT MAKES A GOOD DISTRACTOR?

## 3. Similar in Length to the Key

### Example:

In television viewing of fluoroscopic images, what ensures that the electron scanning beams in the camera and the monitor are exactly synchronized?

- A. Transmission channel
- B. Camera control unit\*
- C. Image storage device
- D. Digital detectors such as charge-coupled devices that provide a direct readout of electrical charge



## WHAT MAKES A GOOD DISTRACTOR?

### 4. Plausible to Some Degree

- To the unprepared or underprepared examinee, there should be enough feasibility in the option for it to be considered.
- Ideally, distractors should be written to represent the compelling, rational, logical *wrong* answers that examinees would come up with if no choices were provided.



## WHAT MAKES A GOOD DISTRACTOR?

### 5. No Tricky Language or Clues

**Tricky language may cause candidates to select incorrect answers.**

- Negative words (such as *not*, *never*)
- Multiple parts to one option (e.g., elements 1 and 2, or 2 and 3, or 3 and 4)
- Jargon, slang, older terminology
- Abbreviations, acronyms
- Ambiguity



### WHAT MAKES A GOOD DISTRACTOR?

## 5. No Tricky Language or Clues

### Example: Multiple and overlapping parts

For ultrasound harmonic abdominal imaging, which of the following combinations of transmit/receive center frequency values is the most appropriate?

- A. 6 MHz/6 MHz
- B. 4 MHz/4 MHz
- C. 2 MHz/4 MHz\*
- D. 2 MHz/3 MHz



### WHAT MAKES A GOOD DISTRACTOR?

## 5. No Tricky Language or Clues

### Better example:

For ultrasound harmonic abdominal imaging, with a transmit center frequency of 2 MHz, what is the most appropriate receive center frequency?

- A. 2 MHz
- B. 4 MHz\*
- C. 6 MHz
- D. 8 MHz



### WHAT MAKES A GOOD DISTRACTOR?

## 5. No Tricky Language or Clues

### Example: Multiple and *overlapping* parts

Three acute radiation syndromes are associated with whole-body exposures in a short period of time with an associated dose level. What grouping with dose level is correct for cerebral vascular (CV) syndrome, hematopoietic (H) syndrome, and gastrointestinal (GI) syndrome?

- A. CV between 0.7 and 10 Gy, H > 50 Gy, and GI between 10 and 100 Gy
- B. CV > 50 Gy, H between 0.7 and 10 Gy, and GI between 10 and 100 Gy\*
- C. CV between 0.7 and 100 Gy, H > 100 Gy, and GI between 0.7 and 3 Gy
- D. CV > 50 Gy, H between 10 and 100 Gy, and GI between 0.7 and 3 Gy

### WHAT MAKES A GOOD DISTRACTOR?

## 5. No Tricky Language or Clues

### Better example:

In cerebral vascular syndrome, what dose level is associated with a whole-body exposure in a short period of time?

- A. < 10 Gy
- B. 12 to 20 Gy
- C. 25 to 50 Gy
- D. > 50 Gy\*

**WHAT MAKES A GOOD DISTRACTOR?**

**5. No Tricky Language or Clues**

**Example: Ambiguous options**

This disease typically affects what age group?

- A. Young adults
- B. Centenarians
- C. Infants
- D. Mature adults
- E. Generation Z



**WHAT MAKES A GOOD DISTRACTOR?**

**5. No Tricky Language or Clues**

**Better example:**

This disease typically affects what age group?

- A. 0 to 5 years
- B. 10 to 15 years
- C. 20 to 25 years
- D. 40 to 50 years
- E.  $\geq 80$  years





## WHAT MAKES A GOOD DISTRACTOR?

### 5. No tricky language or clues

What's wrong with acronyms?



**ABR**  
THE AMERICAN  
BOARD OF RADIOLOGY  
EXCELLENCE • PROFESSIONALISM • PUBLIC TRUST • EST. 1954

## WHAT MAKES A GOOD DISTRACTOR?

### 5. No Tricky Language or *Clues*

Clues may help unknowledgeable but test-wise candidates to select the correct response.

- Vague terms (*might, may, could, can, should*)
- Absolute terms (*always, never*)
- *All of the above/None of the above* options
- Correct response that is grammatically different, longer, or more specific than the other options
- Pairs

**ABR**  
THE AMERICAN  
BOARD OF RADIOLOGY  
EXCELLENCE • PROFESSIONALISM • PUBLIC TRUST • EST. 1954

WHAT MAKES A GOOD DISTRACTOR?

**5. No Tricky Language or Clues**

**Example: Pairs**

Cyclotrons accelerate ions by using which of the following?

- A. Standing waveguide
- B. Low frequency voltage
- C. Pulsed magnetic fields
- D. Pulsed electric fields\*



WHAT MAKES A GOOD DISTRACTOR?

**5. No Tricky Language or Clues**

**Example: Matching Pairs**

Cyclotrons accelerate ions by using which of the following?

- A. Standing waveguide
- B. Traveling waveguide
- C. Pulsed magnetic fields
- D. Pulsed electric fields\*



## Main Topics

- ✓ Item and Test Validity
- ✓ What is an exam item?
- ✓ What makes a good stem?
- ✓ What makes a good key?
- ✓ What makes a good distractor?
- ✓ **Summary**



## Summary

- Keep it focused, clear, and concise.
- Make it linear (follows time).
- Use complete sentences in the stem.
- Ask for the *correct* answer.
- Use clear and accurate phrasing.
- Spell out all abbreviations.



## Writer's Checklist

### I did:

- ❑ Spell out acronyms/abbreviations, unless they are very commonly used in the field.
- ❑ Use accurate phrasing and avoid jargon so that test-takers, including those for whom English is a second language, can completely and accurately understand the question.
- ❑ Ask for the correct, not the “wrong,” answer .
- ❑ Use complete sentence(s) in the stem and ask a direct question.



## Writer's Checklist

### I did:

- ❑ Include complete information and only relevant details.
- ❑ Cover the options to see if I can answer the question without looking at the option list (the cover test).
- ❑ Include 3 to 5 options.
- ❑ Check that options are presented in similar format (e.g., same verb tense, same structure), are independent and not overlapping, and are given in a meaningful order.
- ❑ Keep options grammatically consistent with the stem.



## Writer's Checklist

### I did not use:

- ☐ Negatively worded stems (e.g., *Which of the following is NOT, All of the following are true EXCEPT, Which of the following is FALSE*).
- ☐ *All of the above, None of the above*, or multiple answers (e.g., *Both A and C* or *A, D, and E*) as options.
- ☐ Absolute terms such as *always, never, all, or none*.
- ☐ Imprecise terms such as *seldom, rarely, occasionally, sometimes, few, or many*.
- ☐ Clues such as *may, could, or can*.



The Item Writing Manual of the  
National Board of Medical Examiners is available  
at [nbme.org/publications](http://nbme.org/publications)

## Questions?

