Test Development and Item Writing Guidelines

Roni Reiter-Palmon, Ph.D.

University of Nebraska at Omaha
Understanding the Big Picture

- Content in training must be well-defined and carefully delineated or item writing efforts will be jeopardized

- Content in training must be developed before item writing efforts begin
Understanding the Big Picture

• Training objectives should be developed for each module
  ▫ Training objectives will facilitate module development
  ▫ Training objectives will help guide test development
Understanding the Big Picture

• Content Definition
  ▫ What knowledge should an examinee know after completing the training?
  ▫ What application of the knowledge should an examinee know after completing training?
  ▫ What content from the training is to be tested?
    • The answer to the first two questions guides the answer to this question
Understanding the Big Picture

• Certification and assessment-based certificate examinations have content-defining requirements
  ▫ Examinees who pass the test should demonstrate mastery of the material
  ▫ Test items should adequately reflect the intended domain of content
    • Focus on major aspects of the content (and not minor ones)
      ▫ If the information is not important enough to make the body of the training module, it is not important enough to test on
Understanding the Big Picture

• The number of items should reflect the length of the module, section, and importance of the content
  ▫ Do not have a lot of items on a minor aspect of the content
  ▫ Do not have relatively few items for a major topic
Important Definitions
Question Stem

- The introductory statement that asks a question or poses a problem

- Example:
  - The color of the sky is
    - a) orange
    - b) blue
    - c) pink
    - d) white
Correct Answer

• The single-best correct response to the question

• Example:
  • The color of the sky is
    • a) orange
    • b) blue
    • c) pink
    • d) white
Distractors

• Response options that are either clearly incorrect or not optimal

• Example:
  • The color of the sky is
    • a) orange
    • b) blue
    • c) pink
    • d) white
Types of Multiple Choice Questions

• Basic Knowledge
  ▫ Questions about factual information that require the knowledge of a piece of information

  ▫ Example:
    • What most frequently causes a tennis player to miss the ball completely?
      • a) swinging too early
      • b) swinging too late
      • c) not watching the ball
      • d) gripping the racket incorrectly
Types of Multiple Choice Questions

• Application of Knowledge
  ▫ Using basic knowledge in a specific context

  ▫ Example:
    • Sarah is a split-brain patient who is asked to pack items into a bag with her left hand on one occasion and her right hand on another occasion. She would perform better with which hand?
      • a) right hand
      • b) left hand
      • c) her dominant hand
Types of Multiple Choice Questions

• Analysis, Synthesis, and Evaluation
  ▫ Question based on a hypothetical situation and asks examinee to use knowledge in order to make judgments

  ▫ Example:
    • George is a patient with hemophilia. He is complaining of joint pain and stiffness. As a nurse, what would you recommend to George?
      • a) drink more water
      • b) get exercise
      • c) take daily vitamins
      • d) stay in bed
General Recommendations for Item Writing
General Recommendations

• Recommendation #1
  ▫ Base the questions on matter of facts or consensus expert judgment
  ▫ Do not develop items based on matter of opinion
General Recommendations

- **Bad Example:**
  - Who was the best quarterback in the NFL in 2007?
    - a) Tom Brady
    - b) Drew Brees
    - c) Peyton Manning
    - d) Carson Palmer

- **Better Example:**
  - According to Sports Illustrated, who was the best quarterback in the NFL in 2007?
    - a) Tom Brady
    - b) Drew Brees
    - c) Peyton Manning
    - d) Carson Palmer
General Recommendations

• Recommendation #2

  ▫ Use the single best answer format

  ▫ It will be difficult to develop correct-answer questions
General Recommendations

- **Example:**
  - Which is the most effective painkiller for combating a migraine headache?
    - a) Tylenol
    - b) Motrin
    - c) Advil
    - d) **Imitrex**

- All could help to combat a migraine headache, but the single best answer is Imitrex
General Recommendations

• Recommendation #3
  ▫ Avoid cueing one item with another item
    • All items should be independent of one another
  ▫ Don’t let one item be a clue to answering another item
General Recommendations

• Recommendation #4

  ▫ Don’t test trivial knowledge

  ▫ Test more concrete information instead of abstract information

  ▫ Test what is expected from the nurse to know without consultation with others or resources
General Recommendations

• Recommendation #5
  ▫ Write questions that tap knowledge beyond the level of simple recall
  ▪ Test knowledge and application of knowledge
  ▪ Do not use exact wording from the training to avoid testing recall of familiar words and phrases
General Recommendations

• Recommendation #6
  ▫ Base the questions on important aspects of the training objectives

• Recommendation #7
  ▫ Don’t write intentionally tricky questions
General Recommendations

• Recommendation #8
  ▫ Include as much information in the *stem* and as little in the *response options* as possible

• Recommendation #9
  ▫ Pay attention to spelling and grammar

• Recommendation #10
  ▫ Be concise and be clear
Recommendations for Developing Item Stems
Developing Stems

• Recommendation #1
  ▫ Stem should present a clearly defined problem
  ▫ Examinee should know exactly what is being asked
Developing Stems

- **Bad Example:**
  - Nebraska
    - a) is considered a desert
    - b) has an eagle on the state flag
    - c) is on the East Coast
    - d) has a capital city of Lincoln

- **Better Example:**
  - What is the capital city of Nebraska?
    - a) Grand Island
    - b) Omaha
    - c) Fremont
    - d) Lincoln
Developing Stems

• Recommendation #2

  ▫ Stem should relate to a single concept or problem, rather than several concepts or problems

  ▫ Stem should focus on one clear problem
Developing Stems

- **Bad Example:**
  - What are the similarities and differences between lions and tigers?

- **Better Example:**
  - Which of the following is a characteristic of a lion?
Developing Stems

• Recommendation #3
  ▫ Avoid unnecessary detail or excess verbiage in the stem
  ▫ Strive for conciseness and clarity
  ▫ Minimize the amount of reading in each item
  ▫ Don’t use the stem to:
    • “teach” information
    • Try to confuse or obscure the point of the question
Developing Stems

- **Bad Example:**
  - Hemophilia is almost always found in males. What is the reason behind this?

- **Better Example:**
  - Why is hemophilia most common in males?
Developing Stems

• Recommendation #4

- State the problem in a positively-oriented format; avoid negatively worded items

- Positively-framed stems require the test taker to identify responses that reflect optimal, preferred, or sanctioned policies and procedures

- Avoid the use of “not” and “except” in the stems
Developing Stems

- **Bad Example:**
  - Which of the following is not true about Martin Luther King, Jr.?
    - a) he was a pastor of a Baptist church
    - b) he died of a stroke
    - c) he was a civil rights activist
    - d) he won the noble peace prize

- **Better Example:**
  - In 1968, Martin Luther King, Jr. was assassinated by
    - a) Ralph Abernathy
    - b) Taylor Branch
    - c) James Earl Ray
    - d) Percy Foreman
Developing Stems

• Recommendation #5

  ▫ When using the incomplete sentence format, always put the blank at the end the stem
Developing Stems

- **Bad Example:**
  - A man-made hormone, known as ______, can be used to treat hemophilia.

- **Better Example:**
  - Hemophilia can be treated with a man-made hormone known as ________.
Developing Stems

• Recommendation #6
  ▫ Rather than repeating the same information at the beginning of each option, include the information in the stem
Developing Stems

- **Bad Example:**
  - Hemophilia can cause
    - a) bleeding in the joints
    - b) bleeding in the stomach
    - c) bleeding in the brain

- **Better Example:**
  - Hemophilia can cause bleeding in the
    - a) joints
    - b) stomach
    - c) brain
Developing Stems

- Recommendation #7
  - Don’t provide a clue to the answer through the information in the stem
Developing Stems

- **Bad Example:**
  - Hemophilia is a disorder that interferes with the clotting of
    - a) blood

- **Better Example:**
  - Hemophilia is a disorder that interferes with
    - a) clotting of blood
Developing Stems

• Recommendation #8

  ▫ Include the central idea and most of the phrasing in the stem and not in the options
Developing Stems

- **Bad Example:**
  - What is a yankaeur?
  - a) an oral suctioning tool with a large opening designed to allow effective suction without aspiration
  - b) a medical instrument that is used to obtain a view of the vocal cords and the glottis, which is the space between the cords

- **Better Example:**
  - An oral suctioning tool designed to allow effective suction without aspiration is known as a(n)
  - a) yankaeur
  - b) laryngoscope
Recommendations for Developing Item Options and Distractors
Developing Options and Distractors

• Recommendation #1
  ▫ Include distractors that can identify “red flag” performance
    • e.g., include options that may be appealing to lower performers that would cause harm if done on the job
Developing Options and Distractors

- **Example:**
  - An overweight patient arrives in the office and asks for advice on losing weight. What advice would you give the patient?
    - a) increase exercise
    - b) eliminate all fat intake

- An incompetent examinee may select option *b* because it seems plausible to someone with less knowledge
Developing Options and Distractors

• Recommendation #2
  ▫ All options must refer to the same class of objects, events, procedures, etc.
  ▫ Try to draw the distractors from the larger class of things to which the right answer belongs to
  ▫ All options must be plausible
    • Blatantly absurd, ridiculous options should be avoided
Developing Options and Distractors

- **Bad Example:**
  - Which branch of government is responsible for interpreting if a law is constitutional?
    - a) judicial branch
    - b) executive branch
    - c) olive branch

- **Better Example:**
  - Which branch of government is responsible for interpreting if a law is constitutional?
    - a) judicial branch
    - b) executive branch
    - c) legislative branch
Developing Options and Distractors

• Recommendation #3
  ▫ Provide statements that are true but do not answer the question
Developing Options and Distractors

Example:
- Which of the following foods has the highest fiber?
  - a) black beans
  - b) apples
  - c) whole-grain pasta
  - d) peanuts

All options have fiber, but black beans have the highest fiber
Developing Options and Distractors

• Recommendation #4
  ▫ Cite common judgment errors, misconceptions, or stereotypes that the incompetent examinee may exhibit
Developing Options and Distractors

- **Example:**
  - Which of the following is a cause of kidney stones?
    - a) high blood pressure
    - b) low protein diet
    - c) drinking too much milk

- A common misperception of the cause of kidney stones is drinking too much milk
Developing Options and Distractors

• Recommendation #5
  ▫ Use an important-sounding or technical phrase that would sound plausible to an incompetent respondent
Developing Options and Distractors

- **Example:**
  - A rare, inherited bleeding disorder in which blood does not clot normally is known as
    - a) ankylosing spondylitis
    - b) anorexia
    - c) hemophilia
Developing Options and Distractors

• Recommendation #6

  ▫ All options must be logically parallel or grammatically consistent with the stem
Developing Options and Distractors

- **Bad Example:**
  - The best way to increase test performance is to
    - a) removing poorly written items on the test
    - b) provide a study guide to students
    - c) making the test longer and worth more points

- **Better Example:**
  - The best way to increase test performance is to
    - a) remove poorly written items on the test
    - b) provide a study guide to students
    - c) make the test longer and worth more points
Developing Options and Distractors

• Recommendation #7

  ▫ All options should be homogenous in content and grammatical structure
Developing Options and Distractors

- **Bad Example:**
  - People born with hemophilia have little to none of a
    - a) protein needed for clotting
    - b) antibiotic used to fight diseases

- **Better Example:**
  - People born with hemophilia have little to none of a(n)
    - a) protein needed for clotting
    - b) antibiotic used to fight diseases
Developing Options and Distractors

• Recommendation #8

  ▫ Avoid the use of “specific determiners” such as:
    • Always
    • Never
    • Completely
    • Absolutely
Developing Options and Distractors

- **Bad Example:**
  - Patients with hemophilia should
    - a) always exercise to reduce the symptoms
    - b) never exercise to reduce the symptoms

- **Better Example:**
  - Patients with hemophilia can reduce the symptoms by
    - a) exercising
Developing Options and Distractors

• Recommendation #9
  ▫ *Within each item, keep option lengths as similar as possible*
  ▫ *Keep option as short and concise as possible*
  ▫ *Longer options stand out as unintended clues to right answers*
Developing Options and Distractors

- **Bad Example:**
  - What does research tell us about assigning homework?
    - a) parents should not help with homework
    - b) students do not like homework
    - c) if the homework is relevant to learning outcomes, it is an effective teaching strategy

- **Better Example:**
  - What does research tell us about assigning homework?
    - a) parents should not help with homework
    - b) students do not like homework
    - c) it should be relevant to learning outcomes
Developing Options and Distractors

• Recommendation #10

  ▫ Arrange options in the simplest possible order such as:
    • Logically
    • Chronologically
    • Alphabetically
Developing Options and Distractors

- **Bad Example:**
  - If you buy a $20 shirt that is 50 percent off, how much will the shirt cost?
    - a) $12.50
    - b) $15.75
    - c) $10.00

- **Better Example:**
  - If you buy a $20 shirt that is 50 percent off, how much will the shirt cost?
    - a) $10.00
    - b) $12.50
    - c) $15.75
Developing Options and Distractors

• Recommendation #11
  ▫ Avoid options which overlap
    • Options should be mutually exclusive
  ▫ Options should be distinct and independent
Developing Options and Distractors

- **Bad Example:**
  - Children in what age range are most susceptible to long-term adverse reactions to abuse?
    - a) birth to age 5
    - b) age 3 to age 7
    - c) age 5 to age 10

- **Better Example:**
  - Children in what age range are most susceptible to long-term adverse reactions to abuse?
    - a) birth to age 5
    - b) age 6 to age 10
    - c) age 11 to age 16
Developing Options and Distractors

- **Recommendation #12**
  - Avoid the use of the options *all of the above* and *none of the above*
  - Avoid the use of complex options such as:
    - A and B
    - A but not C
    - C and D, but not A
Developing Options and Distractors

• Recommendation #13
  ▫ The number of options should vary between 3 and 5
    • Aim for 4 options
  ▫ More options are better if they are quality options and contribute to the ability to discriminate between high ability and low ability test takers
  ▫ Develop as many effective options as you can
Developing Options and Distractors

- **Recommendation #14**
  - Provide one answer that is unquestionably, but not obviously correct
  - Distribute the correct answer randomly among the option positions
    - Correct answer should appear an equal number of times in each answer position throughout the entire test
Developing Options and Distractors

• Recommendation #15
  ▫ If the distractors are too difficult to write, consider re-writing the stem
Getting Started
Helpful Ideas for Developing Items

• What is the main idea of _______?

• What is an example of _______?

• What is the cause/effect of _______?

• How does _______ function or operate?
Helpful Ideas for Developing Items

• What is the sequence of events/operations involved in _______?

• What are the strengths/weaknesses of _______?

• How are _______ and _______ different/similar?

• What inferences/conclusions can be drawn about _____?

• Use a case as a starting point – this is especially good for application items
Assessment of Items
Checklist for Reviewing Individual Items

- Has the item been constructed to assess a single written objective?
- Is the item based on a specific problem stated clearly in the stem?
- Does the stem include as much of the item as possible, without including irrelevant material?
- Is the stem stated in a positive form?
Checklist for Reviewing Individual Items

- Are the alternatives worded clearly and concisely?
- Are the alternatives mutually exclusive?
- Are the alternatives homogenous in content?
- Are the alternatives free from clues as to which response is correct?
Checklist for Reviewing Individual Items

- Have the options *all of the above, none of the above*, and other complex options been avoided?

- Does the item include as many functional distractors as are feasible?

- Does the item include one and only one correct or clearly best answer?
Checklist for Reviewing Individual Items

- Has the answer been randomly assigned to one of the alternative positions?

- Is the item laid out in a clear and consistent manner?

- Are the grammar, punctuation, and spelling correct?

- Has unnecessarily difficult vocabulary been avoided?
Checklist for Reviewing Entire Test

- Did you develop learning objectives?
- Are items linked to learning objectives?
- Do items map to the content in the module?
- Does the number of items reflect the importance and amount of material?
- Are there both knowledge and application items?
Checklist for Reviewing Entire Test

- When submitting items indicate the following:
  - Is this a knowledge or application item?

- Which objective is this item intending to measure?
  - Indicate location of content within the module

- Are items sampling the content and learning objectives well?
Post Hoc Item Analyses

• Quality of items will be analyzed post hoc using:
  ▫ Examinee feedback on questionable or confusing items
  ▫ Statistical item analyses such as
    • Inter-item Reliability
    • Item Difficulty
    • Item Discriminability
    • Distractor Analyses


