It is hoped that you will find this information truly valuable and decide to become active with the CPC Item Review team.

Valid test scores are produced when the test measures the specific knowledge and skills that were intended to be measured. The validity of test scores is also affected by item quality. Test items must be representative of the knowledge, skills, and abilities important to the specified content area.

Reliable tests consistently yield comparable scores for examinees with comparable levels of the knowledge and skills measured by the test. Scores are said to be reliable when they are as free from measurement error as possible. The reliability of scores is directly affected by the quality of the items. Items that are ambiguous or inaccurate increase measurement error and, thus, diminish reliability.

A well-constructed multiple-choice test is based on the test content outline. This outlines the content areas of a particular discipline so that all areas of pertinent skills, knowledge, and abilities can be assessed. The test content outline also indicates the relative emphasis, in number or percent of times, that should be given to each content area. Constructing a test that conforms to specified guidelines contributes to the validity of the interpretation of scores derived from it.
Purpose

- The purpose of this tutorial is to introduce the process of writing items (questions) for a certification examination.
- Provide an opportunity for certified optometric staff to become a key member in the design and development of certification examinations.
Objectives

- Participants will:
  - Have an understanding of the process of writing items for a professional certification examination
  - Know how to label the anatomy of an item
  - Learn the advantages of multiple-choice items
  - Be able to construct items using best practices
  - Identify common item flaws

Following completion of this tutorial, participants will be able to identify the many item formats included on examinations, have an understanding of the process of writing items for a professional certification examination, know how to label the anatomy of an item, learn the advantages of multiple-choice items, be able to construct items using best practices, and be able to identify common item flaws.
Prerequisites

- Should have a good knowledge in optometric assisting and/or optometric coding and billing
- Must be CPC certified at any level

In order to be successful, you should have a good knowledge of optometric assisting and/or optometric coding and billing. All participants must be CPC certified at any level.
DISCLAIMER: The example questions shown throughout this tutorial are not and never will be included on any CPC Examination. Example questions are used for demonstration only.

Throughout this tutorial, there are many questions that are used for example purposes. Some will demonstrate the wrong way to prepare a test question and others are what we consider good items. Please note that any question used in this tutorial is not and will never be included on any CPC examination. As noted, they are for demonstration only.
**Multiple Choice**

- Consists of:
  - Stem (question)
  - One Correct Answer
    - Always labeled “A”
  - Three Distractor Answers
    - Always labeled “B, C, & D”
    - Should be plausible but incorrect
Why Multiple Choice?

- Flexibility – allows efficient measurement of a variety of objectives
- Can be used to test many levels of candidates
- Can be used to determine a candidates weaknesses and strengths
- Can provide candidates with feedback
- Can cover a wide range of difficulty levels
- Can include a wide sampling of content or objectives
- A reduced guessing factor when compared to true-false items

So why are multiple choice type items so popular?

To begin with, they are flexible. Multiple choice items allow for efficient measurement of a variety of objectives. Further, they:

- Can be used to test many levels
- Can be used to determine a candidates weaknesses and strengths
- Can provide candidates with feedback
- Can cover a wide range of difficulty levels
- Can cover a wide sampling of content or objectives
- Can reduce the guessing factor when compared to true-false items
Advantages

- Versatility in measuring all levels of cognitive ability
- Highly reliable test scores
- Scoring efficiency and accuracy
- Objective measurement of the test takers achievement or ability

There are many advantages in using multiple-choice items. To list a few, multiple choice items offer versatility in measuring all levels of cognitive ability. The test scores are highly reliable while offering scoring efficiency and accuracy, and they objectively measure the test takers ability.

But where advantages are, disadvantages follow.
Disadvantages

- Difficult and time consuming to construct
- Do not allow the test taker to demonstrate knowledge beyond the options provided
- May encourage guessing as one option will always be right
- Some test takers may misinterpret questions
- Places a high degree of dependence on the candidate’s reliability and the item writer’s writing ability
You might ask, what is a test item? Test items are the building blocks of an exam. They measure one unit of content and a multiple-choice item is composed of three parts.

1. Item Stem (question)
2. Correct Answer
3. Distractors (incorrect responses)
Anatomy of a multiple-choice item

What is chiefly responsible for the increase in the average length of life in the USA during the last fifty years?

A. Compulsory health and physical education courses in public schools.
B. The reduced death rate among infants and young children.
C. The safety movement, which has greatly reduced the number of deaths from accidents.
D. The substitution of machines for human labor.

The multiple-choice item consists of two parts: the stem, which identifies the question or problem and the options (the correct option and three distracters).
General Guidelines for Item Construction

- Strive for clarity and readability
- Have a style guide and be consistent
- Use punctuation at the end of a stem (question mark or a colon)
General Guidelines for Item Construction – Do’s and Don'ts

• If the stem ends with a question mark, do not use a period at the end of the answer options unless they are complete sentences.

• If the stem ends with a colon, place a period at the end of each one because this is the sentence completion format.

• Include in the stem any word(s) that might otherwise be repeated in each option.

• Avoid bias and stereotyping.
General Guidelines for Item Construction – Do’s and Don'ts

- Do not use “a(n)” or “is(are)” combination constructions in the stem. Instead, place the appropriate article in front of each option (e.g., A dog, An aardvark, etc.)
- Each group of answer options should have parallel construction, so that if the sentence completion format is used, each completed sentence would dovetail with the stem.
- If the stem is phrased in the negative, the answer options should not contain negatives.
Constructing the Item Stem

- Example of an open stem:

  APP-CO was founded in:
  A. 1958.
  B. 1959.
  C. 1960.
  D. 1961.

- Example of a closed stem:

  In what year was APP-CO founded?
  A. 1958.
  B. 1959.
  C. 1960.
  D. 1961.

A good stem has clear and direct wording, allowing the examinee to concentrate on figuring out what the answer is. The stem can be presented as either an incomplete statement (open stem) or a complete question (closed stem).
## Constructing the Item Stem

- **Undesirable**
  Which of the following product lines does ABC produce and when did production begin?
  - A. Food processors, 1960
  - B. Can openers, 1962
  - C. Blenders, 1974
  - D. Irons, 1984

- **Desirable**
  Which of the following product lines does ABC produce?
  - A. Food processors
  - B. Can openers
  - C. Blenders
  - D. Irons

Each item should test only one concept. A two-part question does not allow the test developer to know if the examinee knows both parts or only guessed on one part.
Avoid Negatively Phrased Stems

Examples of negatively worded stems:

- All of the following are acceptable ways to answer the telephone EXCEPT:
- Which of the following forms is NOT used for requesting materials?
- Which automobile is LEAST likely to increase gas mileage?

Although research is inconclusive, there is concern that negatively phrased items might confuse examinees, especially those whose first language is not English. The confusion occurs because the examinees must change their normal thought processes to respond to negative items and may overlook the negative aspect and answer the item incorrectly, even though they know the correct response. **Note** that using negatives in both the stem and answer options are never acceptable. **Note** that if a negatively phrased item is used, the negative word should be capitalized, underlined, or bolded.
Avoid Pronouns

A good stem avoids the use of pronouns “it”, “he”, “she”, and “you”.

In the case of “it”, “he”, and “she”, the candidate may be uncertain about the referent.

Questions that ask “What would you do?” can have no right answer, since individual candidates could justify their individual responses.
A tourist in New York City could visit all of the following EXCEPT:

A. Centennial Park
B. Grant Park
C. Central Park
D. Balboa Park

What’s wrong with this Item? The stem is negatively phrased. A more appropriate way to write this item would be, “Which of the following could a tourist visit in New York City?”
Common Item-Writing Flaws

- Where is Atlanta located?
  
  A. In the United States
  B. In the South
  C. In Georgia
  D. In Fulton County

What’s wrong with this Item? There are overlapping options. What would be a better way to word this? Better options are: A. Georgia B. Missouri C. California D. Mississippi
What’s wrong with this Item? A good stem poses a problem to which the correct answer is not likely to change over time. A better way to phrase the stem is, “Who was the President of the United States in 2006?”
Constructing the Options

- **Undesirable**
  What is the most general cause of low individual incomes in the United States?
  A. Lack of valuable productive services to sell
  B. Unwillingness to work
  C. Automation
  D. Inflation

- **Desirable**
  What is the most general cause of low individual incomes in the United States?
  A. A lack of valuable productive services to sell
  B. The population’s overall unwillingness to work
  C. The nation’s increased reliance on automation
  D. An increasing national level of inflation

All options should be similar in content, grammatical structure, and length.
Options should be mutually exclusive and distinctive. Options should not overlap, be the equivalent to, or be contained within any other option.
Avoid Clueing

**Desirable**
Renee Zellweger stars as a show girl in Chicago in which of the following films?
A. Wizard of Oz  
B. Sound of Music  
C. Chicago  
D. Music Man

**Undesirable**
Renee Zellweger stars as a show girl in Chicago in which of the following films?
A. Wizard of Oz  
B. Sound of Music  
C. Chicago  
D. Music Man

Items should not contain clues that could unfairly help examinees who are test-wise but not well prepared.
Items should be recorded and documented. This is an example an item writer might use.

<table>
<thead>
<tr>
<th>Please provide a complete classification for this item.</th>
<th>Correct Answer (mark one) ✓</th>
</tr>
</thead>
<tbody>
<tr>
<td>I ○ II ○ III ○ IV ○ V ○ VI ○</td>
<td>A ○ B ○ C ○ D ○</td>
</tr>
<tr>
<td>A ○ B ○ C ○ D ○</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Month/Year Written:</th>
<th>Illustration? (mark if “yes”) ✓</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item Writer Name (last, first):</td>
<td>Item Writer ID:</td>
</tr>
<tr>
<td>Source Name:</td>
<td>Source Page #:</td>
</tr>
</tbody>
</table>
Items must be classified according to the test content outline. Tests are assembled based on the content classifications. The outline defines the content and relative weight of each content area on the test. If items are not directly related to the content, they cannot be included on the test because the test scores would not be valid.
Linking Items to Test Content Outline

- Can the item be classified according to the test content outline?
- Is the item critical to practice?
- Is the item technically accurate?
- Does the item have only one correct answer?

If you can answer “yes” to all of these questions, the item will likely contribute to valid test scores. Providing evidence that the scores are valid is the most important criterion for any test.
A second level of item scrutiny will assist in providing evidence for the reliability of the test scores. If the items are free from construction flaws, examinees’ scores will be more dependable because a major source of measurement error—ambiguity and confusion within the items—has been removed.
Key Points for Creating Test Items

Ensure that the stem of the item....

✓ contains important content.
✓ is at the appropriate level.
✓ is focused.
✓ avoids “window dressing” (superfluous information).

Ensure that the item options....

✓ have one correct answer.
✓ are plausible.
✓ are homogeneous.
✓ are free of clueing.
✓ are free of overlap.
**Key Points for Creating Test Items**

- Ensure that the items are at the correct cognitive level by identifying the level they are written at, deciding if it is the correct level, and if not, rewriting the item in order to raise the cognitive level to the level required.

- Avoid bias and stereotypes by reviewing each item, and if at all possible, have someone else review the item.
Key Points for Creating Test Items

- Write rationales for all of the items.
- Justify the answer key for all items.
- Link all items to the test content outline and, if possible, have someone who did not participate in the item development do so.