2025 ITEM WRITING GUIDE

ASC

Assessments Systems Corporation (ASC) strives to
positively impact educational and career opportunities
by facilitating the development of fair, valid, and reliable
assessments.

Contents

Introduction

This guide is not meant to be all-inclusive but provides an overview of best practices for writing high-quality test items. It is intended for educators, assessment developers, certification bodies, and organizations involved in assessment development.

Before item writing begins, two essential procedures must take place: conceptualization of assessment and construction of test specifications. Conceptualization involves defining what the assessment will measure—identifying the construct, determining when and how it should be assessed, and establishing the intended interpretations of test scores. Once the construct is clearly defined, a test blueprint is developed to outline the structure of the assessment, ensuring alignment between the test content and its intended purpose. The blueprint serves as a guiding framework, specifying content areas, item distribution, and item formats, setting the direction for item development.

This guide covers the main requirements for item writers, principles of item writing, item review, and validation processes to ensure fairness, reliability, and validity of assessment results. It should be used in conjunction with the respective testing program materials, such as test specifications, exam scoring rubrics, and other guidelines, including the program's style and item format requirements.

Essential Guidelines for Item Writers

Effective item writing ensures that test items accurately measure the intended knowledge, skills, or abilities. Well-constructed items contribute to test validity, ensuring that scores meaningfully reflect examinees' competencies rather than extraneous factors. This section outlines key principles of high-quality item writing, including validity, alignment with objectives, clarity, fairness, and appropriate scoring considerations.

Validity

The most important aspect of test scores is the **validity** of their interpretations. We need the item to specifically measure what is supposed to be measured (called a **construct** in psychometrics), because then the scores on the test will accordingly reflect the construct (**construct-relevant variance**) and not any unrelated traits or aspects of the testing process (**construct-irrelevant variance**). Validity is established through a chain of evidence that connects job analysis, test specifications, item content, test scores, and their interpretation. In the case of professional competency examinations, that chain is: job analysis – specifications – items – scores – interpretations. The test development process should revolve around ensuring linkage within the chain and documenting the linkage as much as possible.

Mapping objectives

An item writer is making the step from test specifications (outline or blueprints) to individual items. Ensuring that items align with test specifications strengthens validity by directly linking assessment content to the intended learning objectives. Therefore, when writing an item, the specification (learning objective, outline point) for which the item is intended must be recorded.

Clarity and conciseness

The item writer must be focused in making sure the content of the item only relates to the piece of knowledge, with no superfluous information, and the examinee responses are designed to differentiate among examinees. Like a scientific experiment that isolates one variable, an effective test item should focus only on the knowledge or skill being measured, avoiding unnecessary complexity.

Understanding examinees

Therefore, one of the most important aspects is to think like an examinee. How would examinees of low, medium, and high ability read, interpret, and answer the item? Obviously, we want examinees of higher ability to be able to recognize the correct response. Examinees of lower ability will not be able to, but not necessarily because they are confused. To write effective items, an item writer must anticipate how examinees with different ability levels will interpret and respond to the question, ensuring the item differentiates appropriately. Items should be of appropriate difficulty; if an item is to differentiate between low and medium examinees, the item writer must conceptualize what constitutes a low and medium examinee.

Anticipating scoring

Also, writers should keep in mind the scoring of the item. For multiple choice items, this is easy. Because the correct answer gets 1 point and the remainders get 0 points, the correct answer should be fully correct while the others are fully incorrect. Unfortunately, this simplification does not lend itself well to assessing complex constructs such as higher-order thinking or psychomotor skills. When developing open-response or performance-based items, item writers should envision the scoring rubric early in the process. For example, if designing a simulated patient scenario for a medical exam, consider how different responses will be assigned points based on accuracy and completeness. If the question involves a conversational speaking response for an English test, how would you algorithmically assign points on a scale of 0 to 5?

Understanding material

Perhaps the most obvious, and therefore possibly overlooked, requirement is that item writers have substantial knowledge of the material. In other words, they must be of high ability themselves. Item writers must possess deep subject matter expertise to create questions that accurately assess higher-level competencies.

Record rationale

Record the reasoning behind the item and the correct response. If the item is used and reviewed a year from now, whoever is reviewing will want to know the rationale. The rationale should document the reasoning behind the item and correct response, including references to textbooks, guidelines, or expert consensus.

The Importance of Validity in Test Development

Developing an examination is a rigorous process that involves several different steps. At a high level, there are four main processes in the test development cycle: a) defining the examination scope and outlining the contents on the examination, b) writing items that align with the examination content specifications while following standardized procedures, c) obtaining preliminary statistical information on these items to generate forms and setting a cut score, and d) administering the exam, analyzing the data, and reporting the final scores to candidates. Figure 1 presents an overview of the relationship between these high-level processes throughout the test development cycle.

Figure 1:
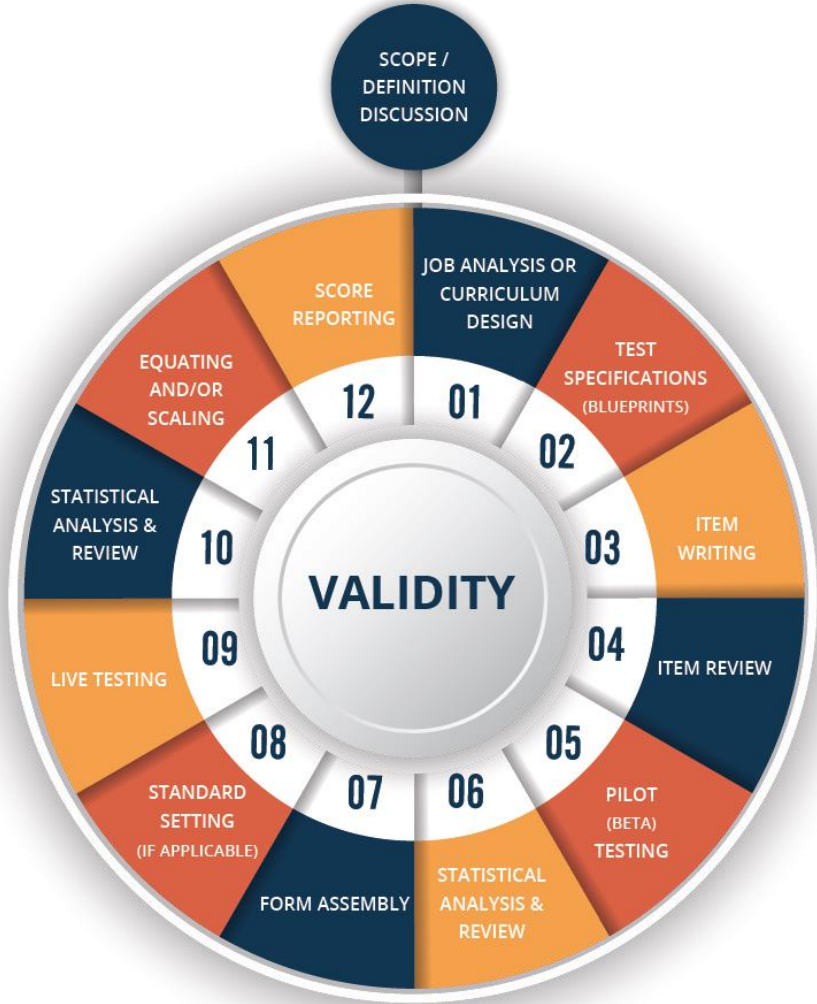
Validity

As can be seen from Figure 1, the entire exam development cycle is driven by **validity**— or the degree to which the examination measures what it claims to measure. Thus, each stage in the development process, from defining the content on the test to reporting scores, adheres to rigorous standards to ensure that the interpretation of the test scores supports the goals of the examination program.

Why Good Item Writing Supports Test Validity

After the initial phase of the exam development cycle, which includes defining the exam purpose and defining the content on the examination, items must be written to support the test content.

When writing items, there are several threats to validity that should be recognized. Some potential threats when writing items include:

- Not accurately reflecting the intended test specifications.
- Failing to align with test content specifications.
- Being unclear or confusing.
- Overlapping multiple content areas inappropriately.
- Introducing systematic bias against specific groups of test takers.

What is an Item?

Structure of an Item

There are four general components of any multiple-choice (MC) item: the stem, the options, the key, and the distractors. Each one of these components plays a specific role and contributes to the overall structure of the item.

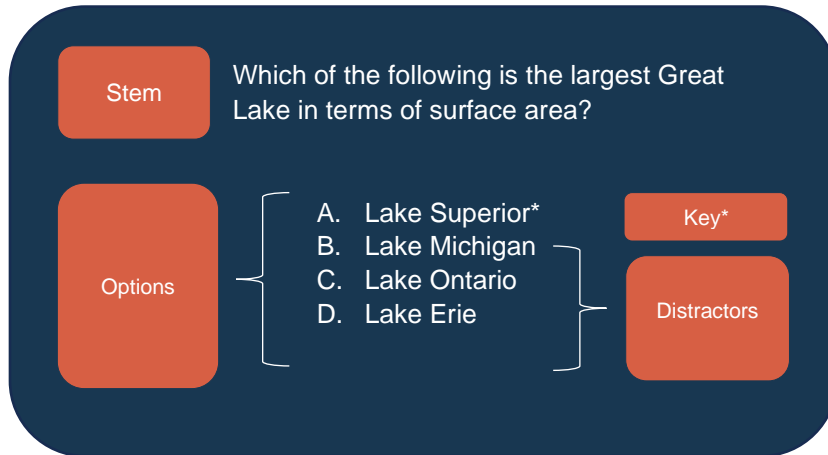**Figure 2.** *Example Single Select, MC Item*

Table 1 provides the definitions of key terminology reflective of each component. The general structure of an MC item is presented in Figure 2.
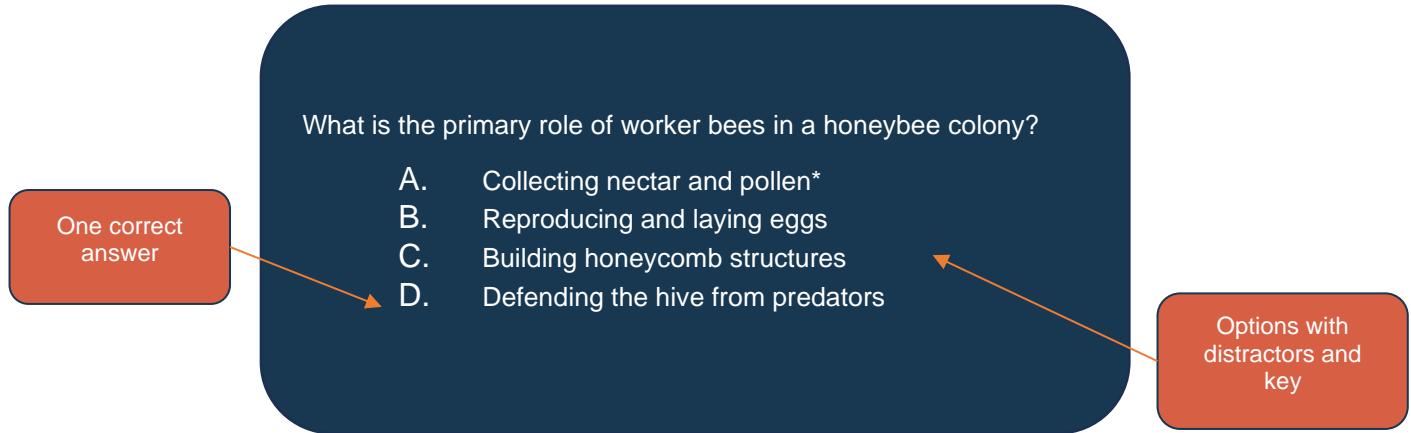
Table 1. *Anatomy of an MC Item*

Term	Description
Stem	The stem is the question that is being asked. The item stem should: • pose a clearly defined problem. • be positively phrased (avoid NOT, EXCEPT). • include all relevant information to answer the question.
Options	The options are the answer choices to the question (for multiple choice items). The item options should: • be similar in length and structure. • fit logically and grammatically with the stem. • **not** include key words from the stem. • **not** contain specific determiners such as ALWAYS and NEVER. • follow a logical order, such as ascending or descending length, numerical value, or chronological sequence, when applicable, to enhance readability and fairness.
Key	The key is the correct answer choice to the question. The key should: • be the only correct answer. • clearly answer the question. • **not** include ALL OF THE ABOVE or NONE OF THE ABOVE. • **not** stand out from the incorrect (distractor) options.
Distractors	The distractors are the incorrect answer choices. The distractors: • should be plausible but incorrect. • be based on common errors or misconceptions about the skill or knowledge. • be similar in content, terminology, and style as the correct answer (key).

Different Item Types

There are several different item types that can be used to formulate items for an exam. Ultimately the type of item used on the test should be decided on given the business objectives, the test design, and the operational structure of the exam. The most common item formats are multiple-choice single select, multiple-choice multiple response, and constructed response items

Single Select Multiple-Choice. Single select multiple-choice items are the most common type of test item as they are easier to write than more complex item types and are easily scored in an objective manner, making them highly reliable.

Figure 3. *Example Single Select, MC Item*

> What is the primary role of worker bees in a honeybee colony?
>
> A. Collecting nectar and pollen*
> B. Reproducing and laying eggs
> C. Building honeycomb structures
> D. Defending the hive from predators

One correct answer

Options with distractors and key

Multiple-Choice Multiple Response (MCMR). Multiple-choice multiple response (MCMR) items are multiple-choice items that have more than one correct answer. Multiple response options look like regular single select multiple-choice items; however, there is more than one correct answer key.

When writing MCMR items, item writers should have clearly defined rules on the number of options as well as the number of correct responses (keys). Additionally, MCMR questions should provide the examinee with information about how many options to choose from the list of options.

There are two primary approaches to MCMR items that can be used. The item can include instructions that state "Select all that apply…," or the item instructions can specify a required number of responses (e.g., "Select two of the following…"). The choice of instructions is closely tied to the scoring model (e.g., weighted options, right/wrong, or partial credit) so the appropriate item writing guidelines are provided to the item writer.

Figure 4. *Example MCMR Item*

What are the THREE stages of the Prenatal Period?

A. Fetal*
B. Feeling
C. Infancy
D. Germinal*
E. Embryonic*
F. Preoperational
G. Preconventional
H. Early Production

Three correct answers

Options with distractors and key
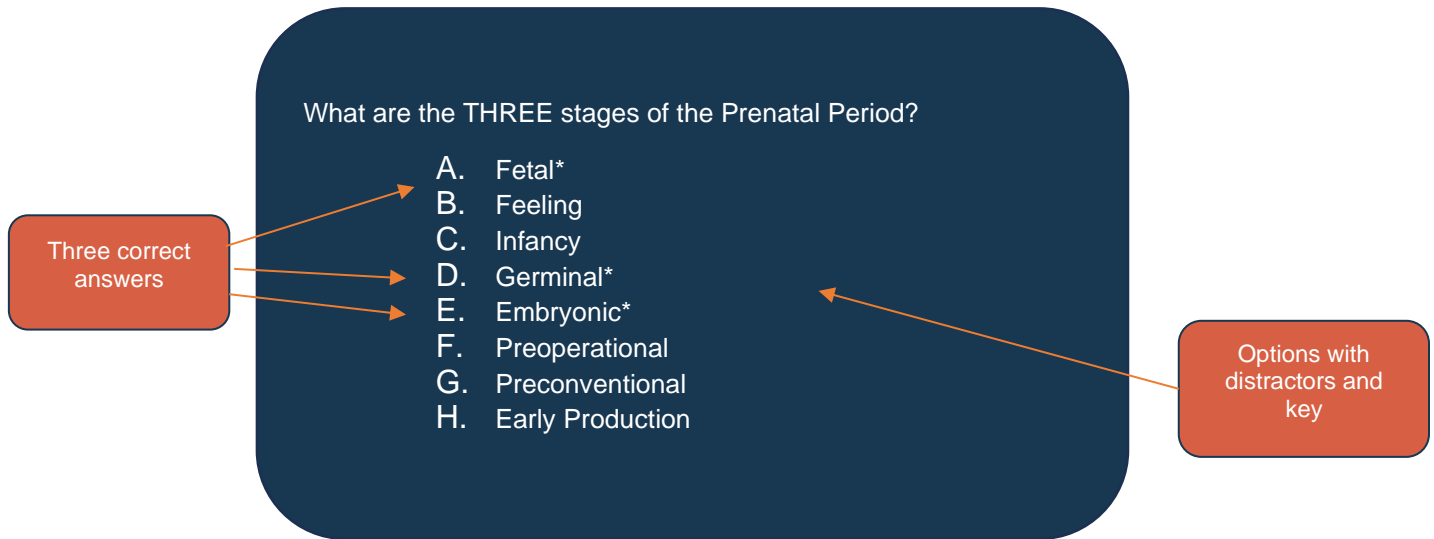
Constructed Response. Constructed Response (CR) items are items that require the candidate to produce the answer instead of simply selecting the answer from a list. The required response can be as simple as a single word or can be as complex as the design of a set of procedures to support the lifecycle of a product. Examples of CR questions include asking a candidate to write an essay using a compare and contrast argument and asking a candidate to write down an equation for some type of problem.
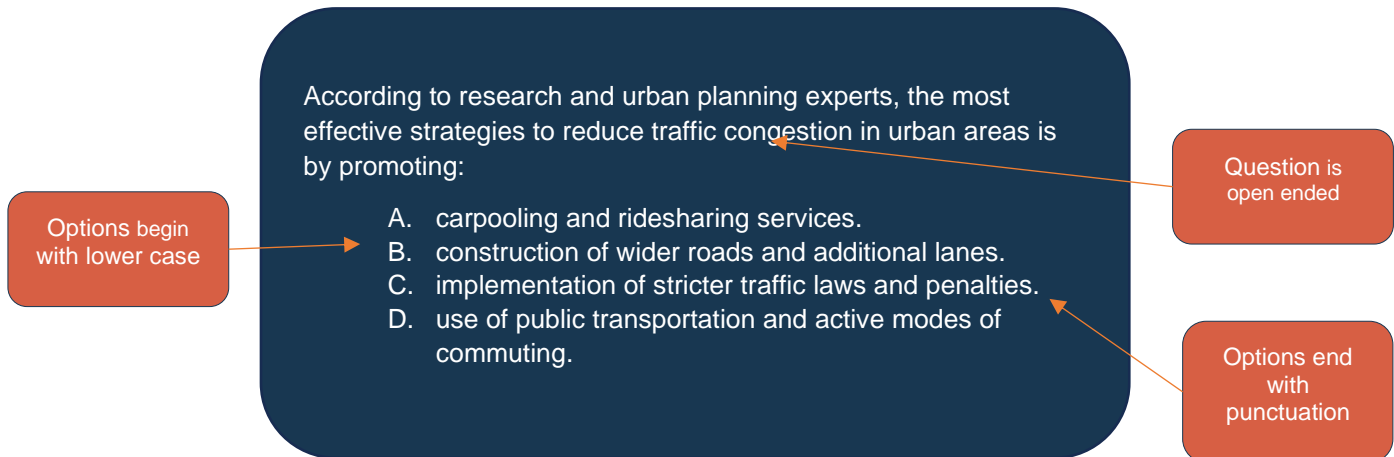
Although CR items may be able to tap into skills and knowledge that may not be assessed with a multiple-choice type item, special consideration must be given to the scoring and cost of constructed response items at the onset of the test development process to ensure that the cost of these types of items adds value to the candidates who take the exam.

Item Formats

Different item writing formats can be used when writing multiple choice items. Each of these item formats are described below with representative examples.

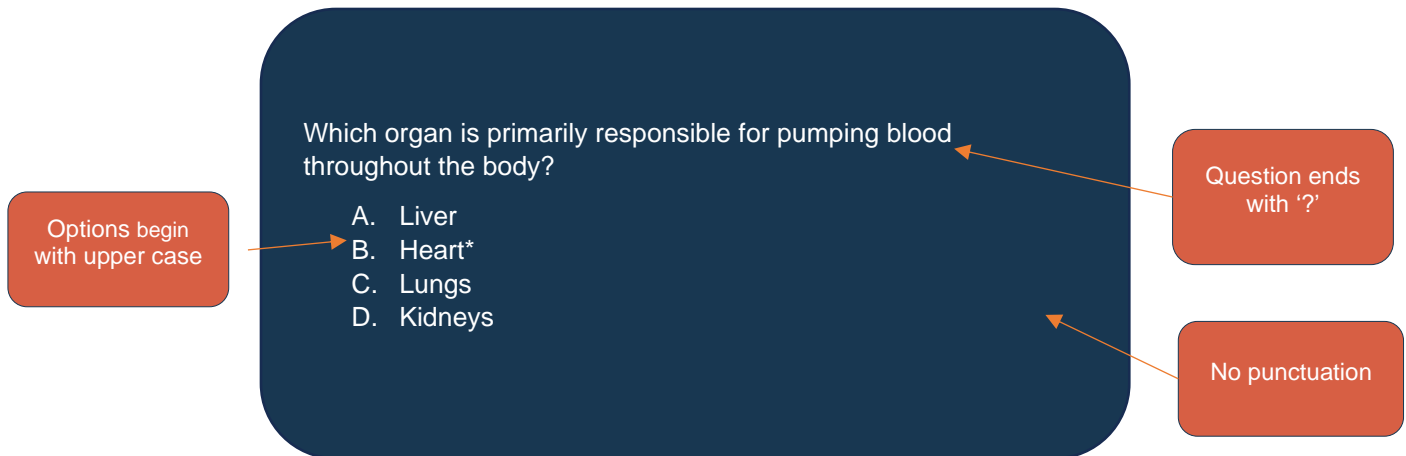Open Stem Items (Sentence Completion). An open stem item, or sentence completion item, is an incomplete statement where each of the answer options completes the question. In an open stem question, note that the question leads into each option, forming a complete sentence. In an open stem question, the item stem ends with a colon, the options begin with a lower-case letter, and end with a punctuation (or end mark).

Figure 5. *Example Open Stem Item*

According to research and urban planning experts, the most effective strategies to reduce traffic congestion in urban areas is by promoting:

- A. carpooling and ridesharing services.
- B. construction of wider roads and additional lanes.
- C. implementation of stricter traffic laws and penalties.
- D. use of public transportation and active modes of commuting.

Options begin with lower case

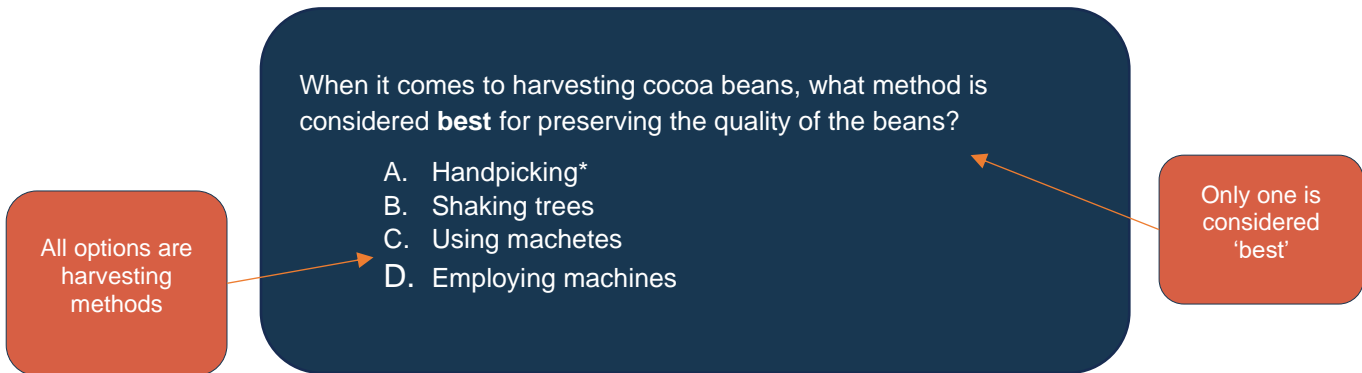Question is open ended

Options end with punctuation

Closed Stem Items. A closed stem item asks a complete question and ends with a question mark. No punctuation is needed in the option choices, and they should begin with capital letters as they are stand-alone statements unless the answer choices are sentences. If the latter is the case, they should end in punctuation.

Figure 6. *Example Closed Stem Item*

Which organ is primarily responsible for pumping blood throughout the body?

- A. Liver
- B. Heart*
- C. Lungs
- D. Kidneys

Options begin with upper case

Question ends with '?'

No punctuation

Best/Most Item Format. The Best/Most item format requires an examinee to evaluate and select the best, or most appropriate response. The information the examinee is required to use to facilitate the evaluation of the correct answer should be included in the stem. The information supporting the correct answer should be objective and supported by a consensus in the respective field.

Figure 7. *Example Best/Most Item*

When it comes to harvesting cocoa beans, what method is considered **best** for preserving the quality of the beans?

    A. Handpicking*
    B. Shaking trees
    C. Using machetes
    D. Employing machines

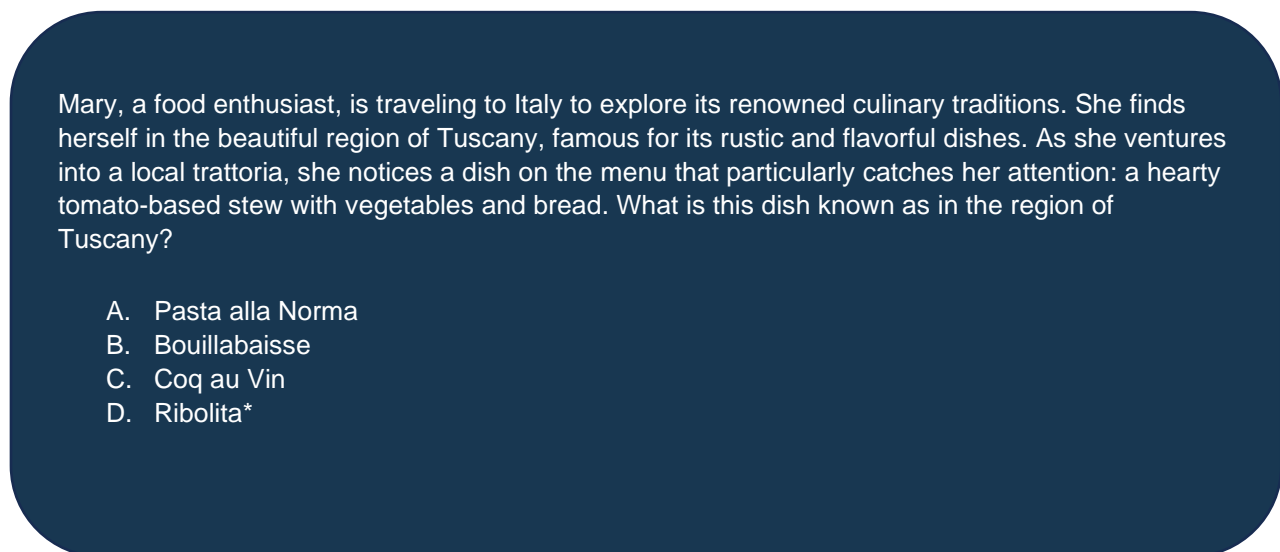All options are harvesting methods

Only one is considered 'best'

Vignettes with Individual Items or Sets of Items. Items can be developed using vignettes. A vignette is a presentation of a realistic scenario that the examinee may encounter in his or her profession. A vignette may form the basis of a single item or a set of items.

Item writers are encouraged to develop items in the form of a clinical vignette. These types of items are beneficial for the test takers taking the exam as they simulate a real-life situation and help test-takers frame their clinical reasoning skills within that context.

Clinical vignettes should NOT be taken directly or even paraphrased from published literature. This is a violation of copyright law.

Figures 8 and 9 present an example of a vignette with an individual item as well as a vignette with a set of items.

Figure 8. *Example Vignette with an Individual Item*

Mary, a food enthusiast, is traveling to Italy to explore its renowned culinary traditions. She finds herself in the beautiful region of Tuscany, famous for its rustic and flavorful dishes. As she ventures into a local trattoria, she notices a dish on the menu that particularly catches her attention: a hearty tomato-based stew with vegetables and bread. What is this dish known as in the region of Tuscany?

    A. Pasta alla Norma
    B. Bouillabaisse
    C. Coq au Vin
    D. Ribolita*

Figure 9. *Vignette with a Set of Items*

Items 1 and 2 refer to the following case.

A 58-year-old patient is brought to the psychiatrist by family members. The family reports that, over the last 2 months, the patient has gradually become increasingly lethargic, forgetful, and irritable. The family reports a change in the patient's personality, and that the patient has begun to make rude sexual remarks in public. The patient has no prior medical, psychiatric, or substance abuse history and takes no medications. Neurological examination reveals a present snout reflex. The Folstein Mini-Mental State Examination (MMSE) is entirely normal (including orientation items). Normal results are also obtained from vitamin B_{12}, folate, VDRL, and electroencephalogram (EEG) testing. Neuropsychological testing reveals impairments in higher adaptive functioning, flexibility of thought, speed of processing, and alternation of sets.

1. Which of the following findings would be expected on a magnetic resonance imaging (MRI) scan of the patient's head?

    A. Frontal lobe atrophy*
    B. Parietal lobe atrophy
    C. Global cortical atrophy
    D. Multiple ring enhancing lesions
    E. Periventricular white matter lesions

2. The patients' most likely diagnosis is:

    A. Pick disease.
    B. Wilson disease*
    C. HIV encephalitis.
    D. herpes encephalitis.
    E. Creutzfeldt-Jakob disease.

How to Begin Writing an Item

How to Write an Item

Items can be framed in a variety of ways and ideas for items can come from a variety of sources. Items should be focused on knowledge that is important for the examinee to know and critical for the examinee to understand to be successful on the respective examination. Items should also clearly align to the test specifications. Any item written that is not aligned to content being measured on the examination for which it is being created is a threat to the validity of the test results.

       Writing the Item Stem. The item stem is the foundation of the item writing process and is generally the first component of the item that is written. Table 2 presents key considerations that item writers should think about when constructing the item stem. It also highlights things to avoid when writing the stem.

Table 2. *How to Write the Item Stem*

The item stem should:	When writing the item stem:
Be focused on only one knowledge concept.	Do **not** write items with multiple content objectives.
Be focused and address a clearly defined problem; the stem should be able to be answered as a stand-alone question without seeing the item options.	Do **not** write items such as "Which of the following is true (or false)?" This type of item poses a non-specific, unfocused question.
Clearly align with the test specifications.	Do **not** write items outside the test specifications.
Be written in common vocabulary that is clear and concise.	Do **not** include extraneous information.
Be written in the positive.	**Avoid** using NOT, LEAST, EXCEPT in the stem. This type of item asks the test taker to find the wrong answer as the key.

      Writing the Response Options. When writing the item options, the item writer should adhere to the test design that is defined by the testing program. For example,

the item writer should confirm the number of options needed for the items on the respective exam, and the number of correct answer keys. Table 3 outlines essential factors to consider when developing item response options.

Table 3. *Writing the Item Response Options*

The item options should:	When writing the item options:
Be homogenous in content (same concepts).	Do **not** write options that subsume each other.
Be parallel in length and structure (similar in length).	Do **not** write items that directly contradict one another.
Be clear and concise, with no extraneous information.	**Avoid** specific determiners, such as ALWAYS and NEVER.
Fit logically and grammatically with the stem.	Do **not** include keywords from the stem.
When writing numeric answer options, answer options should be sorted increasingly.	Do not use NONE OF THE ABOVE or ALL OF THE ABOVE.

Writing the Item Key. When writing the item key, the item writer should make sure that the key is the only correct answer option except when developing MCMR items, as previously discussed. The key should not stand out differently than the other options that are written in grammar, content, or length. Table 4 provides an overview of crucial elements to keep in mind when writing the item key.

Table 4. *Writing the Item Key*

The item key should:	When writing the item key:
Be the only correct answer.	Do **not** write options that include multiple correct answers.
Be similar in length and structure as the other item options.	Do **not** include more information that makes the correct answer (key) obvious to the examinee.
Fit logically and grammatically with the stem.	Do **not** include keywords from the stem.

Writing the Distractors. Creating plausible distractors is the most difficult aspect of creating a high-quality MC item. The best distractors are either accurate statements

that do not meet the full requirements of the problem or incorrect statements that seem correct to a novice professional.

It is helpful to consider the following questions when writing item distractors:

- What is a common error for solving this problem?
- What do individuals usually confuse this concept with?
- What potential mistakes might the individual make while performing this task?
- What are common misconceptions in the field?

Table 5 illustrates key principles for writing distractors, which are often considered the most challenging part of creating an MC item.

Table 5. *Writing the Item Distractors*

The item distractors should:	When writing the item distractors:
Be plausible yet incorrect.	Do **not** make up terms.
Be based on common misconceptions or critical misunderstandings.	Do **not** try to trick the candidates.
Be of similar structure and content as the key and content in the stem.	Make sure that they **cannot be** defensible as the correct answer.

Referencing Items

Items should reflect consensus in the field and should always be referenced with supporting material(s).

      References. For each item, a reference source should be provided. The reference source should include the name of the author, the book edition (if applicable), the year of the publication, and the page number of the supporting evidence of the item key.

      Electronic Roadmap. An electronic roadmap should be provided for online books that are references. This will help examinees easily find the information about the content.

Item Review

Item writers often feel relief after finishing an item writing assignment. However, that is just the first step in a long process of ensuring the quality of each item and the validity of resulting scores.

All items should be reviewed by at least one other subject matter expert (SME) before the item is pilot tested with examinees. The reason for the review process is to ensure validity from a quality control perspective. Irrelevant information can lead to possible challenges by examinees that the item is not valid, namely not focusing on measuring what is supposed to be measured.

The item review process should be managed by the testing program coordinator or vendor that is responsible for the item bank management processes. Item writers and reviewers often benefit from a systematic item review process using an item review checklist that clearly defines to the reviewer what type of things to evaluate. The exemplary checklist provided in Appendix A outlines key criteria for evaluating item content, clarity, structure, fairness, and alignment with test specifications.

General Style Guidelines

ASC has developed these guidelines for writing and reviewing items to facilitate successful item development. It is important to follow these basic procedures so that the item writing process supports the intended examination.

Style

To reduce construct-irrelevant variance, items should be formatted as similarly as possible. All items should have the same font style and size, same number of options when feasible, and similar writing style/structure.

Appropriate Language and Context

Simple and clear language should be used that is consistent with the assessment. Jargon, slang, and specialized vocabulary that is irrelevant for the construct being assessed should be avoided. The maturity and educational level of the examinees should be considered in the item writing process. Language that is inappropriate in tone, such as being patronizing, insulting, elitist, or inflammatory should be avoided.

Avoiding Bias

When preparing assessment items, be sensitive to the possibility of unintentionally placing groups of candidates at an unfair disadvantage. Writing items for an examination requires special attention to the diversity of environments, backgrounds, beliefs, and cultures among test candidates.

To meet this goal, ASC follows specific standards that ensure that test items, the test does not include contain language, symbols, words, phrases, or examples that are generally regarded as sexist, racist, or otherwise potentially offensive, inappropriate, or negative toward any group.

Summary

Following proper item writing procedures is perhaps one of the most critical steps in ensuring that the examination produces scores that can be correctly interpreted as the items on an exam are the foundation of what is being assessed. Items should be fair, clear, and targeted to the exam specifications that the item is written for.

When beginning the item writing process, item writers should have a clear understanding of the type of item they are writing, clearly defined test specifications, and an understanding of their item writing assignment for the testing program.

This guide has presented general item writing guidelines that can help support item writers in the item development process. It has included an overview of the structure of an item, the important characteristics of high-quality items, and an example workflow of the item writing process. Key elements in item development were defined and examples were provided to support item writers in their efforts to create valid and focused items for an examination. Item writers can use this guide as a tool along with other ancillary information (e.g., test specifications and specific testing program styles and rules) to help support them in their item writing efforts.

Appendix A. Item Review Checklist

Review Criteria	Done?
Content Alignment	
All items align with the test specifications and accurately reflect the intended content.	
Items cover the knowledge, skills, or abilities recognized as important for the exam.	
Items are mapped to a specific learning objective or outline point.	
Clarity and Conciseness	
Item stems are clear, concise, and free of extraneous or irrelevant information.	
Items avoid overly complex wording, unnecessary jargon, or specialized language irrelevant to the construct.	
Items are written using simple and clear language appropriate for the target examinee population.	
Item Structure and Format	
The item stem poses a clearly defined problem or question.	
Stems are positively phrased (avoid NOT, EXCEPT, LEAST).	
Items are formatted consistently with the required test design (e.g., single-select multiple-choice with four options).	
The correct answer (key) is the only fully correct response (except in MCMR items, as previously discussed).	
Answer Options	
All distractors are plausible but incorrect, avoiding obvious or overly weak distractors.	
All answer choices are similar in length and structure to prevent unintentional clues.	
Answer options fit logically and grammatically with the stem.	
No answer choices include absolute terms such as ALWAYS or NEVER (unless justified).	
Options follow a logical order (ascending/descending length, numerical value, chronological sequence, etc.).	
No answer choice repeats key words from the stem.	
Bias and Fairness	
Items do not introduce systematic bias toward any specific demographic group.	

Language, symbols, or scenarios are neutral and culturally appropriate unless contextually necessary.	
Gender, race, ethnicity, or age is included only when necessary for answering the question.	
Scoring Considerations	
The correct answer is clearly defensible and based on referenced knowledge.	
For multiple-response items (MCMR), clear scoring rules and guidance are established.	
Open-ended or constructed-response items have a well-defined rubric for scoring.	
References and Documentation	
All items include a rationale or supporting reference for the correct response.	
If graphical material or media is included, appropriate citations are provided.	