## Rules for Writing



## Multiple-Choice Questions



## Arizona's Technical Skills Assessments are designed to certify and document student

 attainment of industry-validated technical knowledge and skills. Delivered online, all assessments are 100 multiple-choice questions that align to program technical standards and range from knowledge and comprehension to application, analysis, and evaluation in cognitive difficulty. Most of the items are operational items that have historically proven to be valid and reliable and the remaining are field-tested items. Students are scored on the operational items not the field-tested items. Those who successfully complete the $60 \%$ pass score receive an industry-endorsed congratulatory letter and certificate.Teachers serve as content experts in the development of assessment items. Committees comprised of secondary teachers convene annually to develop, review, and edit assessment items. The expertise of teachers and their engagement in making connections between curriculum, instruction, and assessment has contributed greatly to the advancement of student achievement. Rules for Writing Multiple-Choice Questions was developed to provide Technical Skills Committees (teachers and facilitators) guidelines for developing good, consistent assessment items.

January 2020

For additional information, contact:
Judy A. Balogh, Team Lead
Technical Standards, Technical Skills Assessments, Career Development Collaboration
Arizona Department of Education / Career and Technical Education
Phone: 602.542.4155; email: judy.balogh@azed.gov

Cathy Reed, Program Specialist
Technical Standards, Technical Skills Assessments, Career Development Collaboration
Arizona Department of Education / Career and Technical Education
Phone: 602.364.0103; email: cathy.reed@azed.gov

# Statement of Professionalism for <br> <br> Technical Skills Assessment Development and Review 

 <br> <br> Technical Skills Assessment Development and Review}

- I understand that the work involved in establishing the Arizona Technical Skills Assessments is confidential and that the items must remain secure.
- I acknowledge my responsibility to follow the procedures set forth by the Arizona Department of Education, Career and Technical Education regarding the development of items and including the surrender of all materials associated with item development at the conclusion of meetings.
- I further understand that if I am found acting indiscreetly with confidential material, I will be dismissed from further involvement in the assessments. I understand this action to be necessary in order to maintain the integrity of the Arizona Technical Skills Assessments.
- I agree to engage with colleagues in developing an ongoing community-of-practice that exemplifies the highest standards of professionalism.

Printed Name of Participant
$\qquad$


CTE Program

Participant's Signature

Learning that works for Arizona
CTE

## General Rules for Writing Multiple-Choice Items

## The item:

1. Is based on a valid skill (measurement criterion).
2. Uses correct grammar, punctuation, and spelling.
3. Is not a trick item that would mislead or deceive students.
4. Uses either the best answer or the correct answer format.
5. Does not have textbook, verbatim phrasing.
6. Is based on important aspects of the content area.
7. Does not contain insensitive content or language.

The stem of the item:

1. Is stated as a question.
2. Is clearly worded and lets the student know exactly what is being asked.
3. Includes enough information to make the question clear.

Additional considerations:

1. If an item is stated negatively, the negative word is capitalized.
2. The phrase "all of the above" is never used.
3. Specific determiners such as "never" and "always" are not used

## Options (distractors):

1. Each item has four options.
2. Options are homogeneous in content and consistent in length.
3. Options are phrased positively, not negatively.
4. There are no clues in grammatical construction.
5. The phrase "none of the above" is never used because the options are randomized.

## Distractors:

1. Are plausible and logical.
2. Incorporate common errors of students.
3. Use familiar yet incorrect phrases.
4. May be true statements but do not correctly answer the item.

## The correct option:

1. Is randomized so that it appears about the same number of times in each possible position for a set of items.
2. Is the one and only correct, or clearly best, answer on which experts would agree.

## Grammar and Formatting Rules

## The stem (question)

- Ensure that the right question is being asked in its simplest form and that it is matched to the correct standard and measurement criterion.
- Improve the clarity and meaning of a question by using present or action conditions (what is; what are) rather than passive or progressive construction (what would be; what should be; what will be).
- Change "which of the following" to "which" or "what" plus a noun (e.g., term, issue, technique, statement, factor, action, purpose, step, etc.); eliminate "of the following" or rewrite the question.
- Capitalize words such as NOT, TRUE, USUALLY, COMMONLY, and PRIMARY for emphasis. Quotation marks and capped letters can be used; however, italics and symbols cannot be used.

The options (distractors)

- Match structure (parallel structure), content, plausibility, and length of distractors.
- Eliminate articles (a, an, and the) introducing options.
- Remove "and" between items listed in a series as an option when plausible.
- Use the logical flow in answering the question to determine whether sentences, phrases, or gerunds (a verb form ending in ing that can serve in place of a noun phrase) are needed.
- Eliminate same word used in options if the word is also used in the stem (e.g., What type of cells are used during tissue culture? a. Meristem eells; b. Leaf eells; c. Stem eells; d. Root eells
- Check for and change key words that appear in the question and are reused in distractors.


## Overall

- Remove unnecessary information/description in both the question and the options.
- Use "that" for a restrictive clause (information essential to the understanding) and "which" for a nonrestrictive clause (information nonessential to the understanding).
- Always use numbers with time, weight, and measurements. Spell out single-digit whole numbers and use numerals for numbers greater than nine. If a numeral is used because one of the numbers is greater than nine, use numerals for all numbers in the question and/or option.


## How to Write Higher-Order Thinking Items

A stem that presents a problem which requires application of principles, analysis of a problem, or evaluation of alternatives focuses on higher-order thinking. A case study that requires knowledge of more than one fact to logically and systematically apply concepts to a problem addresses higher-order thinking. And, item options that involve a high level of discrimination also contribute to higher-order thinking. Following are examples to help you write higher-order assessment items. (Note: The answer to these multiple-choice questions is not always "a.")

Memory-plus Application. Sally's breakfast this morning included one glass of concentrate orange juice, one slice of toast, a small bowl of bran cereal, and a grapefruit. What "whole food" did Sally eat for breakfast?
a. Orange juice
b. Toast
c. Bran cereal
d. Grapefruit

Ability to Interpret Cause-and-effect Relationships. Why does investing money in common stock protect against loss of assets during inflation?
a. It pays higher rates of interest during inflation.
b. It provides a steady but dependable income despite economic conditions.
c. It is protected by the Federal Reserve System.
d. It increases in value as the value of a business increases.

Ability to Justify Methods and Procedures. Technician A says that the basecoat/clearcoat system was designed to provide extra protection against ultraviolet rays and add gloss to the finish. Technician B says that the basecoat colors dry to a flat finish and requires a proper clear coat top coating to achieve correct gloss. Which technician is correct?
a. Both technicians
b. Neither technician
c. Technician B
d. Technician A

Ability to Evaluate Multiple Evidences to Solve a Problem. A nurse is making a home visit to a 75 -year-old man who has had Parkinson's disease for the past 5 years. Which finding has the greatest implication for this patient's care?
a. The client's wife tells the nurse that the grandchildren have not been to visit for over a month.
b. The nurse notes that there are numerous throw rugs throughout the client's home.
c. The client has a towel wrapped around his neck that the wife uses to wipe her husband's face.
d. The client is sitting in an armchair, and the nurse notes that he is gripping the arms of the chair.

Ability to Evaluate Multiple Evidences to Solve a Problem. Tim is in the second grade. He withdraws from his peers, often confuses syllables in words, and loses his place when reading. His teacher has arranged a meeting with his mother to discuss these concerns. Which statement is best for the teacher to say to Tim's mother?
a. Tim needs extra practice reading and writing problematic letters and words at home at least 30 minutes per day.
b. Please discuss the importance of schoolwork to Tim so that he will increase his efforts in classwork.
c. These are possible symptoms of dyslexia so I would like to refer him to a specialist for diagnosis.
d. Please adjust Tim's diet because he is most likely showing symptoms of ADHA due to food allergies.

Complex Thinking, Application of Knowledge, Integration of Material. Two researchers were studying the relationship between amount of sleep each night and calories burned on an exercise bike for 42 men and women. They were interested if people who slept more had more energy to use during their exercise session. They obtained a correlation of .28 , which has a two-tailed probability of .08 . Alpha was .10 . Which is an example of a properly written research question?
a. Is there a relationship between amount of sleep and energy expanded?
b. Does amount of sleep correlate with energy used?
c. What is the cause of energy expanded?
d. What is the value of rho?

## Bloom's Classification for Multiple-Choice Items

This two-dimensional framework focuses on knowledge and cognitive processes that define what students are expected to learn. Specifically, it explores curriculums from three perspectives: learning, curriculum and instruction, and assessment.

|  | The Cognitive Process Dimension |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| The Knowledge Dimension | 1. Remembering <br> Retrieve knowledge from long-term memory (recognizing, recalling) | 2. Understanding <br> Construct meaning from instructional messages using oral, written, and graphic communication (interpreting, exemplifying, classifying, summarizing, inferring, comparing) | 3. Applying <br> Carry out or use a procedure in a given situation (executing, implementing) | 4. Analyzing <br> Breakdown material into its constituent parts and determine how the parts relate to one another and to the overall structure or purpose (differentiating, organizing, attributing) | 5. Evaluating <br> Make judgments based on criteria and standards (checking, critiquing) | 6.-Creatiny <br> Rut alamentstogethar to <br> farm a coherant or <br> functional whola; ruarganiza-alementcinte a new-pattacn-ar structure/ganarating planning, producingt |
| A. Factual Knowledge <br> The basic elements students must know to be able to solve problems, i.e., terminology, specific details, and elements. | List | Summarize | Respond | Select | Check |  |
| B. Conceptual Knowledge <br> The interrelationships among basic elements within a larger structure that enable them to function together, i.e., classifications and categories; principles and generalizations; theories, models, and structures. | Recognize | Classify | Provide | Differentiate | Determine |  |
| c. Procedural Knowledge <br> How to do something, methods of inquiry, and criteria for using skills, algorithms, techniques, and methods, i.e., subject-specific skills and algorithms; subject-specific techniques and methods; and criteria for determining when to use appropriate procedures. | Recall | Clarify | Carry out | Integrate | Judge |  |
| D. Metacegnitive Kinewleolye <br> Knowladga-of cognition in general at-well at awarane 4 and knowladge-ofone's-own cognition, i.e., stratagic knowledye; contaxtual and conditional knowlodga; and solf-knowledga. |  |  |  |  |  |  |
| KEY: | YELLOW <br> Necessary foundational k | CAUTION <br> wiedge and understandings | Target area for item d | FOR THE GOLD <br> ment/summative assessments | GRAY -O <br> Exceeds limitation of multi | UT OF BOUNDS <br> ple-choice assessment items |

Note: The verb in each block is provided as an aid to help individualize and differentiate each classification, i.e., A1 vs. B 1 vs . C 1 , and so forth.
Adapted from: A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives

## Examples of Verbs for the Cognitive Process Dimensions

| 1. Remembering | 2. Understanding | 3. Applying | 4. Analyzing | 5. Evaluating | 6. Creating |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Arrange <br> Define <br> Describe <br> Draw <br> Duplicate <br> Identify <br> Label <br> List <br> Locate <br> Match <br> Memorize <br> Name <br> Order <br> Recall <br> Recognize <br> Relate <br> Reproduce <br> Select <br> State | Characterize <br> Classify <br> Compare <br> Complete <br> Convert <br> Demonstrate <br> Describe <br> Discuss <br> Distinguish <br> Establish <br> Execute <br> Explain <br> Generate <br> Interpret <br> Outline <br> Predict <br> Report <br> Summarize <br> Translate | Administer <br> Apply <br> Change <br> Choose <br> Complete <br> Construct <br> Demonstrate <br> Determine <br> Employ <br> Examine <br> Illustrate <br> Interpret <br> Manipulate <br> Modify <br> Operate <br> Perform <br> Practice <br> Produce <br> Solve <br> Use | Achieve <br> Advertise <br> Analyze <br> Appraise <br> Categorize <br> Classify <br> Compare <br> Contrast <br> Critique <br> Debate <br> Deduct <br> Determine <br> Diagnose <br> Differentiate <br> Examine <br> Explain <br> Identify <br> Investigate <br> Research <br> Specify | Assess <br> Compare <br> Conclude <br> Critique <br> Debate <br> Decide <br> Defend <br> Determine <br> Estimate <br> Evaluate <br> Examine <br> Judge <br> Justify <br> Prioritize <br> Recommend <br> Review <br> Select <br> Support <br> Verify <br> Weigh | Actualize <br> Change <br> Combine <br> Compose <br> Construct <br> Create <br> Design <br> Devise <br> Forecast <br> Formulate <br> Generate <br> Hypothesize <br> Imagine <br> Invent <br> Originate <br> Plan <br> Predict <br> Pretend <br> Produce <br> Propose |

## Item Analysis

Item analysis is the process of examining the performance of individual test items...or more simply put, identifying test items that are not working well. Item analysis consists of three types of statistics or information that we use to improve an item. This activity leads to item reliability (how consistently a measurement of skill or knowledge yields similar results under varying conditions) and validity (the degree to which the instrument measures what it's supposed to measure over time):

1. Difficulty Index. The difficulty index is determined by calculating the number of students who got an item correct ( $p$ value). A large number indicates that many students have learned the content as measured by the item. When none of the students chooses the correct answer, the $p$ value is 0 . When all students choose the correct answer, the $p$ value is 1.0. (Note. The name of this index is counter-intuitive in that there is a measure of how easy the item is not a measure of how difficult the item is.)
2. Discrimination Index. The discrimination index measures the validity of an item. It shows the discrimination between those students who know the content and those who do not know the content. This index is an indication of the extent to which overall knowledge of the content area or mastery of the skills is related to the response on an item. Perhaps the most crucial validity standard for a test item is whether a student got an item correct due to his level of knowledge or ability and not due to chance or test bias. A discrimination value of .3 and above indicates that the item accurately discriminates between those who know and those who do not know the content. High values indicate good items; low indicate poor items.

## NOTE: An item is flagged when it falls below these parameters:

Difficulty Index of .45 with Discrimination Index of 3 OR Difficulty Index of .5 with Discrimination Index of .2
3. Distractor (option) Analysis. By calculating the proportion of students who chose each option, we can identify which distractors are "working" and which distractors are simply taking up space and not being chosen by many students. To eliminate guessing that results in a correct answer purely by chance (which hurts the validity of a test item), we want plausible distractors. Analyses of response options allow us to fine tune and improve items.

## Helpful Hints to Understanding Item Analysis

1. If an item is very easy so that nearly all students answered correctly, the discrimination will be near zero. Extremely easy questions cannot distinguish among students in terms of their performance.
2. If an item is extremely difficult so that nearly all students answered incorrectly, the discrimination will be near zero.
3. The most effective items will have moderate difficulty and high discrimination values. The higher the value of discrimination, the more likely it will be to distinguish between those students who perform well on the test and those who don't.
4. Items with low or negative values of discrimination need to be reviewed for confusing language or an incorrect indicator.
5. A high level of student guessing will result in a discrimination value near zero.

## Worksheet for Multiple-Choice Items



## Check Your Work (or someone else's work)

|  | Yes |
| :--- | :--- |
| ABOUT THE ITEM REFERENCE: | No |
| Are the Standard, Measurement Criterion, and Bloom's Classification identified? |  |
| ABOUT THE ITEM STEM: |  |
| Is the stem stated as a question? |  |
| Is the stem clearly worded without unnecessary words? |  |
| Is the stem without jargon or brand names? |  |
| Is the stem grammatically correct? |  |
| Is the question based on important aspects of the standard/measurement criterion? |  |
| Are words such as NOT, TRUE, PRIMARILY, and BEST capitalized? |  |
| ABOUT THE ITEM OPTIONS (distractors): |  |
| Are there four options? |  |
| Are the options consistent in length (parallel in structure)? |  |
| If the options are complete sentences, are they followed by a period? |  |
| Are the options grammatically correct? |  |
| Is the best or the correct answer identified? |  |

## Instructions for Committee Work

## TASK: DEVELOP ITEMS USING THE BLUEPRINT FOR INSTRUCTION AND ASSESSMENT AND THE INSTRUCTIONAL FRAMEWORK

1. Review these documents before developing new items:

- Industry-validated Technical Standards and Measurement Criteria. This document identifies the knowledge and skills that apply to the practices and processes used in design, manufacture, installation, and/or engagement of a material, product, or assembly. The standards and measurement criteria represent the minimum of what teachers should teach to prepare students for entry-level employment and further training and education.
- Blueprint for Instruction and Assessment. This document illustrates the relationship among program standards, instructional time, and student success on the Technical Skills Assessments. Specifically, blueprints identify clusters of standards in content domains and recommend a percentage of instructional time for each domain which also aligns with the degree to which the content of the standards is tested.
- Instructional Framework. This document explains and expands the content of the measurement criteria. This helps teachers know what they should teach and, as well, guides the multiple-choice items to be written for the Technical Skills Assessments.

2. Review the following guidelines for developing items:

- General Rules for Writing Multiple-Choice Items
- Grammar and Formatting Rules
- How to Write Higher-Order Thinking Items
- Bloom's Classification for Multiple-Choice Items

3. Develop multiple-choice items

- Refer to the Item Bank, Item Status, and Blueprint to determine the need for items.
- Develop items using the Worksheet for Multiple-Choice Items and the Check Your Work document.

The facilitator should collect all items to ensure the integrity of the Technical Skills Assessments.

## Instructions for Committee Work

## TASK: ALIGN ASSESSMENT ITEMS TO NEW/UPDATED STANDARDS AND MEASUREMENT CRITERIA

1. Review the content of the new/updated Standards and Measurement Criteria (alignment document) and the content and formatting of the items in the Item Bank.

- Total number of Items and the breakdown of items by New, Edited, Active, and Operational is shown at the top of the Item Bank.
- Standards: Bold and CAPS
- Measurement criteria: Bold and Upper/Lower Case
- Coding breakdown (example: 43.1.2.14): First number is the program number in the assessment database. You can disregard this number for the alignment task. Second number is the standard. Third number is the measurement criterion. Fourth number is the item number. If an item is missing in the numbering sequence, this means that the item is retired.
- Item Stem: Bold and always in question format
- Options: Four options with answer bold and in "a" position
- Difficulty Index (only shown on the Item Analysis report): Number of students who answered the item correctly (refer to the Item Analysis document for further explanation)
- Discrimination Index (only shown on the Item Analysis report): Degree of difference between those students who know the content and those students who do not know the content (refer to the Item Analysis document for further explanation)
- Item Status and Blooms ID: Refer to Bloom's Classification for Multiple-Choice Items document for an explanation of item classification.

2. The alignment process involves determining the status of each item:

- This is a good item and aligns with a new standard/measurement criterion. (Identify the alignment.)
- With changes, this item aligns with a new standard/measurement criterion. (Identify the alignment.)
- This item is not relevant and does not align with a new standard/measurement criterion. The item should be retired.
- Continue this task until all items have been aligned, marked for modification, and/or retired.

3. All results are recorded by the facilitator on the Master Copies:

- On the Master Copy of the Standards and Measurement Criteria (alignment document) in the column to the left of the measurement criterion, enter the standard number, measurement criterion number, and item number.

| Items from Item Bank | STANDARD 1.0 APPLY SANITATION PROCEDURES |  |
| :--- | :--- | :--- |
| $1.1 .2,1.1 .3$ | 1.1 | Define the concept of HACCP (Hazard Analysis Critical Control Point) |
| 1.6 .3 | 1.2 | Identify major reasons for and recognize signs of food spoilage and contamination |
| $1.3 .2,1.3 .5$ | 1.3 | Identify the most common foodborne illnesses |

- On the Master Copy of the Item Bank to the left of the item, write the number of the new standard and measurement criterion the item aligns with.
- On the Master Copy of the Item Bank, circle any item that needs to be edited. These items can be edited now or later in the space by the item or on the Worksheet for Multiple-Choice Items.
- On the Master Copy of the Item Bank, draw a diagonal line through items that need to be retired.

The facilitator should collect all items to ensure the integrity of the Technical Skills Assessments.

## Instructions for Committee Work

## TASK: REVIEW THE ITEM BANK AND EDIT ITEMS OR DEVELOP NEW ITEMS

1. Review the content and formatting of the Item Bank, Item Status, and the Blueprint for Instruction and Assessment and the Instructional Framework documents.

- Total number of Items and the breakdown of items by New, Edited, Active, and Operational is shown at the top of the Item Bank.
- Standards: Bold and CAPS
- Measurement criteria: Bold and Upper/Lower Case
- Coding breakdown (example: 43.1.2.14): First number is the program number in the assessment database. You can disregard this number for the alignment task. Second number is the standard. Third number is the measurement criterion. Fourth number is the item number. If an item is missing in the numbering sequence, this means that the item is retired.
- Item Stem: Bold and always in question format
- Options: Four options with answer bold and in "a" position
- Difficulty Index (only shown on the Item Analysis report): Number of students who answered the item correctly (refer to the Item Analysis document for further explanation)
- Discrimination Index (only shown on the Item Analysis report): Degree of difference between those students who know the content and those students who do not know the content (refer to the Item Analysis document for further explanation)
- Item Status and Blooms ID: Refer to Bloom's Classification for Multiple-Choice Items document for an explanation of item classification.

2. Review each item for grammar, formatting, clarity, and correctness. Although "not" items have been allowed in the past, consider rewriting these items and removing the negative meaning. Refer to General Rules for Writing Multiple-Choice Items, Grammar and Formatting Rules, and How to Write Higher-Order Thinking Items, and Bloom's Classification for Multiple-Choice Items documents for guidelines about how to write good items. Items should be edited and/or developed on the Master Copy of the Item Report or on the Worksheet for Multiple-Choice Items.
3. Refer to the Item Bank, Item Status, and Blueprint for Instruction and Assessment documents to determine where items are needed.

NOTE to Facilitator: Change item status when appropriate.
All flagged NEW, EDITED, and ACTIVE items that are edited will become or remain EDITED items.
All flagged (and unflagged) EDITED items that are not edited and not retired will become ACTIVE items.
All flagged and unflagged ACTIVE items that are not edited or not retired will become OPERATIONAL items.

The facilitator should collect all items to ensure the integrity of the Technical Skills Assessments.

## Instructions for Committee Work

## TASK: REVIEW THE ITEM ANALYSIS RESULTS AND EDIT, LEAVE, OR RETIRE ITEMS

1. Review the content and formatting of the Item Analysis report. Refer to the Assessment Item Analysis document to better understand item analysis. Note. This report contains all 100 items on the past fall and spring test.

- Total number of Items and the breakdown of items by New, Edited, Active, and Operational is shown at the top of the Item Bank.
- Standards: Bold and CAPS
- Measurement criteria: Bold and Upper/Lower Case
- Coding breakdown (example: 43.1.2.14): First number is the program number in the assessment database. You can disregard this number for the alignment task. Second number is the standard. Third number is the measurement criterion. Fourth number is the item number. If an item is missing in the numbering sequence, this means that the item is retired.
- Item Stem: Bold and always in question format
- Options: Four options with answer bold and in "a" position
- Difficulty Index (only shown on the Item Analysis report): Number of students who answered the item correctly (refer to the Item Analysis document for further explanation)
- Discrimination Index (only shown on the Item Analysis report): Degree of difference between those students who know the content and those students who do not know the content (refer to the Item Analysis document for further explanation)
- Item Status and Blooms ID: Refer to Bloom's Classification for Multiple-Choice Items document for an explanation of item classification.

2. Review each of the "flagged items" and decide to leave, edit, or retire the item. Many factors can contribute to students getting an item incorrect. Some of the most obvious are:

- The question is ambiguous.
- The question has no or more than one correct answer.
- The content was not adequately taught.
- Students were unaware the content was important.
- The question contained cues that misdirected students.
- Students missed an essential element of the question.
- Students were confused by a question's logic (e.g., "not" question).

3. On the Master Copy of the Item Analysis report to the left of the item, indicate the committee's decision to leave or retire items. Edit items on this document or use the Worksheet for Multiple-Choice Items to rewrite an item or to develop a replacement item. Refer to General Rules for Writing Multiple-Choice Items, Grammar and Formatting Rules, and How to Write Higher-Order Thinking Items, and Bloom's Classification for Multiple-Choice Items about how to write new items.

## NOTE to Facilitator: Change Item status when appropriate.

All flagged NEW, EDITED, and ACTIVE items that are edited will become or remain EDITED items.
All flagged (and unflagged) EDITED items that are not edited and not retired will become ACTIVE items.
All flagged and unflagged ACTIVE items that are not edited or not retired will become OPERATIONAL items.

The facilitator should collect all items to ensure the integrity of the Technical Skills Assessments.

NOTES

NOTES

