

Guidelines for Multiple Choice

- Capture essence of item in a single declarative sentence (proposition)
- Step back, reexamine, do the questions focus on what you really want?
- Ask a complete sentence to get the item started
- If you find yourself repeating the same words at the beginning of each response option, reword the stem to move the repetitive material up there
- Be sure there is only one correct or best answer; it is okay to ask students to select a best answer from among a set of answers, all of which are correct
- Word response options as briefly as possible and be sure they are grammatically parallel, not this Why did colonists come to the United States? To escape heavy taxation by their native governments; religion; They sought the adventure of living among Native Americans in the new land; there was the promise of great wealth in the new world; more than one of the above answers instead Why did colonists migrate to the United States? To escape taxation, for religious freedom; for adventure, more than one of the above
- Vary the number of response options presented as appropriate to pose the problem you want your student to solve; try to avoid all of the above, none of the above to fill up space, limit their use; if you want more than one correct response you may want to make the question worth more for example: Which of the labels provided represents a classification category for types of rocks (Identify all correct answers (*)) geologic, metamorphic, sandstone, igneous*
- If you compose a multiple choice item and find that you cannot think of enough plausible incorrect responses, include the item on a test the first time as a fill-in question; as your students respond, those who get it wrong will provide you with the full range of viable incorrect responses you need the next time you use it!

Guidelines for Matching

- State the matching challenge up front with a clear and concise set of directions, specifying what is to be matched
- Keep the list of things to be matched short (maximum of 10)
- Keep the list of things to be matched homogenous
- Keep the list of response options brief in their wording and parallel in construction
- Include more response options than stems and permit students to use response options more than once

Guidelines True/False

- State the matching challenge up front with a clear and concise set of directions
- Keep the list of things to be matched short, maximum number of options is 10
- Keep the list of things to be matched homogenous
- Keep the list of response options brief in their wording and parallel in construction; pose the matching challenge in clear, straightforward language
- Include more response options than stems and permit students to use response options more than once; makes it impossible for students to arrive at the correct response through elimination

- Comparative reasoning: The executive and legislative branches of the U.S. government differ in that the latter is elected directly by the people/Members of executive and legislative branches are both elected directly by the people (true/false)
- Fill in: Election of member of the executive and legislative branches differs in what way?
- Multiple choice: Election of members of the executive and legislative branches differs in what way? Legislators are restricted by term limits; presidents are not; legislators are elected directly; presidents are not; one must register to vote for legislators; not for president

Writing Guidelines

- Writing guidelines
 - Write clearly in a sharply focused manner: Include only material essential to framing the question; be brief and clear
 - Not this: When scientists rely on magnets in the development of electric motors they need to know about the poles, which are? Instead try this: What are the poles of a magnet called: anode and cathode, north and south, strong and weak, attract and repel
 - Ask a question: minimize the use of incomplete statement; not this: between 1950 and 1965 Interest rates increased, interest rates decreased, interest rates fluctuated greatly, interest rates did not change; instead: What was the trend in interest rates between 1950 and 1965: increased only; decreased only, increased and then decreased, remained unchanged
 - Aim for the lowest possible reading level; you do not want to let students' reading proficiency prevent them from demonstrating that they really know the materials; minimize sentence length and syntactic complexity, eliminate unnecessarily difficult or unfamiliar vocabulary
 - Eliminate clues to the correct answer either within the question or across the question; not this: All of these are examples of a bird that flies, except an ostrich, falcon, cormorant, robin (the article an at the end requires a response beginning with a vowel, as only one is offered it must be correct) not this either Which of the following are examples of birds that do not fly? Falcon, ostrich and penguin, cormorant, robin (The question calls for a plural response, only one is offered, it must be correct)
 - If you write the test questions, have a qualified colleague read them through at least once to ensure accuracy
 - Double check the scoring key for accuracy before scoring
 - Kids can make propositions; kids could write questions to propositions; ask a question to get the proposition;
 - State objective, ask kids to come up with proposition (synthesize a variety of answers)

- When we ask students to dip into their knowledge and understanding and to use what they know in novel ways to figure something out, we are asking for reasoning
 - How things are alike, different
 - How something can be subdivided into its component parts or how the parts work together
 - The main idea or theme of a story just read
 - The proper conclusion to be drawn from a science experiment
 - Insights that can be derived from a provided data chart
- We can test student mastery of content knowledge, including what they learn outright and what they learn to retrieve through effective use of reference materials; can also tap a variety of kinds of reasoning and problem solving; including analytical, comparative, and other kinds of inferential reasoning proficiencies

See to proposition

Steps in assessment

- Steps in assessment development
 - ③ • Develop an assessment plan or blueprint that identifies an appropriate sample of achievement
 - Identify the specific elements of knowledge, understanding, and reasoning to be assessed
 - Transform those elements into assessment exercises or test items
- Preparing a blueprint
 - Opportunity for teachers and students to clarify achievement expectations
 - Questions require respondents to do two things: gain access to a specific piece of information, use that knowledge to carry out a cognitive operation
 - Look at table (represents three different kinds of cognitive actions: demonstrate understanding of the content, reason comparatively using elements of the content; use knowledge and reason to classify elements of content)
 - Difference in number of items assigned to each category might reflect amount of instructional time spent on each
 - Amount of materials in each category
 - Relative importance of material in each category for later learning
 - Important relationships among various ideas
- Test and instruction aligned, then students need to have been provided with opportunities to master the essential content and practice with the valued patterns of reasoning
 - When developing the table you can have as many cells as you want; consider the following
 - Natural subdivisions in the materials, such as chapters subsections, etc.
 - Be sure categories have clearly marked limits and are large enough to contain a number of important elements of knowledge within each
 - The categories of content, kinds of reasoning tested, and proportion of items assigned to each should reflect the target priorities communicated to students during instruction, Students can hit any target they can see and that holds still for them!!
 - Can also build your assessment list from a list of instructional objectives: for ex Students will be able to compare and contrast different forms of govt; students will understand a citizen's voting rights and responsibilities

Examples of Common Sources of Bias in Classroom Assessment

1. Problems common to all methods:

A. Potential problems that can occur within the student:

Lack of reading proficiency
Language barriers
Emotional upset
Poor health
Physical handicap
Peer pressure to mislead assessor
Lack of motivation at time of assessment
Lack of testwiseness (understanding how to take tests)
Lack of personal confidence leading to evaluation anxiety

B. Possible problems that can occur within the assessment context:

Noise distractions
Poor lighting
Discomfort
Lack of rapport with assessor
Cultural insensitivity in assessor or assessment
Lack of proper equipment

C. Examples of problems that arise from the assessment itself (regardless of method)

Directions missing or vague
Poorly worded questions
Poor reproduction reduces readability

2. Problems unique to each format:

A. Possible problems with multiple choice tests:

More than one correct response
Incorrect scoring key
Incorrect bubbling on answer sheet
Clues to the answer in the item or in other items

B. Problems with essay assessments

Students lack writing skill
No scoring criteria
Inappropriate scoring criteria
Evaluator untrained in applying scoring criteria
Bias due to stereotypic thinking or knowledge of prior performance
Insufficient time or patience to read and score carefully

C. Potential problems with performance assessment

No scoring criteria
Inappropriate scoring criteria
Evaluator untrained in applying scoring criteria
Bias due to stereotypic thinking or knowledge of prior performance
Insufficient time or patience to observe and score carefully

D. Possible difficulties when using personal communication

Insufficient sample per student
Inaccurate record keeping
Distortions in memory of performance
Bias due to stereotypic thinking or knowledge of prior performance

COMPARISON OF PAPER AND PENCIL ITEM TYPES

Item Type	Used When	Advantage	Limitations
Multiple Choice	There is only one right answer. There are several plausible alternatives to the correct answer.	Can measure a variety of objectives. Easy to score. Can cover lots of material efficiently.	Can't measure extended thinking or performance. Not good for assessing the process by which answers are obtained.
True/False	A large domain of content is to be tested, requiring the use of many test items.	Can ask many questions in a short time. Easy to score.	Can't measure extended thinking or performance Not good for assessing the process by which answers are obtained. Can be trivial or misleading if not written carefully.
Matching	There are many related thoughts or facts; you want to measure association of information.	Can cover lots of material efficiently. Easy to score.	Can't measure extended thinking. Assess identification of an answer rather than production.
Completion	A clear, short answer is required.	Assessing production of a response. Reduces the possibility of guessing. Can cover lots of material efficiently.	Can't measure extended thinking or performance. Takes longer to score.
Essay/Written	Can measure extended thinking. Can assess student ability to organize thoughts and compose an appropriate response.	Assesses production of a response, not just identification of a response. Chance plays little or no part in adequate responding. Can be used to assess the process by which an answer is obtained.	Harder to score reliably. Harder to come up with criteria for scoring. Provides a limited sample of knowledge areas. Inability to write can interfere with ability to show understanding.

GUIDELINES FOR WRITING SOUND PAPER AND PENCIL TEST ITEMS

Consult *Student-Centered Classroom Assessment, 2nd Ed.*, Chapter 8 & 9 for examples and greater detail on guidelines.

General Guidelines (apply to all item formats)

1. Keep wording simple and focused:
 - a. Follow rules of grammar
 - b. Eliminate superfluous material
 - c. Come to the point
 - d. Test mastery of material, not ability to figure out what you're asking
 - e. Aim for lowest possible reading level
2. Ask a question (minimize use of incomplete statements).
3. Avoid providing clues within and between items:
 - a. Avoid use of specific determiners such as "always" and "never"
 - b. Avoid grammatical clues (e.g., "a" and "an", present and past tense, singular and plural, etc.)
 - c. Avoid information clues (e.g., the stem of one item gives away the answer to another)
4. The correct answer should not be obvious to those who have not mastered the materials tested.
5. Highlight critical words such as MOST, LEAST, EXCEPT, and NOT, because they are easily overlooked.

Multiple Choice

1. State the whole question in the test item stem.
2. Keep responses brief and parallel in:
 - a. Length
 - b. Use of specific determiners
 - c. Grammatical construction
 - d. Level of generality
3. Avoid repetition of material in each of the response options.
4. Limit use of "all or none of the above".

5. Be sure there is only one correct or best answer (best answer items should be worded to ask for BEST answer).
6. OK to vary number of response options across items within the same test.

True/False

Make them entirely true or false as stated.

Matching

1. Include only homogeneous items. Do not mix dates, events, names, etc. in a single exercise.
2. Maximum length is 10, shorter is better.
3. Provide more responses than there are things to be matched.
4. Provide directions for the match to be made. Indicate if a response can be used more than once or if an item has more than one match.

Completion or fill-in

1. One blank per item (put blank toward end).
2. Length of blank should not be a clue.

Essay

Exercises

1. Avoid general, all encompassing questions.
2. Goal: test structure of knowledge and/or thinking.
3. Point direction to appropriate response in the essay exercise.
4. Give points and/or time allocation.

Scoring

5. Outline acceptable response in advance.
6. Set policy regarding non-achievement factors, i.e., writing.
7. Score in the blind, if possible.

8. Score all responses to one exercise at a time (faster!).
9. Score holistically or analytically, but with criteria clearly established.

Formatting Test Items

1. Be consistent in the presentation of an item-type.
2. Keep all parts of a test question on one page.
3. Avoid crowding too many questions on one page.

Arranging Test Items

1. Group the questions on the test by item-type--that is, put all completion questions together, all multiple choice, all true/false, and so on.
2. Begin with some very easy questions to give students a chance to "break into" the test.

Writing Directions

1. Write clear, explicit directions for each type of item.
2. State the point value of each item-type.
3. Indicate how the answer should be expressed by the student. For example, should the word TRUE or FALSE be written or simply T or F? Should numbers be rounded to the nearest tenth? Should units, such as months, meters, or grams be included in the answer?

Producing Tests

1. Avoid writing tests on the chalkboard.
2. Type tests for duplication--unless your handwriting is very clear and neat.
3. Proofread the test carefully.
4. Duplicate clear, readable copies.
5. Ask a colleague to review or to take important tests.