Writing _____ Multiple-Choice Test Items

(a) Lots of
(b) Tricky
(c) Effective
(d) Fun
(e) Hard

Learning Goals

The purpose of this session is to help you improve the validity and reliability of your multiple-choice test items. After attending, you should be better able to

1. List several properties of good test items
2. Draft a test item that requires application or evaluation
3. Recognize verbal clues used by test-wise students

Chris Clark, Assistant Director
Kaneb Center for Teaching and Learning
University of Notre Dame
September 17, 2010
Writing Effective Multiple-Choice Test Items

Terminology

• **Stem** – a short text that explains the problem and gives relevant information. It often ends with a question.

• **Options** – possible answers, from which one chooses either as many as requested

• **Distracter** – incorrect option

• **Valid** – measures what it purports to measure or gives results that accurately reflect the concept being measured.

• **Reliable** – provides consistent results. A valid measure must be reliable, but a reliable measure need not be valid

Course Design Oversimplified

1. **Learning Goals** – what students should know and be able to do as a result of the course
2. **Assessment** – how you are going to know if students achieve the learning goals
3. **Activities** – what will happen in classroom and outside

Cognitive Levels

Different levels of knowledge can be tested. Here is one way to categorize them:

• **Recall** – remember specific facts, terminology, principles, or theories (e.g., recollect the characteristics of Impressionist painting).

• **Application** – use knowledge to solve a new problem or analyze a new situation (e.g., apply ecological formulas to determine the outcome of a predator-prey relationship.

• **Evaluation** – derive hypotheses from data or exercise informed judgment (e.g., determine which economic policy a country should pursue to reduce inflation).
EXERCISE: Create a Multiple-Choice Test Item about Notre Dame

STEM

CORRECT ANSWER

DISTRACTORS
1. 

2. 

3. 
Writing Effective Multiple-Choice Test Items

Well-Designed Tests

- Show whether students have achieved the desired learning goals
- Sample material broadly
- Discriminate between more- and less knowledgeable students
- Neither reward "test wise" students nor penalize those with less test-taking skill
- Do not require students to sort out irrelevant complexities

Limitations of Multiple-Choice Testing

- Open to misinterpretation
- Depend on student's reading skills and instructor's writing ability
- Difficult to construct, requiring time and skill
- Instructors often construct fewer test items requiring high cognitive levels because items requiring recall are easier to construct.
- Test-wise students with incomplete knowledge use verbal cues to answer weak questions

Watch out for Verbal Clues!

These give test-wise students a correct answer or eliminate a distracter, reducing validity. Students should not be able to guess the answer from the way the options are written.

- An error in grammar or spelling (distracter)
- One option is longer, more detailed, or more complex (correct)
- Key word appears in the stem and only one option (correct)
- One option in textbook/lecture language (correct), others in everyday language (distracters)
- Two options with the same meaning (distracters)
- Specific determiners: absolute or extreme terms like "always," "never," or "all" (distracter)
- One option with a vague word or phrase like "usually," "typically" or "may be" (correct)
- An option that is clearly implausible or humorous (distracter)
EXERCISE: Test of General Knowledge

1. What led to the formation of the States' Rights Party?
   A. The level of federal taxation
   B. The demand of states for the right to make their own laws
   C. The industrialization of the South
   D. The corruption of federal legislators on the issue of state taxation

2. Who was Charlemagne?
   A. He was a pope in the 10th century
   B. He was a king of the Franks in the 8th century.
   C. He led the Catholic church in the 900s

3. Quantum mechanical tunneling is
   A. What they did at "The Big Dig"
   B. The name of President Bush's dog
   C. The effect of transitioning through a classically-forbidden energy state.

4. The coefficient of correlation found by correlating students' scores on a classroom social studies test with their scores on a standardized social studies test is called a
   A. Index of reliability
   B. Equivalence coefficient
   C. Internal consistency coefficient
   D. Validity coefficient

5. In Spanish, nouns that end in the letter "a" are
   A. Always feminine
   B. Usually feminine
   C. Always masculine

6. In economics, what is elasticity?
   A. It says that transportation costs go up and down along with people's spending habits.
   B. It's the ratio of the proportional change in one variable with respect to proportional change in another variable.
   C. It tells us that when people start to buy a lot of something, businesses will make sure they don't run out of it.
Good Stems

- Involve a single, clearly stated problem. Questions are best, but an incomplete statement may avoid awkward phrasing or convoluted language.
- Are as brief and concise as possible, avoiding undue complexity. Long stems add to reading time, limiting the potential number of items and reducing reliability.
- Do not repeat the same words at the beginning of each option.
- Do not contain extraneous or superfluous details.
- Are stated in positive form. Never put negative options after a negative stem.
- Are understandable without having to be read several times or reading all of the options.
- Do not use phrases lifted directly from text or lecture, thereby creating a simple recall activity. Change the language but use words that are familiar.
- Emphasize crucial words using underlining or capital letters. If you present two similarly worded test item about different things, highlight the differences.

Good Sets of Options

- Are all about the same length
- Follow the grammatical structure of the stem.
- Are listed vertically on separate lines.
- Are 3-5 in number. Don't add weak distracters just to keep a consistent number.
- Randomize the position of the correct answer. Using alphabetical order may do the job.
- Are arranged in logical order, where appropriate – numerically, chronologically, etc.
- Avoid microscopically fine distinctions, unless the ability to make such distinctions is one of your learning goals.
- Avoid "All of the above" – eliminating one distracter immediately eliminates this, too.
- Avoid "Both A & B" – makes it easier to guess the correct answer with partial knowledge.
- Avoid "None of the above" unless the answer requires computation. If you do use it, make it the correct answer to a reasonable number of times.

Good Distracters

- Are incorrect but plausible.
- May be based common errors, helping diagnose student difficulties.
- Use words that are familiar to students.
- May include extraneous information (unlike good stems)
- May be a jargon-ridden statement that is meaningless if one understands the concept.
- If you can't think of at least two good ones, toss the item.
EXERCISE: Rewrite this Stem

"According to Tuckman's model, groups develop through several stages over time. Furthermore, it contradicts Poole's activity-track model which has groups switching among several different linear sequences. Which of the following is not one of the stages identified in Tuckman's model?"
Test items that Involve Higher Cognitive Levels

Illustration – present a diagram, chart, table or figure and ask for interpretation, application, analysis, or evaluation.

Analogy – map the relationship between two items into a different context:

E-mail is to an unmoderated listserv as office hours are to:

a) Class lecture.
b) Class discussion.
c) Review sessions.
d) Tutorials.

Incomplete Scenario – respond to what is missing or needs to be changed. When using a graph or image, lay it out in a different way than students have seen.

What belongs in the empty box in the upper right corner of the diagram?

a) Hardware devices
b) Client Services for Netware
c) Logon Process
d) Gateway Services for Netware

Case Study – a single paragraph with several follow-up questions.

1. Alice, Barbara, and Charles own a small business: the Chock-Full-o-Goodness Cookie Company. Because Charles has many outside commitments and Barbara has a few, Alice tends to be most in touch with the daily operations of Chock-Full-o-Goodness. As a result, when financial decisions come down to a vote at their monthly meeting, they have decided that Alice gets 8 votes, Barbara gets 7, and Charles gets 2-with 9 being required to make the decision. According to minimum-resource coalition theory, who is most likely to be courted for their vote?

   a) Alice
   b) Barbara
   c) Charles
   d) No trend toward any specific person.

2. ...
Quotation – present a contrived quote or a real quote from a newspaper or other published source. Ask for interpretation or evaluation of these quotations.

Premise-Consequence – students identify the correct outcome of a given circumstance.

If nominal gross national product (GNP) increases at a rate of 10% per year and the GNP deflator increases at 8% per year, then real GNP:

a) Remains constant.
b) Rises by 10%.
c) Falls by 8%.
d) Rises by 2%.

Evaluation – describe a situation and pose a problem (perhaps along with a solution) where students must use informed judgment and critical thinking to evaluate and answer correctly.

A student was asked the following question: "Briefly list and explain the various stages of the creative process." As an answer, this student wrote the following:

The creative process is believed to take place in five stages, in the following order: orientation, when the problem must be identified and defined, preparation, when all the possible information about the problem is collected, incubation, when no solution seems in sight and the person is often busy with other tasks, illumination, when the person experiences a general idea of how to arrive at a solution to the problem, and finally verification, when the person determines whether the solution is the right one for the problem.

How would you judge this student’s answer?

a) EXCELLENT (all stages correct in the right order with clear and correct explanations)
b) GOOD (all stages correct in the right order, but the explanations are not as clear as they should be)
c) MEDIOCRE (one or two stages are missing OR the stages are in the wrong order, OR the explanations are not clear OR the explanations are irrelevant)
d) UNACCEPTABLE (more than two stages are missing AND the order is incorrect AND the explanations are not clear AND/OR they are irrelevant)
Memory-Plus Application Example  

Sally’s breakfast this morning included one glass of orange juice (from concentrate), 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

Memory-Plus Application Example  

Which one of the following best illustrates the law of diminishing returns?  
a. The demand for a farm product increased faster than the supply of the product.  
b. The population of a country increased faster than the means of subsistence.  
c. A machine decreased in utility as its parts became worn.  
d. A factory doubled its labor force and increased production by 50 percent.

Ability to Interpret Cause-and-Effect Relationships Example  

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 Example  

Why is adequate lighting necessary in a balanced aquarium?  
a. Fish need light to see their food.  
b. Fish take in oxygen in the dark.  
c. Plants expel carbon dioxide in the dark.  
d. Plants grow too rapidly in the dark.
Developing Complex Application Questions

Devising multiple choice questions that measure higher-level cognitive skills will enable you to test such skills in large classes without spending enormous amounts of time grading. In an evaluation question, a situation is described in a short paragraph and then a problem is posed as the stem of the question. All the rules for writing multiple choice items described above also apply to writing evaluation questions, but students must use judgment and critical thinking to answer them correctly. In the example below (adapted from Welsh, 1978), students must understand the concepts of price inflation, aggregate private demand, and tight monetary policy. They must also be able to analyze the information presented and, based on projected effects, choose the most appropriate policy. This question requires critical thinking and the complex application of economic principles learned in the course.

"Because of rapidly rising national defense expenditures, it is anticipated that Country A will experience a price inflation unless measures are taken to restrict the growth of aggregate private demand. Specifically, the government is considering either (1) increasing personal income tax rates or (2) introducing a very tight monetary policy." If the government of Country A wishes to minimize the adverse effect of its anti-inflationary policies on economic growth, which one of the following policies should it use?

a. The tight money policy because it restricts consumption expenditures more than investment.

b. The tight money policy, since the tax increase would restrict consumption expenditures.

c. The personal income tax increase since it restricts consumption expenditures more than investment.

d. Either the tight money policy or the personal income tax rate increase since both depress investment equally.

Teachers have developed similar questions for analysis and interpretation of poetry, literature, historical documents, and various kinds of scientific data. If you would like assistance in writing this kind of multiple choice question, or if you would like to discuss testing in general, please call the Center for Teaching and Learning for an appointment.

Bibliography


EXERCISE: Draft a Test item that Involves Application or Evaluation

STEM (& QUESTION)

CORRECT ANSWER

DISTRACTORS
1.

2.
Writing Effective Multiple-Choice Test Items

After the Test – Item Analysis

**Discrimination index** – for an item to be a good discriminator most of the top 50% of students should get it right and most of the lower half should miss it. Rewrite an item if equal numbers of students in each half answer correctly or if more students in the lower half than answer correctly.

**Frequency of response** – If no one chooses an option it is not a good distracter.

**NOTE:** Concourse (aka WebCT Vista) or bubble-sheet scoring software should be able to provide these statistics

Test item Matrix

<table>
<thead>
<tr>
<th>Learning goals covered on the test</th>
<th>Totals by Cognitive Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recall</td>
</tr>
<tr>
<td>1.</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td></td>
</tr>
</tbody>
</table>

Practical Suggestions

- Do not write the test in one day.
- Write one or two test items after each class on note cards.
- As a general rule, each item should stand on its own (it should not depend the answer to another question).
- Information in one item should not provide clues to another.
- Test items are easiest to write when there are definite right and wrong answers.
- Have students write items as an assignment.
- Grouping items under headings will improve student performance.
- Ask a colleague to read over your test items to help ensure validity.
- Publisher-provided test items – read each one before using it. Encourage class attendance by also testing material covered only in class.
- First write the stem, then the correct answer, and then the distracters.
Writing Effective Multiple-Choice Test Items

Summary checklist

General practices

☐ Students should not be able to guess the answer by the way the responses are written.
☐ Students should need to take the course to know the answer (not general knowledge)
☐ Use new language, rather than phrases lifted directly from text or lecture.
☐ Offer three to five answer choices.
☐ Distribute correct answers among different letters.
☐ Avoid humorous or ridiculous answer choices

Questions

☐ Avoid complexity that has nothing to do with knowing the answer
☐ Avoid negatively worded statements ("Which is NOT…?" "All of these EXCEPT…")
☐ Put empty spaces at the end, not the middle ("The first US president was George ___").
☐ Use capital letters to emphasize key words like LEAST, BEST, or MOST

Examples of GOOD distracters (wrong answers)

☐ Something that is incorrect but reasonable
☐ A common error that people make on the topic.
☐ A true statement that does not answer the question
☐ A jargon-ridden statement that is meaningless to someone who understands the concept.

AVOID these in correct answers (or use them in distracters!)

☐ The longest answer
☐ One of two opposite answers
☐ Much more scientific-sounding than others
☐ Extreme words: "all," "always" and "never"
☐ "All of the above," "none of the above," and "both a & b"
Sources

http://testing.byu.edu/info/handbooks/14 Rules for Writing Multiple-Choice Questions.pdf


"Writing Multiple-Choice Questions that Demand Critical Thinking." Georgeanne Cooper, University of Oregon. 2007 http://tep.uoregon.edu/resources/assessment/multiplechoicequestions/mc4critthink.html