Designing effective objective test questions: an introductory workshop

CAA Centre

Loughborough University
17 June 1999

Prepared by Colleen McKenna and Joanna Bull

Writing objective test questions

1.1 Question types

The following are examples of some of the question types appropriate for CAA:

Multiple choice questions (MCQs) are the traditional 'choose one from a list' of possible answers.

True/False questions require a student to assess whether a statement is true or not.

Assertion-Reason questions combine elements of MCQ and true-false.

Multiple response questions (MRQs) are similar to MCQs, but involve the selection of more than one answer from a list.

Graphical hotspot questions involve selecting an area(s) of the screen, by moving a marker to the required position. Advanced types of hotspot questions include labelling and building questions.

Text/Numerical questions involve the input of text or numbers at the keyboard.

Matching questions involve linking items in one list to items in a second list.

Sore finger questions have been used in language teaching and computer programming, where one word, code or phrase is out of keeping with the rest of a passage. It could be presented as a 'hot spot' or text input type of question.

Ranking questions require the student to relate items in a column to one another and can be used to test the knowledge of sequences, order of events, level of gradation.

Sequencing questions require the student to position text or graphic objects in a given sequence. These are particularly good for testing methodology.

Field simulation questions offer simulations of real problems or exercises.

Other question types require students to identify and/or manipulate images. Students may be asked to plot a graph, complete a matrix, draw a line or build up an image using parts provided.
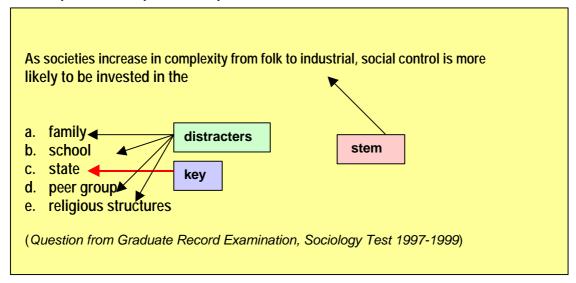
1.2 Multiple choice questions

Parts of a multiple choice question:

A traditional multiple choice question (or item) is one in which a student chooses one answer from a number of choices supplied. A multiple choice question consists of

- a **stem** - the text of the question
- **options** - the choices provided after the stem
- the **key**: the correct answer in the list of options
- **distracters**: the incorrect answers in the list of options

Example of a multiple choice question

As societies increase in complexity from folk to industrial, social control is more likely to be invested in the

a. family
b. school
c. state
d. peer group
e. religious structures

distracters

key

stem

(*Question from Graduate Record Examination, Sociology Test 1997-1999*)

1.3 Suggestions for constructing multiple choice questions

Writing stems

1. **Present a single, definite statement to be completed or answered by one of the several given choices**

   A. Weak question: Polysaccharides

   a. are made up of thousands of smaller units called monosaccharides
   b. are NOT found in the aloe vera leaf
   c. are created during photosynthesis
   d. can be described by the chemical formula: CHHOH

   B. Improved question: Polysaccharides of the plant cell wall are synthesized mainly in

   a. the endoplasmic reticulum
   b. the cytosol
   c. the plasma membrane
   d. the Golgi complex
   e. amyloplasts

   In the top example, there is no sense from the stem what the question is asking. The second example more clearly identifies the question and offers the student a set of homogeneous choices.

   Improved question from GRE - Biochemistry, cell and molecular biology test 1997-1999

2. Avoid unnecessary and irrelevant material

    A. Weak question: Paul Muldoon, an Irish postmodern poet who uses experimental and playful language, uses which poetic genre in "Why Brownlee Left"?

        a. sonnet
        b. elegy
        c. narrative poem
        d. dramatic monologue
        e. haiku

    B. Improved question: Paul Muldoon uses which poetic genre in "Why Brownlee Left"?

        a. sonnet
        b. elegy
        c. narrative poem
        d. dramatic monologue
        e. haiku

The top example contains material irrelevant to the question.

3. Use clear, straightforward language in the stem of the item. Questions that are constructed using complex wording may become a test of reading comprehension rather than an assessment of whether the student knows the subject matter.

Weak example: As the level of fertility approaches its nadir, what is the most likely ramification for the citizenry of a developing nation?

    a. a decrease in the labour force participation rate of women
    b. a dispersing effect on population concentration
    c. a downward trend in the youth dependency ratio
    d. a broader base in the population pyramid
    e. an increased infant mortality rate

Improved question: A major deline in fertility in a developing nation is likely to produce

    a. a decrease in the labour forces participation rate of women
    b. a dispersing effect on population concentration
    c. a downward trend in the youth dependency ratio
    d. a broader base in the population pyramid
    e. an increased infant mortality rate

> *Improved question from GRE - Sociology Test 1997-1999*

4. **Use negatives sparingly. If negatives must be used, capitalize, underscore embolden or otherwise highlight.**

Weak question: Which of the following is not a symptom of osteoporosis?

    a. decreased bone density
    b. frequent bone fractures
    c. raised body temperature
    d. lower back pain

Improved question: Which of the following is a symptom of osteoporosis?

    a. decreased bone density
    b. raised body temperature
    c. hair loss
    d. painful joints

5. Put as much of the question in the stem as possible, rather than duplicating material in each of the options. (Gronlund 1988)

Weak question: Theorists of pluralism have asserted which of the following?

    a. The maintenance of democracy requires a large middle class.
    b. The maintenance of democracy requires autonomous centres of countervailing power.
    c. The maintenance of democracy requires the existence of a multiplicity of religious groups.
    d. The maintenance of democracy requires a predominantly urban population.
    e. The maintenance of democracy requires the separation of governmental powers.

Improved question: Theorists of pluralism have asserted that the maintenance of democracy requires

    a. a large middle class
    b. autonomous centres of contervailing power
    c. the existence of a multiplicity of religious groups
    d. a predominantly urban population
    e. the separation of governmental powers

> ***Question from GRE Sociology test book - 1997-1999.***

Writing distracters

For single response MCQs, ensure that there is only one correct response.

Which of the following texts is considered to represent the pinnacle of modernist achievement?

a. *The Waste Land*
b. *Middlemarch*
c. "Ode to a Nightingale"
d. *Ulysses*
e. *Ethan Frome*

Improved: Which of the following texts represents one of the high points of modernist achievement?

a. *The Waste Land*
b. *Middlemarch*
c. "Ode to a Nightingale"
d. *Ethan Frome*
e. "My Last Duchess"

In the top example, both options a and d could be considered to be correct.

Use only plausible and attractive alternatives as distractors.

Weak question: Dichotic presentation of stimuli and shadowing are often used in the study of

a. brightness constancy
b. sensory scaling
c. shadow dancing
d. cartoons
e. selection attention

Improved question: Dichotic presentation of stimuli and shadowing are often used in the study of

a. brightness constancy
b. sensory scaling
c. illusions
d. depth perception
e. selection attention

In the top example, c and d are not serious distracters.

> **Question from GRE – Psychology, 1997-1999**

3. Avoid giving clues to the correct answer.

       A. Poor question: A fertile area in the desert in which the water table reaches the ground surface is called an

            a. mirage
            b. oasis
            c. water hole
            d. polder

       B. Improved question: A fertile area in the desert in which the water table reaches the ground surface is called a/an

            a. mirage
            b. oasis
            c. water hole
            d. polder

       Example A uses the article "an" which identifies choice b as the correct response. Ending the stem with "a/an" improves the question.

4.If possible, avoid the choices "All of the above" and "None of the above". If you do include them, make sure that they appear as correct answers some of the time.

It is tempting to resort to these alternatives but their use can be flawed. To begin with, they often appear as an alternative that is not the correct response. If you do use them, be sure that they constitute the correct answer part of the time. An "all of the above" alternative could be exploited by a test-wise students who will recognise it as the correct choice by identifying only two correct alternatives.

Similarly, a student who can identify one wrong alternative can then also rule this response out. Clearly, the student's chance of guessing the correct answer improves as they employ these techniques. Although a similar process of elimination is not possible with "none of the above", it is the case that when this option is used as the correct answer, the question is only testing the students' ability to rule out wrong answers, and this does not guarantee that they know the correct one. (Gronlund 1988)

5. **Distracters based on common student errors or misconceptions are very effective.**
One technique for compiling distracters is to ask students to respond to open-ended short answer questions, perhaps as formative assessments. Identify which incorrect responses appear most frequently and use them as distracters for a multiple choice version of the question.

6. **Correct statements that do not answer the question are often strong distracters.**

7. **Avoid using ALWAYS and NEVER in the stem as testwise students are likely to rule such universal statements out of consideration**.

8. **Do not create distracters that are so close to the correct answer that they may confuse students who really know the answer to the question. "Distracters should differ from the key in a substantial way, not just in some minor nuance of phrasing or emphasis." (Isaacs 1994)**

9. **Provide a sufficient number of distracters.**

    You will probably choose to use three, four or five alternatives in a multiple choice question. Until recently, it was thought that three or four distracters were necessary for the item to be suitably difficult. However a 1987 study by Owen and Freeman suggests that three choices are sufficient (Brown 1997). Clearly the higher the number of distracters, the less likely it is for the correct answer to be chosen through guessing (providing all alternatives are of equal difficulty.)

1.4 Extending MCQs

Once you have tackled the common form of MCQs you may wish to try more complicated forms such as multiple true/false and assertion-reason. Examples of these are given in Figures 1 and 2.

Figure1: Example of multiple true/false questions

In the following question, the examiner is assessing whether the student can apply his/her knowledge:

A 45 year old asthmatic woman who has lived all her life in Glasgow presents with a goitre of four years' duration and clinical features suggestive of hypothyroidism. Likely diagnoses include

a.   Iodine deficiency
b.   Dyshormonogenesis
c.   Drug-induced goitre
d.   Thyroid cancer
e.   Auto immune thyroiditis

(Correct answer: true C and E: false A, B and D)

The student has to appreciate that in Great Britain iodine deficiency is not likely to be associated with hypothyroidism, that a 45 year old patient with only a four year history is unlikely to have dyshormonogenisis, that asthmatic patients not uncommonly take iodine containing preparations which may result in a goitre, that hypothyroidism is not usually associated with thyroid cancer and that auto immune thyroiditis typically is found in a middle aged woman with hypothyroidism.
(*From Brown, et al 1997*)

Figure 2: Example of Multiple True/False Questions

In the following question, the student's clinical judgement is assessed:

A 28 year old woman with one child has taken anti-thyroid drugs for 6 months for thyrotoxicosis. She has a friend who has been successfully treated with radio-iodine. She finds she frequently forgets to take her drugs and wants to stop them to have radio-iodine treatment.

a.   She should be told that because of her age radio-iodine is best avoided.
b.   The problems associated with radio-iodine should be discussed with her.
c.   Surgery as a possible alternative should be discussed with her.
d.   She should be advised that some form of further treatment is required.
e.   You should find out more about her friend's treatment.

(Correct answer: true, B, C and D: false, A and E).

 Examples from Harden, R.M. & Dunn, W.G. (1981) *Assessment a Work Manual Dundee Centre of Medical Education* in Brown *et al* 1997

Note: These approaches may be used for testing knowledge and judgement in many subjects.

When grouped together, a series of true/false questions on a specific topic or scenario can test a more complex understanding of an issue. They can be structured to lead a student through a logical pathway (Brown 1997) as in the above example which simulates a medical diagnosis. Such questions may also be useful to the lecturer for diagnostic purposes, because they can reveal part of the thinking process employed by the student in order to solve the given problem.

Assertion-reason

The assertion-reason item combines elements of multiple choice and true/false question types, and allows you to test more complicated issues and requires a higher level of learning.

The question consists of two statements, an assertion and a reason. The student must first determine whether each statement is true. If both are true, the student must next determine whether the reason correctly explains the assertion. There is one option for each possible outcome.

Figure 3: Example of assertion-reason question

Each question below consists of an assertion and a reason. Indicate your answer from the alternatives below by circling the appropriate letter.

	Assertion	Reason	
A	True	True	Reason is correct explanation.
B	True	True	Reason is NOT a correct explanation.
C	True	False	
D	False	True	
E	False	False	

Assertion

1. The blood sugar level falls rapidly after hepactectomy.
                    BECAUSE
    The glycogen of the liver is the principal source of blood sugar.

2. Increased government spending increases inflation under all conditions.
                    BECAUSE
    Government spending is not offset by any form of production.

3. Chloroform has a dipole moment
                    BECAUSE
    The chloroform molecule is tetrahedral.

(Quoted in Brown *et al* 1997, p. 93 based on Matthews 1981.)

Assertion-reason tests can be used to explore cause and effect and identify relationships. When writing assertion-reason questions, keep in mind the following points:

- The reason should be a free standing sentence so that it can be considered separately from the assertion.
- Avoid using minor reasons. These can result in an ambigous question.
- Repeat options A-E in full for each question.
- Use all five options as keys equally.

Multiple response questions

Multiple response questions are a variation of multiple choice in which the student is allowed to choose more than one choice.

Matching

Matching items require students to match a series of stems or premises to a response or principle. They consist of a set of directions, a column of statements and a column of responses.

Figure 4: matching test item

Directions: Column I contains descriptions of geographic characteristics of wind belts. For each statement find the appropriate wind belt in Column II. Record your answer in the appropriate space on the answer sheet. Answers may be used more than once.

Column I

___1. Region of high pressure, calm, and light winds
___2. The belt of calm air nearest the equator.

___3. A wind belt in the northern hemisphere typified by a continual drying wind.
___4. Most of the United States is found in this belt.

Column II

A. Doldrums

B. Horse latitudes

C. Polar easterlies

D. Prevailing easterlies

E. Prevailing westerlies

Example from Ron J. McBeath ed. (1992) *Instructing and Evaluating Higher Education: A Guidebook for Planning Learning Outcomes* (New Jersey: ETP) 207.

Example 2 - matching test item

Directions: Match the quotation in column I with the literary school with which it is associated listed in column II. Items in column two may be used more than once.

Column I
1. You can lead a horse to water but you can't make it
    Hold
    Its nose to the grindstone and hunt with the hounds.
    Every dog has a stitch in time. Two heads? You've been
    Sold
    One good turn. One good turn deserves a bird in the
    hand.

2. I cannot see what flowers are at my feet,
    Nor what soft incense hangs upon the boughs,
    But, in embalmed darkness, guess each sweet
    Wherewith the seasonable month endows
    The grass, the thicket, and the fruit-tree wile;
    White hawthorn, and the pastoral eglantine;
    Fast fading violets cover'd up in leaves;
    And mid-May's eldest child,
    The coming musk-rose, full of dewy wine,
    The murmurous haunt of flies on summer eves.

3. frseeeeeeeeefronnnng train somewhere whistling the
    strength those engines have in them like big giants and the
    water rolling all over and out of them all sides like the end
    of Loves old sweeeetsonnnng the poor men that have to be
    out all the night from their wives and families in those
    roasting engines stifling it was today Im glad I burned the
    half of those old Freemans and Photo Bits leaving things
    like that lying about hes getting very careless

4. Twit twit twit
    Jug jug jug jug jug jug
    So rudely forc'd
    Tereu

5. A perfect Judge will read each Work of Wit
    With the same Spirit that its Author writ,
    Survey the Whole, nor seek slight Faults to find,
    Where Nature moves, and Rapture warms the Mind;

Column II
A. Romanticism
B. Modernism
C. Neo-classicism
D. Post-modernism
E. Humanism
F. Classical realism

Advantages of matching questions

Matching questions are particularly good at assessing a student's understanding of relationships. They can test **recall** by requiring a student to match the following elements:

Definitions - terms
Historical events- dates
Achievements - people
Statements- postulates
Descriptions - principles (McBeath, 1992)

They can also assess a student's ability to **apply** knowledge by requiring a test-taker to match the following:

Examples - terms
Functions - parts
Classifications - structures
Applications - postulates
Problems - principles (McBeath, 1992)

Matching questions are really a variation of the multiple choice format. If you find that you are writing MCQs which share the same answer choices, you may consider grouping the questions into a matching item. Tips for writing good matching questions include:

- Provide clear directions
- Keep the information in each column as homogeneous as possible
- Allow the responses to be used more than once
- Arrange the list of responses systematically if possible (chronological, alphabetical, numerical)
- Include more responses than stems to help prevent students using a process of elimination to answer question.

True/False questions

A true-false questions is a specialised form of the multiple-choice format in which there are only two possible alternatives. These questions can be used when the test-designer wishes to measure a student's ability to identify whether statements of fact are accurate or not.

Example of true/false question

> T F   A poem with the following rhyme scheme could be correctly referred to as an English sonnet: abab cdcd efef gg.
>
> T F   All eukaryotic genes are organized into operons.

True-false questions offer lecturers a very efficient method of testing a wide range of material in a short period of time. They can also be combined within a multiple-choice to create the more complex assertion-reason item. However, true-false questions do have a number of limitations:

- Guessing – a student has a 1 in 2 chance of guessing the correct answer of a question.

- It can be difficult to write a statement which is unambiguously true or false – particularly for complex material.
- The format does not discriminate among students of different abilities as well as other question types.

Suggestions for writing true-false questions:

- Include only one main idea in each item.

- As in multiple choice questions generally, use negatives sparingly.

- Try using in combination with other material, such as graphs, maps, written material. This combination allows for the testing of more advanced learning outcomes. (Gronlund 1988)

- Use statements which are unequivocally true or false.

- Avoid lifting statements directly from assigned reading, lecture notes or other course materials so that recall alone will not permit a correct answer.

- Generally avoid the use of words which would signal the correct response to the test-wise student. Absolutes such as "none", "never", "always", "all", "impossible" tend to be false, while qualifiers such as "usually", "generally", "sometimes" "often" are likely to be true.

Text match response

The text match question requires a student to supply an answer to a question or complete a blank within a brief piece of text, using words, symbols or numbers.

Examples of text match question

a. Tony Blair is the leader of the _____ party.

b. 235 x 23 + (9x5) = _____.

A possible advantage of this question type is that the student must supply the correct answer rather than identify or choose it. The likelihood that the candidate will guess the correct answer is lower than that of a multiple choice question. However, the short answer response questions can be difficult to phrase in such a way that only a single correct answer is possible. Additionally, if you are marking the assessments with computers, spelling errors may disadvantage students who know the right answer. However, with some software, the test designer can identify numerous permutations of the correct answer for which the student will be awarded full marks. For example if "United States" were the correct response to a short answer question, the author of the test could designate full marks to be awarded for all of the following: "United States", "US", "USA" and "United States of America".

References

Brown, G., with Bull, J., and Pendlebury, M. (1997) *Assessing Student Learning in Higher Education.* London: Routledge.

Brown, G. and Pendlebury, M. (1992) *Assessing Active Learning .* Sheffield: CVCP, USDU.

Gronlund, N.E. (1988) *How to construct Achievement Tests.* Englewood Cliffs: Prentice Hall.

Isaacs, G. (1994) "About multiple choice questions" in *Multiple Choice Testing: Green Guide,* No. 16, HERDSA, Cambelltown, NSW, 4-22.

McBeath, R. J. ed. (1992) *Instructing and Evaluating Higher Education: A Guidebook for Planning Learning Outcomes.* New Jersey: ETP.

Owen, S.V. and Freeman, R.D. (1987) "What's wrong with three option multiple items?" In *Educational & Psychological Measurement 47, 513-22.*