

MULTIPLE CHOICE QUESTIONS: AN INTRODUCTORY GUIDE

Elisa Bone and Mike Prosser

Multiple-choice questions (MCQs) continue to be relied upon for the efficient assessment of students' learning in higher education. However, the use of MCQs has not been without criticism. Poor design of MCQs, testing only lower-level learning outcomes¹ such as information recall, and encouraging students to focus on these rote and reproductive approaches to learning are among the criticisms. Each of these criticisms can be mitigated with close attention to the function, purpose and design of MCQs in the curriculum. This Guide focusses on addressing these issues and helping to make the use of MCQs in assessment of learning more valid, reliable and productive.

The first part of the Guide addresses issues of role and function, where we will discuss the outcomes of research and scholarship into these aspects in relation to assessment for student learning.

The second part of the Guide describes the major forms and structures of MCQs and discusses their advantages and disadvantages, with a focus on the design of MCQs for the assessment of student learning.

We finish with a short section on Do's and Don'ts of using MCQs and throughout this guide, our major focus is on the valid and reliable assessment of student learning.

PART 1: Role and Function of MCQs

MCQs can be used to play a wide variety of roles in our teaching and learning. They can be used in lectures and online activities to provide feedback to students. Their use in clickers and polls within large lectures classes and their periodic use in online learning activities can be key to providing immediate formative feedback to students and staff. They can efficiently cover a wide range of topics, and be reliably and objectively marked or scored without the need for a marking rubric. With the development of a large bank of items and appropriate software, they can be used to provide each student with a different test, supporting their use of assessment of learning online.

But, among the criticisms of the use of MCQs, three stand out:

1. They tend to focus on the assessment of low-level knowledge and encourage students to adopt rote learning approaches.
2. The design of items testing high-level knowledge is time consuming and requires expertise and experience.
3. Their use distorts students approaches to study – students may perceive that all they need to do is to memorise enough material and they will pass.

¹ For further information on learning outcomes and their alignment to the curriculum, please, see the MCSHE resource 'Writing Learning Outcomes: A Practical Guide for Academics' by. Available at: <https://melbourne-cshe.unimelb.edu.au/resources/categories/teaching-and-learning/curriculum-design>

A great deal of systematic and anecdotal evidence supports the assumption that MCQs tend to assess low-level knowledge and there are several reasons why this is so. Among them is that the design of items to test high level knowledge and understanding is difficult and time consuming. The design of items that are valid and reliable indicators of high-level knowledge and understanding, with well-designed distractors, requires expertise and experience. It is comparatively much easier to design items to test low-level knowledge and understanding. Indeed, it can be argued that the efficiency in marking items measuring high-level knowledge and understanding may not compensate for the time needed to develop valid and reliable measures of such knowledge and understanding. We will address ways of mitigating these risks later in our list of Do's and Don'ts.

The effect on students' approaches to study also needs consideration^{2,3}. Rote and reproductive approaches to learning are common. The evidence is that students perceive MCQs to test low-level knowledge and understanding. The very structure of MCQs tends to evoke such a perception. Consequently, students tend to adopt 'surface' approaches to studying for MCQ tests, perceiving that if they can remember 'enough stuff' then they can do well. It is difficult to change this perception. Consequently, when well-designed items measuring high-level knowledge and understanding are used, students often do poorly. Suggested ways for mitigating this risk are:

- A. Incorporating MCQs, including those testing higher-level learning, into lecture, tutorial and online programs and clearly articulating expected learning outcomes. This will help students understand what is being assessed.
- B. Using MCQs to *only* test low-level knowledge, whilst using short open-ended questions to test high-level knowledge.
- C. Using MCQs to test both low-level and high-level knowledge within formative assessments, with expected learning outcomes clearly articulated and directed feedback provided.

Suggestion A gives students a mix of low-level and high-level MCQs, asking them to answer the items, then asking them to classify the items as testing low-level or high-level knowledge and finally discussing their responses contributes to both assessment *of*, and *for*, learning. Using this strategy students experience the items and are assisted to reflect on them, building their understanding of the expected levels of achievement in the subject. Incorporating these interactive sessions

Suggestion B incorporates the issue of difficulty in design of high-level items and students' perceptions that they test low-level knowledge. In this approach, relatively easily designed MCQs that are used to test low-level knowledge are supplemented with short open-ended questions to test high-level knowledge and understanding. This approach mitigates against the risk of poorly designed high-level items and is consistent with students' perceptions of MCQs and open-ended questions.

Suggestion C provides a compromised approach and can be useful in cases where, for example, in-depth discussion is not feasible, or assessments are largely online. Feedback can be tailored to discuss both the correct and incorrect alternatives, allowing students to reflect on their reasoning. Feedforward processes can also be incorporated to allow comments and reflections from students, which may in turn be utilised to inform the design of future assessment tasks.

2 Scouller, K. M., & Prosser, M. (1994). Students' experiences in studying for multiple choice question examinations. *Studies in Higher Education*, 19(3): 267–279.

3 Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education*, 35: 453–472.

PART II: Writing MCQs

An MCQ comprises the stem, a question or problem that leads into the list of possible solutions, or *alternatives*, comprising both the *key* (the correct answer) and several incorrect *distractors*.

The *stem* may take the form of a question or statement and may also refer to external or supplementary material, such as a figure or diagram, which students are asked to interpret in selecting the best alternative. Similarly, alternatives can be in the form of single terms, statements, figures or diagrams.

There are several key considerations that need to be applied when writing MCQ items. Taken together, these aim to ensure that MCQ questions are *reliable* – consistently measuring a learning outcome, are valid – testing student learning at the appropriate level, and are focused, with students spending most of their time considering their response to the question, rather than on making sense of extraneous or irrelevant material. So, when writing an MCQ question, one should:

- Make sure all alternatives are plausible and relate in some way to the topic or subject matter. This will increase the question's *reliability* – increasing the chances that students who choose the correct answer meet the learning objective.
- Write alternatives that are as similar as possible in grammar, length, language and form. This will reduce the chances that students may detect hidden clues as to which alternative is the correct one, assisting to increase the question's *validity*.

- Ensure that the stem and alternatives make sense on their own without reference to additional materials and use terms and symbols that are familiar to students. This will focus students' attention on answering the MCQ rather than interpreting its construction, reducing their *cognitive load*. Similarly, use simple sentence structures, avoiding non-standard lingo and jargon, complex grammar and lengthy unnecessary wording.
- Avoid using negative language within the stem, including double negatives within both the stem and alternatives, unless the learning objectives call for such language. Using negative language can both increase a students' *cognitive load*, by requiring them to examine the question more carefully to detect this language, and decrease the question's *validity*, since students who fail to detect the negative language will be less likely to answer the question correctly. If using negative language, ensure it is flagged or highlighted, for example in bold text or capitalisation. Brame (2013) and Chiavaroli (2017) discuss these aspects in more detail.

In the following section, we present several examples of MCQ items, of different structural forms, that aim to assess both low-level learning outcomes such as recall and comprehension, and higher-level learning outcomes such as analysis, evaluation and synthesis.

Examples of common MCQ types

Presented in this section are examples of MCQs extracted from popular textbook question banks, from practice or from the literature. For each question, we present a brief evaluation and point out strengths and weaknesses in the question design. In the appendix, we discuss a broader range of MCQs in more detail. The correct answer for each question is denoted by an asterisk *.

A. MCQs to test lower-level learning outcomes (recall, comprehension)

1. Choose the best answer

From where do most fungi obtain nutrients?

- A. Nonliving organic matter*
- B. Living plants
- C. Living animals
- D. Photosynthesis

- ✓ Stem is posed as a question and relates to general topic.
- ✓ Each alternative answer is plausible.
- ✓ Alternatives are about the same length.
- ① Avoid including concepts that require clarification or imply restrictions or absolutes, e.g. 'only', 'most', 'always'.

2. Fill in the blank (#1)

Autotrophs use _____ as an energy source to drive photosynthesis.

- A. Sunlight*
- B. Hydrogen ions
- C. O₂
- D. CO₂

- ✓ Each alternative answer is a single item.
- ✓ All alternative choices should be familiar terms to the students.
- ① Alternatives are in different forms – both words and molecular formulae. Use consistent formats.
- ✗ The question stem does not stand alone – students will need to evaluate sentence for each alternative.
- ✗ Avoid this format in timed exams and quizzes.

3. Fill in the blank (#2)

Light-dependent reactions in plants proceed in the _____

- A. Thylakoid membrane*
- B. Plasma membrane
- C. Stroma
- D. Cytoplasm

- ✓ Each alternative answer is a single item.
- ✓ Each alternative answer is plausible.
- ✓ All alternative choices use terms familiar to the students.
- ① The question stem is a partial sentence. Writing the stem as a full question is preferable.

4. Complete the statement

A mushroom is _____

- A. The digestive organ of a club fungus
- B. The only part of the fungal body made of hyphae
- C. A reproductive structure that releases sexual spores*
- D. The only diploid phase in the club fungus life cycle

- ✓ Each alternative answer is plausible.
- ✓ Alternatives test students' understanding of concepts.
- ① Avoid including concepts that require clarification or imply limitation/restriction/ absolutes, e.g. 'only', 'most', 'always'.

B. MCQs to test *higher-level learning* (analysis, evaluation, synthesis)

1. Choose the best answer #1 (apply knowledge to context)

Which of the following functions is associated with dense riparian tree cover along narrow streams?

- A. Deposition of fine sediment into stream
- B. Bank destabilisation
- C. Increased productivity of aquatic algae
- D. Filtration of sediment from overland flow*

- ✓ Each alternative answer is plausible and contains familiar terms.
- ✓ Alternatives test students' understanding of concepts.
- ✓ Students need to interpret importance of descriptors in the stem to determine the best answer.

2. Choose the best answer #2 (synthesise knowledge across topics)

An animal's diet consists of tough, abrasive grasses with a low nutrient content and high fibre content. Which of the following combinations of dentition and digestive tracts would be most suitable for processing this diet?

- A. Large incisors, sharp canines, no caecum and a short small intestine
- B. Long canines, few molars, a four-chambered stomach and a long large intestine
- C. Broad molars, sharp incisors, an enlarged foregut containing bacteria, small caecum*
- D. No teeth, a crop, a gizzard containing small stones and a long small intestine

- ✓ Students need to interpret importance of descriptors in the stem to determine the best answer.
- ✓ Students need to draw on knowledge from several areas in answering the question.
- ✓ All alternative answers contain familiar terms.
- ✓ Most alternative answers are plausible but
- ① Option D could be easily excluded by a student with little knowledge of the subject. Try to ensure each alternative is at a similar degree of difficulty.

3. Choose the best answer #3 (evaluate outcomes of a scenario)

In an experimental study, one species of algal grazer was removed from a patch of rocky shoreline and a different species of algal grazer moved into the cleared area, where it did not occur prior to the experimental removal.

This finding shows that:

- A. Predation may be important for zonation of animals across rocky intertidal zones
- B. Disturbance from waves has no effect on animal zonation in rocky intertidal zones
- C. Competition between species may be important for zonation of animals across rocky intertidal zones*
- D. Differences in resources between microhabitats is the only factor influencing zonation of animals across rocky intertidal zones

- ✓ Each alternative answer is plausible and contains familiar terms.
- ✓ Alternatives test students' understanding of concepts and their ability to synthesise knowledge.
- ✓ Students need to analyse scenario within the stem to determine the best answer.
- ① Be careful that alternatives are not too long.
- ① Avoid using negative and absolute language unless appropriate to the intended learning outcomes.

4. Choose the best answer #4 (evaluate validity of an argument)⁴

Consider the question: 'What is meant by the charge that utilitarianism is too demanding'?

Now, consider this answer: 'Utilitarianism requires moral people to respond to important moral concerns such as helping the less fortunate, while allowing immoral people to pursue their careers, family lives and personal projects'.

What is wrong with this answer?

- A. Nothing – that answer is correct
- B. It falsely describes what utilitarianism requires of moral people
- C. It falsely describes what utilitarianism allows of immoral people*
- D. It relies on a false dichotomy between moral people and immoral people

5. Choose the best answer #5 (analyse poetry and identify patterns and relationships)⁵

[poem included here]

The chief purpose of stanza 9 is to:

- A. Delay the ending to make the poem symmetrical
- B. Give the reader a realistic picture of the return of the cavalry
- C. Provide material for extending the simile of the bridge to a final point
- D. Return the reader to the scene established in stanza 1*

- ✓ Each alternative answer is plausible and contains familiar terms.
- ✓ Alternatives test students' understanding of concepts and their ability to synthesise knowledge.
- ✓ Students need to analyse scenarios within both the stem and the alternatives to determine the best answer.
- ① Be careful that the stem is not overly long or complex. These types of questions may not be suited to high-stakes timed exams.
- ① The stem includes negative wording, asking students to consider the flaws in the stated answer. Negative wording, if necessary to include, should be clearly signposted e.g. using bold font or capitalisation.

- ✓ Each alternative answer is plausible and contains familiar terms.
- ✓ Alternatives test students' ability to examine and analyse external material within the context of concepts and theory presented in class.
- ① Be careful that external material is not too complex or will take too long to read and comprehend within the exam time.

4 Green, K. Sample multiple choice questions that test higher order thinking and application. Washington State University Office of Assessment of Teaching and Learning, in McGill University, Workshop: Designing effective multiple choice questions. Available at https://www.mcgill.ca/skillsets/files/skillsets/mcq_handout3.pdf

5 Burton, SJ, Sudweeks, RR, Merrill, PF & Wood, B (1991) How to prepare better multiple-choice test items: Guidelines for university faculty. Brigham Young University Testing Centre <http://testing.byu.edu/info/handbookd/betteritems.pdf>

Quick Guide – Dos and Don'ts of MCQs

DO/DON'T	General strategies	Designing items	Designing alternatives/ distractors
✓ DO	Use familiar language commonly used during classes.	Express the full problem in the stem.	Limit the number of alternatives – no more than 3–5 is best.
✓ DO	Familiarise students with examples of MCQs.	Use clear, unambiguous wording in the question stem.	Make alternatives appealing and plausible.
✓ DO	Write questions throughout the term to avoid a rush at the end.	Ensure the stem is meaningful in and of itself.	Make alternatives grammatically consistent with the stem.
✓ DO	Design questions that can be modified and recycled across cohorts.		Place alternatives in some meaningful order.
✓ DO	Use a mix of question types within a single quiz or exam.		Create alternatives that are a similar length.
✗ DON'T	Use verbal association clues from the stem in the key.	Make question stems too complicated or cluttered with irrelevant information.	Use terms like 'all of the above', 'none of the above' and absolutes such as 'always', 'never', 'all', or 'none'.
✗ DON'T	Use trick questions – questions should test students' knowledge and understanding, not their ability to deconstruct the item.	Use a stem that is incomplete, has gaps or is a partial sentence.	Make answer options too complicated or include overlapping choices (e.g. A or B, A and B).
✗ DON'T	Use negative wording unless learning outcomes require it. Negative wording can unnecessarily confuse students in high-pressure exam situations.	Ask students to choose the incorrect answer from a list of alternatives. This can increase students' cognitive load. It is much more productive and positive to ask students to select the best answer.	

Further Reading

Brame, C. (2013) Writing good multiple choice test questions. Retrieved [May 18, 2020] from <https://cft.vanderbilt.edu/guides-sub-pages/writing-good-multiple-choice-test-questions/>.

Centre for Teaching Excellence (2020, May 18) *Designing Multiple-Choice Questions*, Retrieved from: <https://uwaterloo.ca/centre-for-teaching-excellence/teaching-resources/teaching-tips/developing-assignments/assignment-design/designing-multiple-choice-questions>.

Chiavaroli, N. (2017). Negatively-worded multiple choice questions: an avoidable threat to validity. *Practical Assessment, Research and Evaluation*, 22(3): 1–14.

Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational measurement: issues and practice*, 24(2): 3–13.

Tarrant, M., Ware, J. and Mohammed, A.M. (2009). An assessment of functioning and non-functioning distractors in multiple-choice questions: a descriptive analysis. *BMC Medical Education*, 9: 40.

Appendix – Strategies for Common MCQ Item Types

Form	Description	Strategies	Benefits	Drawbacks	Further considerations
The ‘ choose the best answer ’	<p>The most common and familiar form of MCQ, these present the question stem, several incorrect answers to choose from (<i>distractors</i>), along with the single most correct answer (the <i>key</i>).</p> <p>Asking students to select the ‘best answer’ rather than the ‘correct answer’ mitigates against some distractors having correct elements</p>	<p>Use familiar language commonly used during term</p> <p>Avoid using verbal association clues from stem in the key</p> <p>Ensure the distractors are plausible and effectively test students’ knowledge and <i>understanding</i> of the concepts conveyed in the question</p> <p>Plausible distractor terms could be, for example, common misconceptions on the topic, or numerical answers that are within a plausible range but are not quite correct.</p>	<p>A simple question form that is quick and easy to read for students and easy to grade for teaching staff.</p>	<p>Writing plausible distractors can be difficult – these should be created by content experts.</p> <p>The validity of newly written MCQs can also be evaluated in low-stakes or formative quizzes by analysing student responses.</p> <p>Avoid simply asking students to <i>identify</i> a familiar term from their lectures or textbook.</p>	<p>Choosing how many distractors to include is important. Fewer distractors (2 or 3) can be effective, but only if they are plausible and represent a genuine choice for students (Tarrant et al. 2009; Rodriguez 2005).</p> <p>Too many distractors (more than 3 or 4) may result in students spending substantial time reading, instead of considering the question and choosing a possible answer.</p>
The ‘ complete the statement ’	<p>Students choose from a list of statements that address the question or complete the sentence in the stem</p>	<p>Ensure that only one statement is truly correct, but that others may also be plausible.</p> <p>Alternative statements should be of similar length and complexity.</p>	<p>Choosing the correct statement can help students recognise and address any misconceptions they have with the material.</p> <p>Asking students to consider full statements can also help them integrate their conceptual knowledge, whereas picking a correct single simple answer favours general recall.</p>	<p>Be careful that the statements are not simply copied from the lecture notes or textbook, as this test students’ recall rather than their conceptual understanding.</p> <p>It can be difficult to create plausible alternative statements without veering into ‘trick’ questions that seem correct but have slight differences in wording that trip up students.</p> <p>Be mindful of students with English as a second language when writing questions and try not to include tricky grammar or non-standard lingo.</p>	<p>Reading through question stems and possible answer statements will take longer than assessing a short list of distractors.</p> <p>Think carefully about the best use of the students’ time and increase the difficulty of the question and the number of assigned marks if needed.</p>

Form	Description	Strategies	Benefits	Drawbacks	Further considerations
The 'figure interpretation'	The stem asks students to refer to and analyse supplementary information to answer the question.	Asking students to view a figure, diagram or concept map and show their understanding by filling in gaps in the figure or map, choosing from a list of interpretative statements or making calculations based on the material presented.	Using figures of concepts from lectures, tutorials or other learning activities can help reinforce students' understanding of these concepts. Higher-level learning processes such as evaluation, interpretation and application of knowledge can be assessed in an easy-to-grade format.	These questions take longer to write, read and complete so they usually need to be assessed at a higher level and assigned more marks than a simple distractor. Given their long read-time, they may be more suited to take-home or non-timed quizzes and exams.	Using these types of questions in an online MCQ can be beneficial if the diagram, concept map or figure isn't readily available on the web. Try to create a new or modified figure for the MCQ.
The 'word match'	Students match the terms to their descriptions, or match concepts. Functions in a similar way to the 'most correct statement' but poses single question per statement.	A word match question can help to clarify where misconceptions exist in the material.	Allows students to demonstrate their knowledge of terms and their descriptions. Can be modified to test higher-level conceptual knowledge and applications. All material is relevant to the question, so plausible alternatives or distractors are not required.	Choosing the correct matches may take some time, particularly if the list is long. Each word match is testing a different piece of knowledge and should be allotted a full mark or grade.	Consider increasing the allotted reading time to suit. May be more beneficial in take-home, non-timed or formative quizzes or exams.
The 'fill in the blank'	Students choose words from a list to fill in gaps within a written statement.	The stem can be a description of a concept, with the blanks representing key terms, directions of responses or concepts. Students choose the best term to complete the statement, or the best combination of multiple terms in order.	Tests student's general knowledge of terminology, which can be an advantage in content-heavy subjects and subject themes. Allows students to show understanding of linked concepts within long statements. Versatile – can test knowledge levels from recall to applications.	Presenting multiple blank spaces for students to fill in may confuse students, so that their answers are not a true reflection of their knowledge. Be careful not to overcomplicate the statement and make sure the associations are unambiguous. Avoid using in high-stakes or high-pressure situations.	Several guides (e.g. Brame 2013) advise against using this question form, as it requires students to hold the stem statement in their working memory while evaluating each alternative answer. These questions should be used sparingly and their purpose clearly communicated to the students.

Melbourne Centre for the Study of Higher Education

Level 1, Elisabeth Murdoch Building, Spencer Road

The University of Melbourne

Victoria 3010 Australia

melbourne-cshe@unimelb.edu.au

 melbourne-cshe.unimelb.edu.au