The use of multiple choice question (MCQ) tests for assessing learning in higher education has come under increasing scrutiny because of their strong association with assessing lower order cognition such as the recall of discrete facts. Whilst it is possible to design MCQ tests to assess higher order cognition, this requires considerable skill. The development of effective MCQ tests is a time-intensive activity which entails significant subjective judgement and care to ensure their validity and reliability.

Course designers need to determine whether the use of MCQ tests for summative assessment purposes is justified primarily on educational grounds rather than because they are easily marked. Where their use is warranted, they need to be integrated effectively into assessment design to maximise their potential benefits and minimise their limitations.

The use of MCQ tests for formative assessment and as a stimulus to active and self-managed learning has been shown to have positive effects, both in terms of students’ improved learning performance, and their perceptions of the quality of their learning experience (for example, Velan et al, 2008).

### Benefits

- MCQ test items can provide students with rapid feedback on their learning.
- They can be available 24 x 7 without increasing the marking load.
- They can provide students with an accessible way to review course material, check that they understand key concepts, and receive immediate or timely feedback to help them manage their own learning.
- They can be designed using ‘quiz tool’ software applications that automate presentation and publication formats as well as facilitating quiz administration, scoring, and provision of feedback.
- They can be scored objectively since there is only one right answer, thus eliminating marker bias.
- They can be scored by anyone, or automatically scanned, thereby increasing efficiency, particularly in teaching large cohorts.
- Students’ diverse capabilities as writers do not influence the scores, thus increasing test validity.
- MCQ tests can provide teachers with information about students’ pre-course understanding, knowledge gaps and misconceptions, to help plan learning and teaching approaches.
- MCQ tests can enable a more extensive coverage of the curriculum and learning objectives being assessed than more intensive methods such as extended writing.
- MCQ test items can be progressively developed and accumulated in item banks for re-use in different combinations and settings.

### Challenges and limitations

- MCQs are time-consuming to develop and require skill and expertise to design well.
- They are generally acknowledged to be poor at testing higher order cognition such as synthesis, creative thinking and problem solving.
- Their use for summative assessment has been shown to encourage students to adopt ‘surface’ approaches to learning (Scoulker, 1996).
- MCQ items can be answered correctly by guessing and, if poorly designed, can provide clues to encourage guessing.
- MCQ tests typically provide students with little direction on how to improve their understanding.
- Students’ diverse levels of reading skill influence their capacity to understand MCQ items, regardless of their understanding of the content being assessed, thus reducing test validity.
- Likewise, test validity is reduced by cultural bias which is very difficult to avoid in the design of MCQ items.
Aligning MCQ tests to course learning outcomes

MCQ tests have been shown to favour assessing lower order cognitive processes such as the recall of factual information at the expense of higher level critical and creative reasoning processes. The first step, then, for course designers is to determine whether MCQ tests should be used at all based on the learning objectives and outcomes of their courses.

Where MCQ tests are appropriate to use, are they best used for formative assessment to support students’ self-management of their learning, or for summative assessment to grade the extent of students’ learning at a particular point in time? Importantly, MCQ tests should never constitute the only or major form of summative assessment in university-level courses. When MCQ tests are used primarily for formative learning purposes, assigning a small grade to them nonetheless can provide a signal to students about their importance and value.

Sometimes MCQ tests can provide valuable assessment feedback on students’ recall of the essential facts and concepts on which higher order learning outcomes depend. Where the learning outcomes to be assessed are at a higher cognitive level, then it becomes all the more challenging and resource-intensive to design suitable MCQ tests. The closed-ended nature of MCQ tests makes them particularly inappropriate for assessing originality and creativity in thinking.

Ensuring fairness for all students

As with all assessment forms, the adoption of inclusive design principles should underpin the construction of MCQ tests to ensure success is possible for all students, by accommodating their diversity of abilities, cultural backgrounds, and learning styles and needs.

➢ In the construction of MCQ test items, care needs to be taken to avoid sexual, racial, cultural or linguistic stereotyping to ensure that no groups of students are unfairly advantaged or disadvantaged.

➢ Alternative formats for MCQ-type exams may need to be provided for students with disabilities, for example, adjustments to the time available to take a test, or the provision of assistive technology or readers.

➢ Where timed MCQ tests are conducted online, contingency plans should be identified, for example, in case computer stations in laboratories are not working, or in case of system outages, and any issues arising from time zone differences addressed.

➢ Processes need to be developed in advance so that students can be informed in a timely way about the objectives of MCQ tests, their formats and delivery, and in the case of summative tests, the marking scheme and impact on grades.

➢ To reduce the opportunity for plagiarism, MCQ tests developed and conducted online can incorporate randomised question generation to allow different test items relating to the same content to be randomly provided.

Planning a summative MCQ test

If a MCQ test is deemed appropriate according to the objectives and outcomes for a course, its use must be advised in the course outline provided to students. When planning to incorporate a MCQ test in the summative assessment schedule of a course, course designers need to consider many factors, as outlined in Figure 1.

Constructing MCQ tests

Constructing effective MCQ tests and items takes considerable time and requires scrupulous care, both in the design and the subsequent review and validation stages. This is especially so when MCQ tests are used for high-stakes summative assessment.

For this reason, critically evaluating existing validated tests in order to select one for adoption may prove more efficient than constructing a test from scratch. Rather than taking on the specialist task of designing and validating test items, effort can productively be spent on planning to incorporate an available test into a course, for which numerous decisions need to be taken.

However, in some circumstances the potential benefits for learning may justify the effort. This may be particularly so if the process can be undertaken collaboratively within a department or discipline group, or as a larger project across institutional boundaries, increasing the potential longevity and sustainability of the outcomes.

The progressive development of an item bank can underpin benchmarking processes and the establishment of assessment standards with long-term impacts on assuring quality.
**Figure 1: Factors in planning MCQ test design**

<table>
<thead>
<tr>
<th>Questions</th>
<th>Factors to consider</th>
</tr>
</thead>
</table>
| **When should it be conducted?** | • Timing: early / mid / late / post course?  
• Frequency: regularly, weekly, occasionally, once?  
• If in exam week, timetable clashes with other exams taken by students need to be avoided through formally scheduling the test as an exam. |
| **Where should it be conducted?** | • In a specified physical location, online, by mail?  
• If online, how will different time zones be managed?  
• Can available online software tools facilitate the construction of the MCQ test as well as its conduct, scoring and provision of results and feedback? |
| **Are the costs justified?**   | • How much time will be needed for development and what is the opportunity cost?  
• What is the cost of printing exam scripts?  
• What are the costs of exam venues and invigilation?  
• What are the costs of marking?  
• Is there a more cost-effective method (for example, the use of existing validated MCQ items and item banks)? |
| **How should security be managed?** | • How will the security of exam scripts prior to an exam be assured?  
• Will exams undertaken in computing laboratories be invigilated?  
• How will results information be managed securely?  
• How will results be provided to students cost-effectively but without jeopardising privacy? |
| **How will risks be managed?** | • Which specialist staff / skills may need to be consulted (Learning and Teaching staff, IT support staff, examinations unit staff, disability services)?  
• What is the contingency plan, for example, if there are system outages or non-operational computers in laboratories?  
• If there are students in different time zones, how will  
• Are copyright provisions being observed regarding the use of existing MCQ tests and items? |
| **How will scoring be done?**   | • Will MCQ items in a test be equally weighted?  
• Will non-completed or wrong answers be scored zero? (The use of negative marks intended to reduce the incidence of guessing is not recommended.)  
• Will the mark required for a pass be 50% or will it be raised, even to 100%, in cases justified as testing essential factual knowledge?  
• Will scoring be fully or partly automated (for example, through fully online applications, or based on scanned hard copy scripts)?  
• Whose involvement in scoring needs to be secured (for example, teaching teams, casual demonstrators, students)? |
| **How will feedback be provided?** | • Will feedback be sufficiently timely to support learning improvement (for example, immediate through answer-contingent progression in the test, or within a specified short timeframe)?  
• How will feedback be focused on positive learning improvement rather than negatively on simple notification of the score?  
• How will generic feedback regarding all students’ performance on the MCQ test be able to be used by students in interpreting their own results?  
• Will feedback be provided by teachers or students through peer assessment or self-assessment? |
| **How will quality be assured and improved?** | • How will MCQ test scoring processes be managed, for example, training of scorers, provision of scorer guidelines?  
• Will test items and test banks be developed collaboratively to enhance their validity?  
• Will peer review processes be established to check MCQ tests and items for their alignment with course objectives, for logical sequence, timing, item construction, and so on?  
• How will MCQ tests and items be validated, for the discriminatory value of items?  
• How will the department engage with reviewing and endorsing MCQ use and practices? |
A design framework for planning the MCQ test can help in setting out the overall structure of the test in terms of the relative proportions of test items to assess particular topic areas and types of learning objectives across a spectrum of cognitive demand. As an example, the ‘design blueprint’ in Figure 2 from the Instructional Assessment Resources at the University of Texas at Austin provides a structural framework for planning.

**Figure 2: Design blueprint for MCQ test design**

<table>
<thead>
<tr>
<th>Cognitive domains (Bloom’s Taxonomy)</th>
<th>Topic A</th>
<th>Topic B</th>
<th>Topic C</th>
<th>Topic D</th>
<th>Total items</th>
<th>% of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>12.5%</td>
</tr>
<tr>
<td>Comprehension</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>7</td>
<td>17.5%</td>
</tr>
<tr>
<td>Application</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>15</td>
<td>37.5%</td>
</tr>
<tr>
<td>Analysis</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>10</td>
<td>25.0%</td>
</tr>
<tr>
<td>Synthesis</td>
<td>1</td>
<td></td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2.5%</td>
</tr>
<tr>
<td>Evaluation</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>40</td>
<td>100%</td>
</tr>
</tbody>
</table>

- It is important to design MCQ test items with the most appropriate format for the type of question posed, for example:
  - A single correct answer
  - More than one correct answer
  - True/false (single or multiple correct answers)
  - Matching (for example, a term with the appropriate definition, or a cause with the most likely effect)
  - The best way to complete a given sentence
  - Questions relating to some given prompt material

- Developing MCQ items that assess higher order thinking and reasoning can be particularly challenging, but the use of some prompt material on which to base a cluster of MCQ items can facilitate this effectively, for example, a brief outline of a problem, case or scenario, a visual representation (picture, diagram or table) of the interrelationships among pieces of information or concepts, or an excerpt from published material.

The MCQ items based on the prompt material can be presented in a sequence from basic understanding through to higher order reasoning, including identifying the effect of changing a parameter, selecting the solution to a given problem, and nominating the optimum application of a principle. The addition of some short-answer questions to a substantially MCQ test can be a useful way of minimising the effect of guessing, as students are required to express in their own words the way that they understand, interpret, and analyse problems.

- Students need to be given information well in advance of a MCQ test that explains the purposes (and whether summative or formative), the topics being covered, the structure of the test, whether aids can be taken to the test (for example, calculator, notes, textbooks, dictionaries), how it will be marked, and how the mark contributes to the overall grade. The instructions on the MCQ test need to be clear in explaining the components, their relative weighting, and how much time students are expected to spend on each section so they can optimise their time.

**Assuring the quality of MCQ tests**

Whether MCQ tests are used to support learning in a formative way or for summative assessment, it is important to ensure that the overall test construction and individual items are well-aligned with the objectives for learning. When MCQ tests are used for summative assessment purposes, assuring their validity is all the more critical. These strategies can be used to focus on quality:

- Make use of a basic quality checklist when designing and reviewing the test.
- Take the test yourself, and assume that students may take four times longer than you do to complete it.
- Work collaboratively to develop a MCQ item bank as a dynamic (and growing) repository which can be exploited for formative or summative purposes across the discipline, and which can enable peer review, evaluation and validation.
- Engage in peer review processes to consider whether the use of MCQ tests is educationally justified, to critically evaluate MCQ test design and items, to examine the impacts of using MCQs in the context of the actual learning setting, and to record and disseminate the outcomes of such deliberations for the benefit of students and colleagues.

**Engaging students in active learning with MCQ tests**

- When used formatively, MCQ tests can provide substantial benefits for learning through engaging students in actively reviewing their own learning progress, identifying gaps and weaknesses in their understanding, and consolidating their learning through rehearsal (eg Velan et al, 2008).
MCQ tests can provide a trigger for collaborative learning activities, for example, through discussion and debate about the answers to questions.

Technologies such as electronic voting systems can assist collaborative engagement with MCQ items (Draper, 2009), and students can be asked to score MCQ tests and summarise common areas of misconception within the class.

Activities can also revolve around disrupting the traditional agency of assessment, for example, by requiring students to critique MCQ items in terms of which learning outcomes they are aligned with, or to construct their own MCQ items and prepare explanatory feedback on the right and wrong answers (Fellenz, 2010).

Using technologies to support MCQ tests

Many technology applications are available to enable staff to design and develop MCQ tests. To save time, it is advisable to take advantage of those applications supported by the University as part of the TELT platform, and particularly in high-stakes summative assessment uses of MCQ tests.

Resources


Good Practice Guide in Question and Test Design — PASS-IT Project for Preparing Assessments in Scotland using IT:

IAR (Instructional Assessment Resources) University of Texas at Austin:
http://www.utexas.edu/academic/ctl/assessment/iar/students/plan/method/exams-blueprint.php

UNSW (Technology Enabled Learning and Teaching):
http://telt.unsw.edu.au/blackboard/content/staff/Bb91_designing_tests.cfm?ss=0


References and further reading


http://www.biomedcentral.com/1472-6920/8/52

Acknowledgements

The contributions of staff who engaged with the preparation of this topic are gratefully acknowledged.

The Assessment as Learning Toolkit includes some material adapted from resources developed by the Learning and Teaching Centre at Macquarie University.

Toolkit Online

Visit the UNSW Assessment as Learning Toolkit at

Contact the Toolkit Co-ordinator at
assessment.toolkit@unsw.edu.au