CME

Multiple Choice Question Tests for Educational Assessment

Onkarnath Chattopadhyay*

Overview

The Multiple Choice Question (MCQ) is one of the most popular item formats used in educational assessment.There has been a steady increase in the use of MCQs in higher education to supplement or replace conventional assessment practices due to increase in numbers of students, decrease of resources, modularisation and the increase in availability of computer networks.[1,2,3]

Computer networks enable flexible delivery of MCQs and speed up the marking and the collation of test results. Compared to paper-based MCQs, the use of online computer-assisted assessment (CAA) can significantly reduce the burden associated with testing large student cohorts.[4]

Many teachers believe that only factual information can be tested by MCQs, whereas only essay exams can test higher-order cognitive skills. Poorly written MCQs elicit complaints from students that the questions are confusing. But MCQ tests can test many of the same cognitive skills that essay tests do, and, if the teacher is willing to follow the special requirements for writing MCQs, the tests will be reliable and valid measures of learning.[2,4]

MCQs can be used effectively at almost all levels to measure a wide range of abilities. Well designed MCQs can measure higher end abilities such as analysis and

* MD. Consultant (Public Health), Rural Health Unit & Training Centre (RHU&TC), Singur
Email: onkarchatto@gmail.com

evaluation.[4]

MCQs can be constructed to assess a variety of learning outcomes, from simple recall of facts to Bloom's highest taxonomic level of cognitive skills – evaluation. [4]

Designing multiple-choice questions

MCQ tests are strongly associated with assessing lower order cognition such as the recall of discrete facts. It is possible to design MCQ tests to assess higher order cognition (such as synthesis, creative thinking and problem solving), but skilful drafting of questions is necessary if such tests are to be valid and reliable. This takes time and entails significant subjective judgment.[3,4]

A multiple-choice question (MCQ) is composed of three parts: a stem that identifies the question or problem, and a set of alternatives or possible answers that contain a key that is the best answer to the question, and a number of distracters that are plausible but incorrect answers to the question. Students respond to MCQs by indicating the alternative that they believe best answers or completes the stem.[1,2,3]

In a multiple choice question test, students should be instructed to select the "best answer" rather than the "correct answer". The question should use the same terminology that was used in the course. Verbal association clues from the stem should be avoided in the key. Misleading phrasing and unimportant details should not be given. In general, one should avoid having any negatives in the stem or the options.[7]

While designing stems, it should be ensured that it is possible for the students to answer the question without looking at the options. Information that can be included in the stem should not be repeated in each of the alternatives. Irrelevant information in the stem confuses students and leads them to waste time.[5] The stem and each of the choices should be read aloud to make sure that they are grammatically correct.

While designing alternatives, generally 3-5 alternatives are given. But it is not efficient to include more than four alternatives. For a four-choice question test, 8, 18 and 48 questions guarantee that the probabilities of obtaining a mark above 40 per cent by

pure guesswork are below 5%, 1% and 0.01%, respectively.[9] It is better to avoid two or more correct answers in the choices.

The **distracters** should be similar in length and use similar type of language as the correct solution. The choices should be placed in numerical, chronological or conceptual order. The test should have roughly the same number of correct answers that are a's, b's, c's, and d's (assuming there are four choices per question). The use of 'all of above' or 'none of above' should be avoided. The use of words such as 'always', 'never', 'all', or 'none' should be avoided. Vague words or phrases like "usually," "typically" and "may be" should be avoided. Active voice should be used. The alternatives should be mutually exclusive. It is better to avoid questions of the form "Which of the following statements is correct?" [3,9]

MCQ tests are challenging and time-consuming to create. It is easier to write a few questions each week, perhaps after a lecture when the course material is still fresh in the mind.

The decision to use multiple-choice tests should be based on what the purpose of the test is and the uses that will be made of its results. If the purpose is only to check on factual and procedural knowledge, if the test will not have a major effect on overall curriculum and instruction, and if conclusions about what students know in a subject will not be reduced to what the test measures, then a multiple-choice test might be somewhat helpful -- provided it is unbiased, well written, and related to the curriculum. If they substantially control curriculum or instruction, or are the basis of major conclusions that are reported to the public, or are used to make important decisions about students, then multiple-choice tests may be misleading.

The design of MCQ depends upon what is being evaluated. The question types will differ accordingly:

a.   **Remembering factual knowledge**

- Write the fact as a statement

- Transform the statement into a question

- Ask the student to select the appropriate answer

b. **Understanding conceptual knowledge**

- Which of the following is an example of _____ ?

c. **Applying procedural knowledge:**

- Prepare a short case study or example

- Write a stem that asks the student to demonstrate the use of the procedural knowledge or solve a problem

d. **Analysing conceptual knowledge**

- Present the student with a diagram and ask for analysis—for example, "Given the following chart, which of the following is most likely to occur next?"

e. **Evaluating procedural knowledge**

- Give the student a short case study

- Ask "Which of the following would have been a better plan of action?" [5]

MCQ and Bloom's Taxonomy

The Knowledge and Comprehension categories of Bloom's Taxonomy are fairly simple to measure with MCQs. It lends itself well to questions asking students to "identify", "distinguish", "recognize", "recall", or "classify" something. It is increasingly more difficult to write multiple-choice questions for the higher-order thinking categories of Application, Analysis, and Evaluation. In order to do this, one often has to create complicated stems (providing a reading passage or chart in the question portion) for them to "interpret", "infer", "predict" or "conclude" from. It is impossible to write a multiple-choice question for Synthesis since this category requires the student to create something new.[6]

Analysing the Responses

After the test is given, it is important to perform a test-item analysis to determine the effectiveness of the questions. Most machine-scored test printouts include statistics for each question regarding item difficulty, item discrimination, and frequency of response for each option. This kind of analysis gives the information needed to improve the validity and reliability of the questions.

Item difficulty is denoted by the percentage of students who answered the question correctly, and since the chance of guessing correctly is 25% (in case of 4 options); any item that falls below 30% should be rewritten. One should strive for items that yield a wide range of difficulty levels, with an average difficulty of about 50%. If a question is expected to be particularly difficult or easy, results that vary widely from our expectation merit investigation.

Item discrimination index is derived from the class scores which are divided into upper and lower halves and then compared for performance on each question. For an item to be a good discriminator, most of the upper group should get it right and most of the lower group should miss it. A score of 0.50 on the index reflects a maximal level of discrimination, and items scoring below 0.30 on this index should be rejected. If equal numbers of students in each half answer the question correctly, or if more students in the lower half than in the upper half answer it correctly (a negative discrimination), the item should not be counted in the exam.

Frequencies of responses are examined for the incorrect options under each question, one can determine if they are equally distracting. If no one chooses a particular option, that particular option should be rewritten before using the question again.[11]

Advantages:

MCQ is the most commonly used question type in e-Learning. MCQs provide faster ways of assessing student learning; provide rapid feedback to students and save staff time in marking. An instructional designer prefers MCQs over other question types as they can be scored rapidly and feedback can be given easily. It is an effective way to test a large number of learners, quickly and

effectively.[3,4]

A Multiple Choice test offers more flexibility for assessing a diversity of content .It allows for a precise interpretation for content validity. It is purely objective in scoring and testing. It is easy to be administered both by students and the teacher. Student success does not depend on his/her writing skills. Results can easily be compiled and analyzed to determine patterns of student learning outcomes, level of difficulty of questions, usefulness of questions and follow-up action required.

It can be used with all subject areas. It is versatile in measuring all levels of cognitive skills. It permits a wide sampling of content and objectives. It provides highly reliable test scores. It can be machine-scored quickly and accurately.[2,5]

MCQ test can be used to test a variety of levels of learning, when there are a large number of individuals taking the test, there is time to construct the test items, time is limited for scoring, when it is not important to determine how well individuals can formulate their own answer and when we need to prepare individuals for future assessments that use a similar format.[2,6,7]

MCQ tests can be useful for formative assessment and to stimulate students' active and self-managed learning. They improve students' learning performance and their perceptions of the quality of their learning experience.[4]

Limitations:

Although MCQ tests are widely used, these have recognized limitations. Firstly, it is argued that MCQs promote memorization and factual recall and do not encourage (or test for) high-level cognitive processes. Other researchers, however, maintain that this depends on how the tests are constructed and that they can be used to evaluate learning at higher cognitive levels.

Secondly, the feedback provided through MCQs is usually quite limited as it is predetermined during test construction. Hence there is little scope for personalization of feedback based on different student needs.

Thirdly, the use of MCQs is usually driven by the need for efficiency and rapid feedback rather than by robust pedagogical principles. MCQs require the selection of a correct answer from a set of alternatives, i.e. the recognition of the answer rather than the construction of a response.

In addition, students have no role in setting the goals and standards for MCQ tests, nor are they usually in a position to clarify the test question or its purposes while taking the test. It does not satisfactorily address concerns regarding a more active and participative role of students in assessment processes or development of students' skills to self-regulate their own learning.[4]

MCQ tests inhibit students from expressing creativity or demonstrating original and imaginative thinking. Question design is restrictive, forcing students to fit their understanding into the designer's way of understanding a concept. Success of question depends on suitability of distracters. Longer reading time is required and students with poor reading skills may be disadvantaged, especially under time limit. Some students may guess at answers without understanding them.

Ease of writing low-level knowledge items leads examiners to neglect writing items to test higher-level thinking. They are not useful for measuring students' ability to articulate explanations, display thought processes, organize their thoughts, and generate original ideas and to provide unprompted examples.[2,5]

The quality of MCQ test depends on the examiner's writing ability. It is time consuming to design good questions. It is very easy to construct poor questions, and use of such questions may be much worse than other methods of assessing the same learning outcome

Conclusion:

While preparing and answering the students should learn to think and apply knowledge. Facts and procedures are necessary for thinking, but exclusive reliance on multiple-choice may tend to minimize or eliminate thinking and problem-solving. Therefore, classroom assessments and standardized tests

should not be solely based on multiple-choice or short-answer items. Instead, other well-designed forms of assessment should be implemented and used properly.[12]

Lastly, multiple choice testing is an efficient and effective way to assess a wide range of knowledge, skills, attitudes and abilities. If done well, it allows broad and even deep coverage of content in a relatively efficient way. Although it is true that no single format should be used exclusively for assessment, multiple choice testing still remains one of the most commonly used assessment formats.[13]

REFERENCES:

1.   Cheung, D. How can we construct good multiple choice items? Available from: http://www3.fed.cuhk.edu.hk/chemistry/files/constructmc.pdf accessed on 20.8.2015

2.   Vocational Training Council, Hong Kong. Constructing Quality Multiple Choice Questions for Student Assessment. Available from: http://www.vtc.edu.hk/tlc/webline/weblineMar06/ConstructingMCQsa.pdf accessed on 21.8.2015

3.   UCD Teaching & Learning Resources, Dublin. The Design of Multiple Choice Questions for Assessment. Available from: http://www.ucd.ie/t4cms/UCDTLA0042.pdf accessed on 23.8.2015

4.   Nicol, D. E-assessment by design: using multiple-choice tests to good effect. Journal of Further and Higher Education Vol. 31, No.1, pp. 53–64. Available from: http://www.reap.ac.uk/reap/public/papers/MCQ_paperDN.pdf accessed on 22.8.2015

5.   Center for Teaching and Learning, Johns Hopkins Bloomberg School of Public Health. CTL Teaching Toolkit. Available from: http://www.jhsph.edu/departments/population-family-and-reproductive-health/_docs/teaching-resources/cla-03-multiple-choice-questions-sept-2013.pdf accessed on 24.8.2015

6.	Center for Instructional Technology & Training. University of Florida. Multiple Choice Questions. Available from: http://citt.ufl.edu/online-teaching-resources/assessments/multiple-choice-questions/ accessed on 21.8.2015

7.	Center for Teaching Support & Innovation. University of Toronto. Practical Tips for designing & implementing multiple-choice tests. Available from: http://www.teaching.utoronto.ca/Assets/Teaching+Digital+Assets/CTSI+1/CTSI+Digit al+Assets/PDFs/LCT-Practical+Tips+for+Designing+and+Implementing+MCTs.pdf accessed on 16.8.2015

8.	Centre for Teaching Excellence, University of Waterloo. Designing multiple choice questions. Available from: https://uwaterloo.ca/centre-for-teaching-excellence/teaching-resources/teaching-tips/developing-assignments/assignment-design/designing-multiple-choice-questions accessed on 19.8.2015

9.	Zhao, Y. How to Design and Interpret a Multiple-Choice-Question Test: A Probabilistic Approach. Int. J. Engg Ed. Vol. 22, No. 6, pp. 1281-1286, 2006. Available from: http://pcwww.liv.ac.uk/~yyzhao/papers/2006_ijee_mcq.pdf accessed on 18.8.2015

10.	Heriot-Watt University, Scotland, U.K. Good Practice Guide in Question and Test Design. Available from: http://www.calm.hw.ac.uk/GeneralAuthoring/031112-goodpracticeguide-hw.pdf accessed on 18.8.2015

11.	Center for Teaching and Learning. University of North Carolina at Chapel Hill. Available from: http://www.smu.edu/~/media/Site/Provost/assessment/Resources/MultipleChoices/I mproving%20Multiple%20Choice%20QuestionsUNCCH.ashx?la=en accessed on 22.8.2015

12.	Fairtest. Multiple Choice Tests. Available from: http://www.fairtest.org/multiple-choice-tests accessed on 18.2.2016

13.	Jay Parkes. Multiple Choice Test. Available from: http://www.flaguide.org/cat/mutiplechoicetest/multiple_choice_test1.php accessed on 18.2.2016