Original Article

Constructing Multiple Choice Questions as a Method for Learning

Edward Palmer,[1] *BE (Hons), BSc*, Peter Devitt,[2] *MBBS, MS, FRACS*

Abstract

Introduction: Many different strategies exist to try and encourage students to increase their knowledge and understanding of a subject. This study was undertaken to measure the effect of student-based construction of multiple choice questions (MCQs) as a stimulus for the learning and understanding of topics in clinical surgery. Materials and Methods: The study was carried out at the University of Adelaide, Australia and had 2 components. Fourth-year students were required to provide a case study during a surgical attachment and half of the group was asked to supplement this with MCQs. These students were pre- and post-tested and the effect of the additional intervention (MCQ-construction) measured. Fifth-year students were polled on their preferred methods of learning before and after a learning exercise in which they were asked to undertake a case presentation and create some MCQs. Results: The MCQ questions designed by the students were of a high standard and clearly displayed an understanding of the topic concerned. The 4th-year students in the MCQ construction group showed equivalent outcomes as the case study control group. Students initially ranked MCQ-construction amongst the least stimulating methods of learning, but after the exercise their opinion was significantly more favourable, although still much less than traditional learning methodologies (tutorials, books). Conclusions: Construction of MCQs as a learning tool is an unfamiliar exercise to most students and is an unpopular learning strategy. However, students are capable of producing high quality questions, and the challenge for medical faculties is how best to use this initiative to the students' advantage.

Ann Acad Med Singapore 2006;35:604-8

Key words: Educational measurement, Educational models, Medical students

Introduction

Students in general and medical students in particular are often described as "strategic learners",[1] but in reality many become superficial learners out of necessity when faced with the seemingly boundless volume of material in today's curriculum. The apparent enormity of the task might encourage rote learning and much of what is learnt can be rapidly forgotten once the examination or assessment task is completed. One of the challenges that the teacher faces is how to encourage a deeper understanding of the key areas of the curriculum.

Multiple choice questions (MCQs) have been a favoured method of assessing student knowledge and understanding in the medical field for many years.[2-5] The high reliability and versatility of MCQs and the ability to mark them quickly and efficiently has made them an oft-used tool for teaching and evaluating many different disciplines.[6-8] Online education has relied heavily on the use of MCQs as its main assessment tool as witnessed by the use of learning systems such as Blackboard[9] and Web-CT.[10]

The design of MCQs requires a solid knowledge of the subject being assessed as well as an understanding of the mechanics of making a good question. Thus, creating a series of well balanced, discriminating MCQs is no trivial exercise. The stem and question need to be worded unambiguously and the discriminators need to be appropriate. The knowledge required to effectively set a good quality question is greater than that required to answer one.

Students are often given assignments designed to improve their understanding of the topic they are studying. These types of assignments may vary from problem-solving exercises to project work and essay writing. Asking students to create MCQs based on their learning material may lead

[1] The Centre for Learning and Professional Development, University of Adelaide, Australia
[2] Associate Professor, Department of Surgery, University of Adelaide, Australia
Address for Reprints: Mr Edward Palmer, Evaluation and Survey Service, CLPD, University of Adelaide, Adelaide, South Australia 5005.
Email: edward.palmer@adelaide.edu.au

to a deeper understanding of the study topic than other methods. It has been suggested that if students were to write their own problem-solving exercises, they may benefit by being motivated to study a subject in greater depth.[11]

We have undertaken a study to measure the effect of students constructing their own MCQs on examination performance and learning. The hypothesis to be tested was that when given a topic to research and write on, student understanding would be enhanced if they were asked to construct MCQs based on the material under study.

Materials and Methods

The study was carried out at the University of Adelaide, Australia and involved 4th-and 5th-year medical students. The study had 2 components. The first component (Study 1) involved the 4th-year medical students. Fifty-one students were enrolled in the study. At the start of a 9-week attachment to a surgical clinic, they were briefed on the learning objectives for that part of the course and provided with the details of what were considered the key areas of the surgical syllabus. Students were instructed to produce a case study based on 1 of 5 topics, with some discussion of an area of debate or interest within that topic and perhaps illustrated by their case study. The topics for discussion were colorectal cancer, gallstone disease, carcinoma of the breast, peripheral vascular disease and gastro-oesophageal reflux disease. Students were also asked to identify the learning resources used for the case study.

These students were randomised to 1 of 2 groups. Group A was asked to produce the case report alone and Group B was required to construct 3 MCQs in addition to producing the case report. The Group B students were provided with examples of good and bad MCQs and guidelines on how to construct MCQs. This included guidelines on the construction of a stem sufficient to pass the "cover test" and the avoidance of flaws (negative statements, absolutes, ambiguity). The term "cover test" implies that if the answers are covered up the student should be able to deduce the correct answer from the stem without any prior knowledge of the options. Students were asked to produce MCQs on the topic of their study.

The MCQs constructed by the students were analysed based on their appropriateness as a quality 1-best-answer-from-5 MCQ according to the following 5-point rating scale:

1. Pass the cover test and no item-writing flaws
2. Pass the cover test and 1 to 2 item-writing flaws
3. Cover test dubious and no item-writing flaws
4. Fail the cover test and 1 to 2 item-writing flaws
5. Fail the cover test and more than 2 item-writing flaws

The item-writing flaws were defined as:

- Repetition of part of the stem in an option
- Use of qualifiers within an option
- Complicated or ambiguous stem
- Negative questions not clearly stated
- Use of double negatives
- Absolute options (e.g., never, always, all-of-the-above)

The MCQs were categorised in terms of their educational objectives using Bloom's taxonomy.[12]

Both groups were pre- and post-tested. The pre-test was undertaken at the start of the attachment and the post-test was conducted on the last afternoon of the attachment. The test consisted of 30 MCQs focusing on the 5 topics.

Student opinion was canvassed on the role of constructing MCQs as an adjunct and stimulus to learning and they were asked to rank this method of learning in comparison to other standard methodologies.

In the second component (Study 2), a class of 53 5th-year medical students was provided with a clinical topic to research. Students were asked to research their topic and report back after a period of 1 week, at which point they would provide a 10-minute presentation to their peers and a set of 3 MCQs for their fellow students to attempt to answer. As an aid to their work, the students were provided with a guideline to creating MCQs.

Before beginning the exercise, these students were asked to rank methods of learning about their research topic and they were asked to repeat this ranking after they have made their presentation and presented their MCQs. As a form of motivation, the students were told that their MCQs might either be used in an informal quiz to be held after the group presentations or in the end-of-year examination. After the presentation, students were asked to complete a survey designed to measure their attitude towards creating MCQs as a learning exercise. This exercise was voluntary and students were not assessed on the outcomes.

In both components of this study, students were asked to construct MCQs as an integral part of a learning exercise. In component 1, the students were asked to produce a case report and in component 2, a clinical topic to research. Whilst the exercises were different, the common theme was the use of MCQ-construction as a stimulus to deeper understanding of the materials being studied and learnt.

Analysis of the rankings of student learning preferences in Study 2 was carried out using a 1-way ANOVA test using Tukey's post-hoc test. Correlations of learning preferences were analysed by using Pearson's correlation.

t-testing was used to compare pre- and post-test student scores for the students in Study 1.

Results

In the first component of the study, the 4th-year students

Table 1. The Rating of the MCQs Constructed by 4th-year Students (Study 1)

Rating of MCQs	Assessor 1	Assessor 2
1	45%	45%
2	16%	14%
3	17%	20%
4	18%	16%
5	4%	5%

Table 2. Classification of Study 1 MCQs According to the Cognitive Skills Being Tested

Bloom's categorisation	Percentage of MCQs in this category
Knowledge/Comprehension	75%
Higher order skills – Application/ Analysis/Synthesis/Evaluation	25%

created 95 MCQs which were analysed by 2 assessors and categorised according to the 5-point rating scale. The results are shown in Table 1. Between 59% and 61% of questions were judged as being of sufficient standard for formal use in an MCQ examination. Another 17% to 20% were acceptable after minor alterations. The categorisation of the MCQs using Bloom's taxonomy[12] showed that one quarter of the questions produced were judged to be capable of testing higher order cognitive skills (Table 2).

The questionnaire was completed by 25 students in Group A and 26 students in Group B. Table 3 shows the rankings given by the students in both groups to various learning methods. Students in both groups ranked learning methods similarly with no significant differences between groups. Small group teaching ranked highest and creating MCQs ranked 9th out of the 10 options. Analysis of this ranking showed that students who prefer small-group teaching also prefer large-group teaching (r = 0.55, $P = 0.003$). These students tended not to want assignments that involve doing short essay questions (r = 0.50, $P = 0.007$), writing short essay questions (r = -0.54, $P = 0.004$) or writing essays (r = -0.49, $P = 0.009$). These were the only significant correlations drawn from the data.

Performance in the pre- and post-tests for each group is shown in Table 4. Both groups showed similar improvement in their post-test scores and there was no statistical difference between the 2 groups.

In the second component of the study, 53 students responded to the first survey and 33 students responded to the post-exercise survey. A total of 26 students completed both the first and second sessions of the exercise. The data were analysed on this group alone.

Students created 37 MCQs, which were provided to their

Table 3. Student Rankings of Learning Strategies from Study 1

Learning method	Non-MCQ students Group A	MCQ students Group B	Significance *P* value
Attend a small group teaching session on the topic (<30 students)	1.9/1	2.0/1	ns
From textbooks/journals	3.4/3	3.4/3	ns
Attend a large group teaching session on the topic (>30 students)	4.4/4	4.4/4	ns
Do an assignment on the topic answering short essay questions	4.9/5	4.3/4.5	ns
Do an assignment on the topic answering multiple choice questions	5.2/5	5.1/5	ns
Discuss the topic with fellow students	5.3/5	5.1/5	ns
Write an essay on the topic	5.7/6	6.3/6	ns
Watch a video/DVD on the topic	7.1/7	6.5/7	ns
Create multiple choice questions on the topic	7.8/8	8.2/9	ns
Create short essay questions on the topic	8.3/9	8.3/9	ns

ns: not significant

Table 4. Learning Outcomes for the 4th-year Students in Both Control (no MCQ) and MCQ Groups

	MCQ students Group B	Non-MCQ students Group A	*P* value
Pre-test	12.6	14.0	ns
Post-test	14.9	16.1	ns
Improvement	2.3	2.1	ns
	<0.001	<0.001	

ns: not significant

fellow students. An analysis of the questions indicated that students had created good quality MCQs. The results of the ranking of learning methods are shown in Table 5.

The results show that students ranked small-group teaching sessions very highly, large group teaching, textbooks and fellow students moderately highly, videos and assignments next and other tasks poorly. Testing on pre- and post-exercise rankings indicate that there was a significant difference for the MCQ as a learning exercise option, but not for any other learning method. The ranking for using MCQ creation remained poor after the exercise.

Thirty students responded to the questionnaire on the value of creating MCQs as a learning exercise. The internal reliability of the survey instrument was 0.89 and a scale out

Table 5. Learning Methods Ranked by Students in Study 2

Learning method	Pre-exercise rank (Average/ Median)	Post-exercise rank (Average/ Median)	Significance *P* value
Attend a small-group teaching session on the topic (<30 students)	2.4/1	2.0/1	ns
From text-books/journals	3.6/4	4.1/4	ns
Attend a large-group teaching session on the topic (>30 students)	4.0/4	4.4/3.5	ns
Discuss the topic with fellow students	4.1/4	4.4/4	ns
Do an assignment on the topic answering short essay questions	5.1/5	5.8/5	ns
Do an assignment on the topic answering multiple choice questions	6.0/5	5.5/6	ns
Watch a video/DVD on the topic	6.0/7	6.3/7	ns
Write an essay on the topic	6.8/7	6.9/7	ns
Create multiple choice questions on the topic	8.4/8	7.2/7	0.04
Create short essay questions on the topic	8.5/9	8.0/8	ns

ns: not significant

of 20 (range, 5 to 20) was created. The average result was a scale mark of 12.2 ± 0.5.

Discussion

We have shown that students are quite capable of constructing quality MCQs which fulfil the criteria of being free from flaws and with a stem or lead-in that is clear, free from ambiguity and often able to test more than mere recall of fact. This indicates that the students had a solid understanding of the mechanics of designing MCQs and had obtained knowledge and understanding sufficient to create good discriminators. The large number of quality MCQs is in contrast to Sircar et al[11] who found only 20% of the MCQs designed by undergraduate physiology students to be worthy of inclusion in the departmental question bank. The difference is probably due to the level of experience of the writers of the questions. The students in this study had at minimum 3 years of university experience before beginning this exercise.

Although the students were capable of producing high quality MCQs, the majority of the MCQs created by them tested knowledge and comprehension. The premise of this study is that in order to create MCQs of a high calibre, deeper understanding of the material involved in the question is created. The fact that the MCQs did not test deep learning may suggest that the students did not gain the higher levels of understanding hoped for, but this is a difficult conclusion to arrive at especially as it is not a realistic expectation for students to produce MCQs testing higher order cognitive skills at their first attempt.

The demonstrated ability of students to construct high quality MCQs did not appear to stimulate much enthusiasm amongst the students when they were asked to consider the value of MCQ-construction as a learning exercise. The responses to the questionnaires showed a poor opinion was held of the exercise. While the students in the second component of the study did rate MCQ-construction slightly higher once they had been through the exercise, it still did not draw much interest. This lack of interest was reflected in the almost complete absence of improvement in the objective assessment in the first component of the study, despite the fact that the focus of the end-of-attachment examination on the subject material under study and that almost all the MCQs used in the examination were in the top rating of construction.

Exposing students to MCQs as a teaching exercise should act as a stimulus to learning. The MCQ is perhaps the most common examination format in the medical curriculum, and if the students can understand the process and means of construction of a good MCQ, this should work to their advantage. Not only should it help with actual examinations but, just as importantly, it will teach the difference between mere knowledge acquisition and how that knowledge can be used in terms of comprehension, application, analysis and evaluation.

In reality, exercises of this nature are not likely to be greeted with much enthusiasm as they involve learning methods unfamiliar to many students, and may be perceived as an inefficient use of study time. Both students and staff need to be shown that unfamiliar methods can be beneficial in order for them to be more receptive to new learning initiatives.

It may well be the case that one of the major strengths of an MCQ assessment, i.e. the ability to assess a wide range of material, was one of the weaknesses of the learning strategy in this study. The students were assessed using a 30-question MCQ, which tested a wide range of concepts, yet they only carried out their MCQ design over a small subset of this. The students may have gained a deep understanding of their particular area of study, but they obtained only a superficial understanding of other areas, thus showing no substantial net increase in their overall understanding. This problem may be rectified if students are required to create a larger number of MCQs.

This concept is supported by a study conducted by Brink,[13] where students designed their own version of a 1-hour summative exam. These students were not allowed to write MCQs, but focussed on essay and calculation questions instead. It was concluded that the students could benefit if they could write more questions in order to extend the scope of their understanding. The logistics of such an exercise may become prohibitive with essay questions, although success has been reported in this area,[14] but it may become feasible with sufficient MCQs.

When considering teaching and learning initiatives to adopt in a curriculum it is necessary to balance the time and effort required to produce the initiative and the desired outcome. In this instance, the desired outcome was an improvement of student knowledge and understanding and from that perspective, the exercise did not show any apparent gain. However, there were some unexpected benefits. MCQs are time-consuming to produce. Accuracy of content must be ensured. Relevance and importance of the material must be considered. The structure of each question must be acceptable in terms of an appropriate stem and be free of flaws. Such questions often circulate amongst teachers and committees before they are deemed suitable. Although this was not the intention of the study, the students have produced a large bank of potentially viable MCQs which are perfectly suitable for both formative and summative assessment purposes.

REFERENCES

1. Newble DI, Entwistle NJ. Learning styles and approaches: implications for medical education. Med Educ 1986;20:162-75.
2. Davidson GP, McDougall JA, Roberts FJ. MCQ assessment of introductory clinical course for fourth year medical students. N Z Med J 1978;87:313-4.
3. Fox JS. The multiple choice tutorial: its use in the reinforcement of fundamentals in medical education. Med Educ 1983;17:90-4.
4. Gandhi IS, Ghosh MN. A trial of multiple choice examination in pharmacology. Med Educ 1978;12:287-9.
5. Nnodim JO. Multiple-choice testing in anatomy. Med Educ 1992;26:301-9.
6. Sriram TG, Chandrashekar CR, Moily S, Kumar K, Raghuram A, Isaac MK, et al. Standardisation of multiple-choice questionnaire for evaluating medical officers' training in psychiatry. Soc Psychiatry Psychiatr Epidemiol 1989;24:327-31.
7. Heishman SJ, Singleton EG, Liguori A. Marijuana Craving Questionnaire: development and initial validation of a self-report instrument. Addiction 2001;96:1023-34.
8. Salaniponi FM, Kwanjana JH, Gondwe MK, Harries AD. Use of a multiple choice questionnaire to stimulate TB officers to read the TB manual in Malawi. Int J Tuberc Lung Dis 2004;8:618-20.
9. Blackboard Inc [homepage on the Internet]. Available at: http://www.blackboard.com/us/index.aspx. Accessed 12 April 2006.
10. Web CT [homepage on the Internet]. Available at: http://www.webct.com/. Accessed 12 April 2006.
11. Sircar SS, Tandon OP. Involving students in question writing: a unique feedback with fringe benefits. Am J Physiol 1999;277:S84-S91.
12. Bloom BS. Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain. New York: David McKay Co Inc, 1956.
13. Brink J, Capps E, Sutko A. Student exam creation as a learning tool. Coll Student J 2004;38:262-72.
14. Brown IW. To learn is to teach is to create the final exam. College Teaching 1991;39:153.