Evaluation of Modified Essay Questions (MEQ) and Multiple Choice Questions (MCQ) as a tool for Assessing the Cognitive Skills of Undergraduate Medical Students

Moeen-uz-Zafar-Khan and Badr Muhammad Aljarallah

Assistant Professors, Department of Medicine, Qassim University

Abstract:

Objectives: Developing and testing the cognitive skills and abstract thinking of undergraduate medical students are the main objectives of problem based learning. Modified Essay Questions (MEQ) and Multiple Choice Questions (MCQ) may both be designed to test these skills. The objectives of this study were to assess the effectiveness of both forms of questions in testing the different levels of the cognitive skills of undergraduate medical students and to detect any item writing flaws in the questions.

Methods: A total of 50 MEQs and 50 MCQs were evaluated. These questions were chosen randomly from various examinations given to different batches of undergraduate medical students taking course MED 411-412 at the Department of Medicine, Qassim University from the years 2005 to 2009. The effectiveness of the questions was determined by two assessors and was defined by the question's ability to measure higher cognitive skills, as determined by modified Bloom's taxonomy, and its quality as determined by the presence of item writing flaws.
'SPSS15' and 'Medcalc' programs were used to tabulate and analyze the data.

Results: The percentage of questions testing the level III (problem solving) cognitive skills of the students was 40% for MEQs and 60% for the MCQs; the remaining questions merely assessed the recall and comprehension. No significant difference was found between MEQ and MCQ in relation to the type of questions (recall; comprehension or problem solving x^2 = 5.3, p = 0.07).The agreement between the two assessors was quite high in case of MCQ (kappa=0.609; SE 0.093; 95%CI 0.426 – 0.792) but lower in case of MEQ (kappa=0.195; SE 0.073; 95%CI 0.052 – 0.338). 16% of the MEQs and 12% of the MCQs had item writing flaws.

Conclusion: A well constructed MCQ is superior to MEQ in testing the higher cognitive skills of undergraduate medical students in a problem based learning setup. Constructing an MEQ for assessing the cognitive skills of a student is not a simple task and is more frequently associated with item writing flaws.

Key words: Modified essay question, Multiple-choice question, Bloom's Taxonomy, cognition.

Correspondence

Moeen-uz-Zafar-Khan
Assistant Professor,
Department of Medicine,
Qassim University,
Flat-16, Al muntaza B,
Onaiza, Kingdom of Saudi Arabia
E-mail: drmoeenzafar@yahoo.com

Introduction

The evaluation of the competence of undergraduate medical students is a very critical task, as in the future, these 'to be physicians' have to cater with human lives. [1]

At undergraduate level there are three domains of skills to be evaluated i.e. Cognitive, Affective and Psychomotor. Cognitive domain can be evaluated [2] at different levels including Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation. Modified bloom's taxonomy [3] identified three levels of cognitive domain. In medical education, the major emphasis is on developing and evaluating the level III or problem solving skills, as most of the physician's time is spent in analyzing patient's problems.

Proper cognitive assessment tools reward the students for their higher cognitive skills and abstract thinking. [4] There are various methods to assess the knowledge domain which include Free response examinations (Long Essay Questions, Short answer Questions, Modified Essay questions), Multiple choice questions, Key feature questions, Self-assessment and peer- assessment. Each of these methods has its pros and cons and is addressed to assess different levels of bloom's taxonomy. No single method of evaluation is superior to other and probably a reliable and valid evaluation requires a combination of these methods. [1, 4]

The Multiple choice questions are very popular in evaluation of undergraduate medical students. They are reliable and valid; moreover they are easy to administer to a large number of students. Well constructed MCQs have a greater ability to test knowledge and factual recall but they are less powerful in assessing the problem solving skills of the students. A large proportion of curriculum can be tested in a single sitting. The scoring is very easy and reliable using computer software, but the construction of good MCQs is difficult and needs expertise. Generally MCQs stimulate students to make a superficial and exam oriented study. [5, 6]

Modified essay questions are short clinical scenarios followed by series of questions with a structured format for scoring. They primarily assess the student's factual recall but they also assess cognitive skills such as organization of knowledge, reasoning and problem solving. They also address the writing skills and even ethical, social and moral issues and attitudes. MEQs are more flexible and their value is somewhere between essay type questions and MCQ. But they need to be carefully constructed with provision of model answers and training to avoid inter-rater variability. [7, 8]

In a bird's eye view, examination and evaluation are a source of anxiety and stress for the undergraduate medical students, but in reality it directs the students to study harder and improve their skills. Therefore, it is imperative not only to stimulate their cognitive skills during the teaching but also to frequently examine their higher mental and reasoning skills. [9, 10, 11]

Objectives

Our objectives were

- To compare MCQ and MEQ in their ability to test different levels of the cognitive domain
- To detect item writing flaws in construction of a question

Material and methods

Study design:

Cross-sectional survey

Methods:

50 MCQs and 50 MEQs were chosen randomly from the written examinations delivered to fourth year medical students attending undergraduate internal medicine course at the Department of Medicine, Qassim University. In the 50 MEQs, there were 104 stems. Questions were analyzed individually by two independent assessors as per the 'preformed criteria' to label a question with Level I, II or III of the cognitive domain. Questions belonged to various final and midterm examination from 2005 to 2009.

Ethical approval:

No ethical approval was required as no human subjects were involved in the research

Assessment:

Each question was analyzed separately to

1) Score it according to modified bloom's taxonomy for cognitive skills
Level I Knowledge (recall of information including direct questions asking to check the factual recall, containing words like enumerate; list etc.). Level II Comprehension and application (ability to interpret data; questions including lab data or containing words like analyze). Level III Problem – solving (Use of knowledge and understanding in new circumstances, including scenario based questions which contain case description and lab data asking students to initially make a diagnosis and then to suggest next appropriate investigation; management modalities; counseling etc)

2) Evaluate for any item-writing flaw. Following were regarded as item writing flaws

- Error in formatting, spelling, grammar
- Technical errors
- Double negatives
- Cascading stems
- Absolute options
- Ambiguous stems
- Repetition of information

The analysis was done by each assessor as per the Performa shown in table 1.

Table 1: MCQ /MEQ evaluation form

Question	Bloom's taxonomy	ITEM WRITING FLAWS	YES	NO
	□ Level I: Knowledge recall of information	1.Error formatting		
		2.Error spelling		
		3.Error grammar		
	□ Level II: comprehension and application understanding and being able to interpret data	4.Technical error		
		5.Double negatives		
		6.Cascading stems		
		7.Absolute options		
	□ Level III: problem-solving use of knowledge and understanding in new circumstances.	8.Ambiguity		
		9.Repetition		

Moeen-uz-Zafar-Khan and Badr Muhammad Aljarallah

Statistical analysis:

Software 'SPSS 15' was used for data entry, analysis and interpretation. Kappa test was used to determine the agreement between the two assessors using 'Med Calc' software.

Results

MEQ:

Total of 50 questions with 104 stems were included. Analysis is shown in table 2.

MCQ:

Total of 50 questions were evaluated. Analysis is shown in table 2.

Questions were representing different disciplines of Medicine (graph 1). No significant difference was found between MEQ and MCQ in relation to the type of questions (recall; comprehension or problem solving (x^2 = 5.3, p = 0.07).

The analysis of item writing flaws is shown in table 3.

Table 2 : Classification of Questions according to Bloom's taxonomy

Bloom's Level	MEQ (n=104 stems)		MCQ (n=50)	
	n	Percentage	n	Percentage
Level I	41	39.4%	14	28%
Level II	21	20.2%	6	12%
Level III	42	40.4%	30	60%

Table 3: illustrates the item flaws of both MEQs and MCQs

Item flaw	MCQ	MEQ
1.ERROR FORMAT	1	NIL
2.ERROR SPELLING	NIL	4
3.ERROR GRAMMAR	NIL	NIL
4.TECHNICAL ERROR	2	4
5.DOUBLE NEGATIVES	3	NIL
6.CASCADE	NIL	NIL
7.ABSOLUTE OPTIONS	NIL	NIL
8.AMBIGIOUS	NIL	NIL
9.REPETITION	NIL	NIL
TOTAL	6/50 [12%]	8/50 [16%]

Graph 1: % of Questions according to specialties
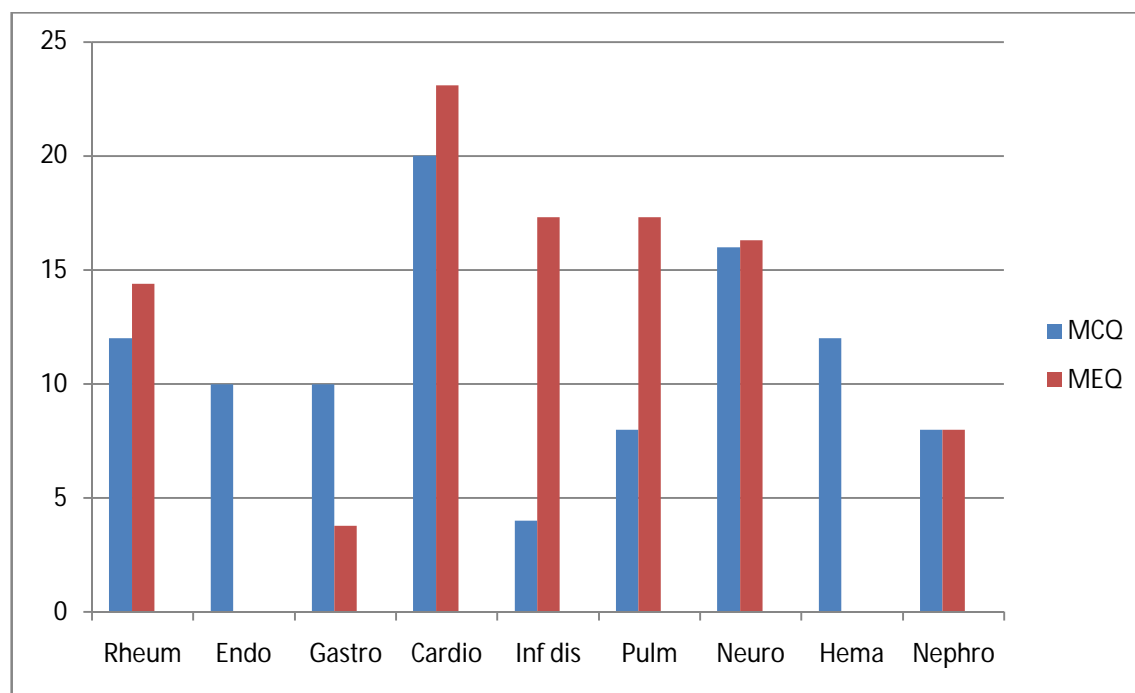
Discussion

The results of this study show that Multiple choice question is a better test of cognitive skills than the Modified essay question as 60% of the MCQs tested the higher cognitive skills and only 40% of MEQs addressed the cognitive level III of modified Bloom's taxonomy.

Although when the proportions of the two forms of questions addressing the different levels of cognitive domain were compared, there was no significant statistical difference among the level of cognition tested, as p-value was 0.07. It is also inferred that constructing an MEQ might be technically more difficult than an MCQ as item-writing flaws in MEQs were 16% as compared to MCQ only 12%.

The results of our study are consistent with Edward JP et al [12] who also found MCQ to be superior in testing the level 3 of cognitive domain. The results do not coincide with those of Irwin WG et al[13] who found MEQ to be superior in testing the highest level cognitive skills but they did not choose the questions randomly rather they compared the whole examination papers of various years. They suggest that both MCQ and MEQ may be designed beforehand to test any particular level of Bloom's taxonomy. Construction of MEQ needs expertise and training, model answers of such questions need meticulous consideration. [14, 15]

It also seems that some specialty examiners tend to put more stress on Level III testing or it is easier to design such questions for some specialties as most of level III questions in both MCQ and MEQ belonged to cardiology.

It is quite understandable that conduction of a proper assessment is not only dependent on the cognitive aspect of question, but there are so many factors which play a role like, reliability, content and construction validity, financial and human resources.

At the department of Medicine all the questions are constructed by the faculty members then these are put forward to a question review committee and after its approval the questions are introduced into the question bank. It was observed during the period of the study that the content and the construction of the questions improved over the years but still individual errors occurred.

The two confounders that played a major role in deciding the level of cognitive domain addressed by each specific type of question were (i) The type of examination to which the question belonged i.e. Midterm or Final (ii) The specialty of the question like cardiology, neurology etc. It may be inferred that in final exams the examiners tend to put the type of questions that address the highest level of cognitive domain (level III) as 80 % of MCQs belonged to the final exams while all MEQs were taken from the Midterm exam papers. This may be one drawback of our study. The total number of questions from which the sample was taken was quite limited which accounts for the small sample size in this study but as more and more questions will be incorporated into the question bank of the department further validation studies will be required.

Conclusion

MCQs were found to be testing the level III of cognitive domain more frequently than MEQs. Training in formulating MCQs and MEQs, is needed to assure achieving level III of cognitive domain and avoid item writing flaws.

References

1. Ebstein RM. Assessment in medical education. N Engl J Med 2007; 356:387-96.
2. Bloom BS (editors). Taxonomy of educational objectives: The classification of educational goals, Book 1. United Kingdom: Harlow, Longman Group; 1956.
3. Anderson LW, Krathwohl DR. Taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. New York, NY: Longman; 2001.
4. Hillard RI. The effects of examinations and the assessment process on the learning activities of undergraduate medical students [dissertation]. Department of education: University of Toronto; 1994.
5. Haladyna TM, Downing, SM, Rodriguez MC. A review of multiple-choice item-writing guidelines. Applied Meas Educ. 2002; 15:309-333.
6. Farley JK. The multiple choice test: writing the questions. Nurse Educator 1989; 14:10-12, 39.
7. Felliti GI, Smith EK. Modified essay questions: are they worth the effort? Med Educ. 1986 March; 20(2):126-32.
8. Felliti GI. Reliability and validity studies on Modified essay questions. Med Educ. 1980 November; 55 (11) 933-41.
9. Wilkinson, T. J. and Frampton, C. M. Comprehensive undergraduate medical assessments improve prediction of clinical performance. Medical Education. Oct 2004; 38: 1111–1116.
10. Hamdy H, Prasad K, Anderson MB, Scherpbier A, Williams R, Zwierstra R, Cuddihy H. BEME systematic review: predictive values of measurements obtained in medical schools and future performance in medical practice. Med Teach. 2006 Mar; 28(2):103-16.
11. Markert RJ. The relationship of academic measures in medical school to performance after graduation. Acad Med. 1993 Feb; 68(2 Suppl):S31-4
12. Edward JP, Peter GD: Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? BMC Medical Education 2007; 7:49:10.1186/1472-6920-7-49.
13. Irwin WG, Bamber JH. The cognitive structure of the modified essay question. Med Educ. 1982 Nov; 16(6):326-31.
14. Rabinowitz HK, Hojat M. A comparison of the modified essay question and multiple choice question formats: their relationship to clinical performance. Fam Med. 1989 Sep-Oct; 21(5):364-7.
15. Des Marchais JE, Vu NV. Developing and evaluating the student assessment system in the preclinical problem-based curriculum at Sherbrooke, Canada. Acad Med. 1996 Mar; 71(3):274-83.