Research Paper

Volume : 2 | Issue : 2 | Feb 2013 • ISSN No 2277 - 8179

Itemized Analysis of Questions of Multiple Choice Question (MCQ) Exam

Medical Science

KEYWORDS : MCQ, Item analysis, Difficulty Index, Discrimination Index

Dr. Kartik A. Patel	25, Savita Society, Nr. Naranpura Railway Crossing, Naranpura, Ahmedabad- 380013 Gujarat, India
Dr. Neeraj R. Mahajan	17, Devpriya Bunglows, Motera Stadium road, Motera, Sabarmati, Ahmedabad-380005 Gujarat, India

ABSTRACT

The present study is a cross-sectional study. The aim of present study was to statistically evaluate the quality of questions asked in an MCQ exam of I MBBS students of Smt. NHL Municipal Medical College, Ahmedabad. By measuring Difficulty Index (p) and Discrimination Index (DI) an examiner can know if an individual question was too difficult or too easy and whether it discriminated better performing students from less performing students. Thus this analysis of questions can be useful for development of better questions in the future. MCQ items having optimum Difficulty Index and Excellent Discrimination should be regarded as best framed questions.

INTRODUCTION

Multiple-choice Questions (MCQ) are widely used for MBBS students in colleges as classroom tests and as entrance test for under-graduate and post-graduate courses. A typical MCQ item consists of a question (stem) and a set of options that consist of possible answers to the question with single best correct answer. A student's task is to select the one option that provides the best answer to the question asked. A distinct advantage of using MCQ items on classroom tests is that grading tends to be quick and without subjective bias of evaluator. Another important advantage is that a well-constructed MCQ test can yield test scores at least as reliable as those produced by a constructed-response test, while also allowing for broader coverage of the topics covered in a course (Bacon, 2003). It is clear that thoughtfully written MCQ items can serve to assess higher-level cognitive processes, although creating such items does require more skill than writing memory-based items (Buckles & Siegfried, 2006; Palmer & Devitt, 2007). One criticism is that the format of MCQ items lets students guess even when they have no substantive knowledge of the topic under consideration (Biggs, 1999). However, Downing (2003) points out that blind guessing is quite uncommon on well-written classroom tests and informed guessing, which is based on a critical consideration of the question and the available options, provides a valid measure of student achievement (DiBattista & Kurzawa, 2011).

Use of MCQ as testing method in MBBS curriculum is increasing. So it becomes very important that quality of questions be maintained too. For that one may follow the widely accepted item-writing guidelines, such as putting the central idea of the question into the stem and avoiding the use of negation whenever possible (Haladyna, Downing, & Rodriguez, 2002). Another way to examine the quality of MCQ items involves analyzing the responses that examinees make, and this is the approach used in the research presented here.

AIMS AND OBJECTIVE

1. To check the quality of MCQ items on the basis of responses of students.
2. To identify properly framed questions and questions those need modifications.
3. To prepare the question bank of properly framed MCQ items.

MATERIAL AND METHOD

As a part of curriculum, 151 students of I MBBS of Smt. NHL Municipal Medical College, Ahmedabad were subjected to MCQ test. There were 50 questions in total all with 4 options including single best option which was considered as correct response. 1 mark was awarded for each correct response. There was no negative mark for incorrect response. No response was considered as incorrect response. All the responses of all the students were noted and quality of questions was analyzed on the basis of responses. We measured Difficulty Index (how difficult the question was for all the takers) and Discrimination Index (how well the question discriminated more knowledgeable students and less knowledgeable students). Most authors suggest that the discrimination coefficient should be at least +0.20 (Ding & Beichner, 2009; Su, Osisek, Montgomery, & Pellar, 2009; Thorndike, 2005). Accordingly criteria were defined for acceptable and unacceptable questions as following.

Difficulty Index (p) was measure as % of correct response from all the students for a particular question.

p = (no. of correct response/no. of total students) x 100 %

Range of p = 0% to 100%

If p is <30% or >70% → Unacceptable (MCQ item needs modification)

If p is between 30% to 70% → Acceptable

If p is between 50% to 60% → Optimum

For Discrimination Index (DI) measurement, merit was prepared on the basis of overall performance of the whole class. Top 50 and Bottom 50 students were identified and DI was calculated.

DI = (no. of correct response in top 50 students – no. of correct response in bottom 50 students)/50

Range of DI = -1 to 1

If DI is < 0.20 → Unacceptable (MCQ item needs modification)

If DI = 0.20 to 0.24 → Acceptable

If DI = 0.25 to 0.34 → Good discrimination

If DI = 0.35 or more → Excellent discrimination

OBSERVATION AND RESULT

No. of questions: 50

No. of students: 151

Table 1: Difficulty Index (p)

Range	No. of Qs
<30% or >70% (Unacceptable)	10
30%-70% (Acceptable)	40
50%-60% (Optimum)	14

Table 2: Discrimination Index (DI)

Range	No. of Qs
<0.20 (Unacceptable)	9
0.20-0.24 (Acceptable)	5
0.25-0.34 (Good discrimination)	16
≥0.35 (Excellent discrimination)	20

DISCUSSION

As seen in Table 1, there were 10 questions with Difficulty Index (p) <30% or >70%. So they required modification before they can be considered as standard questions.

40 questions were within acceptable range of Difficulty Index (p) and out of those 40, 14 questions were of Difficulty Index (p) between 50% and 60%. So they can be considered as optimum as far as difficulty is concerned.

As seen in Table 2, Discrimination Index (DI) of 9 out of 50 questions was below 0.20 and hence unacceptable.

DI of rest of the questions were >0.20 and so acceptable with 16 questions were categorized as having Good Discrimination (DI= 0.25 to 0.34).

20 questions were categorized as having Excellent Discrimination (DI= 0.35 or more)

CONCLUSION

It can be concluded from the present research that Difficulty Index (p) and Discrimination Index (DI) are very nice tool for the assessment of the quality of an MCQ item.

An MCQ item should be considered unacceptable and modified to get difficulty level and discrimination power within acceptable range before it can be included in a standard MCQ bank.

REFERENCE

Bacon, D. R. (2003). Assessing learning outcomes: A comparison of multiple-choice and shortanswer questions in a marketing context. Journal of Marketing Education, 25, 31-36. | Biggs, J. (1999). Teaching for quality learning at university. Buckingham, UK: Society for Research into Higher Education and Open University Press. | Buckles, S., & Siegfried, J. J. (2006). Using in-depth multiple-choice questions to evaluate indepth learning of economics. Journal of Economic Education, 37, 48-57. | DiBattista, David and Kurzawa, Laura (2011) "Examination of the Quality of Multiple-choice Items on Classroom Tests,"The Canadian Journal for the Scholarship of Teaching and Learning: Vol. 2: Iss. 2, Article 4. | Ding, L., & Beichner, R. (2009). Approaches to data analysis of multiple-choice questions. Physical Review Special Topics–Physics Education Research, 5, 020103. | Downing, S. M. (2003). Guessing on selected-response examinations. Medical Education, 37, 670-671. | Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item writing guidelines for classroom assessment. Applied Measurement in Education, 15, 309-344. | Palmer, E. J., & Devitt, P. G. (2007). Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? BMC Medical Education, 7, 49. | Su, W., Osisek, P. J., Montgomery, C., & Pellar, S. (2009). Designing multiple-choice test items at higher cognitive levels. Nurse Educator, 34, 223-227. | Thorndike, R. M. (2005). Measurement and Evaluation in Psychology and Education. Upper Saddle River, NJ: Pearson. |