Original Article

Title: Use of MCQs to assess the problem solving skills in clinical Biochemistry.

Rituparna Maji[1*], Soma Dutta[1], Archana Dhok[2]
1. *Dept. of Biochemistry, Calcutta National Medical College, Kolkata*
2. *Dept. of Biochemistry, Datta Meghe Institute of Medical Sciences, Wardha, Maharashtra*

A B S T R A C T

Introduction: Multiple choice questions (MCQs) are markedly effective assessment tools for assessing the knowledge and skills within a short period. However, their relative efficacy for assessing higher cognitive domain and readiness for problem solving are still inconclusive.

Methods: Effectivity MCQs was analyzed to assess different cognitive levels among 1st year MBBS student as regular formative assessment tools. 5 MCQs with high cognitive domain and 10 MCQs for memory based knowledge were given at the end of each learning session to all students throughout their 1st semester for 4 months. At the end of 1st semester examination, mean scores for each group were calculated for each student and compared with their semester marks obtained in their practical exams.

Results: In comparison to high cognitive MCQs, the average marks obtained by students in memory based MCQs were significantly higher. However, a significant positive association was found between the semester marks and scores of high cognitive MCQs only. On the other hand, students obtaining higher marks in memory based MCQs did not show any correlation with their semester scores of practical exams.

Conclusion: For regular end of the class formative assessments, MCQs become efficient tools only when they contain elements of higher cognition with problem solving approach. Regular use of MCQs with high cognitive domains help in improving the problem solving approach in the practical field   significantly.

INTRODUCTION

In the recent competency based medical education (CBME), the essential role of assessment is to provide effective feedback for developing a problem solving approach in clinical fields including the laboratory medicine. It is done by incorporating the important domains of continuing education, knowledge translation, patient safety, and quality improvement into the course curriculum1-3. Components of an effective formative evaluation are factors like regularity of the whole process of question formulation, conducting the examinations, evaluating the answers, finding out the lacunae in understanding of the students and taking their remedial actions at a regular basis as frequently as possible4 and finally to evaluate the outcome of the training by ascertaining the problem solving skills in clinical field.

Till date, most of the formative assessments are judicious and rational mixes of different types of long answer type questions (LAQ), short answered type questions (SAQs), very short answer type questions (VSAQs) and multiple choice questions (MCQs) with their own merits and demerits5. In the context of time constraint and regularity for the successful implementation of the CBME, MCQs have been more suitable assessment tools for ascertaining the problem solving skills as they are much less time consuming and can be arranged regularly after the end of each class or skill training session with more validity, reliability, feasibility and acceptability6-8. However, poor construction with over representation of the lower levels of cognition in most of the MCQs restrict their successful implementation as effective assessment tools for analyzing and improving different levels of knowledge of the learners9. As most MCQs commonly do not stimulate the students to synthesize or construct their own ideas, they focus mostly on judging the lower cognition levels of the knowledge domain, particularly their capability of rote memory10. Moreover, the common trend to construct most of the MCQs from the lower levels of knowledge domains also hampers the faculty development process by not enabling the faculties and lab instructors to construct MCQs with higher cognitive levels.

With this background knowledge, the research question for the present study was put forward which type of MCQs could be the better assessment tool for assessing the clinical problem solving skills in the field of Biochemistry among the 1st year undergraduate MBBS student and could predict a better outcome in their summative practical evaluations. Accordingly, it was hypothesized that there is a difference in effectivity between the high cognitive domain MCQs and memory based MCQs among undergraduates during their longitudinal formative evaluation. The present research project was designed with following aims and objectives to test this hypothesis.

*Correspondence:

rituparnamaji123@gmail.com
Dr. Rituparna Maji
Asst. Professor, Dept. of Biochemistry, Calcutta National Medical College, Kolkata

Methods

Objectives of the study were i) to evaluate comprehension of the topic by the students after the class by providing them MCQs on both memory based (medium cognition level) and problem based (higher cognition level) domains at end of the class and ii) to correlate these results with the results of their 1st semester practical examination. Accordingly the methodology for the present study was as follows:

Study design: A longitudinal, interventional educational study in a pre-clinical department of an urban medical college during the 1st Semester period of 4 months during February to May 2021.

Sampling & Randomization: 5 MCQs from the high cognitive domain and 10 MCQs from the memory based domain were provided to students on the same topic after end of each class so that the same knowledge can be tested using two different tools. Following the Bloom's criteria 10 questions from each group were constructed from average cognitive domain and 5 questions were constructed as problem solving ones containing higher cognitive domains. The students were asked to answer all questions within 10 minutes at the end of the topic in the class. This process continued after every theory class during the 1st 4 months of phase I MBBS after which the students had their 1st periodical semester practical examination. In the whole process, 211 students in each theory class were assessed using 5 plus 10 i.e. 15 questions each class for 4 classes each month for 4 months. Thus, at the end of the study from each student we received responses for 80 high cognitive MCQs and 160 memory based medium cognitive MCQs.

Ethical issues: The whole proposal was submitted to the institutional ethics committee (IEC). The study was started only after getting the written approval from the IEC. All students were explained about the study design and how to answer the questions within the limited time period. Informed consents from the students was implied as it was a part of routine formative evaluation of the curriculum.

Data collection method: After setting up of required guidelines for formulating appropriate memory based and problem based MCQs for each class, the assessment process after each class was undertaken as described above. Each correct question was credited 1 mark for both types of MCQs. No negative markings were applicable for this study. After each class the marks for each category were obtained. Marks for both groups of MCQs were categorized separately into categories of higher and medium levels of cognition domain. At the end of 4 months, practical marks of the 1st semester (1st Internal Assessment) for each student were collected.

Data analysis plan: Both types of MCQs were provided to the students using google forms with randomized question sequence. MCQs were marked using the automated marking system in the google form. At the end, mean scores of all MCQs from both groups were calculated for all students and were compared by an independent t test for any significance of difference in between them.

Mean scores of both types of MCQs were also obtained separately for each student and were assessed for any significant association with marks obtained in their 1st semester internal assessment practical exam using bivariate correlation analysis.

For all statistical tests, the P value was considered significant at a level of $P < .05$ for a 95% of confidence interval.

Results

Before performing statistical analyses, the data were analysed for their trend of distribution using the Smirnov Kolmogrov test for normal distribution. The results showed that all data were distributed normally ($P > .05$ for all, data not shown in the Table). Therefore, appropriate parametric methods for statistical analyses were used in this study.

The comparison was made firstly between both groups of MCQs for all students (Table 1). It showed that the mean score for all students in memory based MCQs was significantly higher ($P < .001$) than that obtained in for the high cognitive MCQs.

Table 1: Difference between the average marks of memory based MCQs and high cognitive MCQs (percentage) for all students (n = 211).

Type of questions	Mean (percentage) ± SD	P value
High cognitive MCQs	70.2 ± 8.3	< .001*
Memory based MCQs	80.0 ± 6.9	
*Independent t test; P value is considered significant at P < .05 for 95% CI		

In Table 2 the same difference is shown for the students obtaining higher than or equal to 90 percent of marks in the high cognitive MCQs (Group I). The difference is still significant ($P < .001$) in spite of increase in their corresponding average scores.

Table 2: Difference between the average percentage marks of high cognitive and memory based MCQ in students obtaining more than 90 percent in the high cognitive MCQs (Group I, n = 32).

	Mean ± SD	P value
High cognitive MCQs	94.5 ± 2.1	< .001*
Memory based MCQs	96.7 ± 2.6	
*Independent t test; P value is considered significant at P < .05 for 95% CI		

Figure 1: Mean scores of high cognitive and memory based MCQs in all students, (n = 211) and Group I students (scoring more than 90% in MCQs with higher cognitive levels, n = 32)
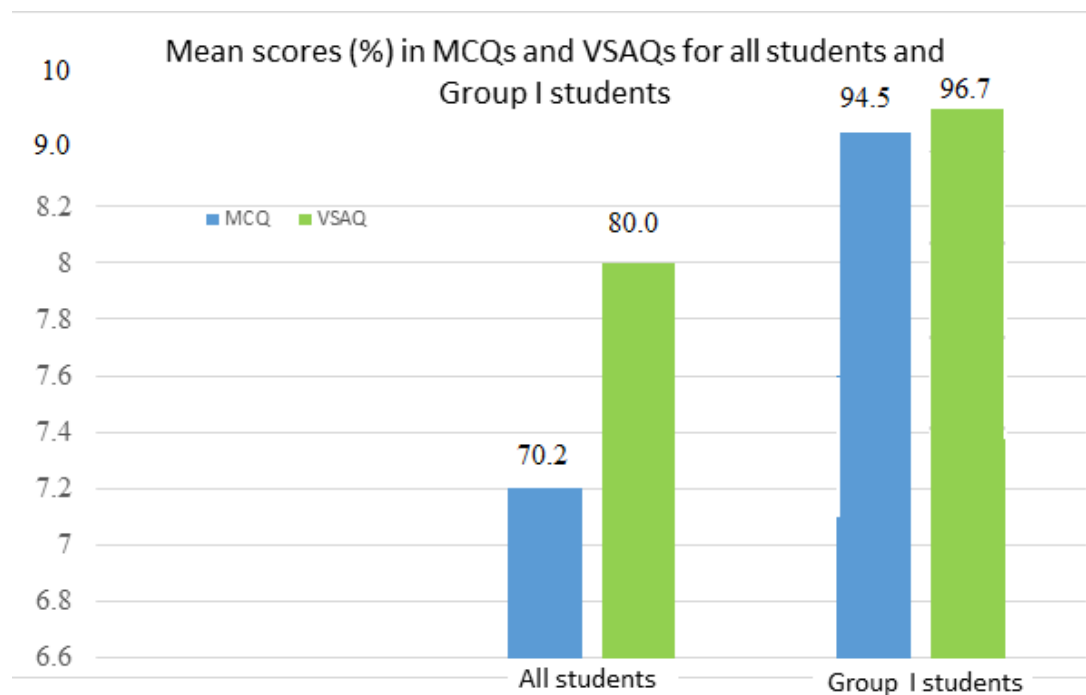
After observing this general trend for scoring better in memory based MCQs for all students, association of this outcome with their marks in their next periodical practical semester examination was analysed using Pearson's bivariate correlation study. Firstly, the correlation was done between the average marks of both groups of MCQs and the semester marks for all students. The results are shown in the Table 3.

Table 3: Correlation between the average marks of high cognitive and memory based MCQs and semester for all students.
Correlations

		High cognitive MCQ	Memory based MCQ	SEMESTER
High cognitive MCQ	Pearson Correlation	1	.170*	.123
	P (Sig.) (2-tailed)		.013	.075
	n	211	211	211
Memory based MCQ	Pearson Correlation	.170*	1	.142*
	P (Sig.) (2-tailed)	.013		.039*
	n	211	211	211
SEMESTER	Pearson Correlation	.123	.142*	1
	P (Sig.) (2-tailed)	.075	.039*	
	n	211	211	211

*. Correlation is significant at the 0.05 level (2-tailed).

The results showed that when all students were considered, their memory based MCQ marks were significantly positively correlated to their semester marks (P = .039). However, no such correlation was observed with their scores obtained in high cognitive MCQs (P = .075). But, results in the Table 4 which analysed the correlation among these parameters for the group I students obtaining more than 90 percent in MCQs with higher cognitive domain, a significantly positive association was found between their MCQ scores and semester marks (P = .049).

Table 4: Correlation between the mean marks of semester with memory based MCQs and MCQs with high cognitive domains in group I students (n = 32).

Correlations

		High cognitive MCQs	Memory based MCQs	Semester practical exam
High cognitive MCQs	Pearson Correlation	1	.393*	.350*
	Sig. (2-tailed)		.026	.049*
	N	32	32	32
Memory based MCQs	Pearson Correlation	.393*	1	.520**
	Sig. (2-tailed)	.026		.002*
	n	32	32	32
Semester practical exam	Pearson Correlation	.350*	.520**	1
	Sig. (2-tailed)	.049	.002	
	n	32	32	32

*. Correlation is significant at the .05 level (2-tailed).
**. Correlation is significant at the .01 level (2-tailed).

Discussion

The present study was aimed to find out the effect of practising regular internal assessments at the end of the laboratory skill training sessions using MCQs with high and medium cognitive domains. The outcome of this practice was analysed by comparing the scores of both groups of MCQs with the periodic quarterly semester skill development examination marks. For the convenience of comparison between the high performers and medium performers in the MCQs a cut off value of 90 percent marks was decided after peer discussion and expert opinions from experienced faculties. Students who secured either equal to or more than 90 percent marks were assigned Group I.

From the results of Table 1 it was evident that the average score of all students for the memory based (80% ± 6.9) were significantly higher (P < 0.001) than their average scores in high cognitive MCQs (70.2% ± 8.3). Furthermore, when the average marks of both groups of MCQs in good performers (Group I) were compared (Table 2) the same trend was observed albeit with increase in their average scores. Interestingly, it was also evident from the Table 2 and Figure 1 that increase in the mean scores of memory based MCQs in the group I students were comparatively lesser than the increase in their mean score of MCQs with high cognitive domain. It suggested that irrespective of the cognitive levels students have a trend of scoring higher in memory based MCQs having low to medium range of cognitive domains. Moreover, as evident from the bivariate correlation analysis (Table 3) the semester scores of all students showed a significant positive correlation with the memory based MCQs (P = .039) only without any significant correlation with the MCQs with high cognitive domain (P = .07). This strongly indicated that irrespective of cognitive classification, regular formative assessments at the end of practical and laboratory skill learning session using low to medium domain MCQs are more high scoring than MCQs with higher cognitive domains for preparing the learners for their decisive summative evaluation.

But, when marks of students who scored significantly better in the MCQs with higher cognitive contents (Group I students) were analysed for their association with their marks obtained in their periodical semester exam it was found to show significant positive correlation (P = .049, Table 4). It is evident from these findings that the association that was insignificant between the scores of MCQs and semester exam for all students of phase I MBBS became significant for the students who scored significantly better in MCQs with higher cognitive domains. As expected, correlation of the semester marks with memory based MCQs was also highly significant in group I students (P = 0.002) and it was in fact, higher than that for the average students (P = .039) as found in the Table 3.

The findings of the present study suggests that MCQs can play the role of effective skill assessment tools only when they have higher contents of cognitive domain. But the major limitation of formulation of MCQs with higher cognitive domains are need of significantly more time, effort, peer validation before use and post validation after tests which restrain their formulation and incorporation frequently in most of the assessments. Formulating good MCQs with higher cognitive domains need more skill, practice and training of the concerned faculties which is still not practiced at many places. It has been reported that almost 90% of the MCQs were from lower cognitive domain in some of the important assessments whereas 60% of them were from recall level in others [11]. Another study has revealed that only 10% of MCQs fulfill all of the criteria for a good, valid and reliable MCQ (Patil et al., 2016). However, if the MCQs are formulated at higher cognitive domains according to the Bloom's criteria they are found to contain less faults and better discriminating ability and vice versa [12]. Thus, it is obvious that most of the MCQs which are constructed from lower domains of knowledge at present cannot perform a successful and effective assessment either formative or summative. Rather they negate the objectives of successful assessment and feedback tools. Using MCQs with higher content and construct validity provide much more cognitive and knowledge assessment skills [13]. This is the explanation of the present observations in this educational project that revealed the essentiality of formulation of MCQs with higher cognitive levels and utilizing them regularly as formative assessment tools. It also explains why the candidates who become well acquainted with MCQs with higher cognitive domain show better performance in their periodical semester examinations for practical and laboratory skills and hence are expected to perform better in their final summative evaluation also.

References:

1. Holmboe ES, Sherbino J, Long DM, Swing SR, Frank JR. The role of assessment in competency-based medical education. Med Teach. 2010;32(8):676-82Available from: https://www.ncbi.nlm.nih.gov/pubmed/20662580.

2. Hauer KE, Boscardin C, Brenner JM, van Schaik SM, Papp KK. Twelve tips for assessing medical knowledge with open-ended questions: Designing constructed response examinations in medical education. Med Teach. 2020;42(8):880-5Available from: https://www.ncbi.nlm.nih.gov/pubmed/31282798.

3. Danilovich N, Kitto S, Price DW, Campbell C, Hodgson A, Hendry P. Implementing Competency-Based Medical Education in Family Medicine: A Narrative Review of Current Trends in Assessment. Fam Med. 2021;53(1):9-22Available from: https://www.ncbi.nlm.nih.gov/pubmed/33471918.

4. Puthiaparampil T, Rahman MM. Very short answer questions: a viable alternative to multiple choice questions. BMC Med Educ. 2020;20(1):141Available from: https://www.ncbi.nlm.nih.gov/pubmed/32375739.

5. Assessment methods in medical education. Int J Health Sci (Qassim). 2008;2(2):3-7Available from: https://www.ncbi.nlm.nih.gov/pubmed/21475483.

6. Case S, Swanson D. Constructing Written Test Questions For the Basic and Clinical Sciences. National Board of Examiners. 2002.

7. Sam AH, Hameed S, Harris J, Meeran K. Validity of very short answer versus single best answer questions for undergraduate assessment. BMC Med Educ.

2016;16(1):266Available from: https://www.ncbi.nlm.nih.gov/pubmed/27737661.

8. Van der Vleuten CPM. The assessment of professional competence: developments, research and practical implications. Adv Health Sci Educ. 1996; 1: 41-67. .

9. Pickering SG. Against multiple choice questions. Med Teach. 1979;1(2):84-6Available from: https://www.ncbi.nlm.nih.gov/pubmed/24479932.

10. Husnain A, Khan A, Khan MU, Hussain FN. Subjective quality of multiple choice questions used in undergraduate courses in orthopedics and other specialties. Pak J Med Sci. 2020;36(7):1618-22Available from: https://www.ncbi.nlm.nih.gov/pubmed/33235585. 24

11. Tarrant M, Ware J. Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessments. Med Educ. 2008;42(2):198-206Available from: https://www.ncbi.nlm.nih.gov/pubmed/18230093.

12. Baig M, Ali SK, Ali S, Huda N. Evaluation of Multiple Choice and Short Essay Question items in Basic Medical Sciences. Pak J Med Sci. 2014;30(1):3-6Available from: https://www.ncbi.nlm.nih.gov/pubmed/24639820.

13. Khan JS, Mukhtar O, Tabasum S, Shaheen N, Farooq M, Irfan MA, et al. Relationship of awards in multiple choice questions and structured answer questions in the undergraduate years and their effectiveness in evaluation. J Ayub Med Coll Abbottabad. 2010;22(2):191-5Available from: https://www.ncbi.nlm.nih.gov/pubmed/21702301.