

# EVALUATION OF MCQs FOR JUDGMENT OF HIGHER LEVELS OF COGNITIVE LEARNING

Mousumi Mukhopadhyay, Kaushik Bhowmick, Sandip Chakraborty, Debes Roy,  
Pradyot Kumar Sen, Indranil Chakraborty

Department of Biochemistry, Medical College Kolkata, Kolkata, West Bengal, India

## ABSTRACT

**Background:** Reliable and valid written tests of higher cognitive function are difficult to produce, particularly for assessment of clinical problem solving. The present study utilized a test instrument MCQs and the objective was to statistically validate the designed MCQs and to assess whether MCQ may be utilized for assessing higher levels of cognitive learning.

**Methodology:** This study was conducted at Department of Biochemistry, Medical College Kolkata, India, in 2009. Undergraduate medical students of second semester were selected for study. MCQs were constructed after addressing the three principal areas viz. preparing preset educational objectives, defining the levels of learning for each objective and writing effective MCQs to test the learning. All questions were pre-validated by peer review and segregated into three classes based on their ability to test the three levels of cognitive learning. Thirty-five items were included, 20 on Recall, 10 on data interpretation and 5 on problem solving.

**Results:** Item analysis measured the difficulty index, discrimination index and distracter effectiveness for each MCQ. The average percentage of correct response for MCQs on Recall, Data interpretation and Problem solving was 56.59, 47.81 and 42.25 respectively. Chi square test showed that these differences were statistically insignificant. This finding justifies that the items were effective in testing higher levels of cognitive skills including problem solving.

**Conclusion:** Well-constructed, peer-reviewed multiple choice questions meet the educational requirements and are advocated for assessing medical students.

**KEY WORDS:** Medical students, Evaluation, MCQs, Cognitive learning.

## INTRODUCTION

Medical Education is more than a job training programme. In addition to the knowledge content, the concepts of professional medical education emphasizes on the development of skills of scientific methods as applied to biological sciences.<sup>1</sup> Accordingly, the undergraduate medical curriculum in India devised by Medical Council of India (MCI) is oriented towards training students to undertake the responsibilities of a physician of first contact who is capable of looking after the preventive, promotive, curative and rehabilitative aspects of Medicine.<sup>2</sup>

The faculty in the medical institutions is entrusted with the responsibility to define the skills of the competent physician, provide experiences that promote acquisition of these skills, and develop assessment tools that provide feedback to students and teachers of what have accomplished. These considerations have become increasingly important in the backdrop of an ever expanding knowledge base and accelerated information output.<sup>3</sup>

The problem is compounded by the fact that we do not have an opportunity to lengthen the duration of medical school. In the light of the above facts, it is evident that medical education is an area where a close interaction between students and teachers is essential. What is required is not an endless list of factual information to be learnt but rather a well planned set of concepts as an introduction to the world of factual knowledge that leads to a behavioral transformation of the student so that when he finishes education, he is a changed person in head, hand and heart. By this we mean, a student must develop problem solving capacity through his acquired skills, guided by cognition and feeling.

But how much basic science does a graduate need to know to be a competent practitioner? The objectives must be derived from what the physician is called upon to do and not from the excess of factual knowledge that the faculty may be biased with. MCI has recommended 240 hours devoted to teaching Biochemistry in the undergraduate syllabus.<sup>2</sup> The goal of teaching Biochem-

istry to undergraduate students is to make them understand the scientific basis of life processes at the molecular level and to orient them towards the application of the knowledge acquired in solving clinical problems.<sup>2</sup> So at the end of the course, the student should be able to recall the data, analyze and interpret investigative data and demonstrate the intellectual skills of solving scientific and clinical problems and decision making.

One of the three major components of educational process is evaluation. It includes measurement plus value judgment. Why do we evaluate? Feedback from evaluations is of help to students in planning further learning and to the faculty in assessing their instructional efforts. Moreover, the society which finally pays the bill for health activity relies on teachers to train doctors to perform a social function corresponding to the health needs and demands of the community. The evaluator must adopt a realistic approach and decide what is essential in a particular situation. Thus, for measurement, value judgment and better educational decision making, construction of effective evaluation protocol is of utmost importance. Unfortunately there is no single ideal evaluation method that can assess the complete set of skills indispensable to physician competence. Consequently, evaluation methods should be chosen from an array of options to satisfy specific evaluation needs. A combination of approaches may provide a more comprehensive assessment.<sup>4</sup>

Principle of evaluation have been traditionally based on Bloom's taxonomy of test development to assess three hierarchical levels of cognitive learning represented respectively by questions on knowledge (recall or recognition of specific information), data interpretation (combined comprehension and application) and problem solving (transferring existing knowledge and skills to new situation).<sup>5</sup> In this respect, it may be pertinent to observe that reliable and valid written test of higher cognitive functions have been difficult to produce particularly for the assessment of clinical and problem solving skills. It has been observed by Palmer and Devitt that Modified Essay Type Questions (MEQ) failed in its role of consistently assessing higher cognitive skills whereas Multiple Choice Question (MCQ) frequently tested more than the mere recall of knowledge.<sup>6</sup> Moreover, the acquiring of problem solving skills may be further compounded by language barrier in case of writing the examination with MEQs. In this case the student may comprehend the problem but may find it difficult to express effectively if they received the school education in the native language medium. Khattak AM et al has similarly stressed the need of

considering the medium of education in school while assessment of performance of the medical students.<sup>7</sup> This led to the formulation of an MCQ as a test instrument in an effort to decrease the significance of this confounding factor in effective assessment.

The present study utilized a test instrument devised in the form of MCQs and was undertaken with dual objectives. The first objective was to statistically validate the designed test MCQs to find out the degree to which they were appropriate, meaningful and useful. Secondly, it was attempted to assess whether MCQ may be utilized as a valid instrument for assessing higher levels of cognitive learning.

### MATERIAL AND METHODS

The study was performed at the Department of Biochemistry, Medical College Kolkata in the year 2009. Undergraduate medical students of second semester who had completed eight months of curriculum were selected for the study. Among the total batch of 159 students, 151 students who appeared for the second semester examination in Biochemistry were included after informing them about the purpose of the study and obtaining their consent. A set of questionnaire devised in the form of MCQs were utilized. The MCQs had been constructed after addressing the three principal areas viz. preparing of preset educational objectives, defining the levels of learning for each objective and writing effective MCQs to test the learning. Items were formulated by the teaching faculty of the department and a question bank was generated. While constructing the MCQs, adequate care was exercised to avoid item-writing flaws like repetition of part of the stem in an option, complicated or ambiguous stem, use of double negatives, etc. All questions in the bank were pre-validated by peer review and segregated into three classes based on their ability to test the three taxonomic levels of cognitive learning according to Bloom's classification. Finally a total of 35 items were included comprising of 20 items on Recall (RC), 10 items on data interpretation (DI) and 5 items on problem solving (PS) in the relative proportion of 20:10:5 as per specification for test construction outlined by Gilbert JJ.<sup>8</sup>

The filled up response sheets were collected and statistically analyzed for post-validation and for computing the results.

### RESULTS

Item analysis measured the difficulty index, and discrimination index for each MCQ. (Table 1)

**Table 1: Results of item analysis for MCQs on Recall, Data interpretation and problem solving.**

Total respondents – 151, No. of High achievers – 47, No. of Low achievers – 43

Category	Difficulty Index (%)			Discrimination Index (ratio)		
	Reference Interval	Interpretation	Observed value	Reference Interval	Interpretation	Observed value
<b>Recall (n = 20)</b>	>70	Easy	7	>0.35	Excellent	10
	30-70	Acceptable	11	0.25 - 0.35	Acceptable	5
	<30	Difficult	2	<0.25	poor	5
	Reference Interval	Interpretation	Observed value	Reference Interval	Interpretation	Observed value
<b>Data Interpretation (n = 10)</b>	>70	Easy	1	>0.35	Excellent	5
	30-70	Acceptable	7	0.25 - 0.35	Acceptable	2
	<30	Difficult	2	<0.25	poor	3
	Reference Interval	Interpretation	Observed value	Reference Interval	Interpretation	Observed value
<b>Problem Solving (n = 5)</b>	>70	Easy	1	>0.35	Excellent	3
	30-70	Acceptable	2	0.25 - 0.35	Acceptable	1
	<30	Difficult	2	<0.25	poor	1
	Reference Interval	Interpretation	Observed value	Reference Interval	Interpretation	Observed value

**Table 2: Correct response (%) obtained for three taxonomic levels of cognitive learning.**

	Recall (n = 20)	Data interpretation (n = 10)	Problem solving (n = 5)
Total of Correct Responses (for all questions)	1709 out of a possible 3020*	722 out of a possible 1510*	319 out of a possible 755*
Average Correct Response (per question)	85.45	72.22	63.83
Average Correct Response (expressed as %)	56.59	47.81	42.25

\* Estimated Total in a probability when all responses by all learners are correct.

**Table 3: Comparison of learner performance at different taxonomic level of cognitive skill.**

Statistical values	Comparison performance of	
	RC and DI	DI and PS
Value of Chi square	1.624	0.727
p-value	>0.05	>0.05

After validation, it was tested whether problem solving skills may be adequately evaluated through the use of MCQs. The average percent-

age of correct response for MCQs on Recall, Data interpretation and Problem solving was 56.59, 47.81 and 42.25 respectively. Results of chi square test showed that these differences were statistically insignificant. This finding justifies that the items were effective in testing higher levels of cognitive skills including problem solving. (Table 2 & 3)

**DISCUSSION**

In the context of the present study, the test MCQs required to be validated in order to ascertain the degree to which they measured the educational objectives sought to be tested. In practice, validity has been described as a simple evi-

dence for inferences made about a test score.<sup>9,10</sup> These evidences may be content related, criterion related or construct related. As content or curricular validity (CV) is appropriately applied to proficiency tests, assessment of CV was performed as the indicator of overall validity in the present study.<sup>11</sup>

Established statistical measures for item analysis were performed and the difficulty index and discrimination index for each MCQ calculated for post-validation. These results for MCQs on Recall, Data Interpretation and Problem Solving are given in Table 1.

The taxonomy of cognitive learning published by Bloom, was described as a hierarchy of knowledge, comprehension, application, analysis, synthesis and evaluation.<sup>5</sup> On adoption of Bloom's taxonomy for test development, educators have simplified the above into three general levels represented by three categories.<sup>12</sup> These three levels include knowledge (recall or recognition of specific information), data interpretation (combined comprehension and application) and problem solving (transferring existing knowledge and skills to new situations). Each MCQ in the evaluation instrument was designed to test at the same level of learning as the objective it was designed to assess.

The statistical evaluation with an objective to compare the performance of the learners across the categories is depicted in tables 2 & 3. Table 2 shows the percentage of correct response obtained in respect to three taxonomic levels relative to an estimated total in a probability when all responses by all learners are correct. The average percentage of correct response for R, DI and PS was 56.59, 47.81 and 42.25 respectively. Though these values apparently differed, results of chi square test showed that these differences were statistically insignificant. This finding may serve to justify that the items were effective in testing higher levels of cognitive skill including PS.

The present study has confirmed the impression reported by Palmer EJ et al and several other workers that MCQs may be utilized to test knowledge as much as they measure higher cognitive skill.<sup>6,13,14</sup>

It has been generally thought that the MCQs focus on knowledge recall but such criticisms may be borne more from poor item construction rather than other contributing factors.<sup>15,16</sup> The present study suggests that if the authors are able to construct MCQs of satisfactory quality, the assessment of various levels of cognitive learning may be better served by this test format. In addition, the reduced effort in marking such an assessment might be of

benefit for institutes with limited faculty and resources.

On the other hand acquiring of problem solving skills are valued component of a good medical education, while an effective assessment process for clinical competence in problem solving abilities have been one of the most difficult areas to measure and quantify.<sup>13</sup> This is because reliable and valid written tests of higher cognitive function are difficult to produce. MCQs have been one of several tools developed to try and assess this skill.<sup>17</sup>

It has been documented that a good MCQ is difficult to write and many may contain item writing and other flaws. The present study has shown that this does not necessarily have to be the case if adequate care is exercised during construction and subsequent moderation of items. Analysis from similar studies have documented that it is indeed possible to produce an MCQ paper that tests a broad spectrum of a curriculum, measures a range of cognitive skills and does so, on the basis of structurally sound questions.<sup>15,16</sup>

It is important to recognize that these results are from one institution only and the processes used to design assessments may not be typical of other institutions. However it was not possible to isolate individual authors, at least a dozen individuals were involved, and there was little variation in the overall Bloom categorization of the questions.

Moreover as validation imparts generalizability, the findings of this study may be transferable to other schools also, especially within the subcontinent with similar socio-economic and academic setup.

## CONCLUSION

Well constructed, peer reviewed multiple choice question meet many of the educational requirements and are advocated for assessing medical students.

## REFERENCES

1. Neville AJ. Basic science and medical education: Dinosaurs, Departments and Definitions- A McMaster View. *The Meducator* 2002; 2: 6-9.
2. Medical Council of India. Regulations on graduate medical education 1997. New Delhi, India. pp 2, 21, 95.
3. Puri D. An integrated problem based curriculum for Biochemistry teaching in medical sciences. *Indian J Clin Biochem* 2002; 17: 52-9.
4. Mandin H. The integration of basic and clinical sciences. *The Meducator* 2002; 2: 2-3.

5. Bloom B, Englehart M, Furst E, Hill W, Krathwohl D. Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain. New York, Toronto: Longmans, Green; 1956.
6. Palmer EJ, Devitt PG. Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? *BMC Med Educ* 2007; 7: 49.
7. Khattak AM, Wazir F, Khan H, Shaukat A, Shah SH. Effect of medium of education during school on performance of students in medical college. *Gomal J Med Sci* 2005; 3: 44-7.
8. Guilbert JJ. Test construction specification table. In: Guilbert JJ ed. Educational handbook for health personnel. WHO Geneva 1991; p. 4.54.
9. Lindsquist EF. In: Educational measurement. Lindquist EF ed. Washington: American Council of Education 1951; p. 141.
10. Singh AK. Validity. Singh AK ed. In: Tests, measurements and research methods in behavioral sciences, Patna, India: Bharati Bhawan Publishers, 2006; p. 93.
11. Anathkrishnan N. Multiple choice questions: facts and fantasies. In: Anathkrishnan N, Sethuraman KR, Kumar S eds. Medical education principles and practice, Pondicherry, India: Alumni Association of NTTCC, JIPMER, Pondicherry, 2000; p. 126.
12. Buckwalter JA, Schumacher R, Albright JP, Cooper RR. Use of an educational taxonomy for evaluation of cognitive performance. *J Med Educ* 1981; 56: 115-21.
13. Marshall J. Assessment of problem-solving ability. *Medical Education* 1977; 11: 329-34.
14. Rabinowitz HK, Hojat MD. A comparison of the modified essay question and multiple choice question formats: Their relationships to clinical performance. *Fam Med* 1989; 21: 364-7.
15. Palmer EJ, Devitt P. Constructing multiple choice questions as a method for learning. *Ann Acad Med Singap* 2006; 35: 604-8.
16. Collins J. Education techniques for lifelong learning: writing multiple-choice questions for continuing medical education activities and self-assessment modules. *Radiographics* 2006; 26: 543-51.
17. Rabinowitz HK. The modified essay question: an evaluation of its use in a family medicine clerkship. *Medical Education*. 1987; 21: 114-8.

**Corresponding author:**

Dr. Indranil Chakraborty  
Block A/2, Flat No. 7  
Uttarayan Housing Estate  
102, BT Road, Kolkata 700035  
West Bengal, India  
E-mail: drindranilchakraborty @rediffmail.com