Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(3), 250-253 | ISSN:2251-6727

Unlocking Cognitive Potential: Validating Multiple-Choice Assessments for Higher-Order Cognitive Skills in Medical Education

Pal M [1], Sukul S [2], Maji B N [3], Sarkar Mahata J [4], Haldar S [4], Das P [4], K. R. Sethuraman[5]

[1:] Associate Professor, Department of Biochemistry, Burdwan Medical College, Burdwan, and MHPE Scholar,Institute of Health Professions Education, Puducherry, Sri Balaji Vidyapeeth, (Deemed to be University) Pondicherry, India

[2:] Assistant Professor, Burdwan Medical College,Burdwan,WB and MHPE scholar,DMIHER,Wardha,Maharashtra

[3:] Assistant Professor, Burdwan Medical College,Burdwan,WB

[4:] Senior Resident, Burdwan Medical College,Burdwan,WB

[5:] Professor, Endowed Chair, Institute of Health Professions Education, Puducherry, Sri BalajiVidyapeeth (Deemed to be University) Pondicherry, India

(Received: 16 March 2025 Revised: 20 April 2025 Accepted: 01 May2025)

KEYWORDS

Medical students, Assessment, MCQs, Cognitive learning

ABSTRACT:

Background: Designing reliable and valid written assessments to test higher-order cognitive skills—especially clinical problem-solving—remains a challenge in medical education. This study aimed to evaluate the effectiveness of multiple-choice questions (MCQs) in assessing different levels of cognitive learning by applying statistical validation methods.

Methodology: Conducted in 2025 at the Department of Biochemistry, Burdwan Medical College, Burdwan, the study involved second-semester undergraduate medical students. MCQs were developed based on clearly defined educational objectives and categorized according to the level of cognitive skill they intended to assess: recall, data interpretation, and problem-solving. All questions underwent peer review to ensure quality. A total of 50 MCQs were used, 25 targeting recall, 15 on data interpretation, and 10 designed to test problem-solving.

Results: Each item was analyzed for difficulty, discrimination, and distractor effectiveness. The average percentage of correct responses was 80 % for recall, 52.9% for data interpretation, and 39.6 % for problem-solving questions. Statistical analysis using the Chi-square test revealed no significant difference between these scores, indicating that the MCQs were effective across different cognitive levels, including higher-order thinking.

Conclusion: Properly designed and peer-reviewed MCQs can serve as reliable tools for evaluating not just factual knowledge but also higher-level cognitive skills like interpretation and problem-solving in medical students.

Introduction

Medical education is a dynamic and multifaceted process, aiming to foster not only the acquisition of knowledge but also the application of that knowledge in real-world clinical settings. Medical students must be proficient in interpreting data, solving complex problems, and making informed decisions under pressure.(1,2) This highlights the importance of assessing higher-order cognitive skills like clinical reasoning and problem-solving, which go beyond rote memorization(3).

In India, Biochemistry forms a key pillar of medical education, laying the groundwork for understanding clinical sciences. Despite its foundational importance, the assessment methods employed to evaluate these higher-order cognitive skills are often inadequate(4,5). Traditional assessments predominantly measure rote

Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(3), 250-253 | ISSN:2251-6727

learning and recall, which fail to assess the ability of students to apply knowledge in clinical settings.

Multiple-choice questions (MCQs) have gained traction in medical education as a reliable method of assessment (5,6,7) While typically associated with testing factual knowledge, well-designed MCQs can be structured to assess various cognitive domains, including recall, data interpretation, and clinical problem-solving. However, the challenge lies in designing MCQs that are capable of evaluating these higher cognitive functions. The primary aim of this study was to validate a set of MCQs designed to assess cognitive learning in undergraduate medical students, focusing on their ability to interpret biochemical data and solve clinical problems.

By using statistical analysis to assess the reliability and validity of these MCQs, this study seeks to demonstrate that MCQs can effectively evaluate both basic knowledge and higher cognitive skills, thus supporting their use in comprehensive medical assessments.

Materials and Methods

The study was conducted in 2025 at Burdwan Medical College, located in Burdwan, West Bengal. The participants included second-semester undergraduate medical students enrolled in the Biochemistry course. Ethical approval was obtained from the Institutional Ethics Committee [Memo No.BMC/2179/1], and informed consent was secured from all participants.

The MCQs used in the study were developed by a team of experienced faculty members from the Department of Biochemistry. These questions were constructed in alignment with the curriculum objectives of the Biochemistry course, ensuring that they adequately represented both fundamental knowledge and higher-order cognitive skills as outlined by Bloom's taxonomy (1). The MCQs were divided into three categories:

Recall: 25 questions focused on basic factual knowledge and recall of biochemical principles.

Data Interpretation: 15 questions designed to assess the ability to interpret biochemical data, such as laboratory results and diagnostic tests.

Problem-Solving: 10 questions aimed at evaluating the students' clinical reasoning and ability to apply biochemical concepts to solve clinical problems.

After developing the MCQs, the questions were reviewed and validated by a panel of peer educators to ensure clarity, relevance, and accuracy. A pilot test was conducted with a small group of students to evaluate the clarity and difficulty of the questions.

The MCQ test was then administered to 151 second-semester students. The test duration was 90 minutes, and each student completed the test individually. Afterward, the responses were scored, and the following statistical parameters were computed for each question:

Difficulty Index (P-value): This was calculated by dividing the number of correct answers by the total number of respondents.

Discrimination Index (D-value): This was calculated by comparing the performance of the upper and lower 27% of students.

Distractor Effectiveness: This was analyzed by examining the frequency with which each distractor was chosen.

The data were then analyzed using SPSS software (version 25.0) to assess the reliability and validity of the MCQs. A Chi-square test was used to compare the performance across the three cognitive domains (recall, data interpretation, and problem-solving) and determine any significant differences in their effectiveness.

Results

the analysis of the MCQs revealed the following findings:

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)**		
Recall (n = 25)	Reference Interval	Interpretation	Observed value	Reference Interval	Interpretation	Observed value
	>70	Easy	8	>0.35	Excellent	11
	30-70	Acceptable	13	0.25 - 0.35	Acceptable	9

Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(3), 250-253 | ISSN:2251-6727

	<30	Difficult	4	<0.25	poor	5
Data Interpre-tation (n = 15)	Reference Interval	Interpretation	Observed value	Reference Interval	Interpretation	Observed value
	>70	Easy	2	>0.35	Excellent	7
	30-70	Acceptable	10	0.25 - 0.35	Acceptable	5
	<30	Difficult	3	<0.25	poor	3
Problem Solving (n = 10)	Reference Interval	Interpretation	Observed value	Reference Interval	Interpretation	Observed value
	>70	Easy	2	>0.35	Excellent	4
	30-70	Acceptable	5	0.25 - 0.35	Acceptable	4
	<30	Difficult	3	<0.25	poor	2

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

	Recall (n = 25)	**Data interpretation (n = 15)**	**Problem solving (n = 10)**
Total of Correct Responses (for all questions)	3020 out of a possible 3775*	1200 out of a possible 2265*	598 out of a possible 1510*
Average Correct Response (per question)	87.28	68.94	66.18
Average Correct Response (expressed as %)	80.0	52.9	39.6

** 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.

Comparison of performance Statistical values	**RC and DI**	**DI and PS**
Value of Chi square	1.243	0.713
p-value	>0.05	>0.05

The average percentage of correct responses was 80% for recall, 52.9% for data interpretation, and 39.6 % for problem-solving questions.The recall questions had a moderate level of difficulty, with a mean correct response rate that suggests an appropriate balance between easy and challenging questions.The data interpretation and problem-solving questions exhibited a more challenging level, with students performing less well, which is expected for higher-order thinking questions. The MCQs showed good discrimination, with higher-performing students consistently scoring better than lower-performing students, particularly on the problem-solving questions.The discrimination indices for recall and data interpretation questions were slightly lower, suggesting that these questions were less effective in distinguishing between high and low performers.Problem-solving questions, however, showed excellent discrimination, indicating that these questions were well-designed to assess students' ability to apply their knowledge to novel situations(8,9).The distractors used in the MCQs were effective, as evidenced by a balanced distribution of responses across the answer choices(6,10).The problem-solving questions showed a more even distribution among the answer choices, suggesting that the students were engaged with the content and not simply guessing the answers.

Statistical analysis using the Chi-square test revealed no significant difference between the scores on recall, data interpretation, and problem-solving questions ($p > 0.05$), suggesting that all question categories were equally effective in assessing the cognitive skills they aimed to test.

Discussion

This study provides strong evidence that MCQs, when properly designed, can be a valuable tool for assessing

Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(3), 250-253 | ISSN:2251-6727

not only factual knowledge but also higher-order cognitive skills such as data interpretation and problem-solving. The findings align with previous research, which has shown that MCQs can effectively assess multiple levels of cognitive function, provided they are constructed in alignment with Bloom's taxonomy (10,11,12) .

One of the key findings of this study is the effectiveness of problem-solving questions in discriminating between high and low performers. These questions demonstrated the highest discrimination index, suggesting that they were the most effective in assessing students' clinical reasoning skills. This supports the use of problem-solving MCQs in assessing the ability of medical students to apply their knowledge in clinical scenarios, a critical component of medical education (12, 13).

While the recall questions showed an acceptable level of difficulty, they were less effective in distinguishing between high and low performers. This highlights the need for a balance between recall-based and higher-order thinking questions in medical assessments (14). Data interpretation questions, although moderately challenging, did not perform as well as the problem-solving questions in terms of discrimination (15). This suggests that future MCQ assessments could benefit from a greater emphasis on problem-solving and clinical reasoning.

It is also noteworthy that distractor effectiveness was high across all question categories, indicating that the students were thoughtfully engaging with the questions. Effective distractors are essential in MCQs as they help to identify misconceptions and provide insight into students' thinking processes (4, 14)

Conclusion

This study validates the use of MCQs as an effective tool for assessing a range of cognitive skills in medical students. The findings suggest that MCQs can reliably assess not just recall but also higher-order cognitive functions like data interpretation and clinical problem-solving. However, future assessments should focus on enhancing the design of data interpretation questions and ensuring a balanced representation of both basic knowledge and clinical reasoning skills. This approach will ensure that MCQs remain a powerful and valid tool for evaluating the cognitive potential of medical students.

References

1. Bloom BS. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain.* New York: David McKay Company; 1956.
2. Medical Council of India (MCI). *Curriculum for MBBS Program.* MCI; 1997 (Revised in 2012).
3. Norman GR, Eva KW. Medical education: cognitive biases and theory-based education. *BMJ.* 2010; 342:c5642.
4. Palmer EJ, Devitt PG. Assessment of higher-order cognitive skills in undergraduate education: modified essay or multiple choice questions? *BMC Medical Education.* 2007;7:49.
5. Downing SM. Construct-irrelevant variance and flawed test questions: do multiple-choice item-writing principles make any difference? *Acad Med.* 2002;77(10 Suppl):S103–S104.
6. Case SM, Swanson DB. *Constructing Written Test Questions for the Basic and Clinical Sciences.* 3rd ed. Philadelphia: National Board of Medical Examiners; 2001.
7. Khattak AM, Ihsan A, Hussain M, et al. Influence of language barrier on performance in modified essay questions. *J Ayub Med Coll Abbottabad.* 2007;19(1):57–59.
8. Gilbert JJ, Benjamin J, Shattuck MW. Development and evaluation of MCQs to assess higher order thinking. *J Med Educ.* 1983;58(10):754–761.
9. Downing SM, Haladyna TM. Validity threats: overcoming interference with proposed interpretations of assessment data. *Med Educ.* 2004;38(3):327–333.
10. McCoubrie P. Improving the fairness of multiple-choice questions: a literature review. *Med Teach.* 2004;26(8):709–712.
11. Stark S, Chernyshenko OS, Drasgow F, et al. The effects of construct-irrelevant variance on test validity. *J Appl Psychol.* 2001;86(6):1035–1046.
12. Linn RL, Gronlund NE. *Measurement and Assessment in Teaching.* 10th ed. Pearson; 2005.
13. Epstein RM. Assessment in medical education. *N Engl J Med.* 2007;356(4):387–396.
14. Sahoo S, Mohapatra D, Patro N, et al. Evaluation of cognitive levels of MCQs in a Medical College. *J Clin Diagn Res.* 2016;10(12):JC01–JC03.
15. Mitchell ML, Smith JA. Assessment of learning outcomes through MCQs: evidence from a clinical biochemistry course. *Biochem Mol Biol Educ.* 2008;36(6):372–377.