

Overview article

DOI: 10.59715/pntjimp.3.2.4

Can Multiple - Choice Questions Measure High - Order Thinking in Medical Education? A Comprehensive Review

Dao Thi Thanh Binh

Center for Medical Education, Internal Medicine Department, Pham Ngoc Thach University of Medicine

Abstract

Multiple choice questions (MCQs) are widely used in medical education as a form of assessment and feedback. However, there is a debate on whether MCQs can effectively measure high order thinking skills, such as critical thinking, clinical reasoning, and problem - solving, which are essential for medical practice. This literature review aims to explore the current evidence and best practices on using MCQs to assess high order thinking in medical education.

Keywords: Multiple choice questions, MCQs, high - order thinking, assessment, medical education.

Received: 20/02/2024

Revised: 14/3/2024

Accepted: 20/4/2024

Author contact:

Dao Thi Thanh Binh

Email:

binhdtt@pnt.edu.vn

Phone: 0913704458

1. INTRODUCTION

High order thinking is defined as “the cognitive processes of analysis, synthesis, and evaluation that go beyond the recall of facts and concepts” [1]. High order thinking involves applying, integrating, and transferring knowledge to novel and complex situations, as well as generating and testing hypotheses, making judgments, and solving problems [2]. High order thinking is crucial for medical students and practitioners, who need to deal with uncertainty, ambiguity, and variability in clinical scenarios, and make informed decisions based on evidence, ethics, and patient preferences [3]. This review helps to investigate the present evidences and optimal approaches for employing Multiple choice questions (MCQs) to evaluate higher-order thinking skills in medical education.

REVIEWS

1. FRAMEWORK AND CONCEPTUAL FOUNDATIONS

Multiple choice questions (MCQs) are a type of objective test item that consists of a stem (a question or an incomplete statement) and a set of options (usually four or five), one of which is the correct answer and the

others are distractors (plausible and the “best answer” rather than the “correct answer”). MCQs have several advantages over other forms of assessment, such as essays, short answer questions, or oral examinations. MCQs are efficient, reliable, valid, and easy to score and administer. MCQs can also cover a large amount of content and assess various levels of cognitive domains [2, 4]. However, MCQs also have some limitations and challenges. One of the main criticisms of MCQs is that they tend to measure low order thinking skills, such as recall and recognition of facts and concepts, rather than high order thinking skills. This is because MCQs often use simple stems that require only factual knowledge or comprehension, and options that are easily recognizable or guessable. Moreover, MCQs may not reflect the authentic tasks and contexts that medical students and practitioners encounter in real life. For example, MCQs may not capture the dynamic and interactive nature of clinical reasoning, which involves gathering information from multiple sources, generating and refining hypotheses, weighing evidence and alternatives, and communicating findings and recommendations [5, 6]. Downing argued that “Effective item writers are trained, not

born ... Without specific training, most novice item writers tend to create poor-quality, flawed, low-cognitive-level test questions that test unimportant or trivial content. Although item writers should be experts in their disciplines, there is no reason to believe that their subject matter expertise generalizes to effective item writing expertise ... It is often helpful and important to provide specific instruction using an item writer’s guide, paired with a hands-on training workshop” [7].

Therefore, there is a need to design and implement MCQs that can effectively measure high order thinking skills in medical education.

2. TYPES OF MCQs FOR ASSESSING HIGH-ORDER THINKING

The most important reason given for using MCQs for examinations was economies of scale related to very large classes and an implicit need for comparing student performances. This section will delineate various formats of MCQs, including single-best answer, extended matching, and key feature questions, focusing on their suitability for measuring high-order thinking. Advantages and limitations of each format in assessing complex cognitive skills will be explored.

Example:

Theme: CBC results of a patient with a pain in the right iliac fossa

	Neutrophils	Lymphocytes	Eosinophils	Basophils
A.	60%	30%	3%	0%
B.	65%	32%	2%	0.3%
C.	80%	15%	4%	0.2%
D.	40%	30%	25%	5%
E.	35%	60%	4%	0.1%

Lead-in:

Based on the white blood cell values provided for each patient in the cells above, select the most likely diagnosis.

Stems:

1. A 21 - year - old male came to the emergency room with a fever, on examination with pressing on this area, coughing, or walking may all make the pain worse.

2.1. Single-Best Answer (SBA) Questions

Single-Best Answer (SBA) questions require the student to select the answer that is the best fit from a set of options. They are versatile in terms of the content and type of questions asked, which can range from factual recall to higher-order thinking.

Advantages: SBAs are easily marked, objective, and reliable. They can be used for both formative and summative assessment and can be used inside or outside of the classroom.

Limitations: The challenge with SBAs is in creating effective questions - ones that are a sufficiently robust assessment of a student’s knowledge base [8].

2.2. Extended Matching Questions (EMQs)

Extended Matching Questions (EMQs) present the student with a theme and scenarios relating to that theme plus a set of potential answers that are the best fit for each scenario. The question asks the student to match the correct option to each scenario.

Advantages: EMQs can assess higher-level learning than the standard MCQ, which usually tests factual recall.

Limitations: The design of EMQs can be complex and time-consuming, and they may be more difficult for students to understand and answer correctly [9].

Answer: C

2. A 16 - year - old girl sudden, watery diarrhea, after eating an allergic food, hospitalized with feeling sick, vomiting, mild fever.

Answer: D

2.3. Key Feature Questions (KFQs)

Focus on the critical steps in resolving a problem or carrying out a procedure. They are

designed to test the application of knowledge and problem-solving skills.

Advantages: Key Feature Questions can effectively assess a student's ability to apply knowledge in a practical context.

Limitations: These types of questions can be challenging to construct and may require a significant amount of time and expertise [10].

Example:

Theme: Acute Coronary Syndrome (ACS):

You encounter a patient with chest pain suggestive of ACS. Describe the essential components of the initial evaluation, risk stratification, and management plan.

Answer: For a patient with chest pain suggestive of ACS:

- **Initial Evaluation:** Assess vital signs, obtain an ECG, and perform cardiac biomarker testing (troponin). Consider risk factors (e.g., age, sex, comorbidities).

- **Risk Stratification:** Use validated risk scores (e.g., TIMI, GRACE) to identify high-risk patients. Consider early invasive management for high-risk individuals.

- **Management:** Administer aspirin, nitroglycerin, and oxygen. Consult cardiologist for further evaluation and potential revascularization.

In conclusion, each format of MCQs has its own advantages and limitations in assessing complex cognitive skills. The choice of question type really depends on what the purpose of your assessment is. It's important to carefully consider these factors when designing MCQs for assessing higher-order thinking.

3. STRATEGIES TO ENHANCE THE VALIDITY OF MCQs

Validity and Reliability of MCQs in Measuring High-Order Thinking: Empirical evidence regarding the validity and reliability of MCQs in measuring high-order thinking will be synthesized. Factors influencing the validity and reliability of MCQs, such as item quality and scoring methods, will be examined to provide insights into their effectiveness as assessment tools.

Several strategies have been proposed and tested in the literature to enhance the validity

and quality of MCQs for assessing high order thinking. These strategies include:

- Using complex stems that present realistic and relevant clinical scenarios or problems that require analysis, synthesis, or evaluation of information [2,7,8].

- Based MCQs that involved clinical vignettes with multiple variables and outcomes; and fact-based MCQs that focused on isolated facts or concepts. The results showed that scenario-based MCQs were more challenging and discriminating than fact-based MCQs, and that they better reflected the students' ability to apply knowledge to clinical situations.

- Using options that are plausible, homogeneous, mutually exclusive, and exhaustive. Plausible options are those that are not obviously wrong or absurd. Homogeneous options are those that belong to the same category or dimension. Mutually exclusive options are those that do not overlap or contradict each other. Exhaustive options are those that cover all possible or reasonable answers [2,4]. For example, a study by Tarrant et al. (2006) examined the impact of option quality on the difficulty and discrimination of MCQs in nursing education. The results showed that MCQs with plausible distractors were more difficult and discriminating than those with implausible distractors. The authors also suggested that using four options instead of five could improve the quality of MCQs without compromising their validity [11].

- Using higher-order verbs or cognitive processes in the stems or options to elicit higher-order thinking skills [1,2]. For example, a study by Palmer et al. (2015) analyzed the cognitive level of MCQs used in pharmacology courses based on Bloom's taxonomy. The results showed that most of the MCQs were at the lower levels of the taxonomy, such as knowledge and comprehension, and that only a few MCQs were at the higher levels, such as application and analysis. The authors recommended using more verbs or phrases that indicate higher-order thinking skills, such as compare, contrast, evaluate, justify, or predict, in the stems or options of MCQs [12].

4. BEST PRACTICES FOR DESIGNING HIGH-ORDER THINKING MCQS

Guidelines for designing MCQs that target high-order thinking skills will be delineated. Strategies for ensuring alignment with intended learning outcomes and enhancing the cognitive complexity of MCQs will be discussed, offering practical recommendations for educators.

Designing MCQs that effectively assess higher-order thinking skills requires careful planning and consideration. Here are some guidelines based on various sources [8,13].

4.1. Align Assessment Items to Learning Objectives

Ensure that your MCQs align with the learning objectives you're trying to assess. This helps ensure that your questions are relevant and measure what they're intended to measure.

4.2. Use Real-World Scenarios

One of the most effective ways to assess higher-order thinking is to use scenario-based questions, particularly ones that simulate real work experiences. This helps test the application of knowledge in a practical context. Context-free version asked totally recall of information, while the context-rich one is stimulating to think and it resembles the real-life situation which the students will face in the future.

4.3. Present Novel Material

Avoid simply testing for recognition or recall. Instead, present novel material that requires students to apply, analyze, or evaluate information.

4.4. Use Verbs Matched to Bloom's Levels

Use verbs in your questions that match the level of thinking you're trying to assess, based on Bloom's Taxonomy. For example, use verbs like "analyze," "evaluate," or "create" for higher-order thinking questions.

4.5. Consider the MCQ Format

Different formats of MCQs can be used to assess different types of thinking. For example, best-answer questions can be used to test critical thinking, while multiple-response questions can show linking relationships.

4.6. Analyze Visuals

You can assess critical thinking skills by asking learners to analyze or interpret

information from visuals, which are provided as part of the question stem or the answer choices.

Remember, designing effective MCQs is a skill that takes practice. Don't be discouraged if your first few attempts aren't perfect. With time and experience, you'll get better at creating questions that accurately assess higher-order thinking skills.

4.7. Quality Distractors

- **Avoid Obvious Incorrect Answers:** Ensure that distractors are not immediately dismissible. They should be plausible and not obviously wrong.

- **Reflect Common Misconceptions:** Good distractors echo typical misunderstandings or mistakes related to the topic. Review your distractors to see if they reflect these common errors.

- **Similar Length and Language:** Make sure distractors are similar in length and complexity to the correct answer. This prevents students from guessing based solely on length or language.

Remember that well-crafted distractors enhance the effectiveness of MCQs by challenging students' thinking and promoting deeper understanding.

5. COMMON CHALLENGES AND CRITICISMS OF USING MCQs TO MEASURE HIGH-ORDER THINKING

5.1. Potential Biases

One of the potential biases associated with using MCQs to measure high-order thinking is that people are significantly more likely to rate an MCQ as higher order if they cannot personally answer it correctly [14]. This bias can skew the perceived difficulty and cognitive level of the MCQs.

5.2. Limitations in Assessing Clinical Reasoning

MCQs, particularly those that are not case-based, may not fully capture the complexity of clinical reasoning, which often involves uncertainty, application of multiple criteria, reflection, and self-regulation. While there are claims that MCQs can assess higher-order cognition, many assertions are problematic because of the difficulty in interpreting what

higher-order cognition consists of and whether or not assessment tasks genuinely lead to specific outcomes.

5.3. Dichotomy Between MCQs and Authentic Clinical Practice

There is a dichotomy between MCQs and authentic clinical practice. MCQs, by their nature, are discrete and often focus on individual aspects of medical knowledge. In contrast, authentic clinical practice is complex and requires the integration and application of a broad range of knowledge and skills [15]. Therefore, while MCQs can be useful for assessing certain aspects of high-order thinking, they may not fully reflect the complexity of real-world clinical reasoning [15].

In conclusion, while MCQs have their place in medical education assessment, it's important to be aware of their limitations and potential biases. A combination of assessment methods, including MCQs, case-based discussions, and direct observation, may provide a more comprehensive and accurate measure of a student's high-order thinking and clinical reasoning skills.

6. INNOVATIONS AND FUTURE DIRECTIONS

Emerging trends and innovations in MCQ design and assessment methodologies are transforming the way we evaluate learning outcomes. Here are some key trends:

6.1. Curriculum and Instructionally-Embedded Assessments

These are assessments that are embedded in high-quality curriculum and administered as part of the flow of instruction. They provide real-time feedback to both teachers and students, allowing for immediate adjustment and improvement.

6.2. Performance Assessments

Performance assessments require a student to demonstrate a skill or conceptual understanding in a new or novel context. They are often used in conjunction with traditional MCQs to provide a more comprehensive picture of a student's understanding.

6.3. Use of Technology

The use of technology in MCQ delivery methods, such as online e-learning platforms

and clickers, is becoming increasingly popular. These technologies allow for instant feedback and can increase student engagement.

6.4. Focus on Higher-Order Cognitive Skills

There is a growing emphasis on designing MCQs that assess higher-order cognitive skills, such as analysis, synthesis, and evaluation. This shift is in response to criticism that traditional MCQs often focus too much on recall of factual information.

6.5. Culturally Responsive Assessments

Assessment systems are being designed to be culturally responsive, sustaining to ensure that every student can see themselves in the assessment experience. This approach recognizes and values the diverse backgrounds and experiences of students.

These trends reflect a broader shift in education towards more holistic, student-centered approaches to learning and assessment. However, it's important to note that while these innovations offer exciting possibilities, they also present new challenges and require ongoing research and evaluation to ensure their effectiveness.

7. CONCLUSIONS

This literature review has discussed the definition, importance, and challenges of high order thinking in medical education, and the advantages, limitations, and strategies of using MCQs to measure high order thinking skills. It is evident that MCQs can be a valid and reliable tool for assessing high order thinking skills if they are designed and implemented carefully and appropriately. However, more research is needed to explore the optimal ways of constructing and using MCQs to measure high order thinking skills in different medical disciplines and contexts. Importantly, not only was student learning and behavior changed by MCQs but also teacher's behavior. There appeared to be an effect of using this form of assessment on both teaching and ease of curriculum change. For example, once MCQs were created as a bank of questions, they framed the curriculum in such a way that they rigidly determined the parameters of what was possible to teach and so teaching could

impact the effectiveness of MCQ assessment in measuring high-order thinking.

Exams based on higher order thinking MCQs can be a viable alternative to other assessment formats, such as essays and written coursework, which are compromised by plagiarism and the availability of essay mills and AI-based content generation systems.

REFERENCES

1. Anderson, L. W., Krathwohl, D. R. A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives: Complete Edition 2001. New York: Longman.
2. Haladyna T. et al. A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. *Measurement in Education* 2010, Volume 15 – Issue 3.
3. Epstein, R. M. Assessment in Medical Education. *The New England Journal of Medicine*, 2007: 356, 387-396. <http://dx.doi.org/10.1056/NEJMra054784>.
4. Lambert W. T. Schuwirth, Cees P. M. Van Der Vleuten. Different written assessment methods: what can be said about their strengths and weaknesses? *Medical Education*, Volume 2004: 38, Issue 9, 974-979.
5. Eva et al. What every teacher needs to know about clinical reasoning. *Medical Education* 2010. doi:10.1111/j.1365-2929.2004.01972.x.
6. Olle ten Cate, Eugène J.F.M. Custers, Steven J. Durning. Principles and Practice of Case-based Clinical Reasoning Education. *Innovation and Change in Professional Education*, 2018, Volume 15.
7. Downing S. M. Twelve steps for effective test development. In S. M. Downing & T. M. Haladyna (Ed.), *Handbook of test development* 2006 (pp. 3-25). London, England: Lawrence Erlbaum Associates.
8. Case S., Swanson D. Constructing written test questions for the basic and clinical sciences, 2014, 3rd edition (revised). Philadelphia, PA: National Board of Medical Examiners. Chapter: Item writing manual.
9. Case S., Swanson D. Teaching and Learning in Medicine 2014. Chapter: Extended-Matching Items: a practical alternative to free-response questions.
10. Marla Nayera, Susan Glover Takahashia, Patricia Hrynchak. Twelve tips for developing key-feature questions (KFQ) for effective assessment of clinical reasoning. *Medical Teacher* Volume 2018, 40, Issue 11.
11. Marie Tarrant, Aimee Knierim, Sasha K. Hayes, James Ware. The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. *Nurse Education Today* 2006, 26: 662–671.
12. Thomas Wanner, Edward Palmer. Personalizing learning: Exploring student and teacher perceptions about flexible learning and assessment in a flipped university course. *Computers & Education*, 2015, Volume 88, Pages 354-369.
13. Abdus Salam, Rabeya Yousuf, Sheikh Muhammad Abu Bakar. Multiple Choice Questions in Medical Education: How to Construct High Quality Questions. *International Journal of Human and Health Sciences* 2020, Vol. 04.
14. Arslaan Javaeed. Assessment of Higher Ordered Thinking in Medical Education: Multiple Choice Questions and Modified Essay Questions 2018. <https://doi.org/10.15694/mep.2018.0000128.1>.
15. Darina Scully. Constructing Multiple-Choice Items to Measure Higher-Order Thinking 2017. Chapter: Practical Assessment, Research and Evaluation.