

Written Assessment Methods: Multiple Choice vs. Short Answer Questions in Biomedical and Medical Education

Ahmed El-Hashash

How to cite this article:

Ahmed El-Hashash. Written Assessment Methods: Multiple Choice vs. Short Answer Questions in Biomedical and Medical Education. 2024;1(2):63-68.

Abstract

Selecting the best assessment method for students in terms of validity and reliability is still a matter of debate among instructors. The written assessments are crucial in modern academic curricula and include multiple-choice questions and short answer questions that are currently important devices to measure students' learning and achievement. Notably, multiple-choice question-based testing has an association with the assessment of lower order cognition in students, including the discrete facts' recall and, therefore, many instructors are questioning their use in the university education. Both multiple-choice questions and short answer questions' testing formats are widely used in the undergraduate education. However, the debate between researches continues as to which of these two different assessments is more effective and reliable, and whether there is a correlation between them. In this article, we will describe the features and limitations of multiple-choice questions and short answer questions' testing formats and their positive and negative effects, and compare these two common assessment methods among biomedical/ medical students.

Keywords: Assessment Methods; Multiple Choice Questions; Short Answer Questions; Biomedicine; Medicine; Education.

Introduction

The curriculum of both biomedical and medical education has rapidly changed in the recent decades, including changes in the assessment methods.¹⁻⁴ Student assessments generally aim to evaluate the student knowledge adequacy with standardization.⁴⁻⁷ Selecting the best assessment method for students in terms of validity and reliability is still a matter of debate among

instructors.⁷⁻⁸ In addition, the lack of data that support one assessment method or another has hindered the selection and modification of the best student assessment method.⁹

Effective assessments are a critical learning tool that is used to determine the fulfilment of the learning.^{3,7} Several assessment methods are currently used to test the knowledge acquired as well as the ability to apply such acquired knowledge by biomedical and medical education students. These assessment methods include oral, written and practical methods, as well as multi-source observations. Each method is well-designed to measure and evaluate a specific learning outcome(s). Deciding which one of these methods as the best assessment method for testing deep and superficial learning among biomedical and medical students is still a matter of debate among different scholars.^{4,10}

Author's Affiliation: Professor, Department of Biomedicine, Stem Cell and Regenerative Medicine, The University of Edinburgh- Zhejiang International Campus, (UoE-ZJU), ZJE Building, Haining, Zhejiang 314400.

Correspondence and Reprint Requests: Ahmed El-Hashash, Professor, Department of Biomedicine, Stem Cell and Regenerative Medicine, The University of Edinburgh- Zhejiang International Campus, (UoE-ZJU), ZJE Building, Haining, Zhejiang 314400.

E-mail: hashash05@yahoo.co.uk

Received on: 11.08.2024; **Accepted on:** 23.11.2024

The assessment methods that are currently used to evaluate and assess the attitude, skill performance, and cognition of biomedical/medical students include written and practical examinations.¹ Many studies have focused on written examination or assessment format, in which students can respond electronically or by using pen and paper.¹¹ The written student examination, including multiple choice questions (MCQs) and short answer questions (SAQs), is currently the most commonly used student assessment format in biomedical/medical institutes. In these examinations, biomedical or medical students can respond in different ways. They can either use a free response format, in which they transcribe from a textbook(s) or respond from a list of suggested answers, a response method named as selected response format.¹² Students can also respond to exam questions freely from what they learnt in their own language(s) before.¹² There are several drawbacks for the written examination or assessment format such as being time consuming, and some errors in marking. Other challenges of this format include difficult or poor student handwriting that makes it difficult to read for a fair assessment, and the need to prepare the marking scheme with enough time in advance to minimize bias and improve this assessment reliability.¹

The cognition of medical undergraduate students is normally assessed using two written formats: selected response format (SRF) or free response format (FRF) that are different from each other.^{1,6} The SRF examples include MCQs and extended matching questions; while FRF examples are SAQs, long answer questions (LAQs), and short essay questions (SEQs). Each of these assessment methods have many benefits and some drawbacks, and there is no consensus among scholars on which method is the best.^{6,13}

The efficacy of MCQ- or SAQ-based testing as a powerful student assessment method is a matter of debate in the last few decades.⁹ Many medical schools worldwide currently use both MCQ- or SAQ-based testing formats in undergraduate exams. The objective of this article is to compare two common assessment methods among biomedical/medical students; MCQs and SAQs.

Multiple Choice Questions (MCQ) Features and Limitations

The MCQ-based test is currently a popular assessment method that measures students' knowledge and has several advantages, including

covering a wider range of the course with many questions that the examiner can use and assessing different student levels of learning outcomes that can range from basic recall to other levels such as application, analysis, and evaluation of information.^{14,15} MCQ-based tests have also other advantages, including their use in a large class size¹⁶, general objective grading¹⁴, and easy and timely feedback, when using a computer-based marking.¹⁷ The MCQs' quality can be evaluated, and the MCQ quality assessment is called item analysis, which aims to collect, summarize, and use information collected from student's responses for assessing the quality of MCQs. The item analyses can also be used for the assessment of which MCQ question(s) can discriminate between slow learners and good students and for showing the difficulty index of each MCQ that are other advantages of MCQ-based tests.^{15,18,19}

MCQs' limitations or disadvantages include a high chance of student guessing and the preparation of students for this kind of assessments will be mainly towards the process of surface learning with largely opting for the crude memorization.^{15,18,19} In sum, MCQ-based testing is currently the norm because of its efficiency, convenience, and reliability between different graders. Several guidelines for writing high quality MCQs have been published recently.^{20,21}

Positive and Negative Effects of MCQ Testing

Several studies have focused on exploring the positive and negative effects of MCQ-based exams. For example, Roediger and Marsh²² have tested the theory of the testing effect, which states that previous student testing can enhance the performance of students on later exams. This research study used the results of multiple other studies to demonstrate the positive consequences of MCQ tests that are probably due to transferring the information positively during student learning.²² Notably, the retrieval timing can affect both long- and short-term retention since it is proposed that the more closely the first exam is taken after learning, the more likely this exam will be a part of short-term retention, and vice versa.²³

MCQ-based testing has also some negative effects. A good example for that is the negative suggestion effect(s), which is associated with the lures' number that are given in each MCQ, in obstructing student performance in later tests.¹⁵ The negative suggestion effect can be defined as the student tendency to believe in a false statement(s).

This negative suggestion effect takes place when the student taking MCQ exam believes a wrong choice to be the correct choice of the testing question, and particularly when there is no feedback from the examiner. Students will, therefore, most likely choose the same incorrect answer(s) on later exams if they choose the wrong answer(s) on the initial exam and without feedbacks from the examiner.¹⁵

In the study of Roediger and Marsh^{22, 24} undergraduates were asked to read passages, which were later used for constructing a 144 MCQ test, with the final cued recall test consisting of 260 questions. The study included 2 independent variables; the passage status (read or unread) and the given possible answer number^(0,2,4,6), while the dependent variable in this study was the number correct answers on the later cued recall test. Roediger and Marsh²² have carried their study within subjects and phases, with the first phase consists of 90 seconds reading of 18 passages, while in the 22-minute-second phase students are asked to complete the MCQ exam. Students are then given a 5-minute filler task, before a 35 minute-cued recall exam. These authors predicted that the performance of student participants will be better in a cued recall test if they took a preliminary MCQ test before²². Remarkably, the results of this study have supported this predication. Thus, participant students, who lack previous MCQ-based test experience have responded correctly 28% of the time on the cued recall test, while participant students with previous MCQ-based test experience have answered 46% of the questions correctly.²²

Students who were giving 2 possible choices have answered correctly 51% of the answers on the cued recall section of the MCQ-based exam. This percentage of the correct answers was reduced to 45% when students were given 4 possible answers and to 43% in MCQ with six possible answers, suggesting that increasing the number of possible answers makes the MCQ more challenging.²² Notably, 75% of the false responses that are recorded by participant students on the cued recall test are lures from previous MCQ-based tests.²² Giving feedbacks on student answers and knowledge is also important in MCQ-based exams since these feedbacks enhance both students' learning of correct information and information retrieval on later testing.²⁴ In conclusion, these studies suggest that previous MCQ-based test experience has a positive effect on the information retention by students, increasing the number of possible answers makes the MCQ more challenging

to students, and instructor feedbacks are critical in MCQ-based exams.

MCQs vs. SAQ-Based Exams

Many studies have compared the multiple-choice (MCQ) and short answer (SAQ)-based examinations and their effectiveness. For example, a study by Funk and Dickson²⁵ has demonstrated that the participant student performance on MCQs is significantly higher than their performance on the same items in the SAQ exam ($p < 0.001$). In this study, Funk and Dickson used the same questions in short-answer and multiple-choice free-text response format, and equated SAQs with recall processes, while MCQs are equated with recognition processes. These authors also concluded that participant students answering MCQs are often in need to both understand and interpret information to correctly recognize the proper answer(s).²⁵ In addition, this study highlighted the significance of MCQ-based testing in enabling the participant students to use various strategies to identify the correct answers such as the recall, recognition, analysis, and other test-taking approaches/strategies, including guessing and/or eliminating the wrong answer(s).²⁵ However, this study has not clarified the argument on whether high performance and marks in MCQ-based testing can really provide accurate information on the levels of student learning or simply overestimate student learning of the course contents. Our conclusion from this study is that passing MCQ-based testing doesn't necessarily imply students' understanding or knowledge of the teaching course and, therefore, MCQs are not enough to measure these understanding or knowledge since students normally learn various strategies to pass MCQ-based testing.

Other studies have compared the performance of two medical student groups, at the schools of medicine and dentistry, in MCQ-based exams with other exam types.²⁶ Participant students in these studies achieved a significantly high score in MCQ-based exams compared to other exams/assessments such as problem-based exams, essay assessments, and viva-voce.²⁶ Interestingly, students performing well in MCQ format of exams are more likely to both achieve a high score in SAQ-based exam²⁷, and successfully pass the final exam(s) with a clearly high grade compared to other assessment forms.²⁸ Therefore, there is a high correlation level between MCQ scores and final exam scores in medical school students.¹⁹ Thus, MCQs format of exams are preferred at medical and dental schools, and MCQs

scores are useful in predicting the grades of final examinations that may correlate well with the final MCQ scores^{19,29} in contrast to other studies showing that MCQs could not predict a higher final exam grade.³⁰ Indeed, students' level (year) of study is a key factor that may play a role in the prediction of the performance of students based on the exam components, probably due to the fact that lower level students adopt a surface learning approach, while higher level students tend to adopt a deep approach to learning.¹⁹

START The strong correlation between SAQs and MCQs was further established by other studies³¹, although a superior reliability and validity were clearly reported for MCQ-based exams with a high level of cognition.^{32,33} In addition, a significant correlation was established between the effectiveness of MCQs and SAQs in the written end-of-clerkship assessments at medical schools.³⁴ These findings will help with avoiding the duplication of different written assessment types and results, leading to the reduction of unnecessary faculty efforts in preparing both SAQs and MCQs to evaluate the same student cognitive skills in one subject.³⁴

Interestingly, studies have explained that gaining a higher score in MCQ-based exam is probably due to the distinction students have to be due to the ease of recalling information observed in these MCQ-based exams.^{35,36} Remarkably, the MCQ can provide a context that facilitates the recalling of information by the participant students^{35,36} and apply a high order of thinking with generally high reliability and validity.^{37,38}

It is worth saying that students normally prepare differently for various types of assessment formats.^{39,40} Therefore, students normally adopt a well-articulated and wider reading in their preparation for MCQs exam format compared to the SAQ-based exam.¹⁹ Another suggestion for achieving higher score in MCQs, however, is the diversion of student attention from regular textbooks to other MCQ resources, including past questions, internet-based sources and MCQ-based textbooks, as documented by student feedbacks.^{41,42,43}

Some studies have demonstrated more advantages of MCQ-based exams due to its ease of scoring (electronically and/or manually), and for being both more objective⁴⁴, and easily assess the skills of problem-solving.³³ However, MCQ-based exams have some reported disadvantages, including their cueing effect and, therefore, there

is a high chance for students to guess the correct answer, ultimately leading to a falsely higher scores than other assessment methods.⁴⁵ Therefore, a careful preparation of high quality MCQ-based exams that can warrant higher student cognition is required, but it is still difficult to achieve and labor-intensive.⁵

Given feedbacks to students has always enhanced the student performance in all types of tests (i.e. MCQs and SAQs). The general consciences are that MCQ-based testing is beneficial to the learning of students and becomes more beneficial when feedbacks on this assessment are given to students.^{15,46} However, it is less likely that MCQ exam results are accompanied with feedbacks since MCQs are normally computer graded. In contrast, SAQ exam results are mostly accompanied with feedbacks since they are instructor graded (i.e. not computer graded). Therefore, participant students who are taking SAQ-based exams with feedbacks can significantly perform better when re-tested later.

Features and Limitations of SAQ-Based Testing

Several studies on the limitations of SAQ-based exams concluded that SAQs are prone to risks of bias and error since SAQ scoring is usually more subjective and time-consuming than MCQs.^{19,44} In addition, student handwritten responses in SAQ-based exams are sometimes not clear enough, illegible and/or ambiguous.^{19,44} Other limitations of SAQs that are not in MCQs, include the need to both structure the grading metrics with enough time in advance and develop the answer key for the assessors in order to minimize the risk of bias in SAQ-based exams.^{19,34,44} Nevertheless, SAQ-format of exams can both reflect the interpretive skills of participant students and provide clear flexibility in students' responses to questions. Other features of SAQ-based exams are their lack of the cueing effect of MCQ-based exams and the ease of using these question types to assess problem solving skills in participant students.^{33,45}

Conclusions

There is no currently single student assessment method that is perfect or with absolute validity and reliability. Therefore, multiple methods, which have a correlation to target the measurement of the same trait⁴⁷, should be used for more efficient and valid student assessments. In addition, since creating exams in basic or pre-clinical subjects is a

time-consuming process that requires instructors' expertise and concentration^{3,6} selecting one of two exam formats that have strong and positive correlations can clearly improve the exam quality and reduce the faculty load by 50% (Farooqui et al., 2018).

The current consciences are that there is a correlation between MCQ and SAQ exam formats in both basic and medical sciences. The student performance on MCQ and SAQ-based exams shows a clearly significant overall correlation in these two assessment methods in medical/biomedical schools.³⁴ Therefore, students performing well in MCQ-based exams are likely to do well in SAQ-based exams.^{5,27,34} A similar compression between MCQ and SAQ modalities was carried out in different medical disciplines such as pharmacology, physiology, anatomy, surgery, and ophthalmology^{5,27,44,48,49} and for physiology and pharmacology in pre-clinical years at medical schools.^{48,49} Future research is needed to further explore MCQ-SAQ correlations by comparing individual student result(s) in each discipline. Using more stringent approaches/methods in future studies will further facilitate the correlations between MCQs, SAQs and other assessment modalities for achieving high quality student assessments in biomedical and medical schools.

Our conclusions and reflections are that not all MCQs are easier than SAQs, and not all SAQs are more difficult than MCQs, and both these assessment methods have features and limitations. Indeed, some research studies have shown that medical and dental students can perform differently in the two components of their examination; SAQs and MCQs. In these studies, the higher level (year of study) and distinction students have higher scores in MCQ-based exams, while the low level students have higher scores in the SAQ-based exams and, therefore, both the exam format and the student's level (i.e. year of study) can affect the final exam grades of medical students.¹⁹

Abbreviations

FRF: free response format

LAQs: long answer questions

MCQs: multiple choice questions

SAQs: short answer questions

SEQs: short essay questions

SRF: selected response format

Compliance with Ethical Standards

Conflict of Interest: No conflict of Interest

Ethical approval: NA

Informed consent: NA

References

- Schuwirth LW, van der Vleuten CP: Different written assessment methods: what can be said about their strengths and weaknesses?. *Med Educ.* 2004, 38:974-9.
- Norman G: Past, present and future. *Perspectives on medical education.* 2012, 1:6-14.
- Zaidi NL, Kreiter CD, Castaneda PR, et al.: Generalizability of competency assessment scores across and within clerkships: how students, assessors, and clerkships matter. *Acad Med.* 2018, 93:1212-7.
- El-Hashash AH, 2019. PBL vs. Non-PBL Approach and Assessment in Biomedical and Medical Education: A Recent Update. *Journal of Global Medical Education and Research.* 2(1), 9-13.
- Mujeeb A, Pardeshi M, Ghongane B: Comparative assessment of multiple choice questions versus short essay questions in pharmacology examinations. *Indian J Med Sci.* 2010, 64:118-24.
- Schuwirth LW, van der Vleuten CP: General overview of the theories used in assessment: AMEE Guide No. 57. *Med Teach.* 2011, 10:783-97.
- Kurtz JB, Lourie MA, Holman EE, Grob KL, Monrad SU. Creating assessments as an active learning strategy: what are students' perceptions? A mixed methods study. *Med Educ Online.* 2019 Dec; 24(1):1630239.
- Cook DA, Beckman TJ: Current concepts for validity and reliability for psychometric instruments: theory and application. *Am J Med.* 2006, 119:7-16.
- Van Bruggen L, Manrique-van Woudenberg M, Spierenburg E, Vos J: Preferred question types for computer-based assessment of clinical reasoning: a literature study. *Perspect Med Educ.* 2012, 1:162-71
- Choudhury B, Freemont A. Assessment of anatomical knowledge: approaches taken by higher education institutions. *Clin Anat* 2017; 30(3):290-9.
- Inuwa IM, Taranikanti V, Al-Rawahy M, Habbal O. Anatomy practical examinations: how does student performance on computerized evaluation compare Reddy ST, Zegarek MH, Fromme HB, Ryan MS, Schumann SA, Harris IB. Barriers and facilitators to effective feedback: a qualitative analysis of data from multispecialty resident focus groups. *J Grad Med Educ* 2015;7:214-9. doi:10.4300/JGME-D-14-00461.1 PMID:26221437