



## Reflections on Using Open-ended Questions

Zahra Abdolreza Gharehbagh<sup>1,2</sup>, Azam Mansourzadeh<sup>3</sup>, Atiyeh Montazeri Khadem<sup>4</sup>,  
\*Masumeh Saeidi<sup>5</sup>

<sup>1</sup>PhD Candidate in Medical Education, Department of Medical Education, Tehran University of Medical Sciences, Tehran, Iran.

<sup>2</sup>Faculty Member, Department of Pediatric Nursing, Faculty of Nursing and Midwifery, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran.

<sup>3</sup>Paris Nanterre University, Paris, France.

<sup>4</sup>MA of Educational Administration, Department of Education, Kerman Branch, Kerman, Islamic Azad University, Iran.

<sup>5</sup>Department of Medical Education, Faculty of Medicine, Tehran University of Medical Sciences, Tehran, Iran.

### Abstract

In open-ended questions, unlike closed-ended ones, the learner has to formulate and present the desired answer. One of the important reasons for using open-ended questions over the years is the low probability of guesswork and cheating in this type of question. It is also stated that these questions go beyond testing the memory and can assess the understanding level and problem-solving ability of students. As professors are highly familiar with designing open-ended questions and their ease of design, this type of question has always been popular among evaluators. However, open-ended questions, especially with an extended response, are not welcomed by learners.

Due to the subjective nature of assessing tests and scoring essays, there is a possibility of errors in the process of correcting these questions, including the Halo effect, Mechanics effect, Order effect, Item-to-item carryover effect, and Test-to-test carryover effect. These errors have to be minimized by measures such as question-by-question correction of all test takers at the same time and without time interval, the anonymity of test sheets, and a scoring rubric. Also, the mentality of the correctors affects the scoring process, reducing the reliability of the test. Answering open-ended questions takes a long time; therefore, fewer questions can be evaluated in each test. This lowers the content validity of tests. Despite the benefits of open-ended questions, they are not suitable for all occasions. In selecting the type of written tests, the homogeneity between the assessment method and the intended educational goal should be considered.

**Key Words:** Advantages, Disadvantages, Essay, Examiner, Open-ended question.

\*Please cite this article as: Abdolreza Gharehbagh Z, Mansourzadeh A, Montazeri Khadem A, Saeidi M. Reflections on Using Open-ended Questions. Med Edu Bull 2022; 3(2): 475-82. DOI: 10.22034/MEB.2022.333518.1054

### \*Corresponding Author:

Masumeh Saeidi, Faculty of Medicine, Tehran University of Medical Sciences, Tehran, Iran.

Email: saeidim@razi.tums.ac.ir@yahoo.com

Received date: Feb. 14, 2022; Accepted date: May.22, 2022

## 1- INTRODUCTION

An important aspect of the teaching and learning process is the extensive measures that evaluate the activities and learning outcomes of students and guide the learning process and their performance. The information needed to measure learning can be obtained in many ways, including tests, questionnaires, grading scales, checklists, research projects, oral exams, laboratory work, homework, interviews, and observation of students' performance in different situations (1).

Written questions are generally divided into two categories: open-ended and closed-ended questions. Before the introduction of closed-ended tests, open-ended questions were one of the few assessment methods used. In open-ended questions, the student must generate the answer in words, phrases, or sentences and write it on the answer sheet. However, in closed-ended questions, the student is given a list of options and must distinguish the correct answer from them (2).

The use of multiple-choice questions began about a hundred years ago. In the following years, this form of question was quickly adopted in different levels and disciplines. The advantages of this type of question include wide content coverage, objectivity, ease of correction, simple question analysis, creating a question bank, transparency, and the possibility of self-learning and independent learning. On the other hand, these questions have limitations such as measuring the extent of knowledge, low face validity, increasing guesswork, and encouraging learners to learn superficially and memorize content. Also, designing a good test question according to the guidelines is time-consuming and difficult. Therefore, despite the many benefits of closed-ended tests, open-ended questions are still used in the assessment of learners (2). Assessment is an essential part of teaching and learning and has a substantial effect on studying

and preparing for exams. Various studies have shown that assessment methods affect learners' study approaches as well as their strategies for preparing for exams. In addition to the overall effect of the test, the type of test is an influential factor in studying and learning approaches (1-3). This study aimed to review the benefits and limitations of open-ended questions in learner assessment.

## 2- MATERIALS AND METHODS

In this review study, a systemic search of electronic databases of Medline (via PubMed), SCOPUS, Web of Science, ProQuest, ERIC, SID, Magiran, CIVILICA, and Google Scholar search engine was performed with no time limit up to February 2022, using the following keywords alone or in combination: "Open-ended question", "Uncued question", "Constructed response", "Closed-ended question", "Selected response", "Cued question", "Advantage", "Disadvantage", "Test", "Exam", "Assessing", "Evaluating", and "Students". The search was performed independently and in duplication by two reviewers, and any disagreement between the reviews was resolved by the supervisor. The database search was done for suitable studies and books. Study abstracts were screened for eligible studies, full-text articles were obtained and assessed, and a final list of eligible studies was made. This process was done independently and in duplication by two reviewers, and any disagreement was resolved by a third reviewer.

## 3- RESULTS

### 3-1. History

Written questions are generally divided into two categories: open-ended and closed-ended-questions. In open-ended questions, the student must generate the answer in words, phrases, or sentences and write it on the answer sheet. However, in closed-ended questions, the student is

given a list of options and must distinguish the correct answer from them. Descriptive questions can be answered by measuring the power of comprehension, in-depth understanding of lesson content, processing answers, and finally, the upper levels of the cognitive domain. Answering these questions requires a deep and systematic thought process (1, 2, 4).

According to Stalnaker et al., a descriptive question must meet the following criteria:

1. Answering these questions requires the production and presentation of answers by the student.
2. The student presents their inference and understanding of the question.
3. Determining the accuracy and quality of the answers requires mental judgment.
4. It is possible to provide answers in various forms; in descriptive questions, the student has the authority to organize and present the material in a way that they believe is reasonable (5).

In open-ended questions, unlike closed-ended ones, the learner has to formulate and present the desired answer (6).

For many years, open-ended questions have been used as the main tool for assessing learners in medical schools around the world. However, numerous problems resulted from the implementation of these questions, including deep dissatisfaction of learners with the time required to answer these questions, low scoring reliability among evaluators, and low coverage of the course content by these questions. In the mid-19th century, with the introduction of a new generation of written tests called multiple-choice questions (MCQ); the use of MCQs to overcome existing problems became widespread. However, it was not long before the limitations of multiple-choice questions were identified, and it was found that despite the many benefits of multiple-choice questions (e.g., high reliability and

the possibility of including a large number of questions in one test), there are still concerns such as the possibility of students guessing the correct answer. It can be said that one of the important reasons that the use of open-ended questions has been maintained over the years is the low probability of guessing and cheating in this type of question. In these questions, a list of options is not provided to the student, so the student must generate the answer and cannot guess the answer by looking at the available options. Another feature of these tests is the effect of the personal opinion of the teachers when evaluating and grading these questions. It is also stated that open-ended questions can be used to assess other skills such as writing style, organization, and presentation and even learners' creativity and innovation in providing answers. In closed-ended questions, this is not possible. These questions can go beyond the level of memorizing and can assess the student's understanding and problem-solving ability. Naturally, this issue should be treated with caution. It is essential to consider that although open-ended questions enable learners to provide concise answers, this does not guarantee that these questions are inherently capable of assessing high levels of the cognitive domain. The form and format of questions do not determine their ability to evaluate the measured levels. It is the content of a question that determines the evaluated level of knowledge. It should be noted that despite the benefits of open-ended questions, these questions are not appropriate for every situation, and when deciding on the choice of written tests, the homogeneity between the assessment method and the educational goal should be considered (1, 2, 7-14).

### **3-2. Types of open-ended questions**

#### **3-2-1. Essay**

In these questions, the learner is asked to produce and provide the answer based on

the learned material. Based on the freedom in providing answers, these questions are divided into extended response and restricted response questions.

### **3-2-2. Modified essay questions**

These are a variety of open-ended questions used to measure problem-solving ability. In this form of questions, the question is asked in the form of a scenario or clinical case, usually asked step-by-step. At each stage, the student is asked to answer a question or questions that address the clinical case from different perspectives.

### **3-2-3. Short answer questions**

In short answer questions, the student is asked to provide their answer in the form of a word or a short and concise phrase (1, 2, 11-14).

## **3-3. Advantages and limitations of open-ended questions**

### **3-3-1. Advantages**

1. Easy design: Instructors are familiar with these types of questions. Designing distractor options is not necessary.
2. Lack of problems in recognition because the list of possible answers is not provided to the test takers.
3. There is no possibility of guessing.
4. Ability to assess high levels of cognitive understanding and creativity: Answering open-ended questions requires the use of thought and reasoning processes.
5. Encouraging deep learning: These questions make it possible to assess the ability to organize between different topics, so the learner needs a deep understanding to answer these questions.
6. Possibility of formative and summative assessments: When the purpose of the test is to enhance the learning of learners, these questions, especially those with short formative answers, can be used during the

training course. It is also common to use these questions in the form of a final exam.

### **3-3-2. Limitations**

1. Lack of versatility: Open-ended questions only allow the assessment of different levels of knowledge. Therefore, when the purpose is to assess performance or attitudinal domains, these questions alone are not sufficient.
2. Time-consuming to answer: It takes a long time to answer these questions, specially modified essay questions.
3. Low validity: The time required to answer these questions and the limited number of questions to be evaluated affect the content coverage and reduce the content validity of the test.
4. Content problem: The problem of limited sampling of this type of question is serious, and it is, therefore, necessary to use short-answer questions instead of extended answers when using these tests to make time for more questions.
5. The effect of subjective judgment on scoring and low objectivity: When there is no clear scoring structure and criteria in descriptive answers, the subjectivity of the correctors may affect the correction process and lead to significant differences in scores and low reliability.
6. Low reliability: In open-ended questions, especially descriptive answers, the possibility of errors is high, and factors other than comprehensive knowledge, such as the writing style, handwriting, and the way of organizing and expressing content, will also affect the scoring process. These factors lead to lower reliability of these questions compared to closed-ended questions.
7. Difficulty of correction: One feature of open-ended questions is that their correction depends on the individual. Unlike multiple-choice questions, where questions can be easily corrected using a

computer, one person must read the answers to open-ended questions and, after comparing with the scoring guides, decide whether the provided answer is acceptable. On the other hand, the more complex the answers to these questions, the more difficult to judge scoring, which requires a higher mental activity of the evaluators in the correction process.

8. Low performance: In general, a large number of open-ended questions are required to have comprehensive content coverage. Adding open-ended questions to a test makes the test impractical and very difficult to administer and correct. This becomes particularly complicated when the number of learners is high or the content and objectives of the training course are large (1, 2, 15-21).

### **3-4. Scoring open-ended questions**

There are generally three ways to correct descriptive questions; analytic, holistic, and primary traits.

**3-4-1. Analytic:** In the analytic method, also called point-scoring, the response pattern is divided into several sections, and then a specific score or point is considered for each section. The use of this method, due to the clarity of evaluation criteria, reduces the impact of the evaluator's mentality and assumptions on the scoring process.

**3-4-2. Holistic:** In the holistic method, the examiner studies the student's total response and then judges its quality. In this method, the evaluator evaluates the student based on their general perception of the answer.

**3-4-3. Primary traits:** In this method, the evaluator evaluates the answers based on the main characteristics that they already have in mind for a correct answer according to the desired subject. Usually, a checklist is designed for this method, and the score is determined based on the number of features provided by the

learner. In this method, a score is also assigned to factors such as the power of expression, the way of presenting the content, and the logical organization of the response (2, 22-26).

### **3-5. Open-ended questions scoring errors**

Due to the subjective nature of the process of correcting and scoring descriptive questions, there is a possibility of errors in scoring these questions. These errors can affect the result of a test.

**3-5-1. Halo effect:** The evaluator's views and opinions about learners can affect the scores. Therefore, test sheets should be scored anonymously to avoid the Halo effect as much as possible.

**3-5-2. Mechanics effect:** In most cases, in addition to the answer content, the scores given by the evaluators are also influenced by the handwriting, punctuation, and volume of the response. Students who do not have good handwriting receive lower grades while they may have provided a correct answer.

**3-5-3. Order effect:** Order effect means that the questions that are scored at the beginning are more desirable from the evaluator's point of view and assigned a higher score; while the questions that are evaluated at the end of the scoring session are assigned a lower score due to the higher expectations of the evaluators.

**3-5-4. Item-to-item carryover effect:** The evaluator's view of how a learner answers earlier questions influences the evaluation of the answers to the questions that follow. In other words, if the evaluator determines that the student's first answer is good and gives it a high grade, it is likely that under the influence of the first question, they will score the next answers of the same learner higher. Therefore, it is recommended that the answers of all students to one question be scored at a time without interruption,

moving on to the next questions for all test takers afterward.

**3-5-5. Test-to-test carryover effect:** Scoring a descriptive test is strongly influenced by the student's performance in tests taken shortly before. If the student received a lower-than-expected score in a previous descriptive test, the subsequent test is considered more favorably (1, 2, 27-33).

### **3-6. Misconceptions about open-ended questions**

**3-6-1. Evaluation of high cognitive domains:** A common misconception about open-ended questions is that these questions are inherently capable of measuring high levels of the cognitive domain. However, in case of improper design, these questions assess low cognitive levels (2).

**3-6-2. The easy design of open-ended questions:** Although designing open-ended questions is easier than multiple-choice questions because it does not require distractor options, it does not mean that designing appropriate descriptive questions requires no effort (1, 2).

**3-6-3. Reduce the likelihood of guessing:** There is a general belief that multiple-choice questions allow the student to answer by guesswork and increase their score, and a similar possibility does not exist in open-ended questions. However, descriptive questions merely change the form of guessing, not eliminate it (34).

**3-6-4. Trying harder to prepare for the exam:** Through the extensive review of existing evidence and studies on the subject, Crook concluded that learners' expectations of cognition and learning the content are more influenced by their reading skills than by the form and type of the test they take. Crook believes that learners prepare for the exam based on the expectations of their teachers than on the type of test they take (2).

Also, exaggeration has the greatest impact on test results when the evaluator has not prepared a response model to correct the answers (35).

## **4- CONCLUSION**

As the process of answering open-ended questions is longer than closed-ended ones, fewer questions can be evaluated in each test. Therefore, the content validity of the test is low. Blueprints and repeated developmental exams can improve the validity of these tests. As fewer open-ended questions can be answered in a test, at one time, the number of questions to be answered is lower than closed-ended questions, reducing the reliability of the test. In scoring these questions, the evaluators' mentality affects the scoring process, further reducing the reliability of the test. On the other hand, students use different study approaches to prepare for the exam. This way, students study more superficially when preparing for multiple-choice exams and more in-depth when preparing for open-ended exams.

As the teachers are highly familiar with the design of open-ended questions and their ease of design, this type of question has always been popular among evaluators. However, open-ended questions, especially extended response questions, are not well received by learners. In general, open-ended questions have many benefits that promote their application. However, these questions are not suitable for every situation, and in choosing these questions, compliance with the objectives and executive considerations are essential.

## **5- AUTHORS' CONTRIBUTIONS**

Study conception or design: MS, and ZA; Data analyzing and draft manuscript preparation: AMK and AM; Critical revision of the paper: MS, and AMK; Supervision of the research: MS and ZA;

Final approval of the version to be published: MS, AM, ZA, and AMK.

**6- CONFLICT OF INTEREST:** None.

## 7- REFERENCES

1. Seif AA. Measurement, Assessment and Evaluation educational. Tehran: Doran; 2021. ISBN: 978-600-6208-71-8.
2. Jalili M, Khabaz Mafinejad M, Gandomkar R, Mortaz Hejri S. Principles and Methods of Student Assessment in Health Professions. The Academy of Medical Sciences Islamic Republic of Iran, 2017.
3. Asle Fattahi, B., Karimi, Y. The Effect of Mindfulness-Based Techniques in Reduction of Students Test Anxiety. *Journal of Instruction and Evaluation*, 2009; 2(6): 7-33.
4. Examples of Open-Ended and Closed-Ended Questions". *yourdictionary.com*. *yourdictionary.com*. Retrieved 17 September 2017.
5. John M. Stalnaker and Ruth C. Stalnaker. Reliable Reading of Essay Tests. *The School Review*, Vol. 42, No. 8 (Oct., 1934), pp. 599-605, Published By: The University of Chicago Press.
6. Ackley, Betty. *Nursing diagnosis handbook : an evidence-based guide to planning care*. Maryland Heights, Mo: Mosby, 2010. ISBN: 9780323071505.
7. Howard Schuman & Stanley Presser. "The Open and Closed Question". *American Sociological Review*. 1979; 44 (5): 692–712. doi:10.2307/2094521. JSTOR 2094521.
8. Worley, Peter. Open thinking, closed questioning: Two kinds of open and closed question". *Journal of Philosophy in Schools*. 2015; 2(2). doi:10.21913/JPS.v2i2.1269. ISSN 2204-2482.
9. Park, J. Constructive multiple-choice testing system. *British Journal of Educational Technology*, 2010; 41: 1054-64. doi:10.1111/j.1467-8535.2010.01058.x
10. Origins and Purposes of Multiple Choice Tests, San Francisco State University, accessed 2016-04-02.
11. Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, 58(4), 1064-1082., doi: 10.1111/ajps.12103
12. Krosnick, J. A., & Presser, S. Question and questionnaire design. In P. V. Marsden & J. D. Wright (Eds.), *Handbook of survey research* (2010; Vol. 2, pp. 263-314). Bingley, UK: Emerald Group Publishing Limited
13. Emde, M., Fuchs, M. Using adaptive questionnaire design in open-ended questions: A field experiment. Paper presented at the American Association for Public Opinion Research (AAPOR) 67th Annual Conference, San Diego, USA, 2012.
14. Züll, C. Open-Ended Questions. *GESIS Survey Guidelines*. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences, 2016. doi: 10.15465/gesis-sg\_en\_002
15. Connor Desai, S. and Reimers, S. Comparing the use of open and closed questions for web-based measures of the continued influence effect. *Behavior Research Methods*, 2018. doi: 10.3758/s13428-018-1066-z.
16. Reja, U., Manfreda, K. L., Hlebec, V., & Vehovar, V. Open-ended vs. close-ended questions in Web questionnaires. *Developments in Applied Statistics*, 2003;19: 159–77.
17. Graesser, A., Ozuru, Y., & Sullins, J. What is a good question? In M. McKeown & G. Kucan (Eds.), *Bringing reading research to life* (pp. 112–141). New York, NY: Guilford, 2010.
18. Çakır, H. and Cengiz, Ö. (2016) The Use of Open Ended versus Closed Ended Questions in Turkish Classrooms. *Open Journal of Modern Linguistics*, 6, 60-70. doi: 10.4236/ojml.2016.62006.
19. Barbara A. Wasik, Annemarie H. Hindman. Realizing the Promise of Open-Ended Questions. *The Reading Teacher* 2013;67(4): 302-11. DOI:10.1002/TRTR.1218.

20. Ronald A. Davidson. Relationship of study approach and exam performance. *Journal of Accounting Education* 20(1):29-44. doi: 10.1016/S0748-5751(01)00025-2.
21. Thomas, P. R., Bain, J. D. Contextual dependence of learning approaches: The effects of open-ended text questions. *of assessments. Human Learning: Journal of Practical Research & Applications*, 1984;3(4), 227–40.
22. Gunter Maris, Timo Bechger. Scoring open-ended questions. *Handbook of Statistics*, Chapter: 27, Elsevier: Editors: C.R. Rao, S. Sinharay; 2007.
23. Matthias von Davier, Lillian Tyack, Lale Khorramdel. Automated Scoring of Graphical Open-Ended Responses Using Artificial Neural Networks. Available at: <https://arxiv.org/ftp/arxiv/papers/2201/2201.01783.pdf>.
24. Nicole Kersting, Bruce Sherin, James Stigler. Automated Scoring of Teachers' Open-Ended Responses to Video Prompts. *Educational and Psychological Measurement*, 2014; 74(6):950-74.
25. Atilgan, H. (2019). Reliability of essay ratings: A study on generalizability theory. *Eurasian Journal of Educational Research*, 2019;19(80): 1–18.
26. Norcini J, Disernes D. The scoring and reproducibility of an essay test of clinical judgment. *Academic Medicine*, 1990; 65 (9): S41-2.
27. Attali, Y. Validity and reliability of automated essay scoring. In: M. D. Shermis, J. Burstein (Eds.), *Handbook of automated essay evaluation: Current applications and new directions* (pp. 181-199). New York, NY: Routledge, 2013.
28. Oleg Sychev, Anton Anikin, Artem Prokudin. *Methods of Determining Errors in Open-Ended Text Questions*. doi: 10.1007/978-3-030-25719-4\_67.
29. Atilgan, Hakan; Demir, Elif Kübra; Ogretmen, Tuncay; Basokcu, Tahsin Oguz. The Use of Open-ended Questions in Large-Scale Tests for Selection: Generalizability and Dependability. *International Journal of Progressive Education*, 2020;16(5):216-27.
30. Engelhard, G. Examining rater errors in the assessment of written composition with a many-faceted Rasch model. *Journal of Educational Measurement*, 1994; 31(2):93.
31. Kim, H. Effects of rating criteria order on the halo effect in L2 writing assessment: A many-facet rasch measurement analysis. *Language Testing in Asia*, 2020;10(1): 1–23.
32. Hafizah Husain, Badariah Baisb, Aini Hussainb, Salina Abdul Samad. How to Construct Open Ended Questions. *Procedia - Social and Behavioral Sciences*, 2012; 60:456.
33. Chase C. Essay Test Scoring: Interaction of Relevant Variables. *J of Educational Measurement*, 1986; 23 (1):33-41.
34. Christian M. Reiner, Timothy W. Bothell, Richard R. Sudweeks. *Preparing Effective Essay Questions* Paperback. ISBN-10:1581070659.
35. Gronlund NE. Assessment of student achievement. ERIC Number: ED417221. SBN-0-205-26858-7. Available at: <https://eric.ed.gov/?id=ED417221>.