ROYAL COLLEGE
OF PHYSICIANS AND SURGEONS OF CANADA

Short-Answer Questions: Guidelines for their Development

Editorial Revisions
July 2017

Educational Research Unit
Unité de la recherché pédagogique

I. Introduction

SAQs are questions that can be answered in a few short words or phrases. Typically, these questions contain words such as "list" or name" suggesting that the answer consists of a series of short responses.

Each SAQ will usually consist of the question, a space for the candidate to provide an answer and the mark allocation for the question. A clinical scenario may be used to provide background information pertinent to the question.

Although this format seems straightforward, there are a few principles that should be applied to improve the quality of the question. When developed correctly, SAQs can be a valid and reliable evaluation method.

II. What are good short-answer questions?

When properly created, the short-answer question (SAQ) format is a RELIABLE, CONSTRUCTED-RESPONSE assessment tool.

- Reliable means that it can be consistently and fairly scored.
- "Constructed response" means that candidates are asked to write down their answers to a given question (RECALL) instead of selecting from a list of provided options like multiple-choice and matching formats (RECOGNITION).

The SAQ format GUIDES CANDIDATES to provide structured, straightforward responses to the questions.

- When given a well-constructed SAQ, candidates understand what is expected of them and how marks are allocated to their answer. The question should be challenging, unambiguous and relevant, and should lead to a concise defined answer.
- Typically, SAQs contain words such as "list" or "name" and ask for a specified number of responses. Words such as “describe, explain or outline” should not be used as they do not clearly indicate what is expected. Visual guidelines such as a series of lines are often provided on the answer sheet to indicate the expected length of the answer.

The creation of a comprehensive, structured MODEL ANSWER is vital to a high-quality item in the SAQ format.

- Model answers should be designed to leave very little room for interpretation for the rater. The model answer must therefore be comprehensive (i.e. it must include all possible responses that a candidate might provide). It must also be correct and anticipate all interpretations of the question.
- Any individual that marks the SAQ should be able to fairly and reliably score the responses and allocate the same marks to **any** given response.

Short-answer questions (SAQs) may be prefaced by a VIGNETTE or SCENARIO in order to put the question into an applicable setting.

- Vignettes provide a good opportunity to test candidates in clinically relevant situations and should reflect common practice.
- Questions should assess a variety of cognitive levels by testing recall of knowledge as well as some application of the knowledge. Problem-solving and other critical thinking skills such as synthesis and evaluation are better evaluated using essays or performance assessments. Due to the complexity of the tasks, these tend to be less reliable and more subjective.

SAQs may consist of several SUB-QUESTIONS that build on a particular content area or vignette.

- Smaller straightforward question often simplify the marking task and lead to reliable scoring while adding depth to the assessment.
- Avoid content specificity. Content specificity occurs when a disproportionate weight of an examination focuses on one specific content area. If a question has too many sub-questions or too many points allocated to its specific content, the question may reduce the content validity of the examination.
- Avoid interdependence of answers. Candidates may receive lower marks if they make a mistake early on but are required to answer sub-questions that build on answers from previous ones.

The goal of good evaluation is to maximize the VALIDITY and RELIABILITY of the assessment tool.

- Validity refers to whether the test assesses what it should assess. This includes issues such as ensuring that the examination looks appropriate (face validity) and that the content is accurate, balanced and appropriate (content validity).
- It is essential that questions display good face validity and content validity so that the assessment is based on the candidate's competence and performance rather than on the perception of the fairness and appropriateness of the examination or the method of assessment.
- Reliability refers to consistency and fairness.

III. Short-answer questions (SAQs) compared to other common written assessment formats.

- There is no perfect assessment format. The challenge of designing assessment materials is to select the most appropriate format and content to assess the material you have chosen to test.
- SAQs use the constructed response format. They are similar to essay-type questions, but can be scored in a more objective manner because they are more directed and concise.
- Multiple-choice questions (MCQs) can be scored in the most objective manner because they use the selected response format.
- SAQs, however, do not display the answers to the questions therefore emphasize ***recall*** rather than ***recognition***.

	MCQs / Matching / True-False	**SAQs**	**Essays / Short Essays**
Response Type	Selected response	Constructed response	Constructed response
Scoring	Highly objective scoring	Objective scoring	Subjective scoring
Structure of response	Highly structured / Choice of provided options	Moderately structured / Some anticipated variation	Low level of structure/ Unstructured
Consistency of scoring	Very consistent	Consistent	Hard to mark consistently
Inter-rater consistency	High inter-rater reliability	Good inter-rater reliability	Poor inter-rater reliability

	MCQs / Matching / True-False	SAQs	Essays / Short Essays
Level of cognitive complexity	Assesses recognition or application of knowledge	Assesses recognition or application of knowledge	Best assesses problem-solving

IV. Question Structure

The following tables outline some common review questions that should be considered guidelines for the development when creating and reviewing SAQs.

a. Does the question follow the short-answer question (SAQ) format?

REVIEW QUESTIONS	COMMENTS
1. Is the question an SAQ and not an essay?	Verbs such as "explain," "outline," "define" and "compare" tend to lead to essays, whereas "list" and "name" lead to SAQs.
2. Does the question clearly outline the problem?	Does the question give an explicit expectation of what information is required in the response?
3. Can the question be interpreted in only one way?	Is there only one way to interpret the question?
4. Does the question specify the number of statements/ideas required in the answer?	List FIVE appropriate medications to ….
5. Will the candidates' answers reflect your intentions (content, interpretation, etc.)?	Will the question elicit a brief and concise response? Does the question indicate the depth/complexity of the content required in the answer?
6. Has the question been simply stated so that there is no benefit to breaking the question into a series of simpler questions?	Avoid using the word "and" to asking two separate questions. Perhaps questions could be separated to ask for the "MOST important" versus "other less important" if that is what is being assessed.
7. Does the question ask for a standard answer rather than an opinion?	Avoid using "would" because what a candidate WOULD do may not be what they SHOULD do.
8. Is the mark allocation clear to the candidate?	Is it straightforward and fair or tricky and unclear?
9. Is the question relevant and appropriate for testing using the SAQ format?	Is the information better assessed using another format? Is the question assessing trivia?
10. Has the question been designed so that it does not give away answers to other questions?	Are clues to the answer given away in the scenario or in other exam questions?

11. Are questions/sub-questions independent of each other?	If the candidate response is incorrect for one question, can subsequent questions still be answered?

b. Scenarios – Does the clinical scenario add to the validity of the question?

REVIEW QUESTIONS	COMMENTS
12. Does the scenario contribute to/establish clinical relevance of the SAQ?	Does the scenario improve the face validity and content validity of the question?
13. Is the length of the clinical scenario appropriate?	Is the scenario too long given the weight of the question? Is it detailed enough to answer the question(s)?

c. Model Answers – Can the model answer be used to reliably score the question?

REVIEW QUESTIONS	COMMENTS
14. If answers call for numbers, have marks been allocated to reflect a suitable margin of error?	If, for example, the answer is 5%, are partial marks given if candidates answer “4.5” or “under 10%”? Markers need a comprehensive approach for numerical questions.
15. Does the model answer allow the markers to score the responses reliably?	Do the answers provided take all possible interpretations of the question into consideration?
16. Are the marks allocated in a fashion that leads to a fair estimate of competence?	How does the scoring of the question contribute to the overall examination? (This question usually requires the “Big Picture” of the overall exam and may not be appropriate when critiquing individual questions.)

Improving SAQs

Example 1

Poor SAQ

You see a 45-year-old man in your office for headaches. The man has six children between the ages of 1 month and 10 years. He is not on any medications. He denies the use of alcohol or recreational drugs. He is not currently sexually active. The headaches started one year ago and have been increasing in frequency. They now occur four to six times per month. With each headache he misses three to four days of work. You suspect he suffers from migraine.

a) List SIX features of migraine.
b) List TWO medications for migraine.

Answer key

a) One-half mark for each (maximum two marks)
- throbbing
- associated with nausea/vomiting
- associated with photo/phonophobia
- hemicranial pain
- preceeding aura
- inability to continue with daily activities
- relieved by sleep

b) One mark for each (maximum two marks)
- amitriptyline
- propranolol
- flunarizine
- valproic acid

Improved SAQ

You see a 45-year-old man in your office for headaches. The headaches started one year ago and have been increasing in frequency. They now occur four to six times per month. With each headache he misses three to four days of work. You suspect he suffers from migraine.

a) List FOUR *clinical* features of migraine. *(2 marks)*
b) List TWO *prophylactic* medications for migraine. *(2 marks)*

This SAQ has been improved by:

a) being more directive and less general so candidates can provide answers you expect
b) removing excess wording that does not contribute to the question. Clinical scenarios appear to make the question more relevant but are optional.
c) indicating the score allocation to the candidate.

Note that the number of responses requested significantly changes the difficulty of the question.

Example 2

Poor SAQ

List the clinical features found in Poland syndrome. (Four marks)

Answer key

Two marks each for a, b and c; One mark each for d, e, f, g and h (maximum four marks)

a. Absence of sternal head of pectoralis major muscle (or complete absence of the muscle)
b. Deficiency or absence of the breast and areola
c. Brachysyndactyly of ipsilateral hand/upper limb shortening
d. Deficiency of subcutaneous fat and axillary hair
e. Abnormalities of costal cartilages and anterior end of ribs
f. Vascular anomalies on affected side
g. Absence or hypoplasia of other thoracic muscle (lattisimus dorsi, pectoralis minor, serratus, infraspinatus, supraspinatus, external oblique)
h. Scoliosis

Improved SAQ

List FOUR clinical features found in Poland syndrome (agenesis of the pectoralis muscle). (Four marks)

Answer key

One mark each for any four (maximum four marks)

- Deficiency or absence of the breast and areola
- Brachysyndactyly/underdevelopment of ipsilateral hand
- Ipsilateral upper limb shortening/underdevelopment
- Deficiency of subcutaneous fat and axillary hair
- Abnormalities of costal cartilages and anterior end of ribs
- Vascular anomalies on affected side
- Absence or hypoplasia of other thoracic muscle (lattisimus dorsi, pectoralis minor, serratus, infraspinatus, supraspinatus, external oblique)*
- Scoliosis

Note: Only give maximum one mark for listing any muscles

This SAQ has been improved by:

a) clarifying the expectations for the candidate
b) focusing the intent of the question (test important clinical knowledge (other congenital anomalies associated with agenesis of the pectoralis muscle) rather than obscure facts (definition of Poland syndrome)).

The answer key is more reliable when decisions can be made ahead of marking time in situations where there may be confusion.

Example 3

Poor SAQ

51 year old man is diagnosed with Barrett's esophagus.

A List 3 treatment options that you would recommend? (6)
B Outline the pathogenesis of Barrett's esophagus. (8)
C List the factors that aggravate gastroesophageal reflux disease. (4)
D List 2 benefits and disadvantages to the long-term health of this patient of enrolling him in a surveillance program. (2)

Improved SAQ

51 year old man is diagnosed with Barrett's esophagus.

A What are 3 treatment options that should be pursued? (6)
B List 4 steps in the pathogenesis of Barrett's esophagus. (8)
C List 2 factors that aggravate gastroesophageal reflux disease. (4)
D List 2 benefits to the long-term health of this patient of enrolling him in a surveillance program. (2)
E List one ongoing clinical issue that needs to be addressed pertaining to the long-term health of this patient of enrolling him in a surveillance program. (2)

This SAQ has been improved by:

- Removing the use of the word "outline" as it implies an essay type question. Verbs such as "outline", "describe", "summarize", etc. should be avoided in SAQs.
- Not asking for opinions. Candidates should not be asked to state their preference for treatment, tests etc. (what would you do?) as such an answer could reflect common practice or available options in a particular environment, rather than the correct approach.
- Stating the number of responses in the question that reflect answer expectations.
- Using direct questions when possible.

Be careful when using the word 'and' because it may indicate that more than one question is being asked. Be clear what is expected.

Appendix

SAQ Quality Checklist Review

1 Does the question clearly outline what is to be evaluated?
2 Does the candidate have clear expectations of what they should write?
3 Can raters consistently mark the question?

Question Structure

- Questions are in-line with their objectives.
- Questions are realistic and relevant for the purpose of the exam.
- Candidates have clear expectations of what they should write for each question/sub-question.
- Questions are direct.
- Questions/sub-questions have a single focus.
- Wording is clear and unambiguous.
- Questions are not short/long essay questions.
- Questions use action verbs such as 'List' or 'Name', when appropriate.
- Words such as 'Discuss', 'Describe', 'Summarize' and 'Outline' are avoided.
- Qualifiers such as 'most likely' and 'best' are used when necessary.
- Questions limit the length of answers or specify a restriction when needed.
- Questions specify the number of elements required.
- Questions that focus on numbers are avoided (if used, they must have a clearly defined marking key and candidates must know how marks are allocated, i.e. if/how many marks are given for the answer/calculation/both).
- Visual guidelines are provided to indicate how long an answer is expected.
- Questions are not dependent on answers to other questions.
- The number of marks for each question, sub-question, and each element required (e.g. list THREE…) are specified.
- The mark allocation reflects the importance of each question.
- Questions are at an appropriate level of difficulty and describe relatively plausible situations.

Model Answers

- Marks are allocated to each sub-question and answer alternative.
- Marks allocated to answer alternatives reflect the value of those alternatives in relation to other answer alternatives (e.g. excellent answer (2), good answer (1.5), acceptable answer (1)…).
- The answer key is designed to ensure reliable scores.
- Marks allocated to questions/sub-questions reflect the importance/difficulty of those questions/sub-questions in relation to the other questions on the exam.
- Model answers are comprehensive.

Scenarios / Vignettes

- Vignettes are not disproportionately long in relation to the value of their questions.
- Cues are not provided in the vignettes.
- Vignettes are at an appropriate level of difficulty and describe relatively plausible situations.