

## EDUCATIONAL RESOURCE

Volume 14 Issue 2 2022

DOI: 10.21315/eimj2022.14.2.9

### ARTICLE INFO

Received: 15-10-2021

Accepted: 26-01-2022

Online: 30-06-2022

# A Guide for Writing Single Best Answer Questions to Assess Higher-Order Thinking Skills Based on Learning Outcomes

Ahmad Fuad Abdul Rahim<sup>1</sup>, Anna Alicia Simok<sup>2</sup>, Shaik Farid Abdull Wahab<sup>3</sup>

<sup>1</sup>Department of Medical Education, School of Medical Sciences, Universiti Sains Malaysia, Kelantan, MALAYSIA

<sup>2</sup>Department of Anatomy, School of Medical Sciences, Universiti Sains Malaysia, Kelantan, MALAYSIA

<sup>3</sup>Department of Emergency Medicine, School of Medical Sciences, Universiti Sains Malaysia, Kelantan, MALAYSIA

**To cite this article:** Abdul Rahim AF, Simok AA, Abdull Wahab SF. A guide for writing single best answer questions to assess higher-order thinking skills based on learning outcomes. *Education in Medicine Journal*. 2022;14(2):111–124. <https://doi.org/10.21315/eimj2022.14.2.9>

**To link to this article:** <https://doi.org/10.21315/eimj2022.14.2.9>

## ABSTRACT

The single-best answer (SBA) question is popularly used in medical education assessment. Writing an SBA which assesses higher order thinking skills (HOTS) is daunting for newcomers as it requires familiarity with learning outcomes (LOs). This guide provides steps to create LOs for HOTS, introduces the SBA, its parts and how the parts relate to the LO. It then provides steps to convert the LO into an SBA. Examples from anatomy, emergency medicine and medical education are provided.

**Keywords:** *Assessment, Single-best answer question, Higher order thinking skills, Learning objectives, Item writing*

## CORRESPONDING AUTHOR

Anna Alicia Simok, Department of Anatomy, School of Medical Sciences, Health Campus, Universiti Sains Malaysia, 16150 Kubang Kerian, Malaysia

Email: [annalicia@usm.my](mailto:annalicia@usm.my)

## INTRODUCTION

The single-best answer (SBA) question, otherwise known as the one best answer (OBA) question, the Type-A question (1) or simply as the multiple-choice question (MCQ) (2), is a paper-based assessment method popular in undergraduate and postgraduate medicine (3). It is an assessment tool for the cognitive domain that is best used to assess problem solving

and higher-order thinking skills (HOTS) (4).

However, writing an SBA question might be daunting for newcomers. Creating an SBA question to assess HOTS requires an understanding of learning outcomes (LOs). Many item-writing guidelines do not emphasise this point or relegate it to a supplemental section deep within their guides.

## OBJECTIVES

To help new SBA-question writers, we describe the steps to write LOs for learning and teaching sessions. We then introduce the SBA question, its parts, how each part relates to a LO and some writing tips. We then present steps to convert LOs to SBA questions that assess HOTS. We provide examples from anatomy (ANAT), emergency medicine (EM) and medical education (ME) to illustrate the guidelines.

## THE STEPS TO WRITE YOUR SESSION-LOs

### Step 1: Identify Your Content Area

People will ask you, as an educator, to provide SBA items (questions) that assess a specific content area. This content area may be a lecture, a tutorial or a seminar you presented. Some institutions might provide you with a blueprint that details the topic area, the number of questions required, etc.

Table 1 shows examples of content areas from ANAT, EM and ME.

**Table 1:** Example content areas from ANAT, EM and ME

Discipline	Content area example
ANAT	Neuroanatomy of hypothalamus, epithalamus and metathalamus
EM	Principles of management in anaphylaxis/hypersensitivity reaction
ME	Principles of assessment – Guidelines for constructive alignment

### Step 2: Determine the LOs for Your Content Area According to Bloom’s Cognitive Taxonomy Levels

After identifying the content area, you need to determine your LOs. LOs are the abilities you expect your students to have after

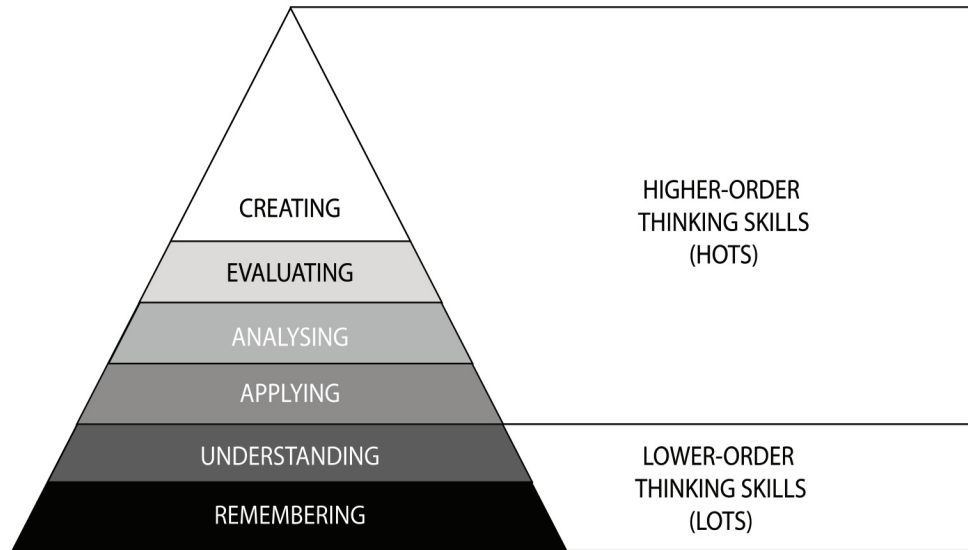
going through your instructions. LOs are commonly presented as verbs related to the cognitive domain, such as list, enumerate, outline or discuss. These verbs make the outcome apparent and easier to assess. For example, how do you assess the outcome “Students will be able to understand the pathogenesis of carpal tunnel syndrome?” How do you ascertain that they have understood it? In contrast, assessment becomes easier if the outcome is “Students will be able to describe the pathogenesis of carpal tunnel syndrome.” Give them a piece of paper and ask them to describe it.

This guideline focuses on cognitive domain (thinking skills) outcomes. It is one of the three domains in education, the other two being the psychomotor (physical skills) domain and the affective (behaviour or attitude) domain. The psychomotor and affective domains are beyond the scope of this guide because they require other forms of assessment.

A course coordinator is usually the person responsible for writing or collating LOs. These might be related to the entire course. To write SBAs, you need LOs for your specific content area. Your predecessor might have left some documentation related to your LOs. If there is none, you need to write the LOs. It would look something like this:

At the end of this class, a student would be able to **interpret** data from clinical trials.

Often, it would begin with “At the end of...”, but this is not compulsory. Within the cognitive domain, there are levels of thinking skills. Figure 1 shows these levels. Bloom et al., who proposed these levels, considered them a taxonomy (5). Beginning with the most basic level, remembering, a learner progresses to higher thinking skills only after mastering the lower levels.



**Figure 1:** Blooms taxonomy of educational objectives (revised) – the cognitive domain showing the division of the levels into LOTS and HOTS (6).

These levels are further divided into two groups: lower order thinking skills (LOTS), comprising the lower two levels, and HOTS, comprising the upper four levels (6). Studies have determined that HOTS LOs are deficient in Health Professions Education (7–9), including in its assessment (10). Therefore, think about the upper four levels when writing your outcomes. Practically, the two topmost levels, evaluating and synthesising, are usually reserved for PhD studies; therefore, we often only focus on the applying and analysing levels for our HOTS LOs.

Not all outcome verbs, however, are suitable for SBA items. For example, “Students will be able to discuss the trends in medical curriculum design” is better assessed in an essay where students write freely on paper. “Students will be able to classify an open fracture using the Gustilo–Anderson classification” is more suitably tested

using an SBA item: We give them a written description of an open fracture, and they choose the correct classification from a list.

Therefore, having well-defined LOs for your teaching ensures that you assess what is important and that you assess them using the proper tools. It is often more challenging to create HOTS outcomes for non-clinical disciplines, but it is not impossible. To help yourself create a HOTS LO, ask, “How will students use this information in their workplaces?” or “How is this knowledge important in real-life situations?”

Table 2 provides the cognitive domain HOTS levels (application and analysis), some suitable verbs for each level, definitions and example LOs. Use these verbs to write your LOs. Note that there is an overlap between the definitions of verbs. The verbs provided are not meant to be authoritative or exclusive. You can look up many other verbs on the internet.

**Table 2:** Bloom’s cognitive domain HOTS levels, suitable verbs and their associated definitions and examples

Level in Bloom Taxonomy	Name	Suitable verbs
3	<b>Application:</b> To use previously learned information in novel situations; solve problems	<p><b>Apply:</b> To use knowledge in a new or practical situation. Examples:                      ANAT: To apply the knowledge of hypothalamus functional anatomy in determining the site of a lesion in the hypothalamus.                      EM: To apply the knowledge of cardiac vascular anatomy to determine the site of cardiac ischemia based on electrocardiography or echocardiography.                      ME: To apply the knowledge of suitable assessment methods for each level of Miller’s pyramid of clinical competencies in choosing an appropriate assessment method for a specific learning situation.</p> <p><b>Ascertain:</b> To determine or decide. Examples:                      ANAT: To ascertain the most likely site of an intracranial lesion by looking at the clinical findings in a patient.                      EM: To ascertain the requirements to stop CPR in a collapsed patient                      ME: To ascertain the relevant threats to validity in an assessment situation.</p> <p><b>Assign:</b> To give a characteristic, role or value to something. Examples:                      ANAT: To assign the subdivision of hypothalamic nuclei based on their specific function.                      EM: To assign the type of shock suffered by a patient based on the clinical features.                      ME: To assign a label of LOTS or HOTS to a learning objective based on the verb used.</p> <p><b>Calculate:</b> To determine the numerical value of something using other numerical data. Examples:                      ANAT: To calculate the concentration of a fixative solution in tissue processing.                      EM: To calculate the compensation of the acid-base balance in a patient using Winters’ Formula.                      ME: To calculate the difficulty index of an item.</p> <p><b>Decide:</b> To select a course of action or a choice after considering several options and other criteria. Examples:                      ANAT: To decide the site of a lesion in the hypothalamus based on the clinical manifestations.                      EM: To decide the best approach in managing a shoulder dislocation.                      ME: To decide on an assessment method after considering the assessment domain and several possible assessment methods.</p>

*(continued on next page)*

Table 2: (continued)

Level in Bloom Taxonomy	Name	Suitable verbs
4	Analysis: To understand the organisational structure of information; see patterns; organise parts	<p><b>Classify:</b> Described above.</p> <p><b>Confirm:</b> To determine something previously uncertain by looking at the evidence. Examples:</p> <p>ANAT: To confirm the anatomical site of hypothalamic injury based on the clinical manifestations.</p> <p>EM: To confirm brain death in severe traumatic brain injury.</p> <p>ME: To confirm the loading of assessment tools in an examination by confirmatory factor analysis.</p> <p><b>Determine:</b> To ascertain, confirm or establish something by looking at evidence or calculating. Examples:</p> <p>ANAT: To determine the anatomical site of an intestinal mucosal tissue sample based on its histological features.</p> <p>EM: To determine the requirement for using an automated external defibrillator during resuscitation.</p> <p>ME: To determine the most likely cause for an item having a low discrimination index</p> <p><b>Diagnose:</b> To determine the presence of a disease by looking at clinical evidence. Examples:</p> <p>ANAT: Not applicable.</p> <p>EM: To diagnose diabetic ketoacidosis</p> <p>ME: Not applicable</p> <p><b>Differentiate:</b> To distinguish between two or more classes or groups of things by looking at their features. Examples:</p> <p>ANAT: To differentiate between supraoptic and tuberal regions of the hypothalamus based on their nuclei and function.</p> <p>EM: To differentiate between sympathomimetic and anticholinergic drug overdose in a patient by their clinical presentation.</p> <p>ME: To differentiate between written assessment tools based on their formats, advantages and disadvantages.</p> <p><b>Discriminate or distinguish:</b> Similar to differentiate.</p>

### Step 3: Write the LOs for Your Session

You can use the format given in the Table 3. You should have three to four LOs for your session, one of them HOTS. Readers are encouraged to consult more detailed guides for writing LOs in the health sciences, such as “Learning objectives in radiology education: why you need them and how

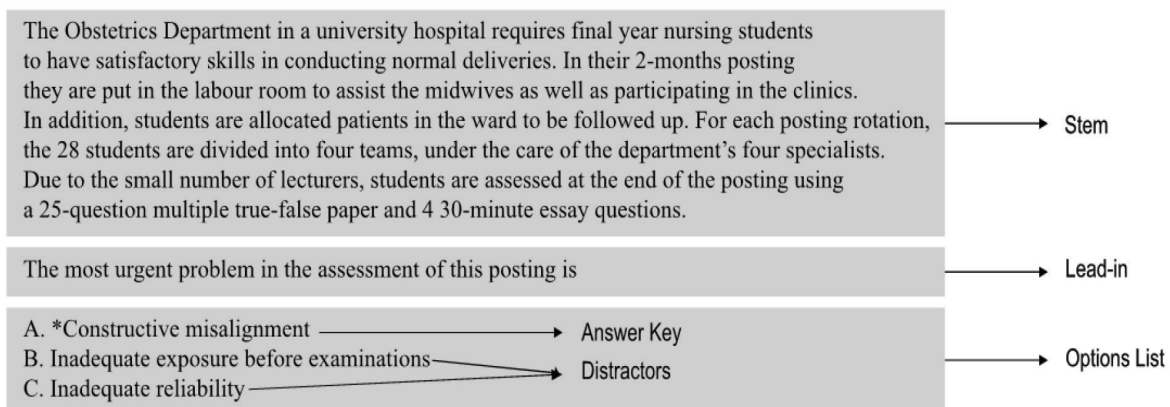
to write them” (11) or “SMART tips for setting teaching objectives” (12).

### THE SBA QUESTION

We will now introduce the SBA (Figure 2) and explain the relevance of each part to your LOs.

**Table 3:** Examples of LOs session from ANAT, EM and ME

Discipline	LOs session
ANAT	<p><b>Describe</b> the location, relations, parts, nuclei, connections and functions of the hypothalamus.</p> <p><b>Describe</b> the location, relations, parts, nuclei, connections and functions of the epithalamus.</p> <p><b>Describe</b> the location, relations, parts, nuclei, connections and functions of the metathalamus.</p> <p><b>Relate</b> the nuclei, connections and functions of the hypothalamus with the pathogenesis of diabetes insipidus.</p>
EM	<p><b>Describe</b> the pathophysiology of anaphylaxis/hypersensitivity reactions.</p> <p><b>Describe</b> the types of anaphylaxis/hypersensitivity reactions:</p> <ol style="list-style-type: none"> <li>Type I</li> <li>Type II</li> <li>Type III</li> <li>Type IV</li> </ol> <p><b>Describe</b> the clinical features of anaphylaxis/hypersensitivity reactions.</p> <p><b>Decide</b> the management for the anaphylaxis/ hypersensitivity reaction based on the patient condition.</p>
ME	<p><b>Describe</b> the concept of assessment utility and its components:</p> <ol style="list-style-type: none"> <li>Validity</li> <li>Reliability</li> <li>Educational Impact</li> <li>Feasibility</li> <li>Acceptability</li> </ol> <p><b>Describe</b> the threats to validity and <b>discuss</b> ways to overcome it.</p> <p><b>Ascertain</b> the relevant threats to validity in an assessment situation and suggest ways to improve it.</p>



**Figure 2:** The parts of the SBA question.

## Stem

A stem is the data presented to a candidate to answer a question. It is usually in the form of a written scenario. If relevant, you can use other stem types, such as diagrams, graphs, photographs or videos (1). A stem should present sufficient data to allow a candidate to perform the cognitive verb in your session LO. For example, if the outcome states that a candidate should decide on a management approach, there should be enough information for him or her to do so.

### Writing tips

For clinical disciplines, base your scenario on actual cases and modify it according to the levels of your candidates. Common and classical presentations are suitable for junior candidates, while atypical presentations suit advanced-level candidates. A candidate must interact with the data in the stem to answer the question in the lead-in. This requirement leads to two issues. First, a stem must be context-rich (13), authentic and related to a candidate's work. It must also be new or novel (14); a very complicated and authentic stem will not invoke HOTS if you have discussed it in your class. Second, avoid interpreting data; write "The pulse rate was 100 per minute" instead of "The patient has tachycardia." Describing data instead of interpreting it increases question discrimination (1). However, you can consider interpreted data for very junior candidates, and the LO is non-clinical.

Similarly, avoid using pathognomonic terminologies for certain conditions if you assess them. For example, avoid writing "rebound tenderness is present" to evaluate a candidate's ability to recognise peritonitis. Instead, write "Sudden removal of the palpating hand results in pain."

For non-clinical disciplines, use situations in which the knowledge is applied. Ask yourself, "How will my students apply the

knowledge of body acid–base balance later when they are working?"

Avoid intentionally misleading the candidate.

After reading your stem and the lead-in, a competent candidate should know the correct answer, even when the options list is covered. This characteristic is called "the cover test" (1) and helps you decide whether sufficient information has been provided in your stem.

Do not worry if your stem appears to be a bit long; long stems and short options are preferred for SBA questions. However, when deciding on the suitable length of the stem, consider the time that candidates have to answer a question.

## Lead-in

SBA is a select-type written question; candidates must select an answer from a list of possible answers (1). A lead-in provides the basis for choosing the answer from a list, e.g., "Which of the following is the likeliest anatomical site of the lesion in this patient?" or "Select the likeliest anatomical site of the lesion in this patient."

You need not include the same verb stated in your LO; just write the lead-in to make a candidate perform the verb when choosing an answer. A likely LO for the lead-in given above is: "The student will be able to determine the anatomical site of a lesion in the central nervous system after considering the clinical signs in a patient."

### Writing tips

Use a complete sentence (e.g., "Which of the following is the likeliest diagnosis?") or an incomplete one (e.g., "The likeliest diagnosis is..."). Incomplete sentences, however, risk ending with "a" or "an," therefore possibly giving away an answer if it is the only one beginning with a consonant or vowel (i.e., grammatical cueing [1]).

Avoid writing lead-ins that are answerable without consulting the stem (i.e., standalone lead-ins). For example, after writing a detailed scenario about a patient with fever, the lead-in asks, “Which of the following is the confirmatory investigation for dengue?” This does not require reading the stem for candidates to answer correctly. The ability to name the confirmatory investigation for dengue is a LOTS outcome, for which no amount of rewriting will correct.

Avoid testing two LOs in an item, such as, “Which of the following is the likeliest diagnosis and management for this patient?” as this tends to be confusing.

### Options List

An options list contains three to five possible answers for a candidate to choose from to answer the question in the lead-in. It includes the answer key, which is the correct answer. A candidate is given marks when he or she selects the answer key. Distractors are options in a list that are incorrect; candidates are not given any marks if they choose any distractor. They are named distractors because they distract incompetent candidates from selecting the answer key.

### Writing tips

Options must match your outcomes. If your candidates must decide on management approaches, all options must consist of management approaches.

All options must be of the same type (homogenous) and of similar length.

Avoid lengthy, complicated options.

Question writers usually think of a correct answer first, and it usually ends up being the first on an options list. Arranging the options in a logical order, such as alphabetical or anatomical, will help reduce the chances of this occurring.

Distractors are as critical as other parts of an SBA question (14–16) but have received relatively little attention in the literature (16). Question writers give a lot of thought when writing a stem, and by the time they write the answer key on the options list, they are exhausted. As a result, the least thought is given to distractors, making them inefficient.

In a good SBA question, the answer key is chosen by competent candidates, and distractors attract the attention of the incompetent ones; this is a discriminating question (14). Distractors work as a team; if any distractor is dysfunctional (chosen by less than 5% of candidates [17]), there are fewer options to choose from, making them easier for incompetent candidates.

Therefore, much thought needs to go into writing distractors. Distractors must be plausible; they appear correct to incompetent candidates (14, 16). Your teaching experience will help you; use common mistakes or misconceptions you know students make. After writing your key, ask yourself, “What do students usually confuse this concept or idea with?”, “What is a common error for solving this problem?” and “What are the common misconceptions in this field?” (18). When you have written a distractor, ask yourself, “What makes this attractive to incompetent candidates?”

There are two issues related to functional distractors. The first is related to the assumed requirement that distractors must have a certain degree of truth; it is just that they are not the best answers among the options. Having partially true distractors is desirable and appears to be associated with HOTS for those answering it. However, we find that writing such distractors is easier for clinical disciplines and more difficult for the basic sciences. Strictly adhering to this requirement is NOT necessary (1, 16, 18) as long as distractors appear attractive to incompetent candidates.



Completely incorrect statements, especially if they test popular misconceptions amongst students, can be good distractors. The latest version of the National Boards of Medical Examiners (NBME) Item-Writing Guide (1), where the OBA question (termed the Type-A question) is discussed in detail, states that “Incorrect options can be partially or wholly incorrect.” Collins (18) wrote, “The best distractors are (a) statements that are accurate but do not fully meet the requirements of the problem and (b) incorrect statements that seem right to the examinee. Each incorrect option should be plausible but clearly incorrect.”

We discourage the use of questions such as the following for assessing HOTS:

The antidote for paracetamol poisoning is:

- a. Lysine acetylsalicylate
- b. N-acetylcysteine
- c. Sodium hydrochloride

Questions like these assess LOTS and encourage surface learning. However, questions that force a candidate to consider the data presented in the stem when answering a question and, most importantly, present plausible and effective distractors should be used, even though the distractors are false.

The second issue is the number of distractors in a SBA question. Experience tells us that question writers can produce one or two good distractors. After this, they run out of ideas, and the falseness of further

distractors becomes increasingly apparent. Accumulated research supports this (19); the ideal number of options seems to be three—one answer key and two distractors (good ones!).

## WRITING YOUR SBA FROM YOUR OUTCOMES

### Convert Your LO into an SBA-question Template

An SBA-question template helps you think about what you need to provide in your SBA for candidates to perform a HOTS verb in your LO. For example, if the verb is “determine”, you need to think: “If I want my students to determine something, then I need to provide them THIS kind of data in the question stem and ask them THIS way in the lead-in.”

Have a look at your LO and fill in the following SBA-question template:

Given a [stem type] that provides [data in the stem], a candidate will be able to [HOTS verb] the correct [related outcome content] and choose it from an options list containing [options description].

Example templates from ANAT, EM and ME are given in Table 4. You can use these templates to write new questions that assess the same outcomes; just replace the stem with a new or modified scenario and the options list.

**Table 4:** Example SBA templates from ANAT, EM and ME

Discipline	LO	Example SBA Template
ANAT	Relate the nuclei, connections and functions of the hypothalamus with the pathophysiology of diabetes insipidus	Given [a written scenario] which provides [relevant signs and symptoms, the results of appropriate investigations, and other pertinent data that points to diabetes insipidus], the candidate will be able to [relate] the correct [lesional site] and choose it from an options list containing [hypothalamus nuclei].

*(continued on next page)*

**Table 4:** (continued)

Discipline	LO	Example SBA Template
EM	Decide the management for the anaphylaxis/hypersensitivity reaction based on the patient condition	Given [a written scenario] which provides [relevant clinical information of a patient having an anaphylactic/hypersensitivity reaction], the candidate will be able to [decide] the correct [management] and choose it from an options list containing [a list of management steps in anaphylaxis/hypersensitivity reaction].
ME	Ascertain the relevant threats to validity in an assessment situation and suggest ways to improve it	Given [a written scenario] which provides [the domain of an assessment, the assessment method, and other relevant data], the candidate will be able to [ascertain] the correct [threat to validity] and choose it from an options list containing [types of threats to validity].

By this time, you can see that absent or unclear LOs make writing effective SBA questions difficult. However, well-written LOs that contain LOTS verbs are unsuitable for SBAs. For example, if students need to list the carpal tunnel content, you just need to ask them to do so. You need not provide any scenarios. Employing an SBA question to assess it is a misuse of an assessment tool.

**Convert Your LO into SBA Question**

It is time to convert the template into an actual SBA question. Guided by the template, write the stem to provide pertinent data, create a suitable lead-in and prepare the options list that contains the answer key and two to four functional distractors.

To help yourself, refer to the previous writing tips. You are encouraged to read detailed guides for the fine art of writing SBA questions, such as the NBME’s Constructing Written Test Questions for the Health Sciences (1), downloadable from its website. Another useful resource is Education Techniques for Lifelong Learning: Writing Multiple-choice Questions for Continuing ME Activities and Self-assessment Modules (18).

Table 5 provides samples of SBA questions from anatomy, EM and ME written based on their stated LOs.

**Table 5:** Examples of SBA questions from ANAT, EM and ME

Discipline	Example
ANAT	<p>A 45-year-old female presented with a weight loss of 5 kg in the last four months. Further questioning revealed that she had increased thirst. She also noticed that she is passing more urine and wakes up at night to pass urine.</p> <p>Her fasting blood sugar was 5.0 mmol/l.</p> <p>Magnetic resonance imaging showed small nodular enhancing lesions seen superior to the optic chiasm.</p> <p>Which of the following is the most likely affected area?</p> <ol style="list-style-type: none"> <li>Paraventricular nuclei of hypothalamus</li> <li>Posterior pituitary</li> <li>Supraoptic nuclei of hypothalamus (Key)</li> </ol>

(continued on next page)

**Table 5:** (continued)

Discipline	Example
EM	<p>A 25-year-old male presents with a sudden onset of shortness of breath, facial swelling and generalised rash. An accompanying friend gives details of eating seafood in a nearby food stall before the onset of symptoms. He has had similar attacks before, but none was this severe.</p> <p>His blood pressure is 90/40 mmHg, heart rate 130 beats/min, respiratory rate 30 breaths/min and temperature 37.8°C. His lips and tongue are swollen, and he has generalised urticaria.</p> <p>Rhonchi are heard bilaterally. Cardiovascular and abdominal examinations are normal.</p> <p>Which of the following is the first-line management for this patient?</p> <ol style="list-style-type: none"> <li>Intramuscular adrenaline 0.5 mg (Key)</li> <li>Intravenous chlorpheniramine 10 mg</li> <li>Intravenous hydrocortisone 200 mg</li> </ol>
ME	<p>In X School of Medicine, students must pass two long cases to pass the final professional examinations. They are assessed independently during the long case by three senior clinicians, who undergo intensive training for the examiner role. There are no other forms of clinical assessment.</p> <p>Which of the following is the most serious threat to validity in this situation?</p> <ol style="list-style-type: none"> <li>Insufficient cases (Key)</li> <li>The lack of authenticity in the assessment task</li> <li>The subjectivity of the assessment method</li> </ol>

### Review Your SBA Questions

Besides adhering to guidelines particular to SBA questions, your questions should conform to standard grammatical and language conventions. Again, we recommend excellent guidelines, such as *Constructing Written Test Questions for the Health Sciences* (1), or resources from general assessment books, such as *Developing and Validating Test Items* (Chapter 6) (14), or those specific to the health professions, such as *Assessment in Health Professions Education* (Chapter 7) (2).

Questions usually undergo a formal review process. If not, have someone look at your questions. A departmental review ensures content accuracy, whereas a central vetting process checks for appropriate difficulty levels and conformity to a prescribed format. An open mind is required to submit

questions for review; we often think that our questions are the epitome of literary creations.

### Evaluate Your SBA Question

“Fire and forget” is a characteristic of guided missiles, not question writers. Nowadays, software that scans SBA-question answer sheets almost always provides item analysis data useful for teaching and question writing. Item analysis data are usually available in the department that carried out the scanning. Ask to see it.

A detailed discussion of item analysis is outside the scope of this guide. In brief, there are three things you look at:

The discrimination index is the extent to which your question was able to differentiate or discriminate between competent and incompetent candidates. Their overall

rank in the examination defines this competency (20). Look for “discrimination index”, “DI”, point biserial index (PBIS) or some other name. The value ranges from –1 through 0 to 1. A good question is discriminating: more competent candidates can answer it compared to incompetent ones. A discrimination index (DI) of 0.2 or more is acceptable (20). Indices of around 0 mean that both competent and incompetent candidates can equally answer a question. A negative DI indicates that more incompetent candidates answered correctly than competent ones. It should make you suspicious of a miskeyed question—that is, a question in which the scanner operator wrongly entered its answer key.

The difficulty index, sometimes called the facility index, proportion correct or *p*-value, is simply the proportion of candidates correctly answering a question, indicated by a percentage from 0 to 100 or a value from 0 to 1 (20). It helps you find the possible causes of a low-discrimination question, either being too easy (having a difficulty index of 80% [0.8] or more) or too difficult (20% [0.2] or less). The values between these two extremes are acceptable. Too-easy questions are easy content-wise or have very weak distractors. Too-difficult questions are attempted poorly, even by competent students; you should recheck the suitability of the question.

The third useful thing is the distractor analysis. This is the percentage of candidates choosing each option, allowing you to know the functionality of each distractor. Most item analysis software provides this information. Distractor analysis is helpful for future assessment and teaching (16). Rewrite or replace dysfunctional distractors, defined as those chosen by less than 5% of candidates (17).

Item analysis indices flag potentially problematic questions, but the final decision belongs to the content experts. Let us say that all candidates can answer a particular question, leading to a difficulty index of

1 (very easy), or that all candidates could not answer it, leading to a difficulty index of 0 (very difficult). If a content expert thinks that the question is fair, well-written and assesses a critical outcome, keep the question despite the unfavourable difficulty index.

Indices are also affected by the number of candidates; low numbers tend to inflate values. More than 200 candidates are needed for a stable index, down to at least about 100 candidates; however, item analysis can still provide you with helpful information, even if you have fewer than 30 candidates (20).

## CONCLUSION

Writing a good SBA question is not easy for the uninitiated; it requires an understanding of the content, creating clear LOs that assess HOTS, knowing how an SBA question works, translating an LO into an SBA, writing skills and writing and teaching experience. We hope this guide has made the fundamental steps of writing one more explicit, if not easier.

## Practice Points

- a. As an SBA-question writer, be clear about the LOs you assess. SBA questions are suitable for evaluating HOTS LOs, but not all HOTS outcomes are optimally assessed using SBAs. Understanding the verbs used to describe HOTS outcomes helps you choose appropriate LOs.
- b. Use authentic and novel scenarios in a stem to assess HOTS.
- c. Write a lead-in to direct a candidate to perform the desired LO verb.
- d. A sufficiently worded stem and lead-in allow a competent candidate to arrive at a correct answer without looking at the options list.

- e. An options list contains homogenous answer choices.
- f. Remember that a well-written SBA question allows competent students to recognise an answer key, while incompetent students are attracted to choosing a distractor. These qualities make a question discriminating, and increase the reliability of a test.
- g. Expend effort to write plausible distractors. In this regard, totally false distractors are allowed if they appear true to incompetent students. Plausible distractors that are partially true are desirable.
- h. Use item analysis to evaluate the quality of your SBA question.

- 5. Krathwohl DR. A revision of Bloom's Taxonomy: an overview. *Theory Pract.* 2002;4(41):212–8. [https://doi.org/10.1207/s15430421tip4104\\_2](https://doi.org/10.1207/s15430421tip4104_2)
- 6. Bissell AN, Lemons PP. A new method for assessing critical thinking in the classroom. *Bioscience.* 2006;56(1):66–72. [https://doi.org/10.1641/0006-3568\(2006\)056\[0066:ANMFAC\] 2.0.CO;2](https://doi.org/10.1641/0006-3568(2006)056[0066:ANMFAC] 2.0.CO;2)
- 7. Légaré F, Freitas A, Thompson-Leduc P, Borduas F, Luconi F, Boucher A, et al. The majority of accredited continuing professional development activities do not target clinical behavior change. *Acad Med.* 2015;90(2):197–202. <https://doi.org/10.1097/ACM.0000000000000543>
- 8. Blanco MA, Capello CF, Dorsch JL, Perry GJ, Zanetti ML. A survey study of evidence-based medicine training in US and Canadian medical schools. *J Med Libr Assoc.* 2014;102(3):160–8. <https://doi.org/10.3163/1536-5050.102.3.005>
- 9. Yeo S. An analysis of verbs used in the course outcomes of outcome-based integrated courses at a medical school based on the taxonomy of educational objectives. *Korean J Med Educ.* 2019;31(3):261–9. <https://doi.org/10.3946/kjme.2019.136>
- 10. Kalasuramath S, Tandon M, Deshpande D, Kumar V. Application of Bloom's taxonomy of verbs to evaluate the cognitive domain in undergraduate medical physiology question papers: a critique. *Int J Res Med Sci.* 2015;3(11):3351–6. <https://doi.org/10.18203/2320-6012.ijrms20151190>
- 11. Webb EM, Naeger DM, Fulton TB, Straus CM. Learning objectives in radiology education: why you need them and how to write them. *Acad Radiol.* 2013;20(3):358–63. <https://doi.org/10.1016/j.acra.2012.10.003>
- 12. Hughes Y. SMART tips for setting teaching objectives. *InnovAiT.* 2017;10(10):614–7. <https://doi.org/10.1177/1755738017720229>

## REFERENCES

- 1. Billings MS, DeRuchie K, Hussie K, Kulesher A, Merrell J, Morales A, et al. NBME item-writing guide – constructing written test questions for the health sciences. Philadelphia: National Board of Medical Examiners; 2020.
- 2. Paniagua M, Swygert KA, Downing SM. Written tests: writing high-quality constructed-response and selected-response items. In: Yudkowsky R, Park YS, Downing SM, editors. *Assessment in health professions education.* 2nd ed. New York: Routledge; 2020. p. 109–26.
- 3. Kilgour JM, Tayyaba S. An investigation into the optimal number of distractors in single-best answer exams. *Adv Heal Sci Educ.* 2016;21(3):571–85. <https://doi.org/10.1007/s10459-015-9652-7>
- 4. Shumway JM, Harden RM. AMEE Guide No. 25: the assessment of learning outcomes for the competent and reflective physician. *Med Teach.* 2003;25(6):569–84. <https://doi.org/10.1080/0142159032000151907>

13. Schuwirth LW, van der Vleuten CP. Different written assessment methods: what can be said about their strengths and weaknesses? *Med Educ.* 2004;38(9):974–9. <https://doi.org/10.1111/j.1365-2929.2004.01916.x>
14. Haladyna TM, Rodriguez MC. Guidelines for writing selected-response items. In: *Developing and validating test items.* New York: Routledge; 2013. p. 89–110.
15. Papenberg M, Musch J. Of small beauties and large beasts: the quality of distractors on multiple-choice tests is more important than their quantity. *Appl Meas Educ.* 2017;30(4):273–86. <https://doi.org/10.1080/08957347.2017.1353987>
16. Gierl MJ, Bulut O, Guo Q, Zhang X. Developing, analyzing, and using distractors for multiple-choice tests in education: a comprehensive review. *Review of Educational Research.* 2017;87:1082–116. <https://doi.org/10.3102/0034654317726529>
17. Haladyna TM, Downing SM. *Functional distractors: implications for test-item writing and test design.* New Orleans, LA: American Educational Research Association; 1988.
18. Collins J. Education techniques for lifelong learning: writing multiple-choice questions for continuing medical education activities and self-assessment modules. *Radiographics.* 2006;26(2):543–51. <https://doi.org/10.1148/rg.262055145>
19. Haladyna TM, Rodriguez MC. Selected-response formats. In: *Developing and validating test items.* New York: Routledge; 2013. p. 61–88.
20. Downing SM, Juul D, Park YS. Statistics of testing. In: Yudkowsky R, Park YS, Downing SM, editors. *Assessment in health professions education.* 2nd ed. New York: Routledge; 2020. p. 70–85.