

# **A guide to successfully writing MCQs**

## EXECUTIVE SUMMARY

### **Introduction**

Compared to other available methods of assessing professionals a web based examination embedded in multiple choice questions (MCQs) is cost effective, legally defensible and a functional basis for objectively documenting the possession of knowledge and clinical reasoning powers. For this method to be acceptable to professionals, the intellectual quality of the MCQs must be of the highest standard. This has the advantage of engendering credibility with both the clinicians sitting the examinations, as well as with National or International regulators.

### **The role of knowledge based assessment (KBA) in postgraduate training**

Regulators wish to foster high professional standards amongst clinical doctors. The characteristics of a safe and caring doctor include not only knowledge, which we recommend is assessed by MCQs, but also skills and professional attitudes. In an age of rapid electronic access to knowledge, the ability of the doctor to **use** knowledge in an appropriate way is as important as the possession of that knowledge. The role of the Knowledge Based Assessment (KBA), or examination, is by definition limited to an assessment of factual knowledge. A KBA may be deemed necessary to demonstrate competence but is unlikely to be sufficient to define a good Doctor.

### **The intellectual context for the use of MCQs**

The contents of MCQs must always relate back to a published curriculum. The MCQs should be seen as a mechanism for ensuring that the candidate possesses knowledge of the whole of the published Curriculum. An important responsibility of the exam setting group is to select from the question bank an appropriate numbers of questions reflecting the relative importance of the different subjects within the curriculum. Similarly question writers generating MCQs for the bank, should therefore be guided to generate questions based on the strategic needs of the bank.

### **The advantages of using MCQ based assessment of knowledge**

1. Marking of the exam can be undertaken electronically.
2. As the bank of MCQs is progressively built up, the costs of assessment of knowledge become predictable and contained.
3. The "correct" answers are predetermined, and do not involve the subjective judgements required in marking narrative, or essay type examinations.
4. Negative marking, where the candidate loses marks for incorrect answers, is not recommended because it may bias against certain personality traits, and does not test for the positive reasoning powers intended to be measured by the test.
5. The lack of dependence on paper-based questions permits substantial flexibility in both the location and the timing of the exam.
6. It is easy to electronically vary the order in which questions are presented to the candidates.
7. The examining body needs to own, and keep electronically secure, the questions and the correct answers.
8. The hardware used in the exams must also be secure to maintain exam security.

## Features of a "good" MCQ

1. A well constructed MCQ consists of a positively worded leading statement or “stem”, followed by a clearly expressed question. The stem will be derived from a specific item within the curriculum. The stem is followed by five possible answers consisting of one agreed correct answer and four wrong answers or “distractors”
2. The question should test concepts of understanding or data evaluation and should avoid simple tasks such as recall or pattern recognition.
3. The stem is positively worded and focuses on a single concept.
4. Within a European context, questions should not relate to specific national requirements.
5. A consistent degree of plausibility in each of the distractors is highly desirable. However, a question where 1 or 2 distractors are weaker than the others may not necessarily function as a “bad” question.
6. The correct answer and the distractors should be of roughly equal textual length.
7. The possible answers should be set out in alphabetical order.
8. There should be an evidence base for determining both which of the answers are correct and which are incorrect. This should be available to the question writer and the question writing group as well as to the candidates.
9. Avoid concepts such as "all of the above", "none of the above", “answers 2 and 3 only” etc.
10. Each question should be self-contained and not refer directly to another question. It should not be possible to deduce the answer of one question from the information presented in a previous or subsequent question.

## **A good question is determined by post hoc analysis of its performance.**

1. In a post-hoc analysis a good question is one which discriminates between good candidates and bad candidates.
2. Bad questions correlate negatively with overall performance. For instance poorly performing candidates answer the question correctly with the best performing candidates giving the wrong answer.

## **The "difficulty" of questions**

It is useful to have a range of difficulty of questions from easy through medium to hard. In assessing a newly submitted question for the bank, the question should not be dismissed simply on the grounds of being too easy or too difficult.

## **Managing the question bank**

In constructing a fair summative examination it is important that the whole breadth of the published curriculum is covered. In addition, different parts of the curriculum will be weighted in terms of importance.

It follows that the number of available questions in the Question Bank relating to each part of the curriculum should be proportional to the predetermined weighting.

## **Setting the pass mark**

Previously used questions, with known high performance, can be used as "marker" questions. The contents of questions should be reviewed in terms of being academically up-to-date. Provided that the bank has good security, the use of "old" questions is valuable in the context of the reproducibility of the exam process across multiple sittings. Different countries and different specialties may use different methodologies for determining the passmark. This is the task of the standard setting group.

### **Technical details**

1. A standard terminology or lexicon for clinical descriptions should be agreed and used and it should be sent in advance to question writers
2. Punctuation, spelling and the use of Capitals should be standardised by the Question Writing Group.
3. Avoid the use of absolutes such as “never”, “always”, “completely”.
4. Avoid the use of excess words which are unrelated to the objective of the candidate determining the correct answer.
5. A standard clinical format of history, clinical examination, results of investigations, diagnostic procedures, and management decisions should always be followed.
6. Abbreviations should be almost always avoided.
7. Distractors should all be structured in line with the stem and the correct answer.

### **Use of images**

1. The need for and use of images will vary between different specialties.
2. Modern information technology allows high definition images to be both stored and transmitted.
3. One good image may be used with different questions.

### **The nature of writing MCQs**

1. Writing MCQs is both a science and an art. Whilst the generation of new questions is essentially a task for the individual, the process of maximising the quality of MCQs is a task for a group of motivated and experienced MCQs experts. All new questions should be reviewed by a question writing group, before being accepted into the “Bank”
2. Face-to-face discussions are an essential component of developing new, high-quality MCQs. Cost considerations may require telephone conferences or video conferences to replace face-to-face meetings of writers.

3. There needs to be a time limit for discussion of each question with the conclusions of "accepted", "rejected", or "back to the author for reworking" agreed in a timely way.
4. The face-to-face format facilitates an international atmosphere and understanding and should aim to encourage individual contributors. Ideally both European and national institutions should recognise MCQ writing for CPD. The quality of questions produced by any individual writer tends to increase with time and experience.
5. Once a question has been accepted it needs to be categorised in terms of its length, its difficulty, and the section of the curriculum to which it refers.
6. All MCQ writers must have access to the current version of the relevant curriculum.
7. In the context of a European exam, writers from different countries, different cultures and different languages must be sourced.
8. Ideally, an electronic template should be employed for writing and then discussing new MCQs. This helps to facilitate uniformity of MCQs writing style, transmission of the questions to other members of the writing group and to facilitate discussion.

### **Setting an exam**

1. A review group should be set up with the specific task of selecting questions for each diet of the exam. This group needs previous experience in MCQ writing, and should be small.
2. The specific objective of the group is to select from the bank a range of questions appropriately distributed throughout the curriculum, with a range of perceived difficulty, and where available, marker questions known to perform well.
3. The use of "old" questions provides a longitudinal standard of the exam process.

4. The basic structure of any exam process should be transparent to national or international regulators invited to recognise the exam.

### **The Standard Setting Group:**

Once the diet of the examination is set the role of the standard setting group is to review each question for 3 reasons. Firstly, transcription errors may have occurred from the question bank and these are easily rectified prior to the examination. Secondly, Medical knowledge constantly evolves and some questions may have become outdated and should be either revised or removed from the examination and the bank. Thirdly, the standard setting group need to determine the overall level of difficulty of the particular diet of the examination and ensure longitudinal consistency between examinations. A commonly used method to do this is that each member of the standard setting group reviews each question and assess the likelihood of a candidate who will just pass the examination overall correctly answering each question (the Angoff score). These scores can then be compared between diets of the examination to ensure consistency and also to ensure that there is a spread of difficulty of questions within each examination diet. The relationship between the examination board and the standard setting group is shown in Figure 1.

### **The Exam Board:**

The examination board is to provide overall governance for the exam and its processes and also to determine the final passmark and thereby which candidates will pass or fail the examination. There are several well defined methods to define the passmark. It is suggested that the same method is used for each diet of the examination. The quality of the examination can be assessed by looking at the spread of marks achieved by all candidates during the examination. The quality and difficulty of each question can additionally be assessed. Questions that have performed badly can be defined by several

different methods including: those where 2 options have been selected by nearly all candidates, those where all options have been chosen equally and those where performance in that question negatively correlates with overall performance during the examination. It is perfectly acceptable to remove such questions from the final analysis in determining the passmark for the examination. In a good examination the number of such questions will be small and therefore the effect on the passmark and which candidates pass and fail the examination will be small.

**For a successful high-quality exam therefore there is a need for:**

1. An MCQs writing group
2. An exam setting group.
3. A standard-setting group who "set" each diet of the exam.
4. An Exam board to set the pass mark after the exam has been completed.