



# A GUIDE TO SUCCESSFULLY WRITING MCQS: EXECUTIVE SUMMARY

UEMS 2015/15

<b>Adopted by UEMS Council 10/04/2015</b>	<b>UEMS-CESMA Guideline</b>	<b>UEMS-CESMA</b>
<b>Authors:</b> Alfred Tenore, Danny G.P. Mathysen, Peter Mills, Mark Westwood, Jean-Baptiste Rouffet, Vassilios Papalois, Owen Sparrow, Zeev Goldik		



# TABLE OF CONTENTS

<b>Introduction .....</b>	<b>4</b>
The role of knowledge-based assessment (KBA) in postgraduate training .....	4
The intellectual context for the use of MCQs .....	5
<b>MCQ formats supported by UEMS-CESMA .....</b>	<b>6</b>
Type A MCQs .....	6
(Modified) Type K MCQs .....	6
Other types of MCQs .....	7
<b>The advantage of using MCQ-based assessment of knowledge .....</b>	<b>8</b>
<b>Points to consider when using MCQ-based assessment of knowledge .....</b>	<b>9</b>
<b>Features of a “good” MCQ .....</b>	<b>10</b>
Post-hoc analysis of performance .....	11
The “difficulty” of questions .....	11
Managing the question bank .....	12
Setting the pass mark .....	12
<b>Potential technical flaws .....</b>	<b>13</b>
<b>Use of images .....</b>	<b>14</b>
<b>The process of writing MCQs .....</b>	<b>15</b>
<b>Setting an assessment .....</b>	<b>17</b>
<b>The Standard Setting Group .....</b>	<b>18</b>



<b>The Assessment Board .....</b>	<b>19</b>
<b>Conclusion .....</b>	<b>20</b>
<b>Related documents .....</b>	<b>21</b>
<b>References .....</b>	<b>22</b>



---

# INTRODUCTION

Compared to other available methods of assessing professionals, an examination comprising multiple choice questions (MCQs) is cost-effective, legally defensible and a functional basis for objectively documenting the possession of knowledge and clinical reasoning ability. 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. MCQ assessments can be organized in a paper-based manner (preferably using an optical/electronic statistical analysis and marking system) or in a completely electronic fashion (either with a web-based or an offline system).

## **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 behaviours. 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 across the breadth 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 topics 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.



---

# MCQ FORMATS SUPPORTED BY UEMS-CESMA

Different types of MCQ formats are available [\[1\]](#). Although in theory, all of these MCQ formats may be used for the organisation of a European postgraduate medical assessment, UEMS-CESMA strongly supports the use of the following MCQ formats.

## **Type A MCQs**

Type A MCQs (also called “single-best answer” MCQs) are predominantly used in European postgraduate medical assessments. Although the chances of obtaining the correct answer by guessing is between 20 per cent (best of five MCQs) and 25 per cent (best of four MCQs), negative marking for type A MCQ’s is ill-advised, since this may bias against certain personality traits, and does not test for the positive reasoning powers intended to be measured by the test.

## **(Modified) Type K MCQs**

Although being abandoned by many assessments, type K MCQs (the so-called multiple independent true/false MCQs) still have their validity, e.g. in assessment formats offered in multiple languages. However, some guidelines on this type of MCQ need to be taken into account. The chances of obtaining the correct answer by guessing are as high as 50 per cent, the reason why negative marking for incorrect answers may be considered. Recent research on this matter has shown that in the case of modified type K MCQs (with a “don’t



---

know” option), the negative effects of negative marking do not apply, in contrast to the situation with the type A MCQs [2-4].

### **Other types of MCQs**

As described above, besides type A and (modified) type K MCQs, other MCQ formats may be utilised [1]. It is important that organisers of a European postgraduate medical assessment note that should they intend to use an alternative MCQ format, all advantages and disadvantages of the MCQ format concerned must be evaluated prior to utilisation.



---

# THE ADVANTAGES OF USING MCQ-BASED ASSESSMENT OF KNOWLEDGE

1. Marking of the assessment can be undertaken electronically, which will reduce measurement errors.
2. As over the years the bank of MCQs will be progressively built up, the costs of assessment of knowledge will become predictable and contained.
3. The “correct” answers of the MCQs are predetermined, and therefore do not involve the subjective judgements required in e.g. marking narrative or essay-type assessments.
4. The potential lack of dependence on paper-based questions permits substantial flexibility in both the location and the timing of the exam.
5. It is easy to vary the order in which questions are presented to the candidates electronically, as a security measure.





---

# **IMPORTANT POINTS TO CONSIDER WHEN USING MCQ-BASED ASSESSMENT OF KNOWLEDGE**

1. The examining body (i.e. the responsible organisers of the assessment, such as the relevant section or board) needs to own, and keep in an electronically secure location, the questions and the correct answers.
2. The hardware used in the assessments must also be secure in order to maintain assessment security.
3. The question bank must be kept secure and up-to-date, taking account of advances in knowledge and changes in medical practice.



---

## 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 answers and the distractors should be of roughly equal textual length.
7. The possible answers (or statements) 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 source of evidence base should be available to the question writer and the question-writing group as well as to the candidates (when preparing for the assessment).
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.

### **Post-hoc analysis of performance**

A good question is determined by post-hoc analysis of its performance (reliable post-hoc analysis results will only become available if the number of candidates participating at the assessment is sufficiently large).

1. In a post-hoc analysis a good question is considered to be one that discriminates between good and poor candidates.
2. Bad questions have a tendency to correlate negatively with the overall performance: e.g. poorly performing candidates answer the question correctly while the best performing candidates tend to give the incorrect 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, if it deals with an essential element of the curriculum.

### **Constructing the question bank**

For a fair summative assessment 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. Hence the structure of each exam should be linked to the curriculum with appropriate weighting (“blueprinted”).

It follows that the number of available questions in the Question Bank relating to each part of the curriculum should be proportional to this 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 specialties may use different methodologies for determining the pass mark. This is the task of the standard setting group. More details on pass mark setting are described the guidelines for organisation of a European postgraduate medical assessment [\[5\]](#).



---

# POTENTIAL TECHNICAL FLAWS

1. A standard terminology or lexicon for clinical descriptions should be agreed upon before being used and it should be sent in advance to question writers.
2. Punctuation, spelling and the use of capitals should be standardised by each examination board's Question Writing Group.
3. Avoid the use of absolutes such as "never", "always", "completely".
4. Avoid the use of excess words that are unnecessary for the candidate to determine 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 avoided unless standardised (e.g. "C" for Celsius temperature).
7. Distractors should all be structured in line with the stem and the correct answer, ensuring that all read well.



---

# 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 for several different questions in the same paper.



---

# THE PROCESS OF WRITING MCQS

1. Writing MCQs is both a science and an art. Whilst the generating 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 MCQ experts. All new questions should be reviewed by members of the question-writing group to ensure clarity, before being accepted into the “Question 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 continuous professional development (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 various countries speaking different languages must be identified, but all questions must be checked by a native speaker of the examination bank language to avoid misunderstanding or ambiguity. A similar process will be needed for all translations into other languages for individual papers.



- 
8. Ideally, an electronic template should be employed for writing and then discussing new MCQs. This helps to facilitate uniformity of MCQ writing style, transmission of the questions to other members of the writing group and to facilitate discussion.





---

# SETTING AN ASSESSMENT

1. An examination panel should be set up by the board with the specific task of selecting questions for each topic to be covered within the assessment. This group needs previous experience in MCQ writing, and should be small.
2. The specific objective of this 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 re-use of marker questions provides a longitudinal standard of the exam process.
4. The structure of any assessment process should be transparent to national or international regulators invited to recognise the exam, as well as to the candidates.



---

# THE STANDARD SETTING GROUP

Once the topics covered within the assessment are selected, the role of the standard setting group is to review each question for the following reasons. Firstly, transcription errors may have occurred from the Question Bank and these are easily rectified prior to the assessment. Secondly, medical knowledge constantly evolves and some questions may have become out-dated and should be either revised or removed from both the assessment and the Question Bank. Thirdly, the standard setting group needs to determine the overall level of difficulty of the particular diet of the examination and ensure longitudinal consistency between assessments. A commonly used method to do this is that each member of the standard setting group reviews each question and assesses the likelihood of a (borderline) candidate who is just competent correctly answering each question (i.e. the Angoff rating). These scores can be used to determine the pass mark, and then be compared between diets of the assessment to ensure consistency and also to ensure that there is a spread of difficulty of questions within each assessment diet.



---

# THE ASSESSMENT BOARD

The assessment board is to provide overall governance for the assessment and its processes and also to determine the final pass mark and thereby which candidates will pass or fail the assessment. There are several well-defined methods to define the pass mark. The same method should be used for each diet. The quality of the assessment can be evaluated by looking at the spread of marks achieved by all candidates during the assessment. The quality and difficulty of each question should also be evaluated. Questions that have performed badly can be identified by several different methods, including those where only two of the options have been selected by nearly all candidates (type A MCQs), those where all options have been chosen equally (type A MCQs), and those where performance in that question negatively correlates with overall performance during the assessment (type A and (modified) type K MCQs). Whilst it may be acceptable to remove such questions from the final analysis in determining the pass mark for the examination, inclusion and return for re-writing would be the usual approach. However, in a good assessment the number of such questions will be small and therefore the effect on the pass mark will be small. Of more importance is the identification of poor questions, where unanticipated ambiguity is identified from the pattern of candidate responses, permitting elimination from the exam.



---

# CONCLUSION

For a successful high-quality assessment the following are required:

1. An MCQs writing group;
2. An examination panel;
3. A standard-setting group that checks the paper(s) for each diet of the assessment;
4. An assessment board to set the pass mark after the examination has been completed.



---

## RELATED DOCUMENTS

- Mathysen D.G.P., Rouffet J.B., Tenore A., Papalois V., Sparrow O., Goldik Z. (2015).  
Guideline for the organisation of European postgraduate medical assessments.  
UEMS-CESMA Publication
- Papalois V., Goldik Z., Mathysen D.G.P., Rouffet J.B., Tenore A., Sparrow O. (2015).  
Guideline for the quality control on behalf of UEMS-CESMA of European postgraduate  
medical assessments. UEMS-CESMA Publication
- Goldik Z., Mathysen D.G.P., Rouffet J.B., Tenore A., Papalois V., Sparrow O. (2015).  
Guideline on examiner selection for European postgraduate medical assessments.  
UEMS-CESMA Publication
- Sparrow O., Mathysen D.G.P., Rouffet J.B., Tenore A., Papalois V., Goldik Z. (2015).  
Guideline on appeal procedures for European postgraduate medical assessments.  
UEMS-CESMA Publication
- Boorman J., Mathysen D.G.P., Noël J.L., Bloch K., Rouffet J.B., Tenore A., Papalois V.,  
Sparrow O., Goldik Z. (2014). Survey on European postgraduate medical assessments by  
the Council for European Medical Specialty Assessments (UEMS-CESMA). *MedEd  
Publish* 3 (42), 1-4
- Mathysen D.G.P., Goldik Z. (2015). On the quality control and importance of European  
postgraduate medical assessments. *Medical Teacher*



## REFERENCES

1. Case S.M., Swanson D.B. (2001). Constructing written test questions for the basic and clinical sciences. National Board of Medical Examiners. Philadelphia (3rd revised edition): [http://www.nbme.org/PDF/ItemWriting\\_2003/2003IWGwhole.pdf](http://www.nbme.org/PDF/ItemWriting_2003/2003IWGwhole.pdf)
2. Mathysen D.G.P., Aclimandos W., Roelant E., Wouters K., Creuzot-Garcher C., Ringens P.J., Hawlina M., Tassignon M.J. (2013). History and future of the European Board of Ophthalmology Diploma examination. *Acta Ophthalmol.* 91 (6), 589-593
3. Mathysen D.G.P., Aclimandos W., Roelant E., Wouters K., Creuzot-Garcher C., Ringens P.J., Hawlina M., Tassignon M.J. (2013). Evaluation of adding item-response theory analysis for evaluation of the European Board of Ophthalmology Diploma examination. *Acta Ophthalmol.* 91 (7), 573-577
4. Mathysen D.G.P. (2013). Statistical evaluation of multiple-choice assessments: Development and validation of a statistical analysis method to monitor the European ophthalmology exit examination, and evaluation of examination questions and methods. Universiteit Antwerpen (ISBN 978-90-572-8414-4)
5. Mathysen D.G.P., Rouffet J.B., Tenore A., Papalois V., Sparrow O., Goldik Z. (2015). Guideline for the organization of a European postgraduate medical assessment. UEMS-CESMA publication