



HOW TO

## How to write short cases for assessing problem-solving skills

L.W.T. SCHUWIRTH<sup>1</sup>, D.E. BLACKMORE<sup>2</sup>, E. MOM<sup>3</sup>, F. VAN DEN WILDENBERG<sup>4</sup>,  
H.E.J.H. STOFFERS<sup>3</sup>, C.P.M. VAN DER VLEUTEN<sup>1</sup>

<sup>1</sup>Department of Educational Research and Development, University of Maastricht, PO Box 616, 6200 MD Maastricht, The Netherlands; <sup>2</sup>Medical Council of Canada; <sup>3</sup>Department of General Practice, University of Maastricht, PO Box 616, 6200 MD Maastricht; <sup>4</sup>Department of Surgery, Academic Hospital, P. Debyelaan 25, 6229 HX Maastricht

**SUMMARY** *In assessment of problem solving the use of short case-based testing is a promising development. In this approach an examination consists of large numbers of short cases each of which contain a small number of questions. These questions are aimed at essential decisions. Writing such cases, however, is not easy. In this article a description of this type of examination is provided. Also strategies and pitfalls are described in writing these cases. These strategies pertain to the selection of essential decisions, the careful writing of cases and questions and the selection of question formats.*

### Introduction

Although in the domain of testing of problem-solving skills long patient case simulations have been favoured for quite some time, there has been a tendency since the mid-1980s towards use of short cases (Bordage, 1987; Graaf, 1988; Van der Vleuten *et al.*, 1994; Schuwirth *et al.*, 1996). The reason for this is that the assumptions on the nature of problem-solving ability underlying the long simulations have proved to be incorrect. The first of these assumptions considered problem-solving ability to be generic, implying that considerable transfer of the problem-solving process would occur from one case to another (Barrows, 1984). The second considered the problem-solving process to be uniform across experts, i.e. that different experts would solve the same problem the same way. The contrary assumptions, however, are more appropriate; problem solving is highly domain-specific (Elstein *et al.*, 1978; Swanson, 1987; Swanson *et al.*, 1987). The score an examinee obtains on one problem is often a poor predictor of the score on any other given case, even within the same domain or on the same topic. Regarding the second assumption it appeared that experts in most cases had difficulty in reaching consensus on the optimal strategy to solve a certain problem, although they did agree on what the correct solution should be (Swanson *et al.*, 1987). The process of problem solving is apparently highly idiosyncratic, even when all experts agree on its outcome.

These findings have had implications for the design of test instruments for the assessment of problem-solving ability. Instead of using a small number of long simulations focusing on the process of problem solving, tests should use

a rather large number of short cases. Each case includes only a small number of problems and focuses on the outcome of the problem-solving process.

In the construction of these short case-based tests, however, it is of paramount importance to take considerable care in the description of the case, the selection of the problems to be asked and a clear and unambiguous description of the questions or decision points to be assessed.

Both the Medical Council of Canada and the University of Maastricht in The Netherlands have extensive experience with this format. The Medical Council has been using it for its national licensing examination since 1992. At the University of Maastricht such an approach has been used for end-of-clerkship examinations since 1993.

This article outlines the most important developmental strategies and pitfalls in writing this type of test material. First, however, a general framework of a case with a question is presented to serve as a general model. Then more specific strategies and pitfalls are described.

### A general framework of cases

A case-based question consists of two (rather obvious) parts: the case and the question(s). Within the case a situation is described or information is presented which the examinee will have to use to solve the problem. In some cases the problem might already be presented within the case text.

#### *Example case 1*

You are a general practitioner. You make a house call on Mrs Van Doorn (65 years old). She consults you because of an acute pain in her abdomen. She tells you further that . . . On physical examination you find . . .

*You wonder whether she should be admitted to casualty immediately or whether it is best to keep her home for the time being and treat her there unless her condition worsens.*

What is the most appropriate action?

- (a) You have her admitted to casualty.
- (b) You keep her at home for the time being.

*Correspondence:* Lambert Schuwirth, Dept of Educational Research and Development, University of Maastricht, PO Box 616, 6200 MD Maastricht, The Netherlands

In other cases it might be easier to present the problem in the form of the following question:

You are a general practitioner. You make a house call on Mrs Van Doorn (65 years old). She consults you because of an acute pain in her abdomen, She tells you further that . . . On physical examination you find . . .

*What is the most probable diagnosis?*

Thus the information in the case is of vital importance. It should contain sufficient information to enable the examinee to make the decisions that are required in the questions, but at the same time should not contain too much irrelevant information to divert the student. Different types of information are included within the case. A patient case, for example, contains both clinical and contextual information. The clinical information reports aspects such as signs, symptoms and findings, but also negative or normal signs, symptoms and findings (i.e. non-abnormal signs). The contextual information reports on aspects that are not the result of the patient's illness but which do have an influence on the decision an examinee has to take. This information consists of aspects such as physical surroundings, gender, age, family circumstances, non-verbal information.

Authenticity of a case is vital and highly dependent on an adequate report of the information. Thus, verbal information is presented 'raw' without any 'chunking' ('respiratory rate of 30/minute' instead of 'tachypnea'); relevant visual and auditory information should be presented directly (e.g. by using multimedia), instead of a textual description.

The above problems are examples oriented to assessing essential decisions. Diagnosis and treatment, however, may not always be the essential problem of a case. Other aspects such as processing information (in history taking or physical examination), (efficient) diagnostic management or informing the patient may be more essential to the case.

Authenticity does not only pertain to the case description but also to the questions asked. Options or questions that are out of context of the case are to be avoided whenever possible. The case writer must be aware of the realistic options that exist in the real-life situation. Some examples may clarify this point. When faced with a severe accident producing two casualties with different traumas and only one ambulance, the decision on which patient to send in first has only two options, but is far from trivial. Deciding which lab test should be ordered to confirm possible hyperthyroidism has only a limited number of options in real life. The recognition of a pattern of symptoms as a certain disease or syndrome often has an 'infinite' number of options in reality. These considerations should be taken into account when determining the format for each question.

### Strategies and pitfalls

This section discusses the strategies and pitfalls in three parts: the case description, the selection of problems and the questions.

#### *Case description*

(1) *Use the representation of real patients.* Three arguments support this strategy. First, real practice provides a rich

source for possible cases and prevents prolonged deliberation on choosing a suitable subject for a case. Second, by randomly selecting patients to form the basis for cases, the test or item bank will cover daily practice more congruently than constructed cases. It will prevent the author from hobby-horse riding or producing only 'exotic cases'. Third, and most important, constructed cases (out of a textbook) that do not relate to genuine patients are often very artificial. Be aware that it is always advisable to preserve the anonymity of your patients when constructing examination material from real case files.

(2) *Ensure that the description of the information is as clear as possible.* In order for the student to process the information the case should be unambiguous and written clearly. It should deal with the exact features that are present in real life. Phrases such as 'possible masses on palpation' constitute non-information, because a mass could be either palpated or not. This doubt may exist in real life, but then the physician will at least have the memory of his/her sensations during the examination. An example of such a flawed case is presented below.

#### *Example case 2*

Mrs Whiteless consults you in your practice because of complaints of a vague pain in the lower abdomen. The pain is "somewhere in my abdomen, doctor". It has been present for quite a while now. The pain has a cramping aspect.

During further history she tells you that her periods are regular, but that her last menstruation was one week later than normal.

On physical examination you find a dubious pressure tenderness in the lower abdomen and you feel a possible mass there. During gynaecological examination a drop of blood is seen on the portio.

What is the most likely diagnosis?

The many flaws in this case may be obvious: the age of the patient is not stated, the exact location of the pain is not given and "a dubious pressure tenderness" is vague, etc. Using more words or sentences to make the case clearer may cost a little extra reading time for the examinee, but will probably decrease total testing time, because the student should be in no doubt about the meaning of the case. An example of a better case presentation might be as follows.

#### *Example case 3*

You are a general practitioner. Mrs Whiteless (28 years old) consults you in your practice because of complaints of pain in the right part of the lower abdomen. It has been present for about 8 days now. In the beginning she barely felt it, but it has increased in the last four days. The pain comes in cramps, occurring about 5–10 times per day, each attack lasting for about 15–30 minutes. During these cramps she feels the urge to lie down and keep as calm as possible.

She tells you further that her period is regular, but that her last menstruation was one week later than normal. On inspection of the abdomen you see no abnormalities except for an appendectomy scar. The peristalsis is normal. On palpation there is a pressure tenderness in the lower right abdomen and you feel a mass barely touching your fingers so you cannot

determine its size. Percussion is slightly painful, but reveals no abnormal sounds. During gynaecological examination you see a drop of blood on the portio, bimanual palpation is painful at the right side (left side is normal). Further gynaecological examination yields no abnormal findings.

What is the most likely diagnosis?

(3) *Provide sufficient realistic clinical information.* Not only should all information that is needed be present in the case but its presentation should be explicit: 'normal routine examination' is often inadequate as a description; the actions should usually be mentioned separately. As mentioned earlier, the disadvantage of the extra reading time is outweighed by the time that would be used by the candidate puzzling over what is meant by the description.

(4) *Provide sufficient realistic contextual information.* To solve a patient case, contextual information is very relevant. Information about previous consultations or the frequency of consultations can be relevant in order for the examinee to come to a judgement of probabilities on a certain diagnosis. This contextual information may also serve to present non-verbal behaviour, such as 'the way the patient presents his complaint gives you the impression that he exaggerates'. It is advisable to report some contextual information as a standard procedure (e.g. gender, age, frequency of previous visits, profession).

(5) *Provide sufficient negative information.* All information that is necessary to answer the question must be in the case. So not only the type of action taken (e.g. liver palpation) and the results (1–2 cm below lower costa, sharp edge), but also the realistic actions not taken or actions that did not show abnormalities should be presented (e.g. 'There is no rebound tenderness'). It is dangerous to assume that every candidate will automatically know whether an action not reported is either not taken or did not yield any abnormal findings.

(6) *Provide information that is not pre-interpreted ('raw').* For efficiency reasons often the information is presented in a pre-interpreted style. This means that some of the interpretation was already done by the case author. In real practice, however, interpretation of findings into larger meaningful 'chunks' is part of the problem-solving process. So instead of reporting a pulse deficit or liver enlargement, it is better to report the pulse rate and the heart rate or the size of the liver to let the examinee draw his/her own conclusions, i.e. to interpret the data for him/herself.

#### *The problem*

(7) *Link the problems directly to the case.* The case should not be presented as an illustration followed by a question that asks for general knowledge. It is essential that case and questions form an inseparable unit. A minimum criterion can therefore be that a candidate should not be able to answer the question without having read the case. A question with that particular flaw is given below.

#### *Example case 4*

Mrs Smith suffers from attacks of heavy cramping pain in the upper abdomen. When such an attack occurs,

she has the urgent need to stand up and walk around. The attacks occur 5–6 times per day and have been present for 4 days now. An attack lasts for about 4 minutes. She has noticed that most of the attacks occurred after she had eaten meals with fatty food. She has taken her temperature, which was 37.3 °C. On physical examination you notice a slight pressure tenderness in the upper right quadrant of the abdomen. Inspection, auscultation and percussion of the abdomen reveal no abnormal findings.

Which of the alternatives is closest to the specificity of an ultrasound scan for detecting gallstones?

- (a) 20%;
- (b) 40%;
- (c) 60%;
- (d) 80%.

The next example appears to solve the problem.

#### *Example case 5*

Mrs Smith suffers from attacks of heavy cramping pain in the upper abdomen. When such an attack occurs, she has the urgent need to stand up and walk around. The attacks occur 5–6 times per day and have been present for 4 days now. An attack lasts for about 4 minutes. She has noticed that most of the attacks occurred after she had eaten meals with fatty food. She has taken her temperature, which was 37.3 °C. On physical examination you notice a slight pressure tenderness in the upper right quadrant of the abdomen. Inspection, auscultation and percussion of the abdomen reveal no abnormal findings.

In this case the positive predictive value of an ultrasound scan for detecting gallstones lies closest to:

- (a) 20%;
- (b) 40%;
- (c) 60%;
- (d) 80%.

Although the question cannot be answered without the case, it is still not aimed at decisions that are essential in practice. In the determination of whether or not to perform an ultrasound scan of the gallbladder, not only the positive predictive value but also the negative predictive value and the costs have to be included in the decision. Therefore, the most realistic way of asking this question is simply to ask whether or not an ultrasound scan of the gallbladder is indicated in this particular case (considering the signs, symptoms, probability of the diagnosis, positive predictive value, negative predictive value, costs, invasiveness, etc). •

(8) *Avoid problems or possibilities that are not present in real practice.* Assessment is of course always an abstraction of real life. This, however, does not necessarily mean that real life is more difficult than an examination. Often possibilities exist in real life that are not present in a test. So, for example, do not focus on aspects that can easily be looked up in real life (such as normal lab values or dosages of drugs). If critical issues of a case are asked, quick referencing does not reveal the answer. In contrast, if the answer can be easily and quickly found in a reference book, the question probably does not relate to a key decision. So allowing the students to use reference books (within the time constraints of the examination) can help you to focus your cases on

essentials. Another approach would be to include reference material such as normal laboratory values within the question or case scenario.

Distracting information may be present in real life. It is advisable to incorporate this in the case. It must be done, though, in such a manner that the student can extract some clues as to whether this information is relevant or not: for example, by describing the tone in which a patient says something or describing other non-verbal communication. While this is somewhat artificial in that the examinee cannot actually hear the tone of the voice of the patient, the case author needs to work within the limitations of the medium of the case and question format, which is often written text. Be cautious when introducing distracting information. This information should not be so confusing as to create an ambiguous case or question.

(9) *Focus on essential problems only.* Most of the problems and decisions involved in solving a case are quite straightforward and follow automatically from others. When for example the decision to admit a patient as an emergency case has been taken, the subsequent problem of whether or not to call for an ambulance is quite trivial. Some decisions, however, are essential for the case, though the distinction between essential and non-essential problems can be difficult.

Some prerequisites for a problem to be essential are:

- The problem must be based on combining the different information parts of the case.

#### Example case 6

Mr Johnson (67 years old) sees you in your practice because he has had a sharp pain in his right shoulder for three weeks. At first he thought that the pain would go away spontaneously, but since it has not, he wants you to take a look. The pain is not related to any movements of the shoulder. It radiates to his back, mainly in the scapular region. Further history taking reveals that he has never been seriously ill. He has been a heavy smoker for almost 45 years. On physical examination you notice a myosis of the right pupil and a ptosis of the eyelid.

What is this combination of the last two symptoms called?

This case provides an example where this prerequisite is not met, because the essential part is not to recognize the Horner's syndrome but to recognize that this is a serious disorder or even that this is probably a Pancoast tumour and should be dealt with accordingly. If only a small part of the information of the case is needed to solve the problem, it is probably not an essential one.

- An incorrect decision must lead to incorrect management of the case.

#### Example case 7

You make a house call on Mrs Van Doorn (65 years old). She suddenly felt a fierce piercing pain in her abdomen about three hours ago. Any movement she makes increases the pain, so she keeps as still as possible. Abdominal examination is barely possible. On auscultation you hear no bowel sounds. Palpation is barely possible, not only because of the tenderness, but also because her abdominal muscles are contracted and cannot be relaxed. Percussion is not possible. She has a temperature of 39.5 °C. She has

had abdominal pains before with periods of fever, but these were different in nature (less sharp) and subsided after a couple of days.

What is the most probable diagnosis?

- diverticulitis;
- appendicitis;
- perforated peptic ulcer.

The exact diagnosis is not essential for effective management in this case. Whether this is diagnosis (a), (b) or (c) does not influence the management in this case. In each of the situations the patient would have to be admitted to casualty as quickly as possible.

- Do colleagues agree with your selection?

It is advisable to present the case to colleagues and ask them to comment on your selection of problems. Different experts often share the same opinion on what essential problems are. In the case of dissent it is recommended to rethink your selection process (Bordage *et al.*, 1995).

(10) *Limit the number of problems to be asked.* If a key feature of the case has been addressed by a question, it should not be pursued any further. It is more efficient to go on to the next (essential/key) problem or even to go on to the next case. The more different cases that can be presented and the more different questions that can be asked (addressing key elements in the problem-solving process), the higher the generalizability will be.

#### The question

(11) *Phrase the questions as clearly as possible.* The rules that apply to normal test questions naturally apply to questions in case-based testing. They should be checked for common flaws. Flaws can work two ways: either they provide cues that might lead an incompetent examinee to produce the correct answer (i.e. lead to false positive results) or they consist of bad phrasing leading the competent candidate astray (i.e. lead to false negative results). Examples of the former are: obvious differences in length of the options in a multiple choice, cueing words such as 'can', 'always', 'never', in multiple-choice or true-false questions, unrestricted open-ended questions (which lead the candidate to write as much as he/she can). Examples of the latter are the use of ambiguous terms (e.g. 'sometimes', 'often') or insufficient indication of the level of detail in an open-ended question. A more detailed description of these flaws goes beyond the scope of this article, but good examples are available in the literature (e.g. [www.nbme.org](http://www.nbme.org)).

(12) *Focus the question on the specific aspects of the problem.* The problem in a case must be defined and presented as clearly as possible. It must be made clear which of the aspects must be considered by the candidates when making a decision.

An example of a question where this is *not* the case is presented below.

#### Example case 8

John Provis (4 years old) has had a fever since yesterday (in the evening 38.9°C). His parents are worried, especially because he also has a severe cough. The general practitioner examines the child and concludes that no

serious illness is present, but that it is a common cold. She advises the parents to give John an Aminocetophin® in the evening to lower the fever and to keep him inside for 3 days.

Is this correct advice?

- (a) Yes
- (b) No

The question cannot be answered because it is not clear what the aim of the question is. One could very well argue that the advice is incorrect since the GP should have informed the parents about the innocence of the disorder, but on the other hand the advice given is not incorrect. The problem can be avoided for example by including the considerations in the case, as is illustrated below.

*Example case 9*

John Provis (4 years old) has had a fever since yesterday (in the evening 38.9°C). His parents are worried, especially because he also has a severe cough. The general practitioner examines the child and concludes that it is nothing but a common cold. She asks herself whether the advice to give an Aminocetophin® in the evening and the advice to keep John inside for 2–3 days would be correct.

The advice concerning the Aminocetophin® is correct:  
true/false

The advice to keep John inside for 2–3 days is correct:  
true/false

Breaking up the different elements in the question will often lessen ambiguity of combined elements.

(13) *Ensure that the answer is defensibly correct, distractors defensibly false.* This suggestion may look obvious but this is not always the case. Often questions are phrased in such a way that some of the boundaries between correct and incorrect answers are not made explicit. The next example may clarify this.

*Example case 10*

Mrs White (45 years old) sees you in your GP practice. She tells you that she has pain in her abdomen. The pain has been going on for one week and it seems to be getting worse.

What questions do you ask?

It may be clear that any answer is defensible. Every decision in practice is a trade-off in which the advantages, probabilities and disadvantages have to be weighed against each other. In this case, giving some more information about the complaint, and then focusing the question to parts of the normal history taking (e.g. family history, digestive tract, urogenital tract) and asking for the selection that is most likely to yield relevant information might resolve the problem.

(14) *Let the content of the question determine the format.* It seems tempting always to use a certain question format and adapt all questions to that particular format, since this is most straightforward in the psychometric sense. However, in terms of authenticity this would not be optimal. Decisions in real life often involve the selection between a limited number of alternatives (such as the decision either to say “I do” at a wedding or not) and sometimes a nearly limitless number of options. To present questions with the same number of realistic options as in real practice may therefore be optimal in preparing the student for his/her future task.

A ‘one format fits all’ approach can be adopted for logistical efficiency or psychometric reasons, but cannot be based on the intrinsic superiority of a particular question format. Even open-ended questions can be lacking in authenticity. A defendant, for example, who is instructed by his lawyer in an open-ended way, will be ill-prepared when faced with ‘yes’ or ‘no’ questions in court. Similarly, a physician who is only used to one type of question for all problems may be sub-optimally prepared for future decisions. To determine what question type to use, the case author should try to determine the number of realistic alternatives that exist in real practice. If this number is very large, an open-ended question appears to be most appropriate, or an extended-menu format where the student is allowed to select several alternatives. If the number of realistic alternatives is limited, a multiple-choice question is preferable.

In the next case, the number of realistic options may be rather large, since the essence appears to be the recognition of the pattern.

*Example case 11*

Mr Brown consults you about his knee. It is swollen and red and it hurts. He has had this for 2 days now and it has gradually become worse. He has not had this before. When you ask him about other complaints, he tells you that he has had a red spot on his left upper leg, about 5–10 cm in diameter. This spot then disappeared and reappeared on the other leg. Two weeks before the spots appeared he spent his holidays in the woods hiking. He had caught a common cold there, which lasted for only 3 days.

What is the most appropriate diagnosis?

In the following, however, the essence of the case is not necessarily to generate a differential diagnosis, but to discriminate between the probabilities of the individual diagnosis.

*Example case 12*

Mr Thomas visits you in your practice because he had chest pain yesterday. He was mowing the lawn, which he had not done for a long time. It was hard work cutting the grass. Suddenly he felt a stabbing crunching pain in his chest on the left side. The pain made it impossible to continue working, so he sat down. After 5–10 minutes the pain decreased, but he had the feeling that the pain was not totally gone until an hour after its onset. A very faint pain remained during that hour. He has never had this before and he is very worried. On cardiac and pulmonary examination (inspection, palpation, percussion and auscultation) you find no abnormalities. Pulse is 80/min (regular) and the blood pressure is 150/80 mmHg.

Which of the following is the most probable diagnosis?

- (a) myocardial infarction
- (b) unstable angina
- (c) stable angina
- (d) pulmonary embolism
- (e) pneumothorax
- (f) hyperventilation
- (g) intercostal neuralgia

(15) *Have your material reviewed by others.* Writing test material is not easy and the quality of the material can easily be negatively influenced by ‘blind spots’ of the author. In the field of scientific research and practical medicine these blind spots are already widely acknowledged and have led to quality control procedures such as cross-checking and review. A

- (1) Use the representation of real patients
- (2) Ensure that the description of the information is as clear as possible
- (3) Provide sufficient realistic clinical and contextual information
- (4) Provide sufficient negative information
- (5) Provide information that is not pre-interpreted ('raw')
- (6) Link the problems directly to the case
- (7) Avoid problems or possibilities that do not exist in real practice.
- (8) Focus on essential problems only
- (9) Limit the number of problems to be asked
- (10) Phrase the questions as clearly as possible
- (11) Focus the question on the specific problem
- (12) Ensure that the answer is defensibly correct, distractors defensibly false
- (13) Let the content of the question determine the format
- (14) Have your material reviewed by others

Figure 1

similar approach should be adopted in the production of high-quality test material: careful review increases the quality of the material.

In Figure 1 a summary of the strategies and tips with respect to case-based items is given.

### Conclusion

Constructing short cases for examinations is not easy. In our experience an average time of 2–3 hours per case can be considered normal. In addition, it has proven to be ill-advised to make up cases without consulting others. Nobody is immune to mistakes, blind spots, etc., (including the authors of this article). It is widely accepted that it is good practice to show manuscripts to colleagues before sending them to a journal and also to confer with colleagues when in doubt about a certain clinical strategy. Since case writing is a difficult task, it is wise to show your test material to others and ask them for comments and criticism.

The nature of the cases and the selection of the essential problems varies of course with the educational context of the test. Tests in undergraduate courses will lead to the selection of different key features from tests in postgraduate education or even in continuing medical education. This will be determined by the expected prior knowledge and experience of the candidates and the specific course goals (e.g. basic sciences versus clinical sciences).

The strategies and pitfalls of this article do not apply to clinical or patient cases only. Short descriptions of physiological or anatomical problems can be used also. Furthermore, clinical cases can be used to ask for basic science problems (Des Marchais *et al.*, 1993; Jean *et al.*, under editorial review; Schuwirth *et al.*, 1993).

A final piece of advice would be the suggestion to the reader to look for possibilities for cooperation with other departments or faculties. Since the production of high-quality test material can be tedious and expensive, cooperation can often lead to a win-win situation. In any case, the use of the short-case approach to measuring problem-solving ability appears to be both viable and desirable.

### Notes on contributors

L.W.T. SCHUWIRTH is an MD working in the area of test development and research for the medical faculty in Maastricht.

DAVID E. BLACKMORE is a medical educator who is the Director of the Evaluation Bureau of the Medical Council of Canada.

E. MOM is a general practitioner, participating in the project on the development of assessment in clerkships.

F. VAN DEN WILDENBERG is a trauma surgeon and coordinator of the clerkship in surgery and the third pre-clinical year at the University of Maastricht.

H.E.J.H. STOFFERS is a general practitioner, participating in the project on the development of assessment in clerkships.

C.P.M. VAN DER VLEUTEN, is a professor in educational psychology and chairman of the department of educational research and development; he is also project leader of the project on assessment and evaluation of the medical faculty of Maastricht University.

### References

- BARROWS, H.S. (1984) A specific, problem-based, self-directed learning method designed to teach medical problem-solving skills, and enhance knowledge retention and recall, in: H.G. SCHMIDT & M.L. DE VOLDER (Eds) *Tutorials in Problem-based Learning*, pp. 16–32 (Assen, Van Gorcum).
- BORDAGE, G. (1987) An alternative approach to PMP's: the 'key-features' concept, in: I. R. HART & R. HARDEN (Eds) *Further Developments in Assessing Clinical Competence, Proceedings of the Second Ottawa Conference*, pp. 59–75 (Montreal, Can-Heal Publications).
- BORDAGE, G., BRAILOVSKY, C., CARRETIER, H. & PAGE, G. (1995) Content validation of key features on a national examination of clinical decision-making skills, *Academic Medicine*, 70, pp. 276–281.
- DES MARCHAIS, J.E., JEAN, P. & NU VIET VU, C. (1993) An attempt at measuring student ability to analyze problems in the Sherbrooke problem-based curriculum: a preliminary study, in: P.A.J. BOUHUIJS, H.G. SCHMIDT & H.J.M. VAN BERKEL (Eds) *Problem-Based Learning as an Educational Strategy*, Vol. 1, pp. 239–248 (Maastricht, Network of Community-Oriented Educational Institutions for Health Services).
- ELSTEIN, A.S., SHULMANN, L.S. & SPRAFKA, S. A. (1978) *Medical Problem-solving: An Analysis of Clinical Reasoning* (Cambridge, MA, Harvard University Press).
- GRAAF, E.D. (1988) Simulation of Initial Medical Problem-solving: a test for the assessment of medical problem-solving, *Medical Teacher*, 10, pp. 49–55.

- JEAN, P., SCHUWIRTH, L.W.T., VAN SANTEN, M. & VAN DER VLEUTEN, C.P.M. (under editorial review). Do problem analysis questions (PAQs) and true/false questions (TFQs) measure different skills?
- SCHUWIRTH, L.W.T., JEAN, P., VAN DER VLEUTEN, C.P.M. & VAN SANTEN, M. (1993). Problem-Analysis Questions, een korte casusvorm voor het preklinische domein [Problem Analysis Questions, a short case format for the preclinical years], in: E. HOUTKOOP, J. POLS, M.C. POLLEMANS, A.J.J.A. SCHERPBIER, & G.M. VERWIJNEN (Eds) *Proceedings of the Derde Gezond Onderwijs Congres*, pp. 104-111 (Veldhoven, Haagse Hogeschool, 's Gravenhage, Netherlands).
- SCHUWIRTH, L.W.T., VAN DER VLEUTEN, C.P.M., DE KOCK, C.A., PEPERKAMP, A.G.W. & DONKERS, H.H.L.M. (1996) Computerized case-based testing: a modern method to assess clinical decision making, *Medical Teacher*, 18, pp. 295-300.
- SWANSON, D.B. (1987) A measurement framework for performance-based tests, in: I. HART & R. HARDEN (Eds) *Further Developments in Assessing Clinical Competence*, pp. 13-45 (Montreal, Can-Heal Publications).
- SWANSON, D.B., NORCINI, J.J. & GROSSO, L.J. (1987) Assessment of clinical competence: written and computer-based simulations, *Assessment and Evaluation in Higher Education*, 12, pp. 220-246.
- VAN DER VLEUTEN, C.P.M., NEWBLE, D.I., CASE, S.M., HOLSGROVE, G., MCCANN, B., MCGRAE, C. & SAUNDERS, N. (1994) Methods of assessment in certification, in: D.I. NEWBLE, B. JOLLY & R. WAKEFORD (Eds) *The Certification and Recertification of Doctors, Issues in the Assessment of Clinical Competence*, pp. 105-125 (Cambridge, Cambridge University Press).