Original research paper

Development and utility of a questionnaire to evaluate the quality of PBL problems

Fadi M. Munshi[1], El Sayed A. El Zayat[2], Diana H. Dolmans[3]

Abstract

***Objective**:* High quality problems are crucial for successful learning in Problem Based Learning (PBL). Many criteria have been described in literature about the quality of PBL problems. The aim of this study was to develop a questionnaire to evaluate the quality of PBL problems and to test its utility.

***Method**:* The questionnaire was developed based on literature review. Both students and tutors used the questionnaire to evaluate PBL problems. It was applied at King Fahad Medical City, Faculty of Medicine in the first block of three consecutive years in the academic year 2008. A total of 12 problems were evaluated by students and the tutors.

***Results**:* A total of 12 tutors and 36 students applied the questionnaire to evaluate 12 PBL problems. At the overall level, the questionnaire gives an impression on the strong and weak aspects of all PBL problems evaluated. The strongest aspect identified by both students and tutors is that the PBL problems rated in this study have a realistic context. The weakest aspect is that the problems do not adapt to students level of prior knowledge. At a more specific level the instrument can be used to identify strong and weak problems and give suggestions for improvement.

***Conclusions**:* It can be concluded that the instrument that has been developed to evaluate the quality of PBL problems provides useful information about strong and weak aspects of PBL problems.

Key words: Problem-based learning

Introduction

It is assumed that learning is an active process of constructing knowledge, rather than a passive process of memorization. Instructional methods should stimulate students by activating relevant prior knowledge related to the knowledge to be attained. They should provide a learning setting that resembles to a large extent the setting in which the knowledge is to be applied in the future. In addition, the learning environment should provide learners with an opportunity to elaborate on their own knowledge. Problem based learning (PBL) is an instructional method that is said to provide the above conditions (Schmidt, 1983).

In PBL, the learning activity starts with a problem, which is a description of a set of phenomena or events in need of explanation in terms of an underlying process, mechanism or principle. Students analyze these problems in small groups. The task of the group of students is to explain the phenomena or events provided in the given problem (Schmidt & Moust, 1998).

The rationale underlining PBL from cognitive psychology is outlined by Schmidt et al in three principles. It depends on the activation of prior knowledge as a key element in the nature and amount of new information that can be processed. This emphasizes attention to adapt the degree of complexity in a PBL problem to students` prior knowledge. Finally, the discussion and elaboration on prior knowledge and on new knowledge creates links between previous and new concepts leading to better retrieval of information from memory.

[1]MD, MSc, Medical Educationalist, Department of Medical Education, Faculty of Medicine, King Fahad Medical City, Riyadh, Saudi Arabia.

[2]MD, Professor of general and vascular surgery and Head of the Clinical Department, Faculty of Medicine, King Fahad Medical City, Riyadh, Saudi Arabia.

[3]PhD, Associate Professor and Educational Psychologist, Department of Educational Development and Research, University of Maastricht, Maastricht, the Netherlands.

Correspondence: F. Munshi, MD, King Fahad Medical City, Faculty of Medicine, Department of Medical Education, P.O. Box 59046, Riyadh 11525, Saudi Arabia; Email: falmunshi@kfmc.med.sa

In a model characterizing the main factors affecting the process of learning in PBL, Schmidt and Gijselaers (1990) showed after extensive research that the amount of prior knowledge, the quality of problems, and the tutor performance were all important features of a problem based curriculum. The most dominant factor that affected the PBL process is the high quality of the problems presented to the students. The key element that attributes to the principles of PBL and to the process of learning is the quality of the problem. Here a question is raised what are the criteria of a high quality problem?

Several papers have been written in which criteria for effective PBL problems are described. Majoor et al (1990) suggested four criteria to be applied in construction of problems. The problem should match the students' level of prior knowledge and motivate students for studying. It should show a clear linkage to the future profession. It must be suitable for analysis and open enough for discussion. Finally, it has to be directive to the block objectives. In another study, seven principles of effective case design were described by Dolmans et al (1997). These principles were based on findings on the nature of learning and cognition. Schmidt et al (1986), developed a questionnaire based on group dynamics literature, to seek the students perspective of the important aspects of problems. The features that were attributes of the problem were goal clarity, openness and concreteness. Marchais (1999) identified nine criteria and ranked them using a series of judges as experts in relation to the weighted importance in designing and evaluating the quality of a PBL problem in a medical curriculum. The criteria defined the important elements in construction of a PBL problem in descending order; openness, autonomy, richness, attractiveness, coverage, arousing curiosity, relevance, comprehensiveness and appropriate vocabulary. A literature review aimed to identify strategies for constructing cases was conducted by Kim et al (2006). The authors concluded five main attributes to case design. The attributes are the case being relevant, realistic, engaging, challenging and instructional. In a study that provided guidelines on generating trigger images for PBL, Azer (2007) emphasized the importance of choosing suitable triggers that are highly authentic; they should be innovative, creative and engaging and chosen appropriately to guide to certain educational objectives.

From the papers cited above, it can be concluded that most authors highly agree with each other on which aspects determine the quality of PBL problems. The aspects that are emphasized in the literature by various authors are: the problem should cover the preset learning objectives and should lead to learning issues that match with the teachers' objectives / fit with block objectives. The problem should provide guidance and stimulating cues that lead to thinking, analysis and reasoning. It should stimulate self directed learning and fit with students' prior knowledge. The problem should show clear links with the future profession and enhance interest in subject matter. These criteria are summarized in Table 1.

Although criteria have been described for effective PBL problems in the literature, the studies cited above do not provide a tool for medical teachers that can be used in educational practice to evaluate the quality of PBL problems. The aim of this study is to develop a questionnaire based on what is known from the literature to measure the effectiveness of a PBL problem.

Methods

The study was conducted in the undergraduate medical curriculum at King Fahad Medical City (KFMC) in the academic year 2008. It was not been submitted to an ethical approval committee within the Faculty of Medicine at King Fahad Medical Study, Saudi Arabia, because such a committee is not yet available. Nevertheless, the data was analyzed anonymously.

The curriculum at KFMC is a six year PBL curriculum in which the first year is a premedical year. A year contains 8-10 blocks. Each block consists of 4-6 problems, with one problem analyzed per week by the Maastricht seven jumps approach. The questionnaire was administered to the first PBL blocks in year 2, 3 and 4. This lead to analysis of diverse problems from three consecutive years with a total of 12 problems.

Each block has 3-5 tutorial groups with different tutors in each group and an average of 8-10 students. The tutor (N = 12) and three random students (N = 36) from each group were asked to fill out a questionnaire to rate each problem. The students who participated in the study volunteered to fill out the questionnaires throughout the block for each problem. The tutors and students were given the option to keep their identities anonymous by not writing their names. The tutors' response rate was 95%, while the students' response rate was 82%.

The questionnaire developed consisted of six factors derived from the literature review on criteria of effective PBL problems, as summarized in the introduction section in Table 1. It covered the six main factors extracted from the literature that describe an effective PBL problem as a problem that: stimulates thinking, enhances analysis and reasoning (3 items), stimulates self directed learning (3 items), leads to studying the intended contents (3 items), enhances interest in subject matter (3 items), is relevant to the

future profession with realistic context (3 items) and matches the level of prior knowledge (3 items). A total of 18 items were formulated. One item was also formulated for the overall rating of the PBL problem.

All items were rated on a five-point Likert scale after the problems were discussed by both the students and the teachers/tutors. Participants were asked to indicate the degree of relevance of the item for each problem (1='strongly disagree' to 5='strongly agree'). Furthermore, four open ended questions were included. The questions gave an opportunity for the participants to describe the strengths and weaknesses they perceive of the problems. These questions also can direct towards areas for improvement. Further comments were welcome. The questionnaire used for this study is included in appendix 1.

The instrument as used for this study was initially tested in a pilot study. Within this pilot-study, the utility of the instrument and the relevance of the items were verified by expert tutors in PBL. They were asked to rate the relevance of each item on a five-point Likert scale (1 = 'strongly disagree' to 5 = 'strongly agree'). They were also asked to give feedback and suggestions for improving the questionnaire and to mention any missing items that could be included. During the pilot-study, the questionnaire was distributed to eight tutors in Maastricht University and 4 tutors in King Fahad Medical City. The reply rate was 100% with a total of 12 ratings received from expert tutors on the principles of PBL. The feedback received from the tutors was analyzed and the common modifications done were rephrasing 6 items and deleting 3 items, resulting in the 18 items described above.

The data was analyzed at two levels. At the overall level, descriptive statistics per item were computed for students and tutors for all the problems together. At a more specific level, descriptive statistics per problem were computed for both students and tutors. Independent groups T-test was used to test whether the opinions of the tutors and the students differed or not. In addition, one weak scoring problem was analyzed more in depth. The standard to interpret each item was 3 or less is insufficient, more than 3 to 3.5 means that improvement is needed, more than 3.5 to 4 is sufficient and more than 4 is good (scale 1-5). While the overall item is on a scale of 1-5, 1 was scored as insufficient and 5 was scored as excellent.

Table 1: An overview of the criteria that are mentioned in the literature on effective PBL problems.

Criteria of Effective Problems \ Authors	Schmidt et al (1986)	Majoor et al (1990)	Dolmans et al (1997)	Marchais (1999)	Kim et al (2006)	Azer (2007)
Stimulates thinking, analysis, & reasoning	X	X	X	X	X	X
Stimulates Self Directed Learning	X		X	X	X	
Leads to studying the intended contents	X	X	X	X	X	X
Enhances Interest in Subject Matter		X	X	X	X	X
Relevance to Future Profession with realistic context	X	X	X	X	X	X
Matches Level of Prior Knowledge		X	X	X	X	

Results

At the overall level, the mean scores and standard deviation were computed for each item for all 12 KFMC PBL problems. The results are shown in Figure 1. The items rated by the tutors ranged from 3.19 (*SD* = 0.84, *N* = 48) to 4.15 (*SD* = 0.71, *N* = 48). The highest scoring item was the problem guiding to the block objectives (7). Two of the highest rated items dealt with the problems being related to the future profession (13, 14). The weak aspects dealt with the problems corresponding to the students` level of prior knowledge (16, 17). The items rated by the students ranged from 3.35 (*SD* = 0.87, *N* = 132) to 3.95 (*SD* = 0.90, *N* = 131). The highest scoring items were that problems had a realistic context and stimulated searching literature (5, 7). Of the lowest scoring items, two items scored low, being the problem adapts to the level of the students` prior knowledge and the problem is in alignment with the curricular material previously covered (16, 17). The third lowest scoring item was the item of the problem providing optimal directions for discussion (2). At the overall level, tutors and students highly agree with each other on weak and strong aspects of the PBL problems analyzed within this study. They both mentioned as a strong feature that the problems fit with the block objectives (7) and as weak features that the problems do not fit with students' level of prior knowledge or that the problems are not in alignment with the curricular materials covered previously (16,17). No statistically significant differences were found when comparing tutors and students scores on the specific items (1-18) except for item 8, $t(176) = 2.23$, $p < .05$.

At a more specific level, the overall quality of each problem was calculated based on the average score on item 19. The results are summarized in Figure 2. Tutors rated the overall quality of the 12 problems in a range of 2.75 (*SD* = 1.26, *N* = 4) to 4.25 (*SD* = 0.50, *N* = 4). Students rated the overall quality of the 12 problems in a range of 3.40 (*SD* = 0.84, *N* = 10) to 4.36 (*SD* = 0.63, *N* = 14). Figure 2 demonstrates from the tutors' opinions that one

Figure 1: The mean score on items 1 to 18 for all 12 problems together, on a scale from 1 to 5, 1 being very irrelevant and 5 being very relevant. The scores of the students and the tutors are given per item.

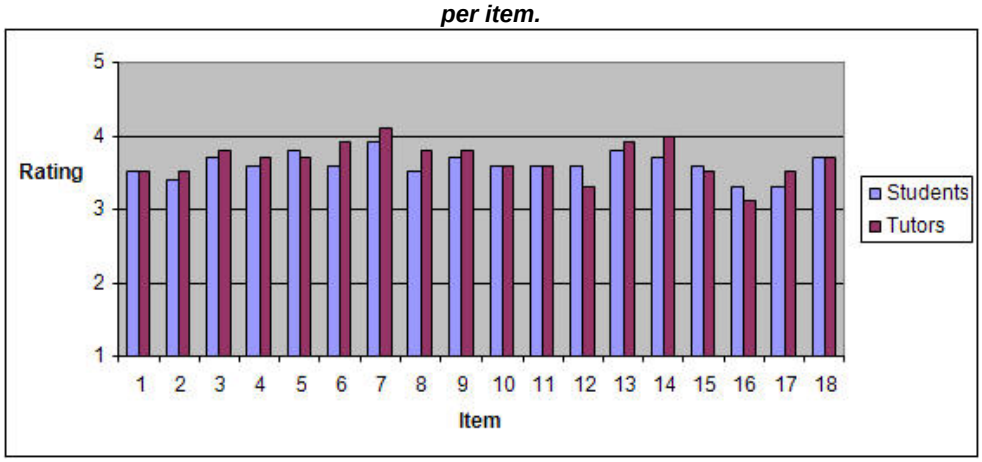
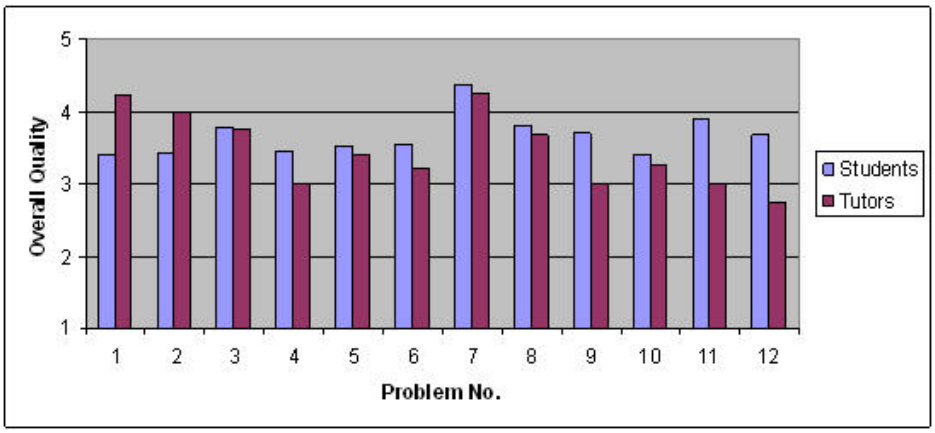


Figure 2: The mean score for the 12 PBL problems on the overall item (evaluate the overall quality of a PBL problem, item 19), on a scale from 1 to 5, 1 being insufficient and 5 being excellent. Both the student and the tutor scores are given.



problem was rated as insufficient (a mean score below 3.0) being problem 12, that six problems needed improvement (a score below 3.5), and that five problems were rated as sufficient (a score above 3.5). Figure 2 also demonstrates the students' opinions. No problem was rated as insufficient by the students (a mean score below 3.0), six problems needed improvement (a score below 3.5), and eight problems were rated as sufficient from the students' and tutors' opinion (a score above 3.5). No statistical significant differences were found when comparing the tutor and student overall rating (item 19) for problem 1 to 12 separately, except for problem 11, $t(9) = -2.28$, $p < .05$ and the overall rating of all the 12 problems together $t(173) = -2.01$, $p < .05$, although the tutor scores in general were somewhat lower than the student scores. Both tutors and students agreed upon the highest rated problem, being problem 7. Students and tutors do not agree with each other on the problem that scores lowest as can be seen in Figure 2. Tutors rated problem 12 as lowest, whereas students rated problem 10 as lowest. Only one problem scored below 3, being problem 12, based on the tutors' rating, (M = 2.75, SD = 1.26, N = 4) but tutors differ of opinion as the standard deviation is quite high. The lowest rated problem by students is problem 10 (M = 3.40, SD = 0.84, N = 10) (Table 4, Figure 2). Although tutors did not rate this problem as lowest, the average score given by the tutors for problem 10 is also quite low (M = 3.25, SD = 0.95, N = 4).

In order to test whether the instrument developed provides suggestions for improvement of low scoring PBL problems, one low scoring problem was analyzed more in-depth, being problem 10 that was rated low by both students and tutors. Students agree with the tutors on the strength of problem 10. Two items were rated less than 3.0 by the tutors for this problem as can be seen in Figure 3. The items are as follows: The problem encourages consulting literature linked to the block objectives (item 9) and the problem is related to a patient not to a disease only (item 15). No item was rated below 3.0 by the students, but the item that was rated low (below 3.5) by both students and tutors is: the problem adapts to the level of the students` prior knowledge (item 16). In the comments on the open-ended question five students stated "the problem is not to our level of prior knowledge". The problem is given as an example in Appendix 2.

Conclusion and discussion

An instrument to evaluate the effectiveness of PBL problems was developed in this study based on a review of the literature on effective PBL problems. Students and tutors applied the questionnaire on twelve problems in an undergraduate medical curriculum in order to test the utility of this instrument.

The instrument provides information on PBL problems on two levels. On the overall level, i.e. when being applied to various problems in a curriculum/block, it can give an idea on the overall quality of the problems and strong and weak aspects of the problems in the curriculum/block. At a more specific level it can be used to detect strong and weak problems and give suggestions on improvement of a single problem.

At the overall level, it can be concluded that students and tutors highly agreed with each other. They agreed that the problems were formulated to guide to one or more of the general objectives. Thus, the problems were designed with strong consideration towards coverage of the learning objectives in the students` and tutors` opinions. Overall no items were rated as insufficient (a mean score below 3.0). One aspect at the overall level needs attention and improvement that both students and tutors agreed upon (a mean score below 3.5). The problems are not well adapted to the level of the students` prior knowledge (16). This weakness has a major impact on the PBL process, because a PBL problem should fit with the level of students' prior knowledge to ensure that students can link the new acquired knowledge to their existing knowledge structures (Schmidt, 1983). This finding is probably caused by the fact that no curriculum map or overall curriculum outline is available in which it is described how the different blocks within the curriculum are linked together. A curriculum map can assist block coordinators to better link the contents covered in the different courses within the curriculum with each other, due to which the teachers obtain a better overview of the level of prior knowledge of the students within a particular unit (Harden, 2001).

On a more specific level, the instrument gives an idea of strong and weak problems. As demonstrated in Figure 3 and 4, a more in-depth analysis of the scores on the separate items for a single low scoring problem, revealed suggestions for improvement of the problem. This was demonstrated for problem 10, which needed improvement in the opinion of both the tutors and students. Tutors identified two main aspects that were insufficient for this problem 10. The first is that the problem does not encourage students to consult literature linked to the block objectives (9). The second is that the problem is related to a disease only and not to a patient (15). The third aspect that needs improvement in both the tutors and students opinion is that this problem 10 does not adapt to the students' level of prior knowledge (16). This example demonstrated that the instrument developed provides feedback for problem designers on the effectiveness of their PBL problems and as such can be seen as an instrument that seems to be useful for practice. Although the tutors and the students agreed with each other about the aspects that needed

improvement, the results also demonstrate that the tutors' and the students' opinions provided supplementary information on some aspects.

Nevertheless, there are some limitations of this study that need to be mentioned. First, only a limited number of problems were analyzed in this study and a limited number of tutors rated the quality of the problems. Further research is needed to test the instruments validity for example by means of an exploratory or confirmatory factor analysis, when more data are collected with this instrument; i.e. when more PBL problems are rated by means of this questionnaire. Second, in each block the problems were developed by the same coordinating team not including the tutors involved in this study. This implies that the problems in a given block have the same structure and backbone. Choosing various problems from different blocks can give a better impression on the overall quality of the PBL problems in a curriculum. Third, so far no data have been collected on whether the feedback is perceived as useful by the block coordinators and designers of the PBL problems. These data should be collected in the future to further test the usefulness of this instrument in practice. Tracking the changes done to the problems and further evaluating the problems by this instrument can give insight in the question whether the evaluation and changes carried out, based on the data collected with this instrument, result in improvement of PBL problems in the teachers' and students' perceptions. Finally, although the instrument needs to be statistically validated and further research is needed, the instrument has been based on a review of the literature on what is known about criteria for effective PBL problems, which can be seen as a strength of this instrument and which implies that the content of the instrument has been validated from the literature.

Figure 3: The mean score on item 1 to 18 for the lowest rated PBL problem (problem 10). The scores of the tutors are given on a scale from 1 to 5, 1 being very irrelevant and 5 being very relevant.

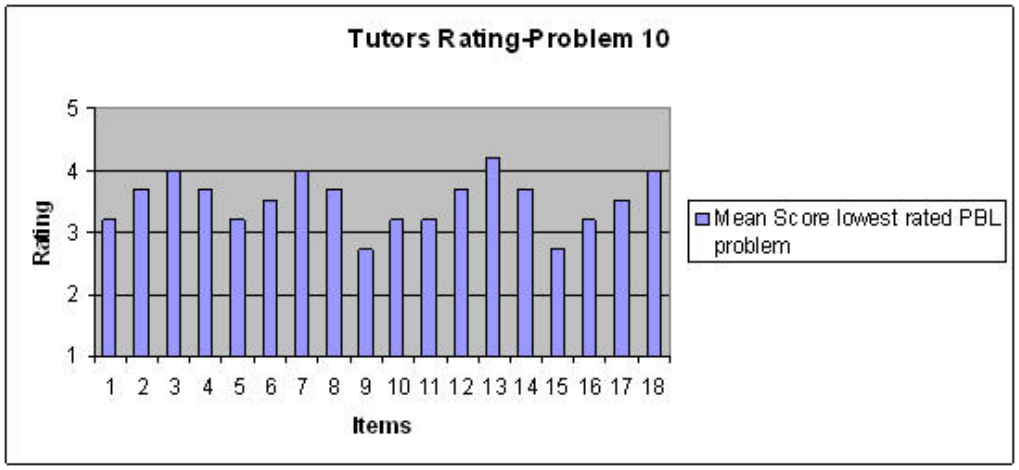
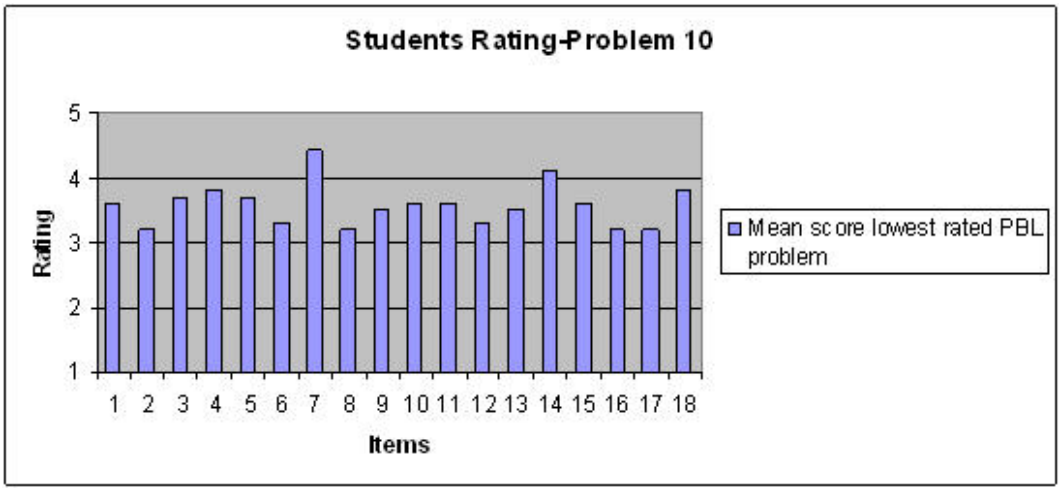


Figure 4: The mean score on item 1 to 18 for the lowest rated PBL problem (problem 10). The scores of the students are given on a scale from 1 to 5, 1 being very irrelevant and 5 being very relevant.



References

Azer, S.A. (2007) Twelve tips for creating trigger images for problem-based learning cases, *Medical Teacher*, 29 (2), pp. 93-97.

Dolmans, D.H.J.M., Snellen-Balendong, H., Wolfhagen, I.H.A.P. & Van der Vleuten, C.P.M. (1997) Seven principles of effective case design for a problem based curriculum, *Medical Teacher*, 19 (3), pp. 185-189.

Harden, R.M. (2001) AMEE guide no. 21: curriculum mapping: a tool for transparent and authentic teaching and learning, *Medical Teacher,* 23, pp. 123-137.

Kim, S., Phillips, W. R., Pinsky, L., Brock, D., Phillips, K., & Keary, J. (2006) A conceptual framework for developing teaching cases: a review and synthesis of the literature across disciplines, *Medical Education*, 40 (9), pp. 867–876.

Majoor, G.D., Schmidt, H.G., Snellen-Balendong, H., Moust, J.C.H. & Stalenhoef-Halling, B. (1990) Construction of Problems for Problem Based Learning, In: Z Nooman, H.G. Schmidt & E.S. Ezzat (eds) *Innovation in Medical Education*, New York: Springer, 114-122.

Marchais, J.E. Des. (1999) A Delphi technique to identify and evaluate criteria for construction of PBL problems, *Medical Education*, 33 (7), pp. 504–508.

Schmidt, H.G. (1983) Problem-based learning: Rationale and description, *Medical Education*, 17, pp. 11-16.

Schmidt, H.G. & Gijselaers, W.H. (1990) Causal modeling of problem-based learning, Paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA.

Schmidt, H.G., Gijselaers, W.H., Nuy, H. & De Grave, W.S. (1986) Dimensions of problems for problem based learning: a factor analytic study, Unpublished manuscript. Maastricht, the Netherlands: University of Limburg.

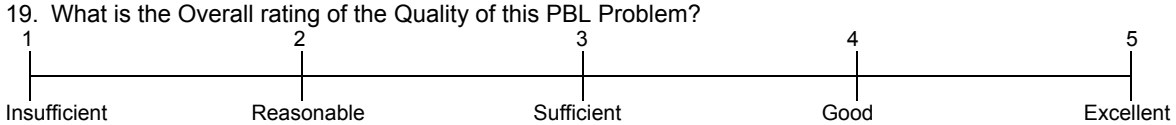
Schmidt, H.G. & Moust, J.H.C. (1998) Processes that Shape Small-Group Tutorial Learning: A Review of Research. Paper presented at the Annual Meeting of the American Educational Research Association.

Appendix 1. Questionnaire to evaluate the quality of PBL problems

	QUESTIONNAIRE						
	Complete one Questionnaire per Problem after conducting the reporting phase. Rate each item on a scale of 1-5 for relevance by marking the box in the appropriate column	Strongly Disagree (1) ▼	Disagree (2) ▼	Neutral (3) ▼	Agree (4) ▼	Strongly Agree (5) ▼
Factor 1: Stimulates Thinking, Analysis and Reasoning						
1	The problem is open enough to sustain discussion					
2	The problem provides optimal directions for discussion (i.e. not too many or too few)					
3	The problem contains appropriate stimulating cues					
Factor 2: Stimulates Self-Directed Learning						
4	The problem stimulates students to formulate their various learning issues					
5	The problem stimulates students to search for relevant literature					
6	The problem stimulates students towards an effective discussion					
Factor 3: Leads to Studying the Intended Contents						
7	The problem is formulated to guide to one or more of the general block objectives					
8	The problem encourages integration of various disciplines					
9	The problem encourages to consult literature linked to block objectives					
Factor 4: Enhances Interest in Subject Matter						
10	The problem is formulated in such a way that it enhances students` interest in the subject matter					
11	The problem is phrased to students` perception of their own environment and culture					
12	The scenario in the problem appears appealing to students					
Factor 5: Relevance to the Future Profession with Realistic Context						
13	The problem shows clear linkage to the future profession					
14	Basic science concepts are presented in a context of a clinical problem					
15	The problem is related to a patient not to a disease only					
Factor 6: Matches the Level of Prior Knowledge						
16	The problem adapts to the level of the students` prior knowledge					
17	The problem is in alignment with the curricular material previously covered					
18	Students are familiar with part of the knowledge necessary for discussing the problem					

Kindly answer the following questions.

19. What is the Overall rating of the Quality of this PBL Problem?



20. What are the *strengths* of this problem?
21. What are the *weaknesses* of this problem?
22. What are the tips for improvement of this problem?
23. Any other comments?

Appendix 2. The lowest overall rated problem by students, Problem 2 in Year 4

Soaad is a Saudi female of 27-year-old with a three-month history of intermittent heat intolerance, sweats, tremors, and severe muscle weakness which has limited her ability to climb stairs. Her appetite has increased remarkably despite weight loss. She is also bothered by the bounding of her heart and some minor difficulty in swallowing. There is a positive family history of thyroid disease, but she denies taking any thyroid medications or having had any radiation to her thyroid gland. Her other medical problems include arthritis which is treated with aspirin 5.4 g/day and DM which is controlled with diet.

Acknowledgment

We would like to thank the faculty members, tutors and other teachers involved, and students from KFMC that participated in this study. The appreciation is further extended to the faculty members in the Department of Educational Development and Research at Maastricht University for participating in the pilot study.