Does problem-based learning lead to deficiencies in basic science knowledge? An empirical case on anatomy

Katinka J A H Prince,[1] *Henk van Mameren,*[2] *Nelien Hylkema,*[1] *Jan Drukker,*[2] *Albert J J A Scherpbier*[3] *& Cees P M van der Vleuten*[4]

Introduction Problem-based learning (PBL) is supposed to enhance the integration of basic and clinical sciences. In a non-integrative curriculum, these disciplines are generally taught in separate courses. Problem-based learning students perceive deficiencies in their knowledge of basic sciences, particularly in important areas such as anatomy. Outcome studies on PBL show controversial results, sometimes indicating that medical students at PBL schools have less knowledge of basic sciences than do their colleagues at more traditional medical schools. We aimed to identify differences between PBL and non-PBL students in perceived and actual levels of knowledge of anatomy.

Methods Samples of Year 4 students in all eight medical schools in the Netherlands completed a questionnaire on perceived knowledge and took part in a computerised anatomy test consisting of both clinically contextualised items and items without context.

Results Problem-based learning students were found to have the same perceived level of anatomy knowledge as students at other medical schools. Differences in actual levels of knowledge were found between schools. No significant effects on knowledge levels were found for PBL schools versus non-PBL schools.

Conclusion The results of this study show that PBL does not result in a lower level of anatomy knowledge than more traditional educational approaches. It remains to be ascertained whether the levels students attain are adequate. Subjects for further study are the desired level of anatomy knowledge at the end of undergraduate medical education and the effectiveness of basic science learning within a clinical context and with repetition over the course of the curriculum.

Keywords education, medical, undergraduate/*standards; problem-based learning/*standards; anatomy/*education; science/*education; curriculum; evaluation studies; questionnaires; Netherlands.

Medical Education 2003;37:15–21

Introduction

Problem-based learning (PBL) is supposed to enhance the integration of students' knowledge.[1] Students use (clinical) problems as a starting point of the learning process and define their own learning objectives in tutorial groups. These learning objectives reflect basic science disciplines as well as clinical disciplines and both fields are studied concurrently. In technical terms, PBL allows for the horizontal and vertical integration of different disciplines. Because learning takes place in a meaningful and authentic context, using clinical cases for example, students learn to connect clinical phenomena to underlying basic science concepts. Evidence from cognitive psychology has shown that integration of knowledge facilitates the storage and later retrieval of relevant information.[2] Integrated knowledge should prepare students better for actual clinical practice.[2]

However, a series of recent studies from the University of Maastricht in the Netherlands, which uses a PBL approach, has given rise to concern. In a qualitative study students reported that they felt deficient in basic science knowledge, particularly in anatomy, when entering clerkships.[3] Students claim that the integrated approach to learning, or, rather, the lack of dedicated courses, causes uncertainty so that students perceive

[1]Skillslab, University of Maastricht, Netherlands
[2]Department of Anatomy/Embryology, University of Maastricht, Netherlands
[3]Institute for Medical Education, University of Maastricht, Netherlands
[4]Department of Educational Development and Research, University of Maastricht, Netherlands

Correspondence: K J A H Prince MD, Skillslab, University of Maastricht, PO Box 616, 6200 MD Maastricht, Netherlands. Tel.: 00 31 4 3388 1978; Fax: 00 31 4 3361 8612; E-mail: k.prince@sk.unimaas.nl

Key learning points

Students at PBL schools experience similar deficiencies in anatomy knowledge to their colleagues at other schools.

When asked to express their perceived anatomy knowledge at the end of Year 4 as a percentage of the overall anatomy knowledge to be mastered by the end of curriculum, the majority of students estimate their knowledge to be less than 50%. This gives cause for concern about knowledge levels attained by the end of the 6-year curriculum.

Problem-based learning students obtain similar scores in an anatomy test to non-PBL students.

themselves as being inadequately prepared for clinical practice. Two earlier survey studies showed similar results: in retrospect Maastricht graduates indicated they had not acquired sufficient knowledge of gross anatomy, and students on clinical rotations expressed a need for more anatomy training both before and during clerkships.[4,5] It is unknown whether students at (more) traditional medical schools experience similar problems. Students in general consider anatomy a very important subject for their future as practising doctors.[6] A study involving graduates of McMaster University, Ontario, Canada, showed that graduates considered insufficient attention was paid to basic sciences in the McMaster PBL curriculum.[7] A number of studies have compared the basic science knowledge of PBL students and non-PBL students. When differences in basic science knowledge are found, these are usually in favour of non-PBL students.[8–11] Generally, the integrated PBL approach seems to be associated with uncertainty and perceived deficiencies in terms of basic science knowledge.

It should be noted that PBL is used to describe many heterogeneous educational activities.[12] It is therefore hard to prove or disprove the claims made by its advocates.[13] In order to objectify the deficiencies perceived by PBL students at Maastricht, we investigated whether PBL and non-PBL students in the Netherlands differ in perceived and real knowledge of anatomy at the start of clinical clerkships.

On the basis of the studies mentioned earlier, clinical anatomy was chosen as the discipline for study. Clinical anatomy is defined as anatomy that is indispensable to a good understanding of the medical physical examination, modern imaging techniques, diagnosis and many invasive and non-invasive procedures. It provides the language of professional communication between physicians. In line with the findings of some of the earlier studies, we expected non-PBL students to show superior knowledge of clinical anatomy. However, given the integrated nature of anatomy learning in PBL, we expected these differences to vanish or even alter in favour of PBL students when knowledge of clinical anatomy was assessed in the context of a clinical problem.

The following research questions were formulated:

1 How do PBL and non-PBL students compare in terms of their perceived levels of knowledge of anatomy?

2 How do PBL and non-PBL students compare in terms of performance in an anatomy test consisting of non-contextual fact-oriented items and clinically contextualised items?

A study was conducted in which a questionnaire and computerised anatomy tests were administered to samples of Year 4 medical students at the eight Dutch medical schools.

Methods

Subjects

The study population consisted of groups of Year 4 students from all eight medical schools in the Netherlands. The study was initiated by the Maastricht Medical School, which uses a PBL approach. The other schools agreed to participate on condition that the results relating to each school would be reported anonymously. Therefore, the schools are referred to by the letters A to H.

Undergraduate medical education in the Netherlands lasts 6 years and consists of 4 years of primarily theoretical education and 2 years of clerkships. Currently, most medical school curricula in the Netherlands are undergoing extensive revision. At Maastricht (school A), the PBL approach has been used in the first 4 years of the medical curriculum since the early 1970s. When we conducted this study, the curricula at schools B to E were not fully problem-based; i.e. there was some degree of horizontal and vertical integration of basic and clinical sciences and problem orientation. Schools F, G and H used more traditional, lecture-based curricula with hardly any vertical integration.

In the Netherlands, students enter medical school directly from high school after they have passed the national final high school examination at the level required for entrance to medical school. Because there are more applicants for medical school than places, a grade weighted lottery system, carried out by a national agency, determines admission. The same agency subsequently allocates students to the different medical schools. Although it has not been strictly proven so far, it is assumed that there are no systematic differences between the students entering the different medical schools. Because we wanted to investigate the effects of different medical school programmes on students' knowledge, we included only students who had spent all of their first 4 years in the same medical school. Moreover, only students who had not started clerkships were included.

All Year 4 students at the eight Dutch medical schools ($n = 1884$) were invited to participate. Due to the limited availability of computers for the test, we restricted the number of participants to approximately 50 students per school, allowing sufficient power for statistical inferences. Students were selected on a first-come, first-served basis and received a small financial incentive. They were given feedback in that they were sent their personal score and the overall national scores in the test.

A total of 424 students participated. Data from seven students were lost due to computer malfunction and six students were excluded because they had already started their clerkships or had attended different schools. Data for 411 students were analysed (22% of all Dutch Year 4 medical students).

It is often the better students who volunteer to participate in a study. To establish whether the participating students were representative of their respective classes, we compared their scores on regular in-training examinations with those of their non-participating fellow students, if available.

Questionnaire

All participating students were asked to complete a short questionnaire before taking the anatomy test. This questionnaire contained questions on demographic characteristics such as age, gender and whether or not the student had assisted in anatomy teaching as an instructor. At several schools some students assist in anatomy practicals as student-instructors. We expected these students to perform better in the test due to higher prior knowledge levels, special interest in anatomy, extra training or a combination of these factors. To answer our first research question, students were asked the following two questions about their perceived level of knowledge of anatomy:

1 If the total amount of anatomy knowledge that you could have acquired at this point in the medical curriculum is 100%, how much of that knowledge do you think you have actually mastered?

2 If the total amount of anatomy knowledge you must have mastered upon graduation is 100%, how much of that knowledge do you think you have mastered at this point in the curriculum?

Students could choose to answer each question in one of four categories, namely, 0–25%, 25–50%, 50–75% and 75–100%.

Anatomy test

In order to test anatomical knowledge using items with and without clinical context, two subtests were constructed. Firstly, more than 50 clinicians from 18 disciplines were asked to generate topics that students are likely to encounter during clerkships and for which they need anatomical knowledge. On the basis of this list, a team of medical practitioners, anatomists and educationalists developed 16 patient cases with 142 accompanying items. In line with the theory that the format of the question should match its content and be based on the number of realistic alternatives that exist in real life,[14] various response formats were used as follows: open-ended items, multiple choice items and true/false type items. The test was reviewed by medical practitioners, who checked the quality of the test, especially the accuracy of patient descriptions, and the answers. A few adjustments were made as a result of these reviews.

The subtest without clinical context consisted of 50 items from an existing examination, the Maastricht Progress Test (MPT).[15,16] The items were selected on the basis of the list of topics described above. All MPT items are presented in the true/false format and undergo an extensive review process.[16] In compiling the two subtests, attention was paid to the distribution of items over the various clinical disciplines and the different regions of the human body.

The items were judged by panels of anatomists and clinicians and an item analysis performed. Six items (four with clinical information and two without clinical information) were removed because they proved to be incorrect or unclear. The final test consisted of 186 items, of which 138 items were presented with clinical context in 16 cases and 48 items were presented without clinical context. Administration of the test took approximately 2 hours. Table 1 shows examples of items from the two subtests.

The test was administered by computer. At seven medical schools, test administration took place in the spring of 2000. At the remaining school clerkships start earlier, in the autumn of Year 4. Therefore, students at that school were tested in the autumn of 2000, immediately before beginning their clerkships.

Table 1 Examples of test items

Example of an item with context
The general practitioner (GP) examines the eye movements of Mrs Ommen. He asks her to look in all directions. Her right eye does not move to the right when she looks in that direction. The GP suspects a paralysis of a specific extra-ocular muscle. Which muscle is involved? *Answer:* Lateral rectus
Example of an item without context
The bronchial arteries are branches of the pulmonary arteries. True or false? *Answer:* False

Statistical analysis

In order to establish whether participating students were representative of their classes, we calculated z-scores of the results of various regular examinations of participating and non-participating students from the same classes. This could only be carried out for schools where such examination results were available. A *t*-test was used to compare the mean z-scores of participating students with those of non-participating students.

We compared age and gender distributions between schools, and the number of student-instructors. The results of the questionnaire and the tests were analysed for student-instructors as a subgroup and for each school, both including and excluding the results of student-instructors.

Students were asked to choose one of four answers to the two questions in the questionnaire (0–25%, 25–50%, 50–75% and 75–100%). Differences between schools were analysed using a Kruskal–Wallis test for non-parametric data with Bonferroni correction for multiple comparisons.

The scores on the anatomy tests were calculated as the percentage of correct answers. Because of the different numbers of questions across cases in the subtest with clinical context, scores were calculated per case and averaged. This way all cases contributed equally to the subtest score.

The reliability of the two subtests was calculated as Cronbach's alpha for internal consistency. For the subtest with clinical context, this was calculated over the 16 cases; for the subtest without context, it was calculated over the 48 items.

One way ANOVA with posthoc analysis was used to test the significance of the differences between the mean test scores in the anatomy tests of the different schools (alpha = 0·05).

Results

Participants

Data for 411 students were analysed. The male/female distribution was similar to that of the entire medical student population in the Netherlands (male = 38·4%; female = 61·6%).[17] The mean age of the students was 22·7 years (range 21–32 years). There were no significant gender or age differences between participants at the various schools.

The comparison between results in the experimental tests and those in regular examinations showed that the participating students from schools A and G were representative of their classes, whereas the participating students from schools C, D, E and H were amongst the better students in their classes. Results of relevant regular examinations were not available for the other two schools.

Of the 411 students, 55 had assisted in anatomy teaching as student-instructors. Their levels of perceived knowledge were higher and they performed much better than the other students. As the number of student-instructors differed between schools and the results did not change when they were left out of the analysis, only the results of the regular students (non student-instructors) are presented in this paper.

Questionnaire

Table 2 shows the responses to the two questions by medical school. Note that 100% represents a different amount of knowledge in the two questions.

The results show that most students thought they had acquired less than half of the anatomy knowledge they could have acquired by the end of Year 4, with no significant differences between schools. Problem-based learning students and non-PBL students had similar perceptions with regard to their levels of anatomy knowledge at the end of Year 4.

The responses to the second question show that 65·2% of all students perceived they had mastered less than half of the knowledge required on graduation (end of Year 6). Students at schools G and H were significantly more positive about their knowledge levels than students at school D. The perceived knowledge level of students at school A, the PBL school at Maastricht, did not differ from that of students at other medical schools.

Table 2 Perceived knowledge (student-instructors omitted) by medical school (A-H)

		Perceived knowledge at the end of Year 4 as percentage of:							
		Knowledge that could have been acquired at the end of Year 4 (Question 1)				Knowledge required at the end of Year 6 (Question 2)			
School	*n*	0–25%	25–50%	50–75%	75–100%	0–25%	25–50%	50–75%	75–100%
A	75	10·8	47·3	39·2	2·7	17·3	49·3	30·7	2·7
B	46	8·7	52·2	34·8	4·3	23·9	39·1	28·3	8·7
C	36	13·9	58·3	27·8	0·0	27·8	41·7	22·2	8·3
D	47	10·6	55·3	29·8	4·3	34·0	55·3	10·6	0·0
E	32	3·1	59·4	34·4	3·1	21·9	37·5	37·5	3·1
F	33	12·1	63·6	24·2	0·0	15·2	54·5	27·3	3·0
G	38	7·9	44·7	47·4	0·0	7·9	28·9	55·3	7·9
H	49	8·2	32·7	57·1	2·0	8·2	53·1	28·6	10·2
Total	356	9·6	50·4	37·7	2·3	19·4	45·8	29·5	5·3

Question 1: only the difference between school H and school F is significant.
Question 2: school D differs significantly from the other schools and school G scores significantly higher than the other schools, with the exception of school H.

Test results

The reliability of the subtests was expressed by Cronbach's alpha for internal consistency. For the subtest with clinical context, alpha = 0·84; for the subtest without clinical context, alpha = 0·29.

The mean score on the items with clinical context was 52·3% ($n = 356$, SD = 8·1, range 34·2–75·4%). The mean score on the subtest without clinical context was 60·3% ($n = 350$, SD = 7·6, range 35·4–85·4%).

Figures 1a and 1b show the means plus 95% confidence intervals of the scores per school. The subtest with clinical context showed significant differences between schools, whereas the subtest without clinical context revealed no such differences.

Students at school G scored significantly higher in the subtest with clinical context than the students at the other schools ($P < 0{\cdot}000$). School D had a significantly lower score than schools A, F and G ($P < 0{\cdot}05$). On the items with clinical context, PBL school A had a score comparable to those of the other schools.

Discussion

We cannot completely rule out the possibility of pre-existing systematic differences between students from different medical schools. However, given the homogeneity of Dutch high school education and the national admission procedure to medical school, significant differences between schools in student levels are unlikely.

Only the students at schools A and G were representative of their classes. The participating students from schools C, D, E and H scored higher than their classmates in regular examinations. Representativeness could not be tested for schools B and F. Students at schools A and G performed better in the anatomy tests than students from the other schools. This suggests that the real differences between schools are likely to lie in the same direction and probably larger than found in this study.

The Maastricht students and the students of two other schools were familiar with the type of questioning in the subtest without context. However, no significant differences were found with this test. Although the subtest with clinical context was constructed in Maastricht, the Maastricht students were not familiar with the type of questioning used in this subtest.

The results for the research question concerning students' perceived knowledge levels showed substantial perceived deficiencies for PBL and non-PBL students alike. The majority of students indicated that they had mastered less than half of the anatomy knowledge they could have mastered by the end of Year 4 and less than half of the knowledge required on graduation. The fact that students perceive that their knowledge levels fall below requirements gives cause for concern. No specific attention is paid to basic sciences, including anatomy, during clerkships in Years 5 and 6. This makes it doubtful whether students will be able to achieve what they regard as the required level of knowledge by graduation.

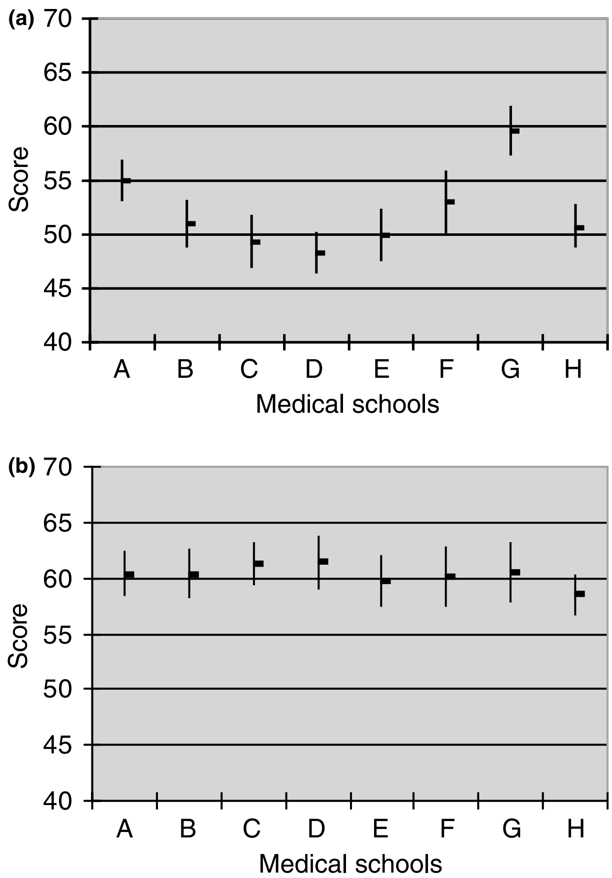

Figure 1a 95% CI of results in subtest with context per school (excluding student-instructors).
Figure 1b 95% CI of results in subtest without context per school (excluding student-instructors).

Regarding the actual knowledge levels of PBL and non-PBL students, no differences between schools were found with the subtest without clinical context. The reliability of this test appeared to be low, indicating that the scores are highly unstable. It is therefore not surprising that no significant differences were found between schools on this test.

Differences between schools were found for the subtest with clinical context. This may be due to the context, but also to the different types of questioning. The fact that PBL students obtained similar scores to the other students suggests that PBL as an educational method does not result in a lower level of anatomy knowledge. Although the best results in the anatomy test were obtained by traditional school G, there was no evidence of better performance by the traditional schools in general.

The study showed no differences in levels of anatomy knowledge between PBL schools and non-PBL schools, but it offers some interesting indications of other relevant educational differences. Although we cannot reveal the identity of the schools, we can highlight some differences in methods of anatomy teaching between the schools. Combined with the study results, these differences may lead us down fruitful avenues for further study of effective basic science teaching.

School G, which scored significantly higher than the other schools in the test with context, schedules twice as much time for anatomy teaching as the other schools do. Both school G and school A, which had the second highest test result in the study, teach anatomy in a clinically meaningful context, in terms of patient problems, diagnostics and therapeutic features. The importance of context in learning is supported by Regehr *et al.*, who identified relevance of the material and the context in which it is embedded as factors which impact positively on retention.[2] It is often assumed that learning in a clinical context is a characteristic of PBL. The fact that learning in context yields good results at school G suggests that more traditional training programmes can also teach successfully in a clinically relevant way.

The better scoring medical schools offer many anatomy features more than once during the preclinical years. There is evidence that repetition of topics has a beneficial effect on learning. Blunt and Blizard showed that initial learning can be strongly reinforced by repetition in later courses.[18]

The study design does not allow any conclusions to be drawn on the basis of these tentative associations between students' results in anatomy tests and teaching methods (time spent on anatomy teaching, repetition and teaching in context). It may be interesting, however, to pursue these relationships in further studies.

Another topic for further study concerns the appropriate level of anatomy knowledge to be attained by graduation. It cannot be ascertained whether students are accurate in perceiving that their knowledge is deficient, because there is no standard available for an absolute interpretation.

In conclusion, the results of this study suggest that students at a PBL school experience deficiencies in anatomy knowledge similar to those of their colleagues at other schools. The discrepancy between what students think they know and what they think they need to know gives cause for concern. No significant differences were found between the anatomy test scores of PBL students and non-PBL students. Future studies might address a standard for anatomy knowledge, and the effects of contextual learning and repetition of topics.

Contributors

KP, HvM, NH, JD and AS developed the instrument of testing. HvM and JD convinced their colleagues in the anatomy departments at other medical schools that it was worthwhile participating in the project. When necessary, AS did the same with colleagues at director level. KP and CvdV analysed the results. All authors contributed to discussion and interpretation of the results. All authors contributed to the writing of the paper.

Acknowledgements

The authors thank Mereke L. B. Gorsira for critically reading and correcting the English manuscript, and the anonymous reviewers for their helpful comments.

Funding

There was no external funding for this paper.

References

1 Barrows HS. A taxonomy of problem-based learning methods. *Med Educ* 1986;**20** (6):481–6.
2 Regehr G, Norman GR. Issues in cognitive psychology: implications for professional education. *Acad Med* 1996;**71** (9):988–1001.
3 Prince KJAH, van de Wiel MWJ, Scherpbier AJJA, van der Vleuten CPM, Boshuizen HPA. A qualitative analysis of the transition from theory to practice in undergraduate training in a PBL medical school. *Adv Health Sci Educ* 2000;**5**:105–16.
4 Drukker J, van Mameren H, Croonen HGHM, Wanders AJB, Verwijnen GM, van der Vleuten CPM. Het oordeel van artsen over het door hen genoten anatomieonderwijs. [Physicians' retrospective evaluation of anatomy teaching.] *Bull Med Onderwijs* 1999;**18**:9–16.
5 van Mameren H, Leiner T, Wanders A, van der Vleuten CPM. Anatomie/embryologieonderwijs aan de Universiteit Maastricht. Het oordeel van studenten geneeskunde. [Students' evaluation of anatomy/embryology teaching at Maastricht University.] *Bull Med Onderwijs* 1997;**16** (2):91–100.
6 Pabst R, Rothkötter HJ. Retrospective evaluation of undergraduate medical education by doctors at the end of their residency time in hospitals: consequences for the anatomical curriculum. *Anat Record* 1997;**249**:431–4.
7 Woodward CA, Ferrier BM. The content of the medical curriculum at McMaster University: graduates' evaluation of their preparation for postgraduate training. *Med Educ* 1983;**17**:54–60.
8 Schmidt HG, Dauphinee WD, Patel VL. Comparing the effects of problem-based and conventional curricula in an international sample. *J Med Educ* 1987;**62**:305–15.
9 Albanese MA, Mitchell S. Problem-based learning: a review of literature on its outcomes and implementation issues. *Acad Med* 1993;**68** (1):52–81.
10 Vernon DT, Blake RL. Does problem-based learning work? A meta-analysis of evaluative research. *Acad Med* 1993;**68** (7):550–63.
11 Verhoeven BH, Verwijnen GM, Scherpbier AJJA *et al.* An analysis of progress test results of PBL and non-PBL students. *Med Teacher* 1998;**20** (4):310–6.
12 Maudsley G. Do we all mean the same thing by 'problem-based learning'? A review of the concepts and a formulation of the ground rules. *Med Teacher* 1998;**20**:323–30.
13 Bickley H. Comparison of problem-based and traditional curricula still not possible. *Acad Med* 1993;**68** (7):545.
14 Schuwirth L. How to write short cases for assessing problem-solving skills. *Med Teacher* 1999;**21** (2):144–50.
15 van der Vleuten CPM, Verwijnen GM, Wijnen WHFW. Fifteen years of experience with progress testing in a problem-based learning curriculum. *Med Teacher* 1996;**18** (2):103–9.
16 Verhoeven BH, Verwijnen GM, Scherpbier AJJA, Schuwirth LWT, van der Vleuten CPM. Quality assurance in test construction: the approach of a multidisciplinary central test committee. *Educ Health* 1999;**12** (1):49–60.
17 VSNU Visitatiecommissie Geneeskunde en Gezondheidswetenschappen. *Onderwijsvisitatie Geneeskunde En Gezondheidswetenschappen.* [Review of undergraduate education at the Dutch Faculties of Medicine and Health Sciences]. Utrecht: Vereniging van Samenwerkende Nederlandse Universiteiten 1997.
18 Blunt MJ, Blizard PJ. Recall and retrieval of anatomical knowledge. *Br J Med Educ* 1975;**9**:255–63.

Received 7 *November* 2001*; editorial comments to authors* 27 *February* 2002*; accepted for publication* 18 *June* 2002