Fifteen years of experience with progress testing in a problem-based learning curriculum

C. P. M. VAN DER VLEUTEN, G. M. VERWIJNEN & W. H. F. W. WIJNEN
University of Limburg, Maastricht, The Netherlands

SUMMARY *This article reports on educational experiences with an assessment method to assess knowledge in a problem-based learning context. This so-called progress test is a comprehensive test sampling knowledge across all content areas of medicine reflecting the end objectives of the curriculum. The test is periodically given to all medical students in the curriculum regardless of their year of training. The format precludes the possibility for students to prepare themselves specifically, therefore preventing the often reported undesirable effects of objective tests such as memorization of facts and interference with tutorial group functioning. The many years of experience indicate that this testing format works effectively. After the introduction of progress tests a number of other, to some extent unexpected, educational advantages became apparent, as well as a few disadvantages. The additional advantages make progress testing also useful in non-problem-based curricula.*

Introduction

Since the introduction of problem-based learning as a new instructional method, test developers have been struggling with assessment of student achievement congruent with the educational principles (Felletti *et al.*, 1983; West *et al.*, 1985; Van der Vleuten & Verwijnen, 1990; Friedman *et al.*, 1994). There is ample evidence that tests and examinations drive student learning (Frederiksen, 1984; Newble & Jeager, 1983; Van der Vleuten *et al.*, 1994). If the educational objectives are not reflected and reinforced by the assessment programme, a 'hidden curriculum' (Snyder, 1971) of assessment objectives will prevail.

In order to accommodate the problem-based learning objectives, test development in problem-based learning has typically attempted to measure higher order cognitive skills, or tried to develop measures for assessing the learning process (Swanson, *et al.*, 1991). Examples of the first approach are the Modified Essay Question (Felletti & Engel, 1980), the Simulation of Initial Medical Problem-solving (De Graaff *et al.*, 1987) and the Problem Analysis Question (Des Marchais *et al.*, 1994); examples of the latter approach are tutor and peer evaluations (De Grave & De Volder, 1984). Instruments reflecting a combination of both approaches are the Triple Jump Exercise (Powles *et al.*, 1981; Chapman *et al.*, 1993) and the Medical Problems examination (Friedman *et al.*, 1994).

Traditional multiple-choice questions were never really considered appropriate for problem-based learning curricula, since they are supposed to measure lower taxonomic levels of knowledge only and they may cause students to 'study for the test' (Swanson *et al.*, 1991). However, this article will report on a multiple-choice test (in this case using true/false questions) called the progress test, which has been specifically developed for a problem-based learning context. The progress test has been used by the Maastricht medical school in The Netherlands since 1977 and considerable experience has been accumulated over the years. In this article the development of the progress test will be described, some outcome data will be presented and, particularly, our experience in terms of its educational value will be discussed. For an overall discussion of the assessment programme referral is made to Van der Vleuten & Verwijnen (1990).

The need for progress testing

The Maastricht medical school started in 1974 as the second school (after McMaster in Canada) to use problem-based learning as a school-wide instructional method. It is a six-year programme with students entering secondary education directly. The first four years of the programme consist of interdisciplinary units of usually 6 weeks. There is an elaborate clinical skills programme integrated with the theoretical curriculum, including short rotations in a variety of clinical settings and ample skills trainings in a Skillslaboratory. The last two years consist of

Correspondence: Professor C. P. M. van der Vleuten, Dept. of Educational Development & Research, PO Box 616, 6200 MD Maastricht, The Netherlands. Email: vandervleuten@educ.rulimburg.nl.

0142-159X/96/04/020103-07

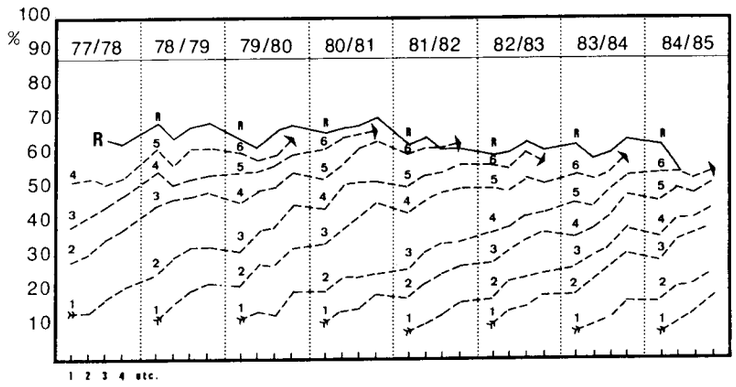

Figure 1. Average year class correct scores on successive progress tests from 1977 to 1985, with solid line (R) representing average scores of national reference groups of recently graduated doctors (in 1977 there were only four classes in the curriculum; the medical school was founded in 1974).

clinical clerkships in the major clinical disciplines, including family medicine (cf. Van der Vleuten & Wijnen, 1990).

Like McMaster, the Maastricht programme heavily emphasizes skills such as self-directedness and open-discovery learning: the students rather than the teachers are responsible for defining learning objectives and for finding the necessary information. This prepares them for lifelong learning. In the absence of national licensing examinations, and due to its deviant curriculum compared with the other medical schools in The Netherlands, the Maastricht medical school was forced to implement a rigorous evaluation system of student achievement, particularly in the knowledge domain. The emphasis was originally placed on end-of-unit examinations. For cost benefit reasons primarily, multiple-choice (true/false) tests were used. Among others, the school introduced end-of-unit examinations. Owing to the large number of students (150 students per class in a six-year curriculum) and because of resource limitations, primarily multiple-choice (true/false) tests were used. Soon, however, it became apparent that these tests had a major steering effect on learning. It became increasingly clear that students prepared themselves specifically for these unit tests (Van Berkel *et al.*, 1995). The tests discouraged individual learning paths and reinforced rote memorization. Instead of defining and presuming their own (or group) learning objectives, the students studied for the expected content of the test. Rather than focusing on understanding and self-directed learning, the approach to assessment taken led to undesirable learning strategies and invited students to 'study for the test'. Despite the educational rationale being well communicated to the students in the tutorial groups, the unit tests rewarded memorization of facts and short-term learning objectives. To increase their chance of academic success (as defined by the unit examinations) the students rapidly started to learn 'how to play the game'.

In order to maintain the educational philosophy, a change in the examination programme was imperative. The solution was sought in cutting the direct relationship between the specific educational programme and the assessment. A new testing method emerged from that idea which came to be known as progress testing.

The progress test

The progress test can best be conceived of as a final examination: a comprehensive examination reflecting the (cognitive) end-objectives of the curriculum. Each progress test consists of approximately 250 (multiple) true/false questions stratified in categories based on the International Classification of Diseases (ICD). It samples knowledge across all disciplines and content areas in medicine relevant for the medical degree [1]. Four times per year the progress test is given to all the students in the curriculum (approximately 900), regardless of their class. For each occasion a newly constructed test is prepared. A single test question may be answered with either true or false, or with an 'I do not know' option (the question-mark option). The latter option is not penalized or rewarded. A correct answer is rewarded with one mark while an incorrect answer is given a negative mark. To discourage guessing a total test score is expressed as the number of correct answers minus the number of incorrect answers. To allow comparison across tests, scores are expressed on a percentage scale.

The freshmen-year students are not able to answer as many questions as the second-year students, who are not able to answer as many as the third-year students and so on. To illustrate what actually happens, a sample of results is depicted in Figure 1. The figure contains average correct scores on the progress test from 1977 to 1984.

This somewhat older sample of data was chosen because it also incorporated scores of a reference group of recently (within one year) graduated doctors from all schools in The Netherlands (group size varied from 50 to about 100 doctors) who were paid to complete the test. The solid top line represents their average score. Because the school started in 1974 only four classes were available in 1977. There are systematic differences between classes (all statistically significant) and average scores increase with time. Mean scores of the sixth-year students are in the vicinity of the reference group scores. After 1984 systematic reference group testing was discontinued owing to the consistency of outcomes and costs involved.

Questions for the progress test are compiled by faculty members from all departments of the medical school. They are of all taxonomic levels and may include facts and

figures or they may contain problem vignettes. Regardless of the taxonomic level of the question, the knowledge addressed should be functional for the student upon graduation. Questions which can only be answered when the answer was memorized the night before have little value in a progress test. All questions are critically reviewed on content, wording and relevance by a review committee consisting of representatives of clusters of disciplines (Van Hessen & Verwijnen, 1989). Virtually all submitted items are reformulated and many are dropped before test administration. All items have a literature reference. After having completed the test, students are given the correct answers and they take the test booklet home. They are invited to criticize individual items when they find conflicting evidence in the literature. Student comments and test statistics are again reviewed by the test review committee and flawed items are dropped. Invariantly approximately 7% of the questions are eliminated following student comments. Final test scores are subsequently calculated. Information on performance, including total scores and detailed profile scores, is sent to students, student mentors (cf. Van der Vleuten & Verwijnen, 1990) and student administration. Mean ICD category and discipline scores per class are reported to departments and curriculum committees. Item scores are reported to the department of origin, including student comments and an item report by the review committee. At present the total item bank includes approximately 15,000 questions.

The pros and cons

Since its introduction in 1977 several surveys have been held to monitor the progress test's functioning and many psychometric investigations were carried out. Instead of reporting and discussing specific data we restrict ourselves to a summary of the most important educational advantages and disadvantages.

The prime motivation to introduce progress tests was to prevent test-directed studying. By cutting the direct relationship with direct previous education by testing the objectives of the entire educational programme (samplewise naturally), it is difficult for students to prepare specifically for the test. In practice this strategy has turned out to be effective. After the introduction of progress tests the unit tests were given formative meaning and student promotion decisions were primarily based on progress tests performance (standards are set in relation to group performance and longitudinal results on progress tests are combined in such a way that progress is rewarded). The impossibility to study for the progress test, in combination with the unthreatening unit examinations, allowed students to rely on their original tutorial group work. There was no longer a need to study for tests directly, and, conversely, individual or tutorial group work was reflected in growth on the progress test. There are still a few students who prepare (e.g. every test invariantly contains questions on the physiology of the heart) or study old questions. These are often the weaker students and their effort seems not to be very effective in achieving significantly better scores.

In summary, the original intent to prevent test-directed studying has indeed been achieved by introducing progress testing in the curriculum (Van Berkel *et al.*, 1995). This desirable effect has been the most important reason for maintaining the progress test as one of the major assessment instruments in the medical school evaluation system. However, since its introduction a number of other and some unexpected advantages have become apparent. We summarize the most important ones:

Focus on functional knowledge

In a system with end-of-unit examinations only, knowledge is easily forgotten. Particularly if the factual load is high and examinations are in competition with each other, students work from examination to examination. Depending on the curriculum structure there is often hardly any longitudinal repetition of content. In a progress testing system there is a constant repetition of topics. For instance, questions related to anatomy, which contents might be learned in year 1, are repeated throughout the entire curriculum, including the clinical clerkships. And, interestingly, knowledge of anatomy still 'grows' during these clerkships. Similar growth patterns emerge for other content areas. Through its repetitive and longitudinal set-up, the progress test emphasizes long-term and functional knowledge rather than detailed and easily forgotten rote memory facts (Semb & Ellis, 1994).

Rich source of information for feedback

As mentioned earlier, the progress test may be taken home by the student and questions may be easily verified through their literature reference. The scores in the progress test can be aggregated to many different sub-scores (i.e. to organ systems, disciplines, clusters of disciplines such as basic science, clinical science and behavioural science) yielding information on knowledge profiles of students. Although the reliability of subscores is low due to the limited number of items, the three-monthly 'X-ray' of the student's knowledge base has proven to be valuable in analysing deficiencies and suggestions for remediation. In addition, the progress test also provides information on what students *do not* know. They can actually assess their state of development in relation to what needs to be achieved at the end of the curriculum. Similarly, average performance on parts of the test provides departments and curriculum committees with valuable information. Growth information on individual items provides individual teachers with useful information.

Early detection of high achievers

Some students, particularly the ones with a relevant prior education (e.g. nursing, physiotherapy) score considerably higher than other students. The progress test immediately identifies these students. This information can be used to plan individual or specialized training. Similarly, it is relative simple to assess the competence level of students coming from other medical schools (e.g. from abroad).

Repeat examinations are no longer necessary

Every progress test is more or less a repeat examination of

a previous progress test. Therefore separate repeat examinations are not necessary. Both from a resource as well as from an educational perspective this is quite an advantage. There is often no end to the (planning of) repeat examinations, and, more importantly, the scheduling of repeat examinations tends to elicit minimal or marginal studying strategies (Starren, 1990).

Organizational insensitivity to curriculum change

A progress test reflects end objectives, not primarily course objectives. Therefore, any change in the curriculum content has no direct consequences for organizing a progress test.

Strong research potential

The cross-sectional and longitudinal design of the progress test provides strong research potential. Programme evaluations at the macro level of comparing curricula can be done easily since the test is not particularly geared towards a specific curriculum. This has been carried out many times in order to compare a problem-based curriculum with other curricula (Verwijnen *et al.*, 1990). Recent international comparisons using Italian and German translations yielded similar outcomes (Albano *et al.*, in press). Curriculum studies at the micro level studying the effectiveness of specific programmes can be carried out using a strong research design with multiple pre-tests and post-tests for a sub-test of interest as well as a sub-test of no interest (as a control). Growth in different knowledge areas may be compared. As an illustration Figure 2 provides a graphical summary of accumulated progress test data from 1983 to 1993 on the 24 possible curriculum measurement points (6 years times four occasions per year) on total test scores (2a) and on basic (2b), clinical (2c) and behavioural science questions (2d). From all available tests mean test scores can be plotted on each of these measurement points. The lines in Figure 2 represent mathematical functions which explain the data best. The function of total scores is completely linear and all others are curved. Although growth remains ongoing throughout the curriculum, basic and behavioural science growth is curved downward, whereas clinical science growth is curved upward. Hence, the linear growth can be explained by curvilinear growth patterns on clusters of disciplines. Although these curvilinear patterns are logical, it is for example interesting to see that basic science knowledge continues to grow after the first few years when most of the basic sciences are offered. In our case, clinical science knowledge exceeds basic science knowledge at about year four. It would be interesting to compare these growth patterns with other problem-based and/or conventional curricula.

Many of the advantages above became clear *after* the introduction of the progress test in our curriculum. The testing concept was as new to us as it probably is for most of the readers of this article.

Naturally, there are disadvantages associated with the progress test. Among them are the following.

Sensitivity of progress test in freshman years

Only a limited part of the test reflects the contents of the first year of the curriculum (approximately 20% of the questions). Therefore, the progress test may be less sensitive for achievement in the freshman year. A study which addressed this issue partly confirmed this problem. The progress test was less sensitive for differences between first-year students at the lower end of the ability scale (Imbos, 1989). Therefore, to discriminate between poorly performing students the progress test is not very appropriate (their identification as such is less of a problem).

Test-difficulty variation

Although the reliability of the progress test is reasonable (within years of training alphas between 0.70 and 0.80; across years usually above 0.95) and test construction is highly standardized, test difficulty variations remain present. Variations in mean performance within years across occasions is substantially larger than variations between different cohorts, indicating true test difficulty variations. Criterion-referenced standard setting without complex psychometric test equation is therefore problematic.

Need for a central organization for test development

The nature of the progress test requires a central organization to develop and administer these tests. This requires substantial resources. A test review committee, in our experience essential for item validity, is time-consuming and costly. For many schools, this may be an impediment [2]. Once such a review system is realized it has major advantages to improve test quality. The remaining disadvantage is the indirect involvement of teachers in the evaluation process.

Identification of (core) content

Since there is no direct link between course and test content, the concrete identification of items in relation to the objectives of the curriculum is problematic. Curriculum objectives only exist in broad terms and provide a limited anchor for judging the relevance of individual items. Therefore, it is often difficult to prevent the incorporation of detailed and specialized knowledge items.

Conclusion

After 15 years of use, the progress test has become a core element of our problem-based learning programme. In our view its philosophy is congruent with problem-based learning in that it emphasizes functional longitudinal knowledge while at the same time self-directed problem-based learning is not impeded. Actually, most of our students do not pay much attention to the progress test. They complete the test in a relaxed way, have a low degree of interest in checking the questions afterwards, and sometimes hardly look at their computer form of profile scores. They use the pass mark as a verification to proceed in the way they are studying. Given the usual anxiety about examinations, this can be considered as another advantage of the format, in

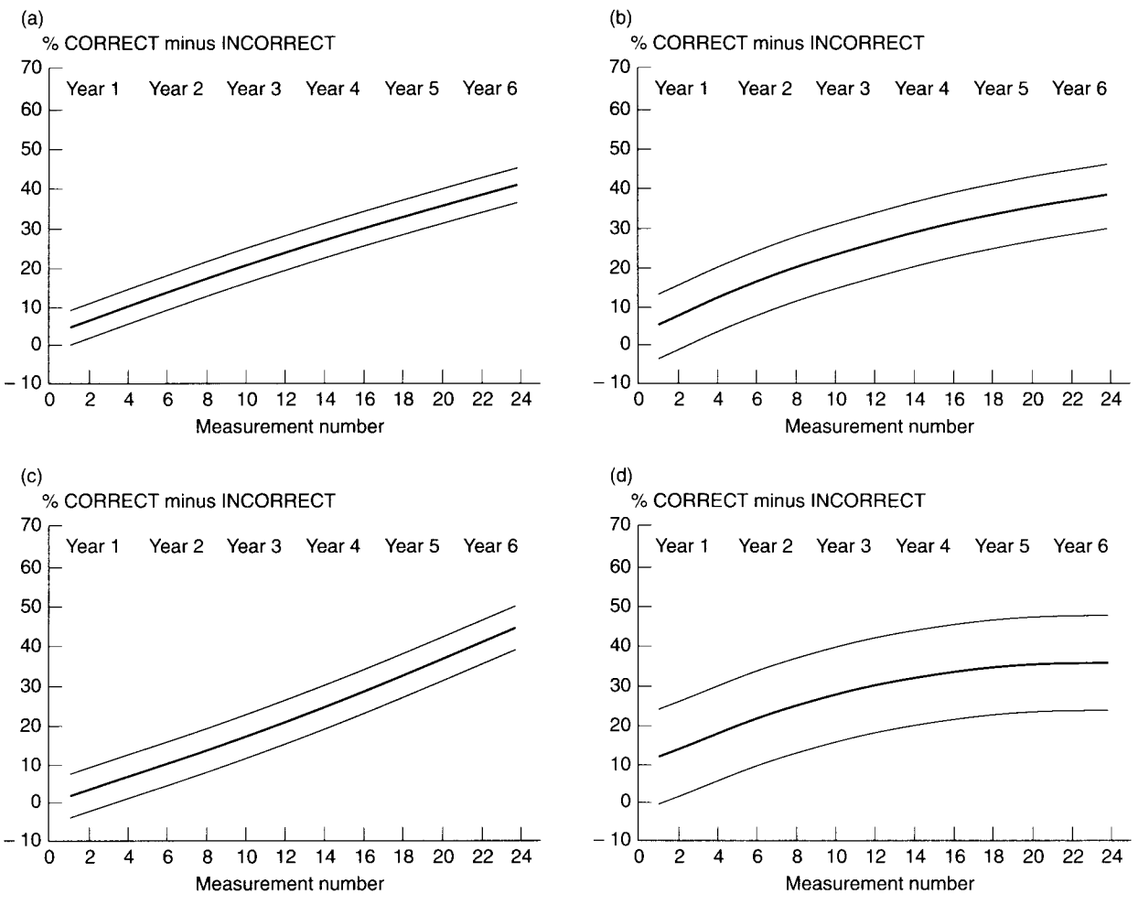

Figure 2. Estimated growth curves of: (a) total progress test scores (correct minus incorrect scores) and 95% confidence intervals based on cumulative data from 1982 to 1994 (function: $Y = 2.49 + 1.87X–0.013X^2$; fit = 96%); (b) basic sciences scores (correct minus incorrect scores) and 95% confidence intervals based on cumulative data from 1982 to 1994 (function: $Y = 3.40 + 2.24X–0.033X^2$; fit = 85%); (c) clinical sciences scores (correct minus incorrect scores) and 95% confidence intervals based on cumulative data from 1982 to 1994 (function: $Y = 0.53 + 1.52X + 0.014X^2$; fit = 95%); (d) behavioural science scores (correct minus incorrect scores) and 95% confidence intervals based on cumulative data from 1982 to 1994 (function: $Y = 9.25 + 2.28X–0.052X^2$; fit = 58%).

line with the non-threatening student-centred philosophy of problem-based learning.

At present, research has well documented the importance of knowledge in clinical competence and problem-solving (Schmidt *et al.*, 1990), and also within a context of problem-based learning (Norman, 1991). The desirable characteristics of the progress test in relation to problem-based learning have also convinced the McMaster medical school, traditionally suspicious of objective tests, to adopt the format as an addition to their tutor evaluations of students (Keane *et al.*, 1994; Blake *et al.*, 1995). The introduction of progress tests into their curriculum had no impact on the learning approach of students (Blake *et al.*, 1994). Within The Netherlands other university programmes use progress tests (Van Berkel *et al.*, 1993), including two other Dutch medical schools (one problem based) and a number of tertiary schools for health science education (physiotherapy, nursing, midwifery), and the postgraduate residency programmes in general practice and paediatrics have successfully applied progress tests (Schuwirth *et al.*, 1994; Van Leeuwen *et al.*, 1995).

The progress test has also been successful in accounting for the quality of our graduates. In the numerous comparisons with reference groups, including students from all classes of other medical schools, doctors after graduation and control groups of non-medical students, our students have proved able to perform as well as students from non-problem-based curricula (while attrition rates at the same time are substantially more favourable for the problem-based programme; Verwijnen *et al.*, 1990). This has been compelling evidence that the quality of Maastricht graduates is equal to that of others, at least in the knowledge domain.

Our progress test consists of true/false questions. This was originally motivated by the relative ease with which they could be constructed. However, the question format is unimportant. The only requirement is the need to sample broadly across content. Time-efficient question formats

are therefore to be preferred. Multiple-choice questions have slightly higher reliabilities per unit of testing time in this regard (Norcini *et al.*, 1985).

Finally, progress tests are developed and used in the context of problem-based learning, but there is no impediment for their use in other educational programmes, particularly in individualized educational programmes. A similar longitudinal testing procedure has been developed, independently from the progress test, in the School of Medicine at the University of Missouri Kansas City (Arnold & Willoughby, 1994). The benefits of longitudinal assessment have proved to be equally valuable in this non-problem-based school. In our view progress tests are also worthwhile exploring in regular conventional curricula.

Notes on contributors

C. P. M. VAN DER VLEUTEN is an educationist and Professor and Chair of Department of Educational Development and Research.

G. M. VERWIJNEN is a General Practitioner and Senior Lecturer in the Skillslab.

W. H. F. W. WIJNEN is an educationist and Professor of the Department of Educational Development and Research.

Notes

[1] The undergraduate programme in The Netherlands leads to an MD degree allowing entrance to specialty training. Independent clinical practice without specialty training is not allowed.

[2] It is not suggested that the Maastricht medical school has more resources than other medical schools (these are all comparable in The Netherlands). The central procedure is realized by saving on peripheral assessment activities: individual teachers or departments do not make examinations.

References

ALBANO, M.G., CAVALLO, F., HOOGENBOOM, R., MAGNI, F., MAJOOR, G.D., MANENTI, F., SCHUWIRTH, L.W.T., STIEGLER, I. & VAN DER VLEUTEN, C.P.M. (in press). An international comparison of cognitive levels of medical students using the Maastricht progress test, *Medical Education*.

ARNOLD, L. & WILLOUGHBY, T.L. (1990) The Quarterly Profile Examination, *Academic Medicine*, 65, pp. 515–516.

BLAKE, J.M., KEANE, D.R., NORMAN, G.R. & MUELLER, C.B. (1994) Introducing progress testing in a traditional problem-based curriculum: impact on student learning, in: *Research In Medical Education (Abstracts)* (Washington: AAMC).

BLAKE, J.M., NORMAN, G.R. & SMITH, E.K.M. (1995) Report card from McMaster: student evaluation at a problem-based medical school, *Lancet*, 345, pp. 899–902.

CHAPMAN, J.A., WESTMORLAND, M.G., NORMAN, G.R., DURRELL, K. and HALL, A. (1993) The structured oral self-directed learning evaluation: one method of evaluating the clinical reasoning skills of occupational therapy and physiotherapy students, *Medical Teacher*, 15, pp. 223–236.

DE GRAAFF., E., POST, G. & DROP, M. (1987) Validation of a new measure of clinical problem-solving, *Medical Education*, 21, pp. 213–218.

DE GRAVE, W. & DE VOLDER, M. (1984) Peer-evaluation and problem-based learning, in: P.A.J. BOUHUIJS, H.G. SCHMIDT & M. DE VOLDER (Eds) *Tutorials in Problem-based Learning* (Assen: Van Gorcum).

DES MARCHAIS, J.E., JEAN, P., DUMAIS, B. & VU, N.V. (1994) 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*, (Maastricht: Network Publications).

FELLETTI, G.I. & ENGEL, C.E. (1980) The modified essay question for testing problem-solving skills, *Medical Journal of Australia*, 1, pp. 79–80.

FELLETTI, G.I., SAUNDERS, N. & SMITH, A. (1983) Comprehensive assessment of final-year medical student performanced based on undergraduate programme objectives, *Lancet*, ii, pp. 34–37.

FREDERIKSEN, N. (1984) The real test bias: influences of testing on teaching and learning, *American Psychologist*, 39, pp. 193–202.

FRIEDMAN, C.P., MURPHY, G.C., SMITH, A.C. & MATTERN, W.D. (1994) Exploratory study of an examination format for problem-based learning, *Teaching and Learning in Medicine*, 6, pp. 194–198.

IMBOS, T. (1980) *Het gebruik van einddoeltoetsen bij aanvang van de studie*, [The use of end-objective assessments in freshmen years], unpublished doctoral dissertation, University of Limburg.

KEANE, D.R., BLAKE, J.M., NORMAN, G., & MUELLER, C.B. (1994) Introducing progress testing in a traditional problem-based curriculum: psychometric properties, in: *Research In Medical Education (Abstracts)*, (Washington: AAMC).

KNOX, J.D.W. (1976) The modified essay question, in: *Medical Education Booklet*, (Dundee: Association for the Study of Medical Education).

LEADER, S.R., FELLETTI, G.I. & ENGEL, C.E. (1979)—Assessment—help or hurdle?, *Programmed Learning and Educational Technology*, 16, pp. 308–315.

MCGUIRE, C.H. & SOLOMON, C.M. (1976) *Construction and Use of Written Simulations*, (Chicago: Phychological Corporation).

NEWBLE, D. & JAEGER, K. (1983) The effect of assessments and examinations on the learning of medical students, *Medical Education*, 17, pp. 165–171.

NORCINI, J.J., SWANSON, D.B., GROSSO, L.J., SHEA, J.A. & WEBSTER, G.D. (1985) Reliability, validity and efficiency of multiple choice question and patient management problem item formats in the assessment of physician competence, *Medical Education*, 19, pp. 238–247.

NORMAN, G.R. (1991) What should be assessed? In: D. BOUD, & G. FELLETTI (Eds) *The Challenge of Problem-based Learning* (London: Kogan Page).

POWLES, A., WINTRUP, N., NEUFELD, V., WAKEFIELD, J., COATES, G. & BURROWS, J. (1981) The triple jump exercise: further studies of an evaluative technique, in: *Proceedings of the 20th Annual Conference on Research in Medical Education* (Washington: American Association of American Colleges).

SCHMIDT, H., NORMAN, G. & BOSHUIZEN, H. (1990) A cognitive perspective on medical expertise: theory and implications, *Academic Medicine*, 65, pp. 611–621.

SCHUWIRTH, L.W.T., SCHRANDER, J.P.P. & VAN DER VLEUTEN, C.P.M. (1994) *Assistententoets kindergeneeskunde: een beschrijving van de psychometrische eigenschappen* [A test for paediatric residents: a description of psychometric characteristics] *Bulletin Medich Onderwijs*, 12, pp. 146–151.

SEMB, G.B. & ELLIS, J.A. (1994) Knowledge taught in school: what is remembered?, *Review of Educational Research*, 64, pp. 253–286.

SNYDER, B.R. (1971) *The Hidden Curriculum* (New York: Knopf).

STARREN, H. (1990) *De beoordeling als hefboom voor onderwijsverbetering* [Assessment as a lever for educational improvement], *De Psycholoog*, 25, pp. 109–113.

SWANSON, D.B., CASE, S.M. & VAN DER C.P.M. VLEUTEN (1991) Student assessment in problem-based learning curricula, in: D. BOUD & G. FELLETTI (Eds) *The Challenge of Problem-based Learning* (London: Kogan Page).

VAN BERKEL, H.J.M., NUY, H.J.P. & GEERLIGS, T. (1995) The influence of progress tests and block tests on study behaviour, *Instructional Science*, 22, pp. 315–331.

VAN BERKEL, H.J.M., SPROOTEN, J. & DE GRAAFF, E. (1993) The development of a progress test for a multi-master program health sciences' curriculum, in: P.A.J. BOUHUIJS, H.G. SCHMIDT

& H.J.M. VAN BERKEL (Eds) *Problem-based Learning as an Educational Strategy*, (Maastricht: Network Publications).

VAN DER VLEUTEN, C.P.M. & VERWIJNEN, G.M. (1990) A system for student assessment, in: C.P.M. VAN DER VLEUTEN & W.H.F.W. WIJNEN, (Eds) *Perspectives from the Maastricht Experience* (Amsterdam: Thesis-publ).

VAN DER VLEUTEN, C.P.M. & WIJNEN, W.H.F.W. (Eds) (1990) *Problem-based Learning: Perspectives from the Maastricht Experience* (Amsterdam: Thesis-publ).

VAN DER VLEUTEN, C.P.M., NEWBLE, D., CASE, S., HOLSGROVE, G., MCCANN, B., MCRAE, C. & SAUNDERS, N. (1994) Methods of assessment in certification, in: D. NEWBLE, B. JOLLY & R. WAKEFORD (Eds) *The Certification and Recertification of Doctors: Issues in the Assessment of Clinical Competence* (Cambridge: Cambridge University Press).

VAN HESSEN, P. & VERWIJNEN, G.M. (1989) *Toetsen getoetst: het beoordelen van toetsvragen in Maastricht* [Tests tested: the review of test items in Maastricht], *Bulletin Medisch Onderwijs*, 8, pp. 100–105.

VAN LEEUWEN, Y.D., POLLEMANS, M.C., MOL, S.S.L., EEKHOF, J.A.H., GROL, R. & DROP, M.J. (1995) The Dutch knowledge test for general practitioners: Issues of validity, *European Journal of General Practice*, 1, pp. 113–117.

VERWIJNEN, G.M., VAN DER VLEUTEN, C.P.M. & IMBOS, T. (1990) Comparing an innovative medical school with traditional schools: an output analysis in the cognitive domain, in: T. KHATTAB, H. SCHMIDT, Z. NOOMAN, & E. EZZAT (Eds) *Innovation in Medical Education: An Evaluation of Its Present Status* (Berlin, Springer-Verlag).

WEST, D., UMLAND, B. & LUCERO, S. (1985) Evaluating student performance, in: A. KAUFMAN (Ed.) *Implementing Problem-based Medical Education* (New York: Springer-Verlag).