Failure of Feedback to Enhance Self-Assessment Skills of General Practitioners

Jacobus J. M. Jansen, Richard P. T. M. Grol, Harry F. J. M. Crebolder, and Jan-Joost Rethans
Department of General Practice
University of Maastricht
Maastricht, Netherlands

Cees P. M. van der Vleuten
Department of Educational Research and Development
University of Maastricht
Maastricht, Netherlands

Background: *Self-directed learning requires accurate self-assessment, but research evidence shows poor validity of self-assessment. Training in self-assessment may improve validity.*
Purpose: *To investigate if repeated personal feedback based on objective knowledge and skill scores enhances self-assessment skills of practicing general practitioners.*
Method: *Participants were general practitioners (n = 60), who received skills training covering 4 clinical skills at 3 months (Group A) or 6 months (Group B) after enrollment in the study. Participants were tested at 3-month intervals with a knowledge test (60 items), a performance-based test (4 stations), and a self-assessment questionnaire (22 items), covering the four different clinical skills. They received personal feedback on the results.*
Results: *At 3 months, mean scores on the self-assessment questionnaire and knowledge test had increased significantly more in Group A compared to Group B, but at 6 months no differences in mean scores remained. Correlations between self-assessment rating and objective scores were low to moderate, with little overall improvement over time.*
Conclusions: *Although self-assessment scores can to some extent be useful in measuring perceived changes in competence in groups, individual self-assessment scores on their own are an invalid source of information concerning competence of practicing physicians, and this does not improve significantly with regular feedback.*

Teaching and Learning in Medicine, *10*(3), 145–151

Self-directed learning is the dominant mode of learning for professionals after graduating. It implies a process in which individuals take the responsibility for diagnosing their learning needs. Accurate self-assessment, the ability of physicians to perceive areas of strength and weakness in their competence or performance, is therefore considered an essential requisite for effective adult learning,[1,2] and the development and maintenance of professional competence.[3,4] Although the importance of self-assessment is widely recognized, research evidence has provided little support for validity of self-assessment in relation to expert ratings or objective tests.[5,6] Correlations between self-assessment and expert ratings among students at different levels of expertise and for various aspects of clinical competence are generally low to moderate.[7–12] It has been argued that these results are a consequence of the absence of specific training in self-assessment skills in medical training programs, and that validity and accuracy of self-assessment may improve with self-assessment training.[5,13]

Personal feedback on actual achievement has proven to be a powerful method of training,[14] but few studies have investigated development of self-assessment skills over time with participants receiving regular feedback on actual scores. Cochran and Spears[15] and Hay[16] found correlations between student and instructor ratings increase from moderate to high. Regular feedback opportunities, including discussion of the completed self-evaluations, were part of the course. However, results were not compared to objective tests, and the high correlations can be explained by either improved self-evaluation skills or, alternatively, as evidence of successful negotiation between students and instructors.[15,16]

Correspondence may be sent to Jacobus J. M. Jansen, MD, Swalm 76, 5032 EM Tilburg, Netherlands. E-mail: jjjansen@knmg.nl

Only two studies, both among undergraduate medical students, have investigated the development of self-assessment over time, comparing self-assessment with objective measures. Arnold et al.[17] compared self-ratings of medical students during 4 years with faculty ratings. Both student and faculty ratings showed annual increases, with the students' increase being smaller. Correlation between self- and faculty ratings decreased to a nonsignificant level among more advanced students, and no correlation was found between self-assessment rating and objective test scores. Students with higher scores on objective tests were more conservative in their self-assessments. Rezler[18] studied the development of self-assessment among medical students during their first 2 years in a problem-based curriculum. Mean ratings increased from 1st to 2nd year, but correlations dropped to nonsignificant levels. Student self-ratings for reasoning and knowledge showed no significant correlation with a knowledge test administered in the 2nd year.

Few studies have investigated self-assessment skills of experienced physicians. Because physicians frequently base their decisions to attend continuing medical education on self-assessment, it is important to know if they can learn to become more accurate in assessing their own educational needs. This study investigated whether repeated personal feedback based on objective knowledge scores and skill scores for technical procedures performed in primary care enhanced self-assessment skills of practicing general practitioners with regard to these procedures.

Method

General practitioners were invited to participate in a continuing medical education course on technical clinical skills as part of an experiment investigating transfer of skills from training to practice environment. Those who agreed to participate were divided into two groups according to their preference of when to take the course. One group (A, the intervention group) took the course 3 months after enrollment, and the other group (B) took the course 6 months after enrollment.

The course covered four different skills: injection technique of the shoulder, ophthalmoscopic control in diabetes, Pap smear, and laboratory examination of fluor vaginalis. The topics were identified by general practitioners as having priority and selected by course providers because proficiency in these procedures was known to be amenable to improvement. The objective of the course was to increase relevant knowledge and proficiency in performance of procedures according to national guidelines for general practice. For each skill the training was given in small groups (4–8 persons) by two trainers experienced in the area. The content of the training was based on national guidelines for general practice and included discussion of the guidelines with supervised hands-on practice of skills forming the core of each training session. Total training time was 3 hr, with 1 hr for injection technique of the shoulder and ophthalmoscopy, and half an hour for Pap smear and laboratory examination of fluor. Satisfaction of the participants was measured directly after the course using a questionnaire.

To evaluate the cognitive effect of the training, a 60-item multiple choice test was used to measure relevant knowledge. The content of the knowledge test was directly based on the content of the course and using national consensus guidelines.[19] With an Objective Structured Clinical Examination using trained observers and detailed checklists, with 23 to 33 items for each skill based on the content of the course, proficiency was assessed in the four different clinical skills. Manikins were used for shoulder injection (Limbs & Things® shoulder model) and Pap smear (Schultz®). For fluor vaginalis a specimen of fluor was used, and trained real patients were used for ophthalmoscopic control in diabetes mellitus.

A self-assessment questionnaire was developed, based on the content of the course, and consisted of 22 items to be scored on a 7-point Likert scale. As an illustration, sample items of the different formats are shown in Figure 1.

The knowledge test and self-assessment questionnaire were administered to both groups at the start (enrollment), 3 months after enrollment, and 6 months after enrollment. The performance-based (skill) test was administered to the intervention group (Group A) 3 months after enrollment and to both groups 6 months after enrollment.

All participants received personal detailed written feedback on their scores. The correct answers were provided, allowing review of errors, and individual scores for each procedure were compared with results of the peers, indicating whether scores were low, average, or high. Written educational information was provided, with the feedback, reviewing correct performance, with step-by-step guidelines for the different procedures.

Scores on the test formats for the different skills were aggregated to total scores and converted to a percentage of the maximum score. To evaluate self-assessment in relation to objective scores the differences between self-assessment scores and knowledge or skill scores were calculated after transforming original scores into z-scores to adjust for differences in average scores. Subgroups of low, intermediate, and high scores were constructed by equally dividing participants over the three subgroups with group assignment according to percentile score. A t test was used to compare mean scores between the intervention (A) and control (B) groups, and a paired t test was used to compare mean scores between the different testing intervals (enrollment, 3 months after, *Self assessment questionnaire*

		dominance	*educational need*
		1. almost no dominance; 2. some dominance, insufficient integration; 3. some integration, insufficient for practical use; 4. sufficient dominance for practical use; 5. sufficient dominance for practical use, some routine; 6. good dominance and routine; 7. excellent dominance	1. none; 2. small; 3. some but less than average; 4. average; 5. considerable, more than average; 6. large; 7. very large
1.	anatomy of the shoulder joint	1 2 3 4 5 6 7	1 2 3 4 5 6 7
2.	Physical Ex.of the shoulder joint	1 2 3 4 5 6 7	1 2 3 4 5 6 7
3.	Differential Diagnosis of shoulder complaints	1 2 3 4 5 6 7	1 2 3 4 5 6 7
4.	injection technique articulatio glenohumeralis	1 2 3 4 5 6 7	1 2 3 4 5 6 7
5.	injection technique bursa subacromialis	1 2 3 4 5 6 7	1 2 3 4 5 6 7
6.	injection technique subacromial space	1 2 3 4 5 6 7	1 2 3 4 5 6 7
7.	injection technique articulatio acromio-clavicularis	1 2 3 4 5 6 7	1 2 3 4 5 6 7

Knowledge test

item 16 — Pain in the shoulder with irradiation into the hand is more frequently caused by a problem in the neck rather than the shoulder.(true)

item 17 — Inflammation of the shoulder joint capsule causes disturbances *both* in active and passive examination of shoulder movements. (true)

item 18 — Restriction of the passive horizontal adduction of the arm during physical examination of the shoulder is an indication of osteo-arthritis of the acromio-clavicular joint. (true)

item 29 — A correct insertion point for injection of the glenoid cavity is about 4 cm below the angle between spina scapulae and acromion. (false)

item 30. — During injection of the glenoid cavity from behind the correct direction of the needle is towards the top of the coracoid process. (true)

Performance based test (part of scoring grid)

	not/incorrectly performed	correct performed
Choice of materials and preparation for injection		
- correct syringe	☐	☐
- adequate needle	☐	☐
- adequate dosage of lignocaine	☐	☐
- adequate dosage of corticosteroid	☐	☐
- disinfection of skin and palpating fingers	☐	☐
Injection technique		
- correct insertion site	☐	☐
- correct angle of insertion	☐	☐
- correct depth	☐	☐
- aspiration before injection	☐	☐

Figure 1. *Examples of items on the self-assessment questionnaire, knowledge test, and performance-based test for shoulder injection technique.*

Table 1. *Knowledge, Skill, and Self-Assessment Scores for Groups A and B at 3-Month Intervals*

Internval	Knowledge			Skill			Self-Assessment		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
At Enrollment									
Group A	30	51.5	7.9		Not administered		30	50.0	10.8
Group B	29	49.0	6.7		Not administered		30	50.6	10.4
3 Months After Enrollment									
Group A	30	80.1	6.4[a,b]	30	75.7	9.3	30	65.9	9.9[a,b]
Group B	30	66.3	11.2[b]		Not administered		30	53.2	13.4
6 Months After Enrollment									
Group A	19	81.5	9.0	24	81.8	5.6	29	70.3	10.5
Group B	30	79.8	9.5[b]	30	80.0	7.2	30	67.6	10.4[b]

Note: All entries expressed as percentage of maximum score.
[a]$p < .001$ between Groups A and B. [b]$p < .001$ between scores at enrollment versus 3 months after enrollment or 3 months after enrollment versus 6 months after within groups.

Table 2. *Correlations Between Knowledge, Skill, and Self-Assessment Scores*

Interval	*N*	Knowledge	Self-Assessment
At Enrollment			
Knowledge	59		0.19
3 Months After Enrollment			
Knowledge	60		0.46[a]
Skill	30	0.24	0.06
6 Months After Enrollment			
Knowledge	49		0.21
Skill	54	0.20	0.46[a]

[a]$p < .001$.

and 6 months after). Chi-square (or Fisher exact) was used to analyze differences in percentages between groups. One-way analysis of variance (with Student Neuman Keuls as post hoc multiple comparison method) was used to evaluate the influence of nominal and ordinal characteristics on scores. For interrater reliability of the performance-based test, intraclass coefficients were used.[20] Bivariate correlations were expressed as Pearson product moment coefficients.

Results

Scores were available for 60 participants, equally divided between Group A and B. However, various participants had incomplete data. At the start, one knowledge test scoring sheet from Group B was not returned. At 6 months, 24 participants of Group A were tested on skill, and only 19 knowledge test scores were available; all but one filled out the self-assessment rating sheet. Subgroup analysis of nonparticipants versus participants at 6 months revealed no indication of selection bias. Personal characteristics (age, sex, years of experience) and practice characteristics of both groups showed no statistically significant differences. Satisfaction of participants was high about course content (97%), assessment (89%), and feedback (93%), and not significantly different between both groups.

Scores on knowledge and self-assessment showed no statistically significant differences between groups at the start (Table 1). Both groups showed improvement in these scores after 3 months, but improvement was significantly higher in Group A. At 6 months, after Group B had also received the training, no significant differences in scores remained between the two groups.

No significant influences of personal and practice characteristics on knowledge scores, skill scores, and self-assessment ratings were found, except for city practices, whose doctors had higher self-assessment scores (but not higher knowledge or skill scores) at 3 and 6 months.

The interrater reliability for the performance-based test was .80 at 3 months and .83 at 6 months. Correlations between self-assessment rating and knowledge score were low at the start, increased to moderate at 3 months, and declined again at 6 months to the same level as at the start. The correlation between self-assessment and skill was very low at 3 months and increased to moderate levels at 6 months (Table 2).

Accuracy of self-assessment was additionally assessed by comparing the mean standardized difference between self-assessment and knowledge score or skill score, after equally dividing participants in three subgroups (low, intermediate, and high) using the total scores on the knowledge test and skills test. The results are given in Table 3. The mean standardized difference scores between self-assessment and knowledge score or skill score of the low-scoring group varied between 0.33 and 0.74 above the overall mean, whereas the mean difference for the high scoring group ranged between 0.30 and 0.72 below the overall mean. Variance within groups (expressed as standard deviations) was large compared to differences between groups.

Discussion

The increase in mean self-assessment scores among the participants, and the different patterns of increase

Table 3. *Difference Between Self-Assessment Score and Scores on the Knowledge and Skills Test for Low, Medium, and High Achievers*

Interval	N	Low Achievers		Medium Achievers		High Achievers		F
		M	SD	M	SD	M	SD	
At Enrollment								
Knowledge	59	0.720	.88	–0.16	0.84	–0.53	0.85	< 0.001
3 Months After Enrollment								
Knowledge	60	0.43	1.22	–0.12	0.86	–0.31	0.75	< 0.05
Skill	30	0.74	1.00	–0.02	0.56	–0.72	0.85	< 0.01
6 Months After Enrollment								
Knowledge	49	0.68	0.78	0.26	0.67	–0.84	0.86	< 0.001
Skill	54	0.33	1.06	–0.04	1.02	–0.30	0.85	n.s.

Note. All scores after Z-transformation of difference between self-assessment score and knowledge test score or skills test score. Group assignment according percentiles: low achievers: scores $< p$ 33.3; medium achievers have scores less than $p = 33.3$ to $p = 66.7$; high achievers have scores greater than $p = 66.7$.

of the intervention group compared to the control group, together with corresponding changes in knowledge score, can be considered supportive for validity of self-assessment at the group level. Growth in self-evaluation scores during training have been reported in various studies.[16–18, 21–24] However, in isolation, this is hardly compelling evidence for the validity of self-evaluation.

The results of correlation analysis suggest some improvement of self-assessment over time, first for knowledge and later for skill. This can be interpreted as an indication that participants shifted from predominantly "knows how" (at 3 months) "to shows how" (at 6 months),[25] and thus became somewhat more accurate in their self-assessment of proficiency concerning the technical clinical skills. Nevertheless, self-assessment was generally a poor predictor of competence.[26] Only 20% to 25% of variance on self-assessment was explained by the scores on the objective tests, and no substantial effect of repeated detailed personal feedback could be demonstrated.

Other factors apparently heavily influenced the relation between self-assessment and competence. We found no indications of significant contributions of personal characteristics (age or sex), professional (experience), or practice characteristics, with exception of rural–urban differences at 3 and 6 months. These results are consistent with other research.[17,27,28] Some authors have argued that self-assessment is strongly influenced by noncognitive attributes, such as self-representation and personality.[6,7,17,23,29] These factors might account for the large unexplained variability of self-assessment in this study.

The finding that high achievers tended to underestimate their knowledge or performance and low achievers tended to overestimate is consistent with the results from other self-assessment studies.[5,7,10,15,23,30–32] Various explanations have been forwarded for this intriguing phenomenon. One possible explanation is that the physicians were evaluating themselves according their ambition rather than their actual performance,[7] implicating that high achievers may compare themselves to more stringent standards.[17] Alternatively they may have viewed their performance as not being quite as good as it was influenced by their internal self-representations developed early in life.[23] However, the consistent finding could also be interpreted as yet another proof of poor validity of self-assessment. As the statistical chance for low achievers to overestimate their knowledge and performance is greater than for high achievers, these results may in fact underline that self-assessment is of questionable validity.

This study has some limitations. The participants are a rather small group of physicians who volunteered to submit themselves to an experiment including assessment. Although we found no indication that nonparticipants at 6 months differed from their peers, the missing knowledge scores from 11 participants and missing skill scores from 6 participants in the intervention group at 6 months are a potential source of bias in this study. Another limitation was that participants were not very familiar with self-assessment, so scores may reflect a large error component. Nevertheless it would have been likely to expect that some learning would occur during the course, especially in this highly motivated group, resulting in improved correlation between self-assessment scores and objective scores, and this effect was not observed. The use of pre- and postintervention self-assessments can be criticized because internal standards of participants may change as a result of the intervention; instead the use of retrospective pre- and post-self-assessments has been recommended as more consistent with objective measures.[33,34] However, the skills covered in this study were all familiar to participants. Therefore a misunderstanding or misconception of the skill or concept is not likely. Moreover the use of retrospective pre- and post-self-assessment would not have changed the self-assessment score at 6 months, after having received the training, and this score also correlated rather poorly with objective measures. It might be argued that feedback procedures as used in this study were not adequate to realize changes.[13] However such feedback procedures have been proven to be effective to bring about change in competence and performance,[35] and self-assessment scores at the group level did show changes reflecting changes in objective scores, but at the individual level self-assessment failed to become more accurate. So we believe that the findings in this study indeed support the minor influence of feedback on self-assessment, consistent with research indicating that self-assessment is more closely related to generalized self-attributions and only minimally influenced by external feedback.[36]

This result has important implications for a voluntary continuing medical education system in which selection of activities is based on individual preferences. Although it is obvious that motivation is crucial in adult learning,[37] the use of objective measurement must provide the basis to bring subjective and objective learning needs closer to each other, to enhance rational choice of continuing medical education topics.[38–40]

In conclusion, although self-assessment scores at the group level can be useful to some extent in measuring perceived changes in competence, individual self-assessment scores on their own are an invalid source of information concerning competence of practicing physicians, and this does not improve significantly with regular feedback. Therefore objective tests should have a much larger place as a basis for individualized continuing medical education.

References

1. Knowles MS. *The modern practice of adult education: From pedagogy to andragogy.* Chicago: Follett, 1980.
2. Jennett P, Jones D, Mast T, Egan K, Hotvedt M. The characteristics of self-directed learning. In DA Davis, RD Fox (Eds.), *The physician as a learner: Linking research to practice* (pp. 47–65). Chicago: American Medical Association, 1994.
3. Schön DA. *The reflective practitioner: How professionals think in action.* New York: Basic Books, 1983.
4. Schön DA. *Educating the reflective practitioner: Towards a new design for teaching and learning in the professions.* San Francisco: Jossey-Bass, 1987.
5. Boud D, Falchikov N. Quantitative studies of student-self-assessment in higher education: A critical analysis of findings. *Higher Education* 1989;18:529–49.
6. Gordon MJ. A review of the validity and accuracy of self-assessments in health professions training. *Academic Medicine* 1991;66:762–9.
7. Stuart MR, Goldstein HS, Snope FC. Self-evaluation by residents in family medicine. *Journal of Family Practice* 1980;10:639–42.
8. Scalabassi SE, Woelfel SK. Development of self-assessment skills in medical students. *Medical Education* 1984;18:226–31.
9. Kolm P, Verhulst S. Comparing self- and supervisor evaluations. A different view. *Evaluation & the Health Professions* 1987;10:80–9.
10. Cushing AM, Jolly BC, Dacre JE, Hitman G, Griffiths S, Southgate L. Critical self-appraisal and examiner, patient and student self-ratings of communication skills in OSCE. In AI Rothman, R Cohen (Eds.), *The Sixth Ottawa Conference on Medical Education* (pp. 145–8). Toronto: University of Toronto, 1994.
11. Furman GE, Colliver JA, Galofré A. Student self-ratings and standardized patient ratings of a medical interview. In AI Rothman, R Cohen (Eds.), *The Sixth Ottawa Conference on Medical Education* (pp. 142–4). Toronto: University of Toronto, 1994.
12. Jansen JJM, Tan LHC, Van der Vleuten CPM, Van Luijk SJ, Rethans JJ, Grol RPTM. Assessment of competence in technical clinical skills of general practitioners. *Medical Education* 1995;29:247–53.
13. Gordon MJ. Self-assessment programs and their implications for health professions training. *Academic Medicine* 1992;67:672–9.
14. Patrick J. *Training: Research and Practice.* London: Academic, 1992.
15. Cochran SB, Spears MC. Student self-assessment and instructors' ratings: A comparison. *Journal of American Dietetic Association* 1980;76:253–5.
16. Hay JA. Investigating the development of self-evaluation skills in a problem-based tutorial course. *Academic Medicine* 1995;70:733–5.
17. Arnold L, Willoughby TL, Calkins EV. Self evaluation in undergraduate medical education: The longitudinal perspective. *Jounal of Medical Education* 1985;60:21–8.
18. Rezler AG. Self-assessment in problem-based groups. *Medical Teacher* 1989;11:151–6.
19. Grol R. National standard setting for quality of care in general practice: Attitudes of general practitioners and responses to a set of standards. *British Journal of General Practice* 1990;40:361–4.
20. Kramer MS, Feinstein AR. Clinical biostatistics. LIV. The biostatistics of concordance. *Clinical Pharmacological Therapeutics* 1981;20:111–23.
21. Bleys FC, Gerritsma JGM, Netjes I. Skills development by medical students and the influence of prior experience: A study using evaluation by students and self-assessment. *Medical Education* 1986;20:234–9.
22. Woolliscroft JO, Palchik NS, Dielman TE., Stross JK. Self-evaluation by house officers in a primary care training program. *Journal of Medical Education* 1985;60:840–6.
23. Woolliscroft JO, Tenhaken J, Smith J, Calhoun JG. Medical students' clinical self-assessments: Comparisons with external measures of performance and the students' self-assessments of overall performance and effort. *Academic Medicine* 1993;68:285–94.
24. Day SC, Cook EF, Nesson HR, Wolf MA, Goldman L. A learning-curve approach to the self-assessment of internal medicine training. *Journal of Medical Education* 1984;59:672–5.
25. Miller GE. The assessment of clinical skills/competence/performance. *Academic Medicine* 1993;65:S63–7.
26. Norman GR. Defining competence: A methodological review. In VR Neufeld, GR Norman (Eds.), *Assessing clinical competence* (pp. 15–35). New York, Springer, 1985.
27. Strober Escovitz E, Cohen DG. An analysis of gender differences in students clinical competence self-assessments. In W Bender, RJ Hiemstra, AJJA Scherpbier, RP Zwierstra (Eds.), *Teaching and assessing clinical competence* (pp. 110–7). Groningen, The Netherlands: Boekwerk, 1990.
28. Farmer E. Changes in self assessment of communication skills of Australian family medicine programme trainees compared with a group of hospital peers. In RM Harden, IR Hart, H Mulholland (Eds.), *Approaches to the assessment of clinical competence* (pp. 121–5). Dundee, Scotland: Centre for Medical Education, 1992.
29. Kegel-Flom P. Predicting supervisor, peer and self-ratings of intern performance. *Journal of Medical Education* 1975;50:812–5.

30. Richards BF, Philip EB, Frye AW, Philip JR. Integrating self-assessment in an OSCE for pre-clinical medical students. In IR Hart, RM Harden, J Des Marchais (Eds.), *Current developments in assessing clinical competence* (pp. 338–46). Montréal: Can-Heal, 1992.
31. Nathan RG. Using an in-training examination to assess and promote the self-evaluation skills of residents. *Academic Medicine* 1992;67:613.
32. Morton JB, Macbeth WAAG. Correlations between staff, peer and self assessments of fourth-year students in surgery. *Medical Education* 1977;11:167–70.
33. Levinson W, Gordon G, Skeff K. Retrospective versus actual pre-course self-assessments. *Evaluation & the Health Professions* 1990;13:445–52.
34. Skeff KM, Stratos GA, Bergen MR. Evaluation of a medical faculty development program: The comparison of traditional pre/post an retrospective pre/post self-assessment ratings. *Evaluation & the Health Professions* 1992;15:350–66.
35. Wensing M, Grol R. Single and combined strategies for implementing changes in primary care: A literature review. *International Journal for Quality in Health Care* 1994; 6:115–32.
36. Gordon MJ. A review of the validity and accuracy of self-assessments in health professions training. *Academic Medicine* 1991;66:762–9.
37. Mann K, Ribble J. The role of motivation in self-directed learning. In DA Davis, RD Fox (Eds.), *The physician as a learner: Linking research to practice* (pp. 69–88). Chicago: American Medical Association, 1994.
38. Sibley JC, Sackett DL, Neufeld V, Gerard B, Rudnick KV, Fraser W. A randomized trial of continuing medical education. *New England Journal of Medicine* 1982;306:511–5.
39. Premi J. Individualized continuing medical education. In DA Davis, Rd Fox (Eds.), *The physician as a learner: Linking research to practice* (pp. 201–16). Chicago: American Medical Association, 1994.
40. Tracey JM, Arroll B, Richmond DE, Barham PM. The validity of general practitioners' self-assessment of knowledge: Cross sectional study. *British Medical Journal* 1997;315:1426–8.

Received 12 February 1997
Final revision received 21 January 1998