Exploring family physicians' reactions to multisource feedback: perceptions of credibility and usefulness

Joan Sargeant,[1] Karen Mann[2] & Suzanne Ferrier[1]

PURPOSE Physician performance is comprised of several domains of professional competence. Multisource feedback (MSF) or 360-degree feedback is an approach used to assess these, particularly the humanistic and relational competencies. Research studying responses to performance assessment shows that reactions vary and can influence how performance feedback is used. Improvement does not always result, especially when feedback is perceived as negative. This small qualitative study undertook preliminary exploration of physicians' reactions to MSF, and perceptions influencing these and the acceptance and use of their feedback.

METHODS We held focus groups with 15 family physicians participating in an MSF pilot study. Qualitative analyses included content and constant comparative analyses.

RESULTS Participants agreed that the purpose of MSF assessment should be to enhance practice and generally agreed with their patients' feedback. However, responses to medical colleague and co-worker feedback ranged from positive to negative. Several participants who responded negatively did not agree with their feedback nor were inclined to use it for practice improvement. Reactions were influenced by perceptions of accuracy, credibility and usefulness of feedback. Factors shaping these perceptions included: recruiting credible reviewers, ability of reviewers to make objective assessments, use of the assessment tool and specificity of the feedback.

CONCLUSION Physicians' perceptions of the MSF process and feedback can influence how and if they use the feedback for practice improvement. These findings are important, raising the concern that feedback perceived as negative and not useful will have no or negative results, and highlight questions for further study.

KEYWORDS clinical competence/*standards; feedback; physicians/*psychology; attitude of health personnel; focus groups; pilot projects.

*Medical Education 2005; **39**: 497–504*
doi:10.1111/j.1365-2929.2005.02124.x

[1]Continuing Medical Education, Dalhousie University, Canada
[2]Division of Medical Education, Dalhousie University, Canada

Correspondence: Ms Joan Sargeant, Continuing Medical Education, Dalhousie University, Clinical Research Centre, 5849 University Avenue, Halifax, NS B3H 4H7, Canada. Tel: (902) 494 1995; Fax: (902) 494 1479; E-mail: joan.sargeant@dal.ca

INTRODUCTION AND LITERATURE REVIEW

Physician performance is multifaceted, comprised of a number of domains of professional competence. Among these, knowledge and basic skills have traditionally been more adequately taught and assessed than the integrative, relational, affective and reflective aspects of professional competence.[1] Recently, 360-degree or multisource feedback (MSF) has been introduced as a formative assessment and quality improvement approach to enhancing performance in multiple domains of practice.[2,3] A questionnaire-based process using self-assessment and medical colleague, co-worker and/or patient reviewers, MSF can be an effective means of assessing residents and physicians[4–9] and has stimulated practice improvement.[10,11]

However, performance assessment is not simply an objective, cognitive process. The following quotation, although dated with respect to gender, reminds us of its affective component: 'Performance appraisal touches on one of the most emotionally charged activities in work life – the assessment of a man's

Overview

What is already known on this subject

Multisource feedback (MSF) can be an effective means to assess physician performance and stimulate practice improvement; however, research has shown that responses to such feedback is quite varied and can influence how the feedback is used.

What this study adds

This exploration of physicians' reactions to MSF suggests that perceptions of the credibility, accuracy and usefulness of MSF, influenced how, and if, this small group of experienced physicians used their feedback for practice improvement and continued learning.

Suggestions for further research

Future studies could explore the following questions: will responses of a larger, purposeful sample be similar to these? How can we better understand the influence of negative perceptions of feedback upon its acceptance and use? Will reviewer and physician education about MSF enhance perceptions of objectivity? How can the feedback and the process of providing feedback be improved to enhance practice improvement and continued learning?

contribution and ability. The signals he receives about this assessment have a strong impact on his self-esteem and subsequent performance'.[12] Reactions to performance assessment vary and can influence how the feedback is used. Although the intent is to improve performance improvement does not always result, particularly when feedback is perceived as negative.

Our understanding of these responses is informed by research in organisational psychology. An analysis of over 600 management performance appraisal studies showed that about one-third of the managers improved their behaviour, one-third stayed the same and one-third decreased their performance. Managers receiving negative feedback often became discouraged and not motivated to change.[13] Assessment results which compared peer rankings tended to generate some loss of positive feelings on the part of those not in the upper part of the distribution, even for those ranked 'satisfactory' compared to those who excelled.[14] Individuals being assessed may not see the need for change when unfavourable performance feedback is inconsistent with self-perceptions.[15] More recently, a study of MSF demonstrated that negative ratings were not seen as accurate or useful, and resulted in negative reactions such as anger and discouragement and not in increased awareness.[16] However, MSF has also been found to be useful, and managers who focused on positive aspects were more likely to improve their own performance.[17,18]

Various factors appear to influence acceptance of assessment feedback. For managers, these included perceptions of familiarity with reviewers, objectivity of assessment, credibility of process, relevance of competencies being assessed and manner in which feedback is delivered.[2] Similarly, medical residents' receptivity to feedback in the clinical setting was influenced by sender credibility, in turn influenced by method of feedback delivery, feedback content and perceptions of sender characteristics.[19]

Within medicine, MSF studies have shown familiarity with colleague and co-worker reviewers to be positively associated with scores; i.e. those knowing physicians well rated them higher.[4,5,9,20] Moreover, physicians' agreement with feedback decreased as their scores decreased[20] and they believed that reviewer bias reduced the credibility of their feedback.[21] What stimulated these responses? Do these responses influence use of feedback? This small qualitative study was undertaken to explore questions such as these and, specifically, physicians' reactions to the MSF process and feedback, and perceptions influencing these and the acceptance and use of feedback.

METHODS

Study background

In 2002, the College of Physicians and Surgeons of Nova Scotia (NS) undertook a pilot study of MSF using the standardised process and questionnaires of the College of Physicians and Surgeons of Alberta programme, Physician Achievement Review. The goal of this programme is formative review and quality improvement.[22,23] The NS target population was family physician volunteers. Using the programme

guidelines for reviewer selection, physicians identified 8 medical colleague and 8 co-worker reviewers, and randomly selected 25 patient reviewers. Reviewers completed the questionnaires by rating the physician on multiple items, and participating physicians completed similar self-assessment questionnaires. Questionnaires used 5-point Likert scales, with an additional 'unable to assess' option. Colleague and co-worker forms also asked how well the rater knew the physician. The 142 participants received their feedback by mailed confidential report presenting, in tables and graphs, their individual mean scores compared to the NS aggregate mean scores, for each item and by competency. Scores above the 90th percentile received 'commendation' flags, and below the 10th percentile, 'information' flags. The report also compared self-assessment scores with medical colleague scores. Additionally, participants and reviewers completed brief programme evaluation questionnaires soliciting their opinions about the MSF process.[20]

Study design

We used focus groups with MSF physician participants to explore their responses to the assessment. Focus groups enable participants to describe experiences and perceptions meaningful to them and through discussion with others, reflect and respond to those of others, and potentially create new understanding.[24]

We sent invitation letters to the 113 of the 142 volunteer physicians who completed evaluations to participate in the focus groups. Of 42 from across the province who volunteered, 15 were able to participate in three focus groups, others being prevented by distance, scheduling and funding constraints.

We sent the study questions to participants beforehand, to encourage reflection upon their experiences before the focus groups. We held the focus groups in December 2002, facilitated by two members of the research team. Each was 1–1.5 hour-long, audio-recorded and transcribed. The study was approved by the Dalhousie University Research Ethics Board.

Data analysis

Focus group questions addressed participants' general expectations of the MSF process, overall response to their feedback and to specific reviewer groups and domains assessed, use of feedback for practice improvement and learning, and overall opinion of the process as a tool for both assessing and enhancing physician performance. Analysis proceeded at two levels using accepted qualitative procedures. Using content analysis, the 2 researchers facilitating the focus groups independently coded data by question and developed categories describing physicians' reactions.[24] They compared findings and resolved any differences. Then, using a constant comparative approach[25] 1 researcher compared and contrasted data in and among categories, to understand linkages and relationships and determine central themes. Although the majority of responses were positive, participant reactions to medical colleague and co-worker feedback in particular appeared to be influenced by 2 central concepts: perceived accuracy of the data and credibility of the process and perceived usefulness of their feedback. The second researcher reviewed these conceptual decisions. To confirm them, we carefully re-analysed the data. To further ensure trustworthiness of the analytical process, we compared focus groups findings with participant and reviewer evaluation results.[24]

RESULTS

Participants were office-based family physicians, including 5 women and 10 men. Six were from urban areas (population > 30 000) and 9 from rural communities, 8 of these with populations of 5000 or less. They represented 7 of the 9 provincial health districts. Of note, this was an experienced group of family physicians, having graduated from medical school an average of 23 years ago (range 10–39 years).

Reactions to MSF

Participants agreed that the overall purpose of a review process such as MSF should be practice enhancement and quality improvement, i.e. to identify ways to help individual physicians improve their practice. However, a few also expressed distrust and concern that it was not being used for quality improvement but to 'weed out bad apples'. There was strong consensus that patients as consumers should participate in the MSF process and participants generally agreed with patient feedback.

Responses to feedback received from medical colleagues and co-workers varied. Although the majority viewed this generally as positive and helpful a few did not, and others questioned particular aspects. Feedback perceived as being strongly negative generally

evoked emotional responses, including anger and discouragement. Being surprised, i.e. receiving feedback that differed from self-perceptions, seemed to contribute to this. The following quotations illustrate the scope of responses:

> 'It was pretty much reasonable throughout, so I must say I didn't have a huge big surprise. … I kind of got what I expected.' (Rural 1).

> 'People probably already guessed I wasn't very happy with it! I just sent it around to a half dozen of my colleagues, all of whom sent me letters back saying, ''This is completely off the wall; out to lunch; that's not how we feel about you at all.'' … So where's the validity!' (Rural 2)

The two central emerging concepts, i.e. that these reactions were influenced by perceptions of (1) accuracy and credibility of the data and process and (2) usefulness of feedback, are discussed in the following sections.

Perceptions of accuracy of data and credibility of MSF process

Three factors contributed to perceptions of accuracy and credibility. These included selecting unbiased yet informed reviewers, reviewers' ability to observe and effectively assess them, and use of the assessment tool.

Selecting unbiased yet informed reviewers

Selecting co-worker and medical colleague reviewers was problematic for many participants. The challenge appeared to be a balancing act, a tension between selecting reviewers who knew them well enough to respond to the questionnaire items and provide a broad perspective of their practice, while avoiding those who knew them so well that responses might be biased. The following quotation describes this:

> 'you have to have someone who can adequately evaluate you… it may be our responsibility to choose people who know you well enough. But the other side of that coin, is, you don't want to influence by choosing someone you know. So that's a real difficult issue, and I don't know what the solution is. …I don't know how to get from one extreme to the other.' (Rural 1)

Strategies to manage this tension varied. Some felt strongly that it was important to choose colleagues and co-workers who knew them sufficiently well. One participant added that he also chose the ones he 'knew for a long time and knew what their personalities are like …the ones who would actually do an honest job' (Urban 1).

Conversely, others felt that it was important to avoid biased responses. One participant did this by asking his staff to randomly select reviewers. Two others deliberately did not invite colleagues and co-workers they knew well:

> 'I work in a group of six doctors, but I particularly didn't want to use all those five doctors who work with me because I felt if this was a proper assessment I wanted a broader view of what was going on.' (Urban 2)

These 3 physicians were disappointed by their results: 2, because feedback was more negative than anticipated, and 1, insufficient medical colleagues responded to provide a rating for that section.

Reviewers' ability to observe and assess effectively

Central to participants' concerns about recruiting medical colleague and co-worker reviewers was the opportunity for reviewers' to observe their performance, to make objective assessments. This ability appeared key to perceptions of credibility and accuracy. It seemed dependent upon practice context and specific performance items being assessed.

Surprisingly, participants' practice setting (rural or urban) or type of practice (solo or group) did not independently determine participants' ability to recruit reviewers who knew them well enough to assess them. Instead, it seemed to be the individual practice context and professional relationships which determined this. Working regularly in a hospital or nursing home seemed to enhance the number and richness of relationships for some. The following quotations describe varying relationships:

> 'we have a core of nurses that I work with more hours of the week than I care to recount and the same with our local pharmacists, there's essentially only three in the area that I deal with. So you do have much more of a working relationship with them …' (Rural 1)

> 'it's easier for people who are in a community hospital who have a group of people they refer to and deal with every day, to find cohorts who are able to assess you.' (Rural 3).

In contrast, a participant in practice with another physician in a city where family physicians do not have general hospital privileges, described her professional relationships as follows:

> 'Just a comment on the assumption that because you live in ''X'' city there's lots of co-workers. The problem here is that when I send patients with a prescription, they go to one of, I don't know, hundreds of pharmacies and there's no one pharmacist that I have a relationship with. There's no one physiotherapist, there's no one nurse.' She summarised by saying, 'Actually, I've heard the comment that ''there's nothing as isolated as a city GP''.' (Urban 3)

Another urban physician reflected:

> 'There's so many specialists that we don't always have a close relationship with them. So you don't know how well they actually know you.' (Urban 4).

Some participants suggested that distance need not be a barrier to assessing performance. They agreed that consultants to whom they referred patients, and communicated solely through referral letters and telephone, could come to know them well enough over time to assess many aspects of their performance. There was the belief that: 'they get a pretty good sense of you from your notes…' (Urban 5).

There was consensus across groups that some components of care were infrequently observed and hence more difficult to assess. One of these was the ability to assess family physicians' 'management of patient psycho-social issues', although others suggested that even this was possible to assess at a distance, conveyed through 'the flavour of the [referral] letters' (Rural 5). Another component was record-keeping. One participant questioned the score received for this: 'It was a real slap in the face and I don't know how anybody would really know what my medical records were like!' (Urban 6). Others questioned colleagues' and co-workers' ability to assess 'technical competence'. While physicians whose practice included direct patient care in hospitals or nursing homes described being frequently observed by others, physicians who were solely office-based did not have opportunities for others to observe them.

Participants suggested reviewers should be encouraged to use the 'unable to assess' option on the scale, if they were unable to observe the behaviour or had insufficient data to assess it objectively. They wanted to be assured that assessments were made as objectively as possible.

Use of the assessment tool

Lack of clarity about interpretation and use of the Likert scale caused some participants to question the objectivity and accuracy of assigned scores, especially if scores were low. Participants who had also been reviewers for others expressed particular reservations: 'I would never presume to judge other physicians because I don't know where they're coming from' (Rural 3), or 'you want to positively evaluate people. You don't want to fail them' (Rural 4). They were also concerned about the tendency for reviewers to not discriminate among items and score all the same: 'It's a lot easier to put a ''3'' if you don't want to think about it very much' (Urban 6), or 'if you're a little ambivalent you put ''3'' ' (Rural 1).

Additionally, the concept of 'average', and what constituted 'being average', appeared unclear. There was general consensus that on a 5-point scale, '3' should be average, '4' above average and '5' exceptional. Most felt they were average in their performance. However, the aggregate means of scores reported were '4' or above for most items. Hence, if one was scored as '3' it generally fell below the average and created distress.

Perceived usefulness of the feedback

Usefulness of the feedback appeared directly related to credibility and accuracy, and specificity of the feedback. Physicians who agreed with their feedback appeared more accepting of its accuracy and were more likely to consider practice improvements based on that feedback. One participant commented that she 'knew her medical colleague reviewers well and hence trusted their assessments' (Urban 4). Another described the process as: 'trying to look at it as a critique as opposed to people being critical. In other words, it was a positive event and the results are to be looked at that way too' (Rural 6). A few gave examples of changes: improving communication with consultants, purchasing a cell phone to increase access for patients, reflecting upon patient information needs following diagnostic tests. Conversely, three who responded negatively questioned the credibility of medical colleague and co-worker feedback. Two of these suggested that they would now conduct their own independent practice reviews to see if others agreed with the MSF report, before making any changes to practice.

Equally important, these 3, plus a fourth who reported being depressed by negative feedback, did not believe the feedback was specific enough to be helpful. This was a grave concern for them. For example, the participant concerned about the low score for medical records questioned:

> 'But so far as I know none of my colleagues have ever seen my records so I'm left with this sort of pile of cotton wool that's, you know, a big criticism, and what do I do about it?' (Urban 6)

The participant who was depressed described her reactions:

> 'For about a week I was really depressed and the lowest score I had was, ''Is aware of own shortcomings'', on which I got a 3.4. and you know, that's not useful at all. I thought, what shortcomings am I not aware of? There wasn't anything in here to help me, I have no idea where this came from or whether somebody really has a problem with me. Or did several people just think I was average and it kind of evened out lower than the others? …There's this fuzzy cloud over me, and I have no way of knowing where it's coming from or whether there's anything I can do about it. and that's why I felt very depressed. '(Urban 3)

Although physicians who agreed with their feedback did not describe it as being unhelpful, they agreed that more specific feedback would be more useful. Suggestions included adding reviewer comments and examples, especially for items scored at the extremes of the scale. There was strong consensus that using MSF in combination with chart review, as used in the current provincial office peer review, would create a more robust and objective assessment process.

DISCUSSION

This preliminary study describes the reactions of 15 experienced family physicians to multisource feedback and explores perceptions influencing these reactions. Participants' perceptions of MSF ranged from positive through negative, reflecting results of research in organisational psychology.[2,16] Several responding negatively did not agree with their feedback nor were inclined to use it for practice improvement. These findings are important, raising the concern that feedback perceived as negative and not useful will have no or negative results.[16] Further study of the influence of physicians' negative reactions to MSF upon acceptance and use of feedback for practice enhancement will be helpful.

Similar to other studies, responses to MSF were shaped by perceived objectivity of the assessment process and credibility of reviewers.[2,16,19] Credibility was influenced by participants' perceptions of reviewers' familiarity with their practice and ability to make objective assessments. Interestingly, participants who intentionally did not select reviewers who knew them well, received disappointing results, illustrating the difficulty of discriminating between reviewers being 'adequately familiar' and being biased. MSF was developed for use in organisations where working relationships provided adequate opportunity to observe and assess others, while family physicians work in varied settings. Their working relationships may require further exploration to determine characteristics of credible reviewers. Education of participants and reviewers about the MSF process and tools may also enhance credibility of the process[2] and reduce common measurement errors such as the 'halo' effect and central tendency.[26,27]

Effective assessment feedback is specific in nature and participants agreed that their MSF report, provided by mailed document and including only numerical scores, was often inadequate to identify needed improvements. This was distressing for some. Adding written comments to explain numerical scores could increase specificity and providing feedback in person could clarify questions and ambiguities. Making available a mentoring service to support physician feedback, reflection, learning and change, can increase acceptance and use of feedback.[1,28,29] As Epstein and Hundert[1] suggest, 'good assessment is a form of learning and should include guidance and support to address learning needs' (p. 229). These approaches deserve further investigation.

The study is limited by the small volunteer convenience sample preventing full exploration of concepts. However, consistent with the literature, participants described a broad range of reactions to MSF and findings support earlier results of evaluations of the NS and Alberta MSF programmes.[20,21]

CONCLUSIONS

These results suggest that perceptions held by a small group of experienced family physicians of the credibility, accuracy and usefulness of MSF influenced how, and if, they used their feedback for practice

improvement and continued learning. Questions raised by this study include: will responses and perceptions of a larger, purposeful sample be similar to these? How can we better understand the influence of negative perceptions of feedback upon its acceptance and use? Will reviewer and physician education about MSF enhance perceptions of objectivity? How can the feedback and the process of providing feedback be improved to enhance practice improvement and continued learning?

Contributors: JS collaboratively designed the research proposal and developed focus group questions, conducted the focus groups, conducted data analysis, and wrote the manuscript. KM collaboratively designed the research proposal and identified focus group questions, reviewed focus group results and the analytical process and assisted with writing the manuscript. SF assisted with the development of focus group questions, the conducting of the focus groups, data analysis and writing the manuscript.
Acknowledgement: we acknowledge review and critique of the manuscript by Cees van der Vleuten PhD, University of Maastricht and Jocelyn Lockyer PhD, University of Calgary.
Funding: this study was funded by the College of Physicians and Surgeons of Nova Scotia, Halifax, Nova Scotia, Canada.
Conflicts of interest: there are no conflicts of interest.
Ethical approval: the study was approved by the Research Ethics Board of Dalhousie University.

REFERENCES

1 Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA* 2002;**287**:226–35.
2 Bracken DW, Timmreck CW, Church AH. *The Handbook for Multisource Feedback*. San Francisco, CA: Jossey-Bass 2001.
3 Tornow WW, London M, CLL. Associates. *Maximizing the Value of 360-Degree Feedback*, 1st edn. San Francisco, CA: Jossey-Bass 1998.
4 Ramsey PG, Wenrich MD, Carline JD, Inui TS, Larson EB, LoGerfo JP. Use of peer ratings to evaluate physician performance. *JAMA* 1993;**269**:1655–60.
5 Wooliscroft JO, Howell JD, Patel BP. Resident–patient interactions: the humanistic qualities of internal medicine residents assessed by patients, attending physicians, program supervisors and nurses. *Acad Med* 1993;**68**:680–7.
6 Weinrich MD, Carline JD, Giles LM, Ramsey PG. Ratings of the performances of practicing internists by hospital-based registered nurses. *Acad Med* 1993;**68**:680–7.
7 Violato C, Marini A, Toews J, Lockyer J, Fidler H. Feasibility and psychometric properties of using peers, consulting physicians, co-workers, and patients to assess physicians. *Acad Med* 1997;**72** (Suppl. 10):S82–4.
8 Hall W, Violato C, Lewkonia R, Lockyer J *et al.* Assessment of physician performance in Alberta: the physician achievement review. *CMAJ* 1999;**161**:52–7.
9 Lipner RS, Blank LL, Leas BF, Fortna GS. The value of patient and peer ratings in recertification. *Acad Med* 2002;**77** (Suppl. 10):S64–6.
10 Fidler H, Lockyer JM, Toews J, Violato C. Changing physicians' practices: the effect of individual feedback. *Acad Med* 1999;**74** (702):14.
11 Lockyer J. Multisource feedback in the assessment of physician competencies. *J Contin Educ Health Profess* 2003;**23** (1):4.
12 Thompson PH, Dalton GW. Performance appraisal: managers beware. *Harvard Business Rev* 1970;**48**:149–57.
13 Kluger AN, DeNisi A. The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychol Bul* 1996;**119**:254–84.
14 Pearce JL, Porter LW. Employee responses to formal performance appraisal feedback. *J Appl Psych* 1986;**71** (2):211–8.
15 McCall MW, Lombardo MM. Off the track: why and how successful executives get derailed. Greenboro, NC: Centre for Creative Leadership 1983.
16 Brett JF, Atwater L. 360 Feedback. accuracy, reactions, and perceptions of usefulness. *J Appl Psych* 2001;**86** (5):930–42.
17 Atwater LE, Ostroff C, Yammarino FJ, Fleenor JW. Self–other agreement: does it really matter? *Personnel Psychol* 1998;**51**:577–98.
18 Walker AG, Smither JW. A five-year study of upward feedback: what managers do with their results matters. *Personnel Psychol* 1999;**52**:393–423.
19 Bing-You RG, Paterson J. Feedback falling on deaf ears: residents' receptivity to feedback tempered by sender credibility. *Med Teach* 1997;**19**:40–4.
20 Sargeant JM, Mann KV, Ferrier SN *et al.* Responses of rural family physicians and their colleague and co-worker raters to a multi-source feedback process: a pilot study. *Acad Med* 2003;**78** (10):S42–S44.
21 Alberta's Physician Achievement Review (PAR) Program: A review of the first three years. College of Physicians and Surgeons of Alberta. Available at: http://www.cpsa.ab.ca/collegeprograms/attachments/par_evaluation.pdf. Accessed 30 April 2004.
22 Alberta's Physician Achievement Review (PAR) Program. College of Physicians and Surgeons of Alberta. Available at: http://www.cpsa.ab.ca/collegeprograms/par_program.asp. Accessed 28 April 2004.
23 Nova Scotia Physician Achievement Review (NSPAR) Program. College of Physician and Surgeons of Nova Scotia. Available at: http://www.nspar.ca/. Accessed 28 April 2004.
24 Rice PL, Ezzy D. *Qualitative Research Methods*. Victoria, Australia: Oxford University Press 2000.
25 Strauss A, Corbin J. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 2nd edn. Thousand Oaks, CA: Sage 1998.

26 Johnson D, Cujec B. Comparison of self, nurse, and physician assessment of residents rotating through an intensive care unit. *Crit Care Med* 1998;**26**:1811–16.
27 Norcini JJ. Peer assessment of competence. *Med Educ* 2003;**37** (6):539–43.
28 Pringle M, Wilson T, Grol R. Measuring 'goodness' in individuals and healthcare systems. *BMJ* 2002;**325** (7366):704–7.
29 London M. *Job Feedback: Giving, Seeking and Using Feedback for Performance Improvement*. Mahwah, NJ: Lawrence Erlbaum Associates 1997.

Received 5 May 2004; editorial comments to authors 22 June 2004; accepted for publication 22 July 2004