



# Cross-cultural challenges for assessing medical professionalism among clerkship physicians in a Middle Eastern country (Bahrain): feasibility and psychometric properties of multisource feedback

Ahmed Al Ansari, Khalid Al Khalifa, Mohamed Al Azzawi, Rashed Al Amer, Dana Al Sharqi, Anwar Al-Mansoor & Fadi M Munshi

**To cite this article:** Ahmed Al Ansari, Khalid Al Khalifa, Mohamed Al Azzawi, Rashed Al Amer, Dana Al Sharqi, Anwar Al-Mansoor & Fadi M Munshi (2015) Cross-cultural challenges for assessing medical professionalism among clerkship physicians in a Middle Eastern country (Bahrain): feasibility and psychometric properties of multisource feedback, *Advances in Medical Education and Practice*, 10, 509-515, DOI: [10.2147/AMEP.S86068](https://doi.org/10.2147/AMEP.S86068)

**To link to this article:** <https://doi.org/10.2147/AMEP.S86068>



© 2015 Al Ansari et al. This work is published by Dove Medical Press Limited, and licensed under Creative Commons Attribution – Non Commercial (unported, v3.0) License



Published online: 11 Aug 2015.



Submit your article to this journal [↗](#)



Article views: 67



View related articles [↗](#)



View Crossmark data [↗](#)

# Cross-cultural challenges for assessing medical professionalism among clerkship physicians in a Middle Eastern country (Bahrain): feasibility and psychometric properties of multisource feedback

Ahmed Al Ansari<sup>1-3</sup>

Khalid Al Khalifa<sup>1</sup>

Mohamed Al Azzawi<sup>1</sup>

Rashed Al Amer<sup>1</sup>

Dana Al Sharqi<sup>4</sup>

Anwar Al-Mansoor<sup>5</sup>

Fadi M Munshi<sup>6</sup>

<sup>1</sup>Department of General Surgery, Bahrain Defense Force Hospital,

<sup>2</sup>Surgical Department, Arabian Gulf University, <sup>3</sup>Medical Education Department, Royal College of Surgeons in Ireland - Bahrain,

<sup>4</sup>Department of Internal Medicine,

<sup>5</sup>Department of Dietetics and Nutrition, Bahrain Defense Force Hospital, Riffa, Kingdom of Bahrain;

<sup>6</sup>College of Medicine, King Fahad Medical City, King Saud Bin Abdulaziz University for Health Sciences, Riyadh, Kingdom of Saudi Arabia

**Background:** We aimed to design, implement, and evaluate the feasibility and reliability of a multisource feedback (MSF) system to assess interns in their clerkship year in the Middle Eastern culture, the Kingdom of Bahrain.

**Method:** The study was undertaken in the Bahrain Defense Force Hospital, a military teaching hospital in the Kingdom of Bahrain. A total of 21 interns (who represent the total population of the interns for the given year) were assessed in this study. All of the interns were rotating through our hospital during their year-long clerkship rotation. The study sample consisted of nine males and 12 females. Each participating intern was evaluated by three groups of raters, eight medical intern colleagues, eight senior medical colleagues, and eight coworkers from different departments.

**Results:** A total of 21 interns (nine males and 12 females) were assessed in this study. The total mean response rates were 62.3%. A factor analysis was conducted that found that the data on the questionnaire grouped into three factors that counted for 76.4% of the total variance. These three factors were labeled as professionalism, collaboration, and communication. Reliability analysis indicated that the full instrument scale had high internal consistency (Cronbach's  $\alpha$  0.98). The generalizability coefficients for the surveys were estimated to be 0.78.

**Conclusion:** Based on our results and analysis, we conclude that the MSF tool we used on the interns rotating in their clerkship year within our Middle Eastern culture provides an effective method of evaluation because it offers a reliable, valid, and feasible process.

**Keywords:** MSF system, interns, validity, generalizability

## Introduction

The regular evaluation of medical professionals is an essential step in the process toward improving the quality of medical care provided by any health care institution.<sup>1</sup> Such evaluations point to both the strengths and weaknesses found amongst the health care staff.<sup>2</sup> Thorough evaluation of medical interns is critical as this is the starting point of their professional careers. The feedback received from these evaluations will address the areas of improvement they might need to address with more attention.<sup>3</sup>

Multisource feedback (MSF), also called the 360-degree evaluation, is an evaluation process by which various clinicians fill out surveys to evaluate their medical peers and colleagues. This provides feedback from individuals other than the attending and supervising physicians.<sup>4</sup> This type of assessment uses raters from a variety of groups who interact with trainees. Though this type of evaluation process may seem subjective,<sup>5</sup>

Correspondence: Ahmed Al Ansari  
Department of General Surgery, Bahrain  
Defense Force Hospital, Off Waly Alahed  
Avenue, PO Box 28743, West Riffa,  
Kingdom of Bahrain  
Tel +973 1776 6666  
Email [ahmed.alansari@bdfmedical.org](mailto:ahmed.alansari@bdfmedical.org)

several studies have showed that this type of assessment is reliable and valid provided that enough assessors are used to eliminate the bias factor.<sup>6</sup>

MSF is advocated as a feasible approach to evaluating physicians on their interpersonal relationships, and interactions in particular.<sup>7</sup> MSF is gaining acceptance and credibility as a means of providing doctors with relevant information about their practice, to help them monitor, develop, maintain, and improve their competence. MSF is focused on clinical skills, communication, and collaboration with other health care professionals, professionalism, and patient management.<sup>8</sup> As such, it has been argued that all physician evaluation processes include MSF as a key component.<sup>9</sup>

Although MSF has been applied across many clinical settings using a variety of different medical specialties, this type of assessment has not been used in graduate medical education, for those rotating in the clerkship year, nor has it been used in the Middle Eastern culture.

We used MSF to evaluate the medical interns at Bahrain Defense Force (BDF) Hospital. The goal of this study was to assess the feasibility, reliability, and validity of the MSF process for the clerkship year specifically and in the Middle East generally.

## Methods

### Study settings and participants

This study was conducted in the BDF hospital, a military teaching hospital in the Kingdom of Bahrain. The BDF hospital is a military hospital with 450 beds, 322 physicians and dentists, and 1,072 nurses and practical nurses. Annually, 32,462 patients are seen as inpatients, along with over 347,025 outpatients seen in the clinics. The internship year is a full academic year that undergraduate medical school students must complete after finishing medical school. The internship year program for the study year class consisted of the following rotations: 3 months in Surgery, 3 months in Internal Medicine, 2 months in Pediatrics, 2 months in Obstetrics and Gynecology, 1 month in Emergency Medicine, and 1 month an elective of the student's choice. A total 21 interns (who represented the total population of the interns for the year), who had finished their medical school programs and started their 1-year clerkship rotation in our military teaching hospital were assessed in this study. The study included nine males and 12 females. Sixteen interns had graduated from the Royal College of Surgeons of Ireland (RCSI) in Bahrain, two had graduated from medical schools in Egypt, one from Sudan, one from Saudi Arabia, and one had graduated from Yemen.

## Raters

Research has shown that when physicians choose their own raters, the resulting evaluations are not significantly different from evaluations by raters selected by a third party.<sup>10</sup> Furthermore, studies evaluating how familiar the physician was with his/her rater found that familiarity did not alter the ratings significantly.<sup>11</sup> Despite these findings and due to the cultural differences in the Middle East, we hypothesized that the self-selection of raters may influence the results of the evaluations; therefore, we decided in this study to select raters randomly for each intern. The basic criterion was that the intern must know and have worked with his/her individual raters for a minimum of 2 months. Different groups of raters were selected for this study. Each intern was evaluated by three groups of raters: eight medical intern colleagues, eight senior medical colleagues (such as chief residents and consultants from different departments), and eight coworkers from different departments. An independent administrative team was formed to carry out the responsibility of distributing the evaluation materials and collecting them in sealed envelopes. Each envelope consisted of the evaluation instrument and a guide for the implementation of the MSF and the purpose of the study. As this was the first time that MSF has been implemented in the hospital, it was necessary to clarify several points for the raters, as some of these points may be essential for their accurate rating of interns. We clarified in the guideline form that the results would be used to improve the interns' performance and not for their promotion or selection into the residency training programs. Also, the main purpose of this research was stated as: to assess the feasibility, reliability, and validity of implementing the MSF system in the organization. It was also described that each participating intern would receive feedback from his/her evaluations to point to individual areas that needed improvement. Each evaluator was sent the forms that he/she was required to complete. Two weeks later, an e-mail was sent as a reminder from the administrative team. After another 2 weeks, raters who did not submit their envelopes received a phone call from the administrative team as a second reminder.

## Statistical analysis

A number of statistical analyses were undertaken to address the research questions posed. Response rates and the time required to fill out the questionnaire were used to determine feasibility for each of the respondent groups.<sup>7,12</sup>

For each item on the survey, the percentage of the "unable-to-assess" answers, along with the mean and standard deviation, was computed to determine the viability of the

items and the score profiles. Items for which the “unable-to-assess” answer exceeded 20% on a survey might be in need of revision or deletion according to previously conducted research.<sup>12</sup>

We used exploratory factor analysis to determine which items on each survey were suitable to group together (ie, become a factor or scale). In this study, using individual-physician data as the unit of analysis for the survey, the items were intercorrelated using Pearson product moment correlations. The correlation matrix was then broken down into its main components, and these components were then rotated to the normalized varimax criterion. Items were considered to be part of a factor if their primary loading was on that factor. The number of factors to be extracted was based on the Kaiser rule (ie, eigenvalues >1.0).<sup>13</sup>

The factors or scales established through exploratory factor analysis were used to establish the key domains for improvement (eg, professionalism), whereas the items within each factor provided more precise information about specific behaviors (eg, maintains confidentiality of patients, recognizes boundaries when dealing with other physicians, and shows professional and ethical behavior). Physician improvement could be guided by the scores on factors or items.

This analysis made it possible to determine whether the instrument items were aligned into the appropriate constructs (factors) as intended. Instrument reliability (stability) was assessed. Internal consistency/reliability was examined by calculating the Cronbach's coefficient for each of the scales and for each factor separately. Cronbach's coefficient is widely used to evaluate the overall internal consistency for each instrument as well as for the individual factors within the instruments.<sup>14</sup>

This analysis was followed by a generalizability analysis to determine the generalizability coefficient ( $Ep^2$ ) and to ensure there were ample numbers of questions and evaluators to provide accurate and stable results for each intern on each instrument. Normally, an  $Ep^2=0.70$  suggests data are stable. If the  $Ep^2$  is below 0.70, it suggests that more raters or more items are required to enhance stability.

The G analyses were based on a single-facet, nested design with raters nested within the doctors who were being assessed, using the formula:<sup>15</sup>

$$Ep^2 = \frac{\text{Physician (var comp)}}{\text{Physician (var comp) + Error (var comp)}} \quad (1)$$

where “var comp” is the component variability. Although this type of design does not allow for estimation of the interaction

effect of raters with the doctors they are evaluating, it does allow for determination of the  $Ep^2$  of raters. We further conducted a D study where we estimated the  $Ep^2$  for one to ten raters.

## Ethical approval

The research was approved by the research ethics committee of the BDF hospital. Written consent was obtained from the interns and verbal consent was obtained from raters. The study was conducted from January 2014 to May 2014.

## Instrument

A modified instrument was developed based on the Physician Achievement Review (PAR) instrument, which was used to assess physicians in Alberta.<sup>16</sup> The focus of the instrument is the assessment of professionalism, communication skills, and collaboration. To establish face and content validity, a table of specifications was constructed, and a working group was involved in developing the instrument. Expert opinion was taken into consideration as well. The instrument consisted of 39 items: 15 items to assess professionalism, 13 items to assess communication skills, and eleven items to assess collaboration. The instrument was constructed in a way that can be applied by different groups of people, including interns, senior medical colleagues, consultants, and coworkers.

The items on the instrument had a five-point Likert response scale, where 1= among the worst, 2= bottom half, 3= average, 4= top half, and 5= among the best, with the additional option of “unable-to-assess”. After the committee had developed the questionnaires, they were sent to every physician whose work fit the profile for episodic care, for feedback. The questionnaires were modified upon receipt of that feedback.

## Results

A total of 21 interns, nine males and 12 females, who represented the total number of interns rotating in the BDF hospital for the year 2014, were assessed. A total of 16 interns had graduated from the RCSI in Bahrain, two from a medical school in Egypt, one from Sudan, one from Saudi Arabia, and one from Yemen. Each intern was evaluated by three different groups. Group 1 consisted of the medical intern colleagues, group 2 included the senior medical colleagues (chief residents/consultants), and group 3 comprised coworkers.

The total number of collected forms was 314, including 105 surveys from coworkers, 93 surveys from medical intern colleagues, and 116 surveys from senior medical colleagues (chief residents/consultants). Each intern was evaluated by

different number of forms, ranging between 13 and 19 forms for each intern.

The total mean response rates were 62.3%, and the time needed to complete each questionnaire was 7 minutes, which shows the feasibility of the survey.<sup>7,12</sup> Most of the questions were answered by the respondents. There were only four questions (27, 28, 36, and 38) that exceeded 20% for the response “unable-to-assess” by the raters. Those questions may need to be reviewed, revised, or deleted in future implementation. However, even with the elimination of those four questions, the reliability of our survey remains high (Cronbach's  $\alpha = 0.98$ ).

The whole instrument was found to be suitable for the factor analysis (Kaiser–Myer–Olkin [KMO] = 0.953; Bartlett test significant,  $P < 0.00$ ). The factor analysis showed that the data on the questionnaire could be grouped into three factors that represented 76.4% of the total variance: professionalism, collaboration, and communication (Table 1).

The reliability analysis (Cronbach's  $\alpha$  reliability of internal consistency) indicated that the instrument's full scale had high internal consistency (Cronbach's  $\alpha = 0.98$ ). The reliability measure of the factors (subscales) within the questionnaire also had high internal consistency reliability (Cronbach's  $\alpha \geq 0.91$ ). G study analysis was conducted employing a single-facet, nested design. The  $Ep^2$  was 0.78 for the surveys. Also shown in Table 2 is a D study, where we estimated the  $Ep^2$  for one to ten raters: for one rater,  $Ep^2 = 0.30$ , for eight raters,  $Ep^2 = 0.78$ ; and for ten raters,  $Ep^2 = 0.81$ .

## Discussion

In this study, we evaluated the applicability of questionnaire-based assessments, such as the MSF, of the interns who rotated in our military teaching hospital during their clerkship year. To our knowledge, this was the first study to combine feedback from medical intern colleagues, senior medical colleagues, such as chief residents, consultants, and coworkers, and from medical support staff for the assessment of interns in the clerkship year. In addition, we believe that this was the first study investigating the feasibility, reliability, and validity of the MSF system in the Middle East.

In this study, a set of MSF questionnaires to assess interns (by fellow interns, senior medical colleagues, and coworkers) was developed and evaluated. Also, the aim was to assess the feasibility and reliability of the instruments and to explore the evidence of validity. Interns were assessed on a number of aspects of practice that the regulatory authority and the physicians themselves believed to be important. Although not designed to specifically assess

Accreditation Council for Graduate Medical Education or CanMEDS competencies, the items and factors used allowed us to evaluate some aspects of both sets of competencies, respectively.<sup>17</sup> However, to develop a tool that fully assessed either set of competencies would have required the addition of new items and the retesting of the instrument, as well as its factors.

This type of assessment is feasible in the contextual setting described because this was demonstrated by the adequate response rates. Although this may be partly explained by the fact that this type of assessment is a new form of evaluation for both the BDF hospital and Middle Eastern culture, the received response rates were consistent with the response rates that were found in other studies.<sup>18,19</sup>

Initial evidence for the validity of the instruments has been found, recognizing that establishing validity is a process that cannot be proven from a one-time study. Almost all of the questions were answered by the responding physicians and coworkers. However, there were items on all questionnaires that many of the respondents were unable to assess; these items must be reviewed. Some may be amenable to modification, while others may need to be deleted.

Our exploratory factor analyses showed that items did, in fact, group together into factors in the same way as predicted by the table of specifications. Regulatory authorities tend to focus on improving both professionalism and collaboration, especially amongst young physicians.<sup>20</sup> As such, the factors identified in our results provided a larger picture of the areas of improvement for our physician population as a whole, whereas the individual items used pointed to more specific feedback. Each physician received descriptive data (both means and standard deviations) on the scales and individual items, for him/herself as well as for the group as a whole.

Additional work will be required to examine the validity of the instruments. For example, it would be useful to determine whether physicians who received high scores on the MSF evaluation are also high performers on other assessments that examine performance more objectively. Similar work to identify the criterion validity was examined by other researchers in different settings. Yang et al found a medium correlation between the 360 degree evaluation and the small scale Objective Structured Clinical Examination (OSCE) ( $r = 0.37$ ,  $P < 0.05$ ). This finding supports the criterion validity of the MSF.<sup>21</sup>

There are limitations to this study. This study focused on interns in one hospital in the Kingdom of Bahrain. We cannot currently conclude whether our results would be



**Table I** Descriptive statistics, item analysis, and factor analysis

Q		N	M	SD	% UA	Self (mean)	SD	Factors identified by factor analysis		
								Comm	Colla	Profe
Q1	Maintains confidentiality of patients	267	4.40	0.73	12.1	4.55	0.66			0.78
Q2	Recognizes boundaries when dealing with other physicians	285	4.38	0.76	9.20	4.50	0.79			0.82
Q3	Recognizes boundaries when dealing with other health care professionals	285	4.40	0.75	9.20	4.75	0.62			0.79
Q4	Shows professional and ethical behavior	291	4.46	0.77	7.30	4.77	0.44			0.68
Q5	Is punctual, and performs tasks in a time-appropriate manner	283	4.27	0.88	9.80	4.58	0.67		0.66	
Q6	Is able to handle situations in a professional manner and exhibits self-control, avoiding emotional outbursts in stressful situations	273	4.16	0.84	13.0	4.00	0.71		0.53	
Q7	Respects patient's autonomy and right to be involved in his/her own management	276	4.39	0.71	12.1	4.77	0.44		0.57	
Q8	Is reliable and responsible when performing his duties	288	4.30	0.86	8.28	4.77	0.56		0.72	
Q9	Is honest and handles his/her duties in a dignified manner	284	4.40	0.74	9.55	4.77	0.44		0.68	
Q10	Accepts constructive criticism and develops goals for improvement	265	4.13	0.85	15.6	4.46	0.66		0.65	
Q11	Respects cultural, individual, and role differences including age, sex, race, religion, disability, language, sexual orientation, and socioeconomic status	281	4.51	0.66	10.5	4.84	0.37			0.63
Q12	Follows institutional policies and procedures	285	4.32	0.79	9.23	4.61	0.65		0.73	
Q13	Arrives on time to scheduled appointments and hospital activities	279	4.23	0.88	14.3	4.53	0.66		0.77	
Q14	Manages health care resources efficiently	263	4.27	0.73	15.9	4.33	0.65			0.49
Q15	Leads with respect and fair treatment of colleagues	279	4.37	0.72	11.1	4.61	0.51			0.56
Q16	Communicates efficiently and in a clear, understandable fashion with colleagues within his/her team	285	4.40	0.75	9.23	4.60	0.76	0.64		
Q17	Communicates efficiently and in a clear, understandable, and compassionate way with patients	275	4.39	0.76	12.4	4.53	0.77	0.69		
Q18	Allows the patient to elaborate about his condition	273	4.34	0.76	16.2	4.38	0.76	0.72		
Q19	Communicates efficiently and in a clear, understandable, and compassionate way with patients' families	261	4.28	0.80	16.8	4.33	0.77	0.69		
Q20	Communicates clearly and effectively with other health care workers eg, nurses	281	4.33	0.77	10.5	4.46	0.76	0.59		
Q21	Explains what is being done for the patient during examination or procedures	255	4.23	0.77	18.7	4.38	0.87	0.68		
Q22	Communicates purpose and results of investigations to patients well	251	4.21	0.78	12.0	4.10	0.83	0.73		
Q23	Follows up appropriately and in a timely manner on patients' hospital course	259	4.22	0.83	17.5	4.30	0.75		0.56	
Q24	Communicates management options to patients in a clear, understandable way, taking into account the patients' opinions	264	4.24	0.78	15.9	4.23	0.92	0.70		
Q25	Displays empathy in dealing with patients by eye contact and verbal responses	274	4.30	0.78	12.7	4.46	0.76		0.57	
Q26	Summarizes the information given for the patient in small quantities, with concrete explanations and understandable language	252	4.23	0.79	19.7	4.23	0.72	0.69		
Q27	Maintains calm in emergency situations, in order to communicate information clearly to his/her seniors	235	4.19	0.78	25.1	4.10	0.79	0.60		
Q28	Communicates accurate patient information to physicians from other departments when required to do so	247	3.27	0.80	21.3	4.23	0.83	0.69		
Q29	Manages to work well as part of a health care team	286	4.32	0.81	8.90	4.61	0.65	0.61		
Q30	Facilitates the learning of medical colleagues and coworkers	254	4.29	0.79	19.1	4.39	0.65	0.68		
Q31	Collaborates well with nurses and other health care workers	280	4.32	0.81	8.90	4.53	0.77	0.65		
Q32	Concerned about the safety of patients and coworkers	269	4.51	0.82	14.3	4.69	0.63		0.60	
Q33	Coordinates patient care efficiently	269	4.30	0.77	14.2	4.61	0.66		0.61	
Q34	Collaborates with other health care workers in order to achieve optimal patient care	276	4.23	0.84	12.1	4.69	0.63		0.65	
Q35	Participates in a system of call in order to provide care for patients	273	4.32	0.83	13.0	4.75	0.62		0.65	
Q36	Provides appropriate guidance and help to team members on regular basis	247	4.24	0.83	21.3	4.50	0.67		0.61	
Q37	Takes on extra work, when appropriate, to help the team	269	4.13	0.93	14.3	4.46	0.66		0.69	
Q38	Enables the team to achieve agreements for team process and collaborative completion of assignment	247	4.22	0.81	21.3	4.50	0.67		0.67	
Q39	Participates fully in collaborative process and fulfilled team agreements	256	4.22	0.80	18.4	4.58	0.67		0.68	

**Abbreviations:** Colla, collaboration; Comm, communication skills; M, mean; N, number of respondents; Profe, professionalism; Q, question; SD, standard deviation; Self, self-assessment; UA, unable to assess.

**Table 2** G study. The number of raters required and the generalizability coefficient

Number of raters	Ep <sup>2</sup>
1	0.302
2	0.464
3	0.565
4	0.634
5	0.684
6	0.722
7	0.752
8	0.776
9	0.796
10	0.812

**Note:** Coefficients were based on a D study.

**Abbreviation:** Ep<sup>2</sup>, generalizability coefficient.

applicable to different interns in other countries in the Middle East. Another limitation of the current study was the limited number of participants, using 21 interns in total.

## Conclusion and future research

Determining the validity and the reliability of the MSF system in assessing interns was a challenge in our institution. The MSF instruments for interns provide a feasible way of assessing interns in their clerkship year and of providing guided feedback on numerous competencies and behaviors. Based on the present results, the currently used instruments and procedures have high reliability, validity, and feasibility. The item analyses, reliability, and factor analyses all indicate that the present instruments are working generally very well in assessing interns rotating through the hospital in their clerkship year. Replicating this study in other Middle Eastern settings may provide evidence to support the validity of the MSF process.

## Author contributions

AAA and KAK contributed to the conception and design of the study. MAA, RAA, and DAS worked on the data collection. AAA, AAM, and FMM constructed the blueprint and table of the specifications for communication, collaboration, and professionalism, and constructed the questionnaire. AAA, KAK, and FMM contributed to the data analysis and interpretation of the data. AAA and AAM contributed to the drafting of the manuscript. AAA gave the final approval of the version to be published. All authors contributed toward data analysis, drafting and critically revising the paper, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

## Disclosure

All authors report no conflicts of interest in this work.

## References

1. Lockyer J. Multisource feedback in the assessment of physician competencies. *J Contin Educ Health Prof.* 2003;23(1):4–12.
2. Southgate L, Hays RB, Norcini J, et al. Setting performance standards for medical practice: a theoretical framework. *Med Educ.* 2001;35(5):474–481.
3. Fidler H, Lockyer JM, Toews J, Violato C. Changing physicians' practices: the effect of individual feedback. *Acad Med.* 1999;74(6):702–714.
4. Lockyer JM, Violato C, Fidler H. The assessment of emergency physicians by a regulatory authority. *Acad Emerg Med.* 2006;13(12):1296–1303.
5. Ferguson J, Wakeling J, Bowie P. Factors influencing the effectiveness of multisource feedback in improving the professional practice of medical doctors: a systematic review. *BMC Med Educ.* 2014;14:76.
6. van der Vleuten CP, Schuwirth LW, Scheele F, Driessen EW, Hodges B. The assessment of professional competence: building blocks for theory development. *Best Pract Res Clin Obstet Gynaecol.* 2010;24(6):703–719.
7. Archer JC, Norcini J, Davies HA. Use of SPRAT for peer review of paediatricians in training. *BMJ.* 2005;330(7502):1251–1253.
8. Violato C, Lockyer J, Fidler H. Multisource feedback: a method of assessing surgical practice. *BMJ.* 2003;326(7388):546–548.
9. Violato C, Lockyer JM, Fidler H. Changes in performance: a 5-year longitudinal study of participants in a multi-source feedback programme. *Med Educ.* 2008;42(10):1007–1013.
10. Ramsey PG, Wenrich MD, Carline JD, Inui TS, Larson EB, LoGerfo JP. Use of peer ratings to evaluate physician performance. *JAMA.* 1993;269(13):1655–1660.
11. Lipner RS, Blank LL, Leas BF, Fortna GS. The value of patient and peer ratings in recertification. *Acad Med.* 2002;77(10 Suppl):S64–S66.
12. Violato C, Lockyer JM, Fidler H. Assessment of psychiatrists in practice through multisource feedback. *Can J Psychiatry.* 2008;53(8):525–533.
13. Violato C, Saberton S. Assessing medical radiation technologists in practice: a multi-source feedback system for quality assurance. *Can J Med Radiat Technol.* 2006;37(2):10–17.
14. Lockyer JM, Violato C, Fidler H, Alakija P. The assessment of pathologists/laboratory medicine physicians through a multisource feedback tool. *Arch Pathol Lab Med.* 2009;133(8):1301–1308.
15. Brennan RL. *Generalizability Theory*. New York, NY: Springer-Verlag; 2001.
16. Violato C, Lockyer J. Self and peer assessment of pediatricians, psychiatrists and medicine specialists: Implications for self-directed learning. *Advances in Health Sciences Education.* 2006;11:235–244.
17. Frank JR, Danoff D. The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. *Med Teach.* 2007;29(7):642–647.
18. Wenrich MD, Carline JD, Giles LM, Ramsey PG. Ratings of the performances of practicing internists by hospital-based registered nurses. *Acad Med.* 1993;68(9):680–687.
19. Allerup P, Aspegren K, Ejlersen E, et al. Use of 360-degree assessment of residents in internal medicine in a Danish setting: a feasibility study. *Med Teach.* 2007;29(2–3):166–170.
20. Chapman DM, Hayden S, Sanders AB, et al. Integrating the Accreditation Council for Graduate Medical Education Core competencies into the model of the clinical practice of emergency medicine. *Ann Emerg Med.* 2004;43(6):756–769.
21. Yang YY, Lee FY, Hsu HC, et al. Assessment of first-year post-graduate residents: usefulness of multiple tools. *J Chin Med Assoc.* 2011;74(12):531–538.

**Advances in Medical Education and Practice****Dovepress****Publish your work in this journal**

Advances in Medical Education and Practice is an international, peer-reviewed, open access journal that aims to present and publish research on Medical Education covering medical, dental, nursing and allied health care professional education. The journal covers undergraduate education, postgraduate training and continuing medical education

including emerging trends and innovative models linking education, research, and health care services. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <http://www.dovepress.com/advances-in-medical-education-and-practice-journal>