

Multisource Feedback for Residents: How High Must the Stakes Be?

OLLE TEN CATE, PHD
JOAN SARGEANT, PHD

Miller's pyramid of competence (FIGURE) provides a useful 4-level model for the assessment of learners. The "does" level at the top of the pyramid is the real world. It is the most authentic and valid level above the "knows," "knows how," and "shows how" levels.¹ At the top of the pyramid, above the clouds of the structured educational world of tests and simulations, assessment focuses on what our trainees do in practice. How do we evaluate this, and what information can we provide them to guide further learning and professional development?

Here is where multisource feedback (MSF), or 360° evaluation, has a unique potential to contribute to residents' professional development. MSF, developed in the business world, is a younger member of the family of assessment tools of medical competence. Original research was conducted in the United States in the early 1990s.² Since then MSF has been tested in the evaluation of residents and physicians in a number of countries, including the United Kingdom, Canada, the Netherlands, and Denmark, and has shown potential to become an established and valuable assessment tool.

What Can MSF Add to the Assessment of Clinical Competence?

Clinical competence has been defined as "the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions, values, and reflection in daily practice for the benefit of the individual and community being served." This powerful definition³ delineates the multifaceted nature of competence, as well as its context-dependent nature.⁴ "Doing well" here and now may be different from doing well in other contexts, and doing well in the eyes of peers may differ from doing well in the eyes of patients, managers, or other colleagues. Hence, doing well cannot be captured in a standardized single measure, or evaluated from one perspective or by one person. MSF, or 360° evaluation, is the combined

evaluation of a person by multiple individuals that have different working relationships with the person, through use of questionnaires and a compiled feedback report. For trainees, reviewers generally include peers, supervisors, other health professionals, and patients.

Building a balanced image of a medical trainee, including strengths and weaknesses in an array of competencies, requires a broader view than most traditional methods of assessment. Competence is multifaceted and includes domains that are difficult to capture by traditional methods, among them professionalism, health advocacy, practice-based learning ability, scholarly approach, teamwork ability, and interpersonal skills.⁵⁻⁷ Many of these aspects cannot be evaluated by a single focused observation or by standardized assessment in a simulated environment. Miller's pyramid's highest level, "does," is not an assessment setting. It is like the light in the refrigerator. We want to trust that it works—that is, that the light turns off when we close the door—but we cannot observe that property. We need observers from inside to inform us. MSF has the potential to inform a resident and a program director about observed behavior over a longer period of time under natural circumstances and by "insiders," staff not otherwise responsible for formal judgments about trainees. Studies have shown that MSF can contribute particularly to the development of behavior in more intangible domains of competence, such as communication, timeliness of completing tasks, taking responsibility,⁸ and professionalism.⁹

Developing High-Stakes MSF Instruments—And the Other Side of the Coin

MSF was originally designed to be formative in purpose—to provide feedback to individuals to increase their awareness of their performance, and to guide learning and improvement. Used in this manner, concerns for the degree of rigor of the reliability and validity of the process are not as high as for MSF employed for summative purposes.

During the past decade, a more high-stakes version of MSF has evolved in medical education. Researchers have investigated psychometric and other properties of MSF that would enable its use in high-stakes summative decisions. This has included identifying the domains being assessed, developing items and scales, identifying and preparing appropriate reviewers, making decisions regarding inclusion of narrative comments and providing feedback, and

Th. J. (Olle) ten Cate, PhD, is Professor of Medical Education at the Center for Research and Development of Education, University Medical Center, Utrecht, the Netherlands; and **Joan Sargeant, PhD**, is Professor in the Division of Medical Education, Dalhousie University, Halifax, Nova Scotia, Canada.

Corresponding author: Th. J. ten Cate, PhD, Center for Research and Development of Education, University Medical Center, Utrecht, HB 4.24, PO Box 85500, 3508 GA Utrecht, the Netherlands, +31.88.7557010, t.j.tencate@umcutrecht.nl

DOI: <http://dx.doi.org/10.4300/JGME-D-11-00220.1>

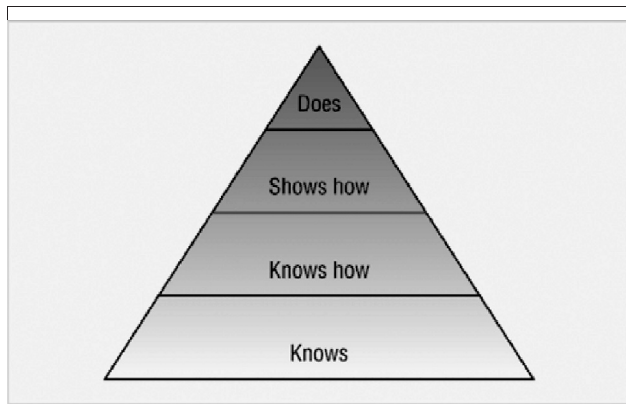


FIGURE | MILLER'S PYRAMID FOR CLINICAL COMPETENCE

From: Norcini J. Work based assessment. *BMJ*. 2003;326:753.

maintaining confidentiality.¹⁰ This is laudable work. However, collecting valid information from multiple sources and feeding this information back in a structured manner that meets the rigorous standards required for high-stakes summative feedback pose a number of challenges. First, a certain number of reviewers are required to provide reliable and anonymous feedback: generally about 8 to 20, depending on the degree of variability in the reviewer group, the number of groups, and the number of questionnaire items. Logistically, there are multiple challenges in managing the recruitment of reviewers, data collection, report generation, and feedback in a confidential and organized manner. If this is done via paper, there is a lot of paper involved that must be collected and organized without mistakes. In addition, there are a substantial number of privacy issues. Personal information for feedback purposes is confidential and must be treated this way. All participants must be informed and understand their role in the procedure. Observers must be trained to be a credible and reliable source of information about residents.

The paper by Richmond et al¹¹ in this issue reports challenges in the implementation of the Assessment of Professional Behaviors, an MSF program of the US National Board of Medical Examiners. Programs withdrew from this pilot study when they realized the extent of what was required to implement the MSF assessments. This is not a new finding. In a surgical residency program, Weigelt et al¹² found that the burden of work to carry out MSF did not match the benefits because average scores did not substantially add to information from existing assessment tools. To use MSF for high-stakes evaluations much preparation is needed, including the training of many evaluators. However, if the emphasis is not on summative decisions and average scores, but rather on formative narrative feedback for growth, a more informal MSF

approach may yield added rich information without a heavy burden of obligations on clinical staff.

A Formative Route for MSF

If the MSF goal is primarily enhancing trainees' self-awareness of their own performance to enable improvement, the emphasis can be placed on the kind of feedback information that best facilitates these activities. Credibility and specificity of feedback are critical ingredients.¹³ Although most, if not all, MSF programs use numeric rating scales, useful for establishing psychometric qualities and making comparisons with others, evidence indicates that scores may be less helpful for feedback. In contrast, narrative comments are invaluable in raising self-awareness and guiding improvement.¹⁴ We believe MSF in medical education has its predominant strength when used for formative feedback and not for summative purposes. Also, facilitation by an engaged supervisor appears to be beneficial in helping the trainee to understand the feedback and make improvements.¹⁵

Formative MSF for Residents—An Example

In 2008 Utrecht University made available an easy-to-use, web-based MSF instrument for residents,¹⁶ targeted to Dutch residency programs (www.multisourcefeedback.nl). The instrument is designed to support the preparation for residents' annual progress interviews with their program director. It is recommended by a Dutch national authority for graduate medical education (the Central Bureau for Education in Healthcare), which also funded its development.

The Utrecht instrument is simple to access and use. Interested program directors registered with the Royal Dutch Medical Society can apply for an account. Next, a personal website is created that provides space to register residents, their e-mail addresses, and the date for closure of the feedback process, usually 1 to 3 months later, when a report will be generated. The program director starts the procedure by listing all residents who must receive MSF. The residents receive an e-mail that requests the e-mail addresses of multiple observers in 3 categories: medical colleagues (6 or more), other health care colleagues (6 or more), and patients (10 or more). The first two groups will receive an e-mail with a link to a confidential MSF subsite that contains a questionnaire with space for narrative comments: "tops" to stress positive observations, and "tips" for improvement. Patients are asked to participate following a clinical encounter, either after hospital or ambulatory care. Residents can observe the response process online and send automatic reminders, but they cannot identify or access individual responses. The procedure stops at the preset closure date; a report is

generated and automatically sent to the resident and program director. All quantitative data are summarized in a small table, categorized according to the CanMEDS framework, and followed by a long list of “tips” and “tops.” Whereas the source of comments remains confidential to the resident, the program director may, for a limited time, identify respondents if desired. The program director and the resident can discuss the report in the resident’s progress interview.

An evaluation of this use of MSF is currently underway and will include assessment of its psychometric properties. Since its launch in November 2008, the service has been used by more than 110 program directors to evaluate more than 700 residents in 38 hospitals and 33 disciplines, with more than 11 000 respondents. Participants receive training only via instructions provided by text on the website and supplemented in some cases by an oral explanation upon request. Informal feedback has been positive. The most common observations are that the procedure is user friendly, nonbureaucratic, and information rich, particularly because of the included “tips” and “tops” comments.

Conclusion

We propose that the primary goal of MSF should be learner improvement. Focusing on the logistic and personnel burdens of MSF for valid high-stakes decisions may distract from the many strengths of this approach for assessing and providing constructive feedback on clinical performance. It would be unfortunate if programs were to decide not to use MSF because of a logistical quagmire, whereas a simple

formative approach is likely to yield invaluable feedback. We should continue to explore this promising method.

References

- 1 Miller GE. The assessment of clinical skills/competence/performance. *Acad Med*. 1990;65(9):563–567.
- 2 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.
- 3 Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA*. 2002;287(2):226–235.
- 4 ten Cate TJO, Snell L, Carraccio C. Medical competence: the interplay between individual ability and the health care environment. *Med Teach*. 2010;32(8):669–675.
- 5 Lockyer J. Multisource feedback in the assessment of physician competencies. *J Contin Educ Health Prof*. 2003;23(1):4–12.
- 6 Frank JR, Danoff D. The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. *Med Teach*. 2007;29(7):642–647.
- 7 Accreditation Council for Graduate Medical Education. Outcome Project. <http://www.acgme.org/outcome/>. Accessed August 20, 2011.
- 8 Brinkman WB, Geraghty SR, Lanphear BP, et al. Effect of multisource feedback on resident communication skills and professionalism: a randomized controlled trial. *Arch Pediatr Adolesc Med*. 2007;161(1):44–49.
- 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 Lockyer JM, Clyman SG. Multisource feedback (360 degree evaluation). In: Holmboe ES, Hawkins RE, eds. *Practical Guide to the Evaluation of Clinical Competence*. 1st ed. Philadelphia, PA: Mosby-Elsevier; 2008:75–85.
- 11 Richmond M, Canavan C, Holtman MC, Katsufraakis PJ. Feasibility of implementing a standardized multisource feedback program in the graduate medical education environment. *J Grad Med Educ*. 2011;3(4):XX–XX.
- 12 Weigelt JA, Brasel KJ, Bragg D, Simpson D. The 360-degree evaluation: increased work with little return? *Curr Surg*. 2004;61(6):616–626.
- 13 Sargeant J, Mann K, Sinclair D, Van der Vleuten C, Metsmakers J. Challenges in multisource feedback: intended and unintended outcomes. *Med Educ*. 2007;41(6):583–591.
- 14 Sargeant J, Eva KW, Armson H, et al. Features of assessment learners use to make informed self-assessments of clinical performance. *Med Educ*. 2011;45(6):636–647.
- 15 Sargeant J, Mcnaughton E, Mercer S, Murphy D, Sullivan P, Bruce DA. Providing feedback: exploring a model (emotion, content, outcomes) for facilitating multisource feedback. *Med Teach*. 2011;33(9):744–749.
- 16 Horsman MA, ten Cate TJ. Guide multisource feedback for the resident [in Dutch]. *Neth J Med Educ*. 2010;29(2):S1–S52.