Canadian Medical Education Journal

Commissioned Scientific Reports

Multisource feedback: an overview of its use and application as a formative assessment

La rétroaction multisource : un aperçu de son application comme outil d'évaluation formative

Jocelyn Lockyer,[1] Joan Sargeant[2]

[1]Department of Community Health Sciences, Cumming School of Medicine, Alberta, Canada; [2]Continuing Professional Development and Medical Education, Faculty of Medicine, Dalhousie University, Nova Scotia, Canada

Correspondence to: Jocelyn Lockyer, Professor Emerita, Department of Community Health Sciences, Cumming School of Medicine, 3330 Hospital Drive NW, Calgary AB. Canada T2N 0L1; e-mail: Lockyer@ucalgary.ca; phone: 403 630 2101

Published ahead of issue: March 28, 2022; published: Aug 26, 2022. CMEJ 2022, 13(4) Available at https://doi.org/10.36834/cmej.73775

Abstract

Multisource feedback (MSF), often termed 360-degree feedback, is a formative performance assessment in which data about an individual's observable workplace behaviors are collected through questionnaires from those interacting with the individual; data are aggregated for anonymity and confidentiality; the aggregated data, along with self-assessment if available, are provided to the individual; and the recipient meets with a trusted individual to review the data and develop an action plan. It is used along the continuum of medical education. This article provides an overview of MSF's utility, its evidence base and cautions.

Résumé

La rétroaction multisource (RMS), ou rétroaction 360 degrés, est une évaluation formative du rendement dans laquelle des informations sur les comportements observables d'un individu dans son lieu de travail sont recueillies par le biais de questionnaires auprès de ceux avec qui il interagit. Après avoir été agrégées pour garantir l'anonymat et la confidentialité, ces données, et l'auto-évaluation s'il y a lieu, sont remises à la personne évaluée. Accompagnée d'une personne de confiance, elle les examinera et élaborera un plan d'action, qui sera utilisé tout au long du continuum de sa formation médicale. Cet article présente un aperçu de l'utilité et des fondements de la RMS, ainsi que quelques mises en garde.

Introduction

This manuscript was based on a White Paper submission to the Medical Council of Canada (MCC). The MCC established an Assessment Innovation Task Force to inform the MCC regarding what it should be assessing, how, when, and why. The request recognized that this would require consideration of what the MCC was uniquely positioned to offer to Canadian medical care, how its efforts would ideally fit within the Canadian medical education landscape, and what was missing in its current activity. To that end, the Task Force approached a variety of stakeholders in search of white papers on a variety of assessment topics that could be used to inform the complexity underlying how licensing processes fit (or could fit) into modern healthcare and healthcare education. Multisource Feedback (MSF) was one of the papers solicited. As will be shown, MSF has been used occasionally in Undergraduate Medical Education (UGME), more frequently in Post Graduate Medical Education (PGME) and most commonly for physicians in practice where it is used as part of continuing professional development (CPD) as an approach to facilitate continuing competence or certification.

What is MSF?

Multisource Feedback (MSF) is a formative performance assessment with a four-stage process whereby (1) data about an individual's observable workplace behaviors are collected through questionnaires from those interacting with the individual; (2) data are aggregated for anonymity and confidentiality; (3) the aggregated data, along with

self-assessment if available, are provided to the individual in a report; and (4) the recipient meets with a trusted individual to review the data and develop an action plan.[1,2] The goal of MSF is to provide performance feedback in a manner which enables the recipient to reflect upon it and use for ongoing development.

What is the evidence base for MSF?

MSF provides formative assessment. Assessments of reliability demonstrate it doesn't achieve a level appropriate to make summative decisions based on the scores and narrative from this assessment tool. As with all assessment tools used in medicine, evidence for validity and reliability are temporal and contextual. This means that as new MSF tools are developed or adapted for different contexts, the psychometrics for each should be examined using Kane's evidence for validity.[3,4] Nonetheless, MSF can be part of the data used for decision-making[5] and may trigger other assessments or monitoring.

There are several reviews demonstrating evidence for validity for use of the data (scores/ratings) with practicing physicians;[6,7] however Stevens' et al.[7] caution from their review of eight systematic reviews:

> *This review has demonstrated that the evidence base supporting the statistical and psychometric properties of MSF is sufficient. The internal structural validity of MSF has been repeatedly tested, with feedback instruments often demonstrated to be statistically reliable methods of performance assessment. What is also apparent, although the size of the evidence base is smaller, is that the results of MSF assessments often correlate highly with other WBA [Workplace Based Assessment] methods. Finally, sufficient evidence also exists to demonstrate that MSF is a feasible method of assessing medical performance in terms of cost, time, and response rates. We have also shown however that validity evidence is currently lacking in three areas: (1) how best to ensure that MSF tools measure what they intend to measure (content validity); (2) how best to maximize positive impact on practice (consequential validity); and (3) how to ensure that the process of assessment delivery is rigorous, robust, and free from bias (response process validity).*[7p267]

Despite Stevens' et al.[7] caution, it must be noted that their review was a review of reviews and as such doesn't capture all facets of validity evidence. As noted by Lockyer et al.,[8] content validity, consequential validity and response process validity have been examined for two UK and one Canadian instrument drawing on criteria for good assessment.[9] Recent studies have examined consequential validity and demonstrated the importance of providing quantitative and qualitative data from MSF instruments, along with a facilitated discussion to ensure outcomes.[10-13] While there are other approaches, including providing a peer with a script for discussion[10] , the R2C2 model, which involves a facilitator building a relationship, exploring reactions to the feedback, exploring the content of the feedback and coaching for change,[14] has demonstrated its effectiveness as an approach to facilitated feedback.[11,12]

MSF may be combined with other tools.[5,15] As noted by van der Meulen,[15] while MSF tools can provide robust performance information, no instrument is able to capture the whole complex construct of physicians' professional performance, particularly when decisions are required for promotion, remediation, or suspension/termination.

There are publications examining aspects of validity in PGME[5] as well as with medical students.[16-20]

For these reasons, ongoing evaluation of MSF initiatives are critical to ensure that the MSF tools being used can demonstrate evidence of validity and reliability for the specific context in which they are being used and the data and approaches to feedback discussions and follow-up are optimized to ensure utility for the end-user. Despite the fact that MSF is a formative assessment, attention to the four stages is important. Each stage is important with individual characteristics and requirements. Hence, evidence is required to support each of the four stages of MSF—namely data collection, data aggregation, data provision to recipient and a discussion with a trusted individual to address data and develop an action plan. A lack of rigour in any stage can result in failure of the process and impact on participants acceptance and use of the data.

Why is MSF necessary?

MSF has been used to assess all CanMEDS Roles, either implicitly or explicitly,[21,22] those more commonly and perhaps more appropriately assessed are Communicator, Professional, and Collaborator. There are practical reasons for this. These Roles can be readily observed by other health professionals and patients and are not readily assessed using other approaches. Other tools are available for assessing other Roles. For example, performance in the Medical Expert Role is more readily and accurately assessed using through multiple choice questions, direct observation of procedures, OSCE's and performance audit.[23] There are specialized instruments for the

assessment of Leader.[13,24,25] Health advocate and Scholar Roles are less easily observed and rated using an MSF format. The inclusion of items that can't be observed has the potential to undermine the overall value for the recipient and can lead to 'guessing' by raters. When those assessed receive feedback on performance which they don't believe their raters can observe, or data they don't believe their raters have access to, they question its validity and report itself.

Communicator, Professional, and Collaborator Roles are critical to patient safety.[26-29] It is known that poor communication with patients, colleagues and coworkers and unprofessional behavior negatively influence patient care and outcomes. For example, inadequate handover communication results in errors that impact patient safety. Similarly, unclear or inappropriate communication in any way compromises patient care (e.g., unclear discharge plans, communication with the health care team, and explanation and planning with patients). Likewise, unprofessional behavior that prevents open and respectful communication by all involved in a patient's care can compromise safe care. Conversely, good communication and collaborative skills handled in a professional way will enhance patient safety.

As noted earlier, MSF's primary use to date appears to be for practicing physicians. The goal is to guide their ongoing development and creation of learning plans through reflection upon and comparison of feedback data received from various sources.[30] In Canada, the Medical Council of Canada has transformed the former Physician Achievement Review Program developed by the College of Physicians and Surgeons of Alberta into the Medical Council of Canada's MSF program, MCC 360.[31] The new instruments assess and provide feedback on Professional, Collaborator, and Communicator Roles. It is currently being used by medical regulatory authorities as part of continuing competence programs on a volunteer or mandatory basis in some provinces (British Columbia, Alberta, and Manitoba), The Practice Enhancement Program for Saskatchewan Physicians (PEPSask) as well as by hospitals (The Ottawa Hospital and North York General Hospital).[31] There are other MSF programs available to physicians, through their health systems or privately, as well. For example, PULSE 360 is used to assess leadership skills across the USA and Canada.[13,32] Internationally, MSF is also being used in PGME (residency) programs.[5,35-39] In Canada, for the most part, the initiatives appear to be localized, being developed within programs for that program and/or that University.[35-38]

While MSF has been used less frequently in UGME, two important initiatives have been reported. There is a broadly based instrument being used in Germany that assesses several facets of competence including responsibility, teamwork, empathy, active listening to patients and other competencies that has good evidence of validity.[19,20] In addition, the MCC is currently engaged in a UGME pilot project involving four Canadian universities in a study with up to 25 clinical clerks per school. The rater groups were modified from the MCC 360 for practicing physicians to include self, patient/caregivers, and a mix of supervisor, resident, and peers. The MCC UGME surveys focus on Communicator, Collaborator, and Professional Roles and include three types of questions for raters: frequency (never to always and unable to assess), yes/no options and open-ended questions. Other published applications with medical students include assessments of interprofessional competencies,[16] communication skills[17] and professionalism.[18]

What are the opportunities and cautions with MSF?

MSF provides a unique opportunity to provide feedback related to the Communicator, Professional and Collaborator Roles. Due to the interpersonal nature of these roles, assessing and providing feedback on specific behaviours included in each is often challenging. Direct observation is most frequently used, yet the idiosyncratic nature of direct observation and infrequent reporting, can lead to gaps in accurate assessment. Nonetheless, Regulatory Authorities and health systems continue to identify Communication, Professionalism, and Collaboration as areas receiving larger proportions of formal patient complaints than other roles[40-42] suggesting a continuing need to adopt MSF as an approach to continuously monitor and improve physician skills in these areas. Facilitated discussions provide physicians and medical students with an opportunity to gain a further understanding and interpretation of their data reports and use the data to develop an action plan.[11-13,43]

However, there are a number of cautions to consider when deciding to include MSF amongst the tools for assessment and feedback.

MSF is expensive in terms of the opportunity costs associated with other methods that might be used as well

as the direct and indirect costs associated with implementing a new program. Leaders, raters, and those assessed need to be assured that there will be a potential gain in terms of the information provided to inform decision making.

There are real costs as well. MSF requires a stable infrastructure to sustain. Systems can be developed on an ad hoc or pilot basis. In the longer term, implementation requires leadership to obtain full support from people in the organization to sustaining the program. It also requires participant commitment to an orientation of those being assessed and those doing the assessment to ensure all parties have consistent ideas related to the purpose, rationale, questions, scales, and use of the data. Without a stable infrastructure and commitment, it can be difficult to sustain rigorous MSF initiatives[44]

There are human resources issues to consider as well in terms of the numbers of assessors and the availability of trained individuals for de-briefing sessions. In assessing whether an instrument produces reliable results (i.e., reliability coefficient G > 0.70), the number of items on the questionnaires and numbers of assessors are analyzed. Research across a number of studies and systematic reviews suggests that this can be attained with 5 to 11 physician assessors, 10 to 20 non physician assessors, and 25 to 50 patient assessors.[6,45-47] Too many raters adds workload and runs the risk that the raters haven't observed the physician to do an accurate assessment of performance. Too few raters may impact on confidentiality and anonymity of results.

Facilitated discussions are critical to guide the physician or learner in the co-development of the learning change plan and its follow-up.[11-13] These studies and other publications[1,2,30,43,48,49] have demonstrated the importance of discussion and follow-up and are informing best practices for MSF. The de-briefer or facilitator requires training and may be someone appointed for this purpose for medical students and residents so that they are at arms-length from other evaluation and assessment processes or it may be a preceptor or supervisor. For licensed physicians, it may be a division/department head for licensed physicians, or in the case of a medical regulator, it may be another physician who has received training for this role.

Another caution is that MSF can be 'gamed.' As Alofs et al.[34] noted, if learners and raters perceive the instrument will be used for decision making and not for development, raters selected for MSF may deliberately adapt their scoring or comments compromising the value of the feedback as a formative tool for growth and development. Preventing such bias requires that MSF development and implementation ensure attention is paid to communication throughout the development and implementation of data collection, reporting and de-briefings.

Conclusion

MSF has the potential to offer medical students, residents, and licensed physicians invaluable information about their workplace performance, which is then available for them to use for ongoing development. The goal is continued learning and improvement through application of the feedback. In order to achieve this goal though, particular attention needs to be paid when developing and implementing MSF to ensure (1) high quality instruments that are psychometrically robust, provide evidence for validity and will provide data to guide recipient behavior are created, (2) data are aggregated and anonymized to ensure confidentiality of respondents, (3) the individual receives their data in an easy-to-understand format and (4) the individual is engaged in a facilitated, reflective discussion of their data and guided in using them to develop an action plan, ideally with later follow-up to confirm progress.

Conflicts of Interest: The authors do not have any conflicts of interest.
Funding: None

References

1. Lockyer JM, Sargeant J, *Introduction to MCC 360: a multi-source feedback initiative*. Prepared for Medical Council of Canada, Ottawa, Canada, Nov 17, 2017. https://mcc.ca/media/MCC-360-MSF-Guidelines.pdf [Accessed Oct 12, 2021]
2. Lockyer J, Multi Source Feedback, in ES Holmboe, RE Hawkins, SJ Durning (eds), *A practical guide to the assessment of clinical competence, 2nd Ed*. Elsevier. 2018. Paperback ISBN: 9780323447348; eBook ISBN: 9780323448949. https://www.elsevier.com/books/practical-guide-to-the-evaluation-of-clinical-competence/holmboe/978-0-323-44734-8
3. Kane MT. Validating the Interpretations and Uses of Test Scores. *J Educ Meas*. 2013; 50: 1-73. https://doi.org/10.1111/jedm.12000
4. Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ*. 2015;49(6):560-75. https://doi.org/10.1111/medu.12678.
5. Moonen-van Loon JM, Overeem K, Donkers HH, van der Vleuten CP, Driessen EW. Composite reliability of a workplace-

based assessment toolbox for postgraduate medical education. *Adv Health Sci Educ Theory Pract*. 2013;18(5):1087-102. https://doi.org/10.1007/s10459-013-9450-z.

6. Donnon T, Al Ansari A, Al Alawi S, Violato C. The reliability, validity, and feasibility of multisource feedback physician assessment: a systematic review. *Acad Med*. 2014;89(3):511-6. https://doi.org/10.1097/ACM.0000000000000147.
7. Stevens S, Read J, Baines R, Chatterjee A, Archer J. Validation of multisource feedback in assessing medical performance: a systematic review. *J Contin Educ Health Prof*. 2018;38(4):262-268. https://doi.org/10.1097/CEH.0000000000000219.
8. Lockyer J. Multisource feedback: can it meet criteria for good assessment? *J Contin Educ Health Prof*. 2013;33(2):89-98. https://doi.org/10.1002/chp.21171.
9. Norcini J, Anderson MB, Bollela V, et al. 2018 Consensus framework for good assessment. *Med Teach.* 2018;40(11):1102-1109. https://doi.org/10.1080/0142159X.2018.1500016.
10. Francois J, Sisler J, Mowat S. Peer-assisted debriefing of multisource feedback: an exploratory qualitative study. *BMC Med Educ.* 2018;18(1):36. https://doi.org/10.1186/s12909-018-1137-y.
11. Pooley M, Pizzuti C, Daly M. Optimizing multisource feedback implementation for Australasian physicians. *J Contin Educ Health Prof.* 2019;39(4):228-235. https://doi.org/10.1097/CEH.0000000000000267.
12. Arabsky S, Castro N, Murray M, Bisca I, Eva KW. The influence of relationship-centered coaching on physician perceptions of peer review in the context of mandated regulatory practices. *Acad Med*. 2020; 95 (11S Association of American Medical Colleges Learn Serve Lead: Proceedings of the 59th Annual Research in Medical Education Presentations):S14-S19. https://doi.org/10.1097/ACM.0000000000003642.
13. Hu J, Lee R, Mullin S, Schwaitzberg S, Harmon L, Gregory P, Elkin PL. How physicians change: multisource feedback driven intervention improves physician leadership and teamwork. *Surgery*. 2020;168(4):714-723. https://doi.org/10.1016/j.surg.2020.06.008.
14. Sargeant J, Lockyer J, Mann K, et al. facilitated reflective performance feedback: developing an evidence- and theory-based model that builds relationship, explores reactions and content, and coaches for performance change (R2C2). *Acad Med*. 2015; 90(12):1698-706. Https://doi.org/10.1097/ACM.0000000000000809.
15. van der Meulen MW, Boerebach BC, Smirnova A et al. Validation of the INCEPT: a multisource feedback tool for capturing different perspectives on physicians' professional performance. *J Contin Educ Health Prof*. 2017;37(1):9-18. https://doi.org/10.1097/CEH.0000000000000143.
16. House JB, Franko LR, Haque F, Cranfod JA, Santen SA, Variation in assessment of first-year medical students' interprofessional competencies by rater profession, *J. Interprof. Educ. Pract*. 2021 (24); 100424. ISSN 2405-4526 https://doi.org/10.1016/j.xjep.2021.100424
17. Lai MMY, Roberts N, Mohebbi M, Martin J. A randomised controlled trial of feedback to improve patient satisfaction and consultation skills in medical students. *BMC Med Educ.* 2020;20(1):277. https://doi.org/10.1186/s12909-020-02171-9.
18. Emke AR, Cheng S, Chen L, Tian D, Dufault C. A Novel Approach to Assessing Professionalism in Preclinical Medical Students Using Multisource Feedback Through Paired Self- and Peer Evaluations. Teach Learn Med. 2017;29(4):402-410. https://doi.org/10.1080/10401334.2017.1306446. Epub 2017 May 12. PMID: 28498003.
19. Prediger S, Fürstenberg S, Berberat PO, Kadmon M, Harendza S. Interprofessional assessment of medical students' competences with an instrument suitable for physicians and nurses. BMC Med Educ. 2019;19(1):46. Https://doi.org/10.1186/s12909-019-1473-6. PMID: 30728006; PMCID: PMC6364398.
20. Prediger S, Schick K, Fincke F, Fürstenberg S, Oubaid V, Kadmon M, Berberat PO, Harendza S. Validation of a competence-based assessment of medical students' performance in the physician's role. BMC Med Educ. 2020;20(1):6. Https://doi.org/10.1186/s12909-019-1919-x. PMID: 31910843; PMCID: PMC6947905.
21. Lockyer JM, Violato C, Fidler H. The assessment of emergency physicians by a regulatory authority. Acad Emerg Med. 2006;13(12):1296-303. Https://doi.org/10.1197/j.aem.2006.07.030. Epub 2006 Nov 10. PMID: 17099191.
22. 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):1301-8. Https://doi.org/10.5858/133.8.1301.
23. Holmboe ES, Durning SJ, Hawkins RE, Practical guide to the evaluation of clinical competence 2nd ed, Philadelphia: Elsevier 2018. https://www.elsevier.com/books/practical-guide-to-the-evaluation-of-clinical-competence/holmboe/978-0-323-44734-8
24. Gregory PJ, Robbins B, Schwaitzberg SD, Harmon L. Leadership development in a professional medical society using 360-degree survey feedback to assess emotional intelligence. Surg Endosc. 2017;31(9):3565-3573. https://doi.org/10.1007/s00464-016-5386-8.
25. Rourke J, Bornstein S, Vardy C, Speed D, White T, Corbett P. Evaluation of and Feedback for Academic Medicine Leaders: Developing and Implementing the Memorial Method. *Acad Med*. 2017;92(11):1590-1594. https://doi.org/10.1097/ACM.0000000000001722.
26. Desmedt M, Ulenaers D, Grosemans J, Hellings J, Bergs J. Clinical handover and handoff in healthcare: a systematic review of systematic reviews. *Int J Qual Health Care.* 2021;33(1):mzaa170. https://doi.org/10.1093/intqhc/mzaa170
27. Bonds RL. SBAR Tool Implementation to advance communication, teamwork, and the perception of patient safety culture. *Creative nursing*, 2018; 24(2), 116–123. https://doi-org.ezproxy.lib.ucalgary.ca/10.1891/1078-4535.24.2.116
28. Etherington C, Wu M, Cheng-Boivin O, Larrigan S, Boet S. Interprofessional communication in the operating room: a narrative review to advance research and practice. Communication interprofessionnelle en salle d'opération: un compte rendu narratif pour faire avancer la recherche et la pratique. *Can J Anaesth*. 2019;66(10):1251-1260. https://doi.org/10.1007/s12630-019-01413-9

29. Chladek MS, Doughty C, Patel B et al. The standardisation of handoffs in a large academic paediatric emergency department using I-PASS. *BMJ Open Qual*. 2021;10(3):e001254. https://doi.org/10.1136/bmjoq-2020-001254.
30. Narayanan A, Farmer EA, Greco MJ. Multisource feedback as part of the medical board of Australia's professional performance framework: outcomes from a preliminary study. *BMC Med Educ*. 2018;18(1):323. https://doi.org/10.1186/s12909-018-1432-7. Erratum in: BMC Med Educ. 2019 Mar 4;19(1):73.
31. Medical Council of Canada, *MCC 360,* https://mcc.ca/assessments/mcc360/ [Accessed Feb 15, 2022].
32. 360 PULSE Program: quality at work https://pulseprogram.com. [Accessed Feb 15, 2022].
33. Saedon H, Salleh S, Balakrishnan A, Imray CH, Saedon M. The role of feedback in improving the effectiveness of workplace based assessments: a systematic review. *BMC Med Educ*. 2012; 12:25. https://doi.org/10.1186/1472-6920-12-25.
34. Alofs L, Huiskes J, Heineman MJ, et al. User reception of a simple online multisource feedback tool for residents. *Perspect Med Educ*. 2015 (2):57-65. https://doi.org/10.1007/s40037-015-0173-0.
35. Yama BA, Hodgins M, Boydell K, Schwartz SB. A qualitative exploration: questioning multisource feedback in residency education. *BMC Med Educ*. 2018;18(1):170. https://doi.org/10.1186/s12909-018-1270-7.
36. G, Bandiera G. Multisource feedback and self-assessment of the communicator, collaborator, and professional CanMEDS roles for diagnostic radiology residents. *Can Assoc Radiol J.* 2014;65(4):379-84. https://doi.org/10.1016/j.carj.2014.04.003.
37. Beaudoin PL, Labbé M, Fanous A, et al. Teaching communication skills to OTL-HNS residents: multisource feedback and simulated scenarios. *J Otolaryngol Head Neck Surg*. 2019;48(1):8. https://doi.org/10.1186/s40463-019-0329-8.
38. Castonguay V, Lavoie P, Karazivan P, Morris J, Gagnon R. Perceptions of emergency medicine residents of multisource feedback: different, relevant, and useful information. *Ann Emerg Med*. 2019;74(5):660-669. https://doi.org/10.1016/j.annemergmed.2019.05.019.
39. Buis CAM, Eckenhausen MAW, Ten Cate O. Processing multisource feedback during residency under the guidance of a non-medical coach. *Int J Med Educ*. 2018; 23;9:48-54. https://doi.org/10.5116/ijme.5a7f.169d.
40. Wenghofer EF, Campbell C, Marlow B, Kam SM, Carter L, McCauley W. The effect of continuing professional development on public complaints: a case-control study. *Med Educ*. 2015;49(3):264-75. https://doi.org/10.1111/medu.12633.
41. Chaulk D, Krueger C, Stang AS. A retrospective review of physician-related patient complaints from a tertiary pediatric hospital. *Pediatr Qual Saf*. 2019 13;4(1):e136. https://doi.org/10.1097/pq9.0000000000000136.
42. Saha R, Kabanovski A, Klejman S, Margolin E, Buys YM. Patients' complaints involving ophthalmologists in the province of Ontario, Canada: a 5-year review. *Can J Ophthalmol*. 2020;55(3 Suppl 1):22-26. https://doi.org/10.1016/j.jcjo.2019.08.007.
43. Brennan N, Bryce M, Pearson M, Wong G, Cooper C, Archer J. Towards an understanding of how appraisal of doctors produces its effects: a realist review. *Med Educ*. 2017 Oct;51(10):1002-1013. Https://doi.org/10.1111/medu.13348.
44. Lockyer J, Sargeant J, *Implementing multisource feedback*, in D Boud and E Molloy (Eds), Feedback in Higher and Professional Education, London, Routledge, 2013. ISBN: 9781135107468. https://www.routledge.com/Feedback-in-Higher-and-Professional-Education-Understanding-it-and-doing/Boud-Molloy/p/book/9780415692298
45. Moonen–van Loon JM, Overeem K, Govaerts MJ, Verhoeven BH, van der Vleuten CP, Driessen EW. The reliability of multisource feedback in competency-based assessment programs. *Acad Med*, 2015; 90 (8), 1093-1099. Https://doi.org/10.1097/ACM.0000000000000763.
46. Al Khalifa K, Al Ansari A, Violato C, Donnon T. Multisource feedback to assess surgical practice: a systematic review. *J surg educ*, 2013; 70(4), 475–486. https://doi-org.ezproxy.lib.ucalgary.ca/10.1016/j.jsurg.2013.02.002
47. Andrews JJ, Violato C, Al Ansari A, Donnon T, Pugliese G. *Assessing psychologists in practice: Lessons from the health professions using multisource feedback*. Professional psychology: research and practice. 2013;44(4):193.
48. Sargeant JM, Holmboe ES, *Feedback and coaching in clinical teaching and learning.* ES Holmboe, RE Hawkins, SJ Durning (eds), A practical guide to the assessment of clinical competence, 2nd Ed. Elsevier. 2018. Paperback ISBN: 9780323447348; eBook ISBN: 9780323448949. https://www.elsevier.com/books/practical-guide-to-the-evaluation-of-clinical-competence/holmboe/978-0-323-44734-8
49. Ashworth N, de Champlain AF, Kain N, A review of multi-source feedback focusing on psychometrics, pitfalls and some possible solutions, *Springer Nature Social Sciences*, 2021 I:24. https://link.springer.com/article/10.1007/s43545-020-00033-1