Doing well, getting better

Citation for published version (APA):

Bindels, E. (2021). Doing well, getting better: facilitating physicians' reflection on their professional performance. [Doctoral Thesis, Maastricht University]. Ipskamp Printing BV.
<https://doi.org/10.26481/dis.20210222eb>

Document status and date:

Published: 01/01/2021

DOI:

[10.26481/dis.20210222eb](https://doi.org/10.26481/dis.20210222eb)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

The background of the entire cover is a stylized illustration of a spiral staircase. The staircase is depicted in a light teal or green color, winding upwards. A silhouette of a person is shown walking on the stairs, positioned towards the right side of the frame. The overall aesthetic is clean and modern, with a focus on geometric shapes and a limited color palette.

DOING WELL GETTING BETTER

*Facilitating physicians' reflection
on their professional performance*

ELISA BINDELS

DOING WELL GETTING BETTER

*Facilitating physicians' reflection
on their professional performance*

ELISA BINDELS

ISBN: 978-94-6421-205-1

The research reported here was carried out at

in the School of Health Professions Education

and

at the Amsterdam University Medical Centers, Amsterdam, the Netherlands

© Elisa Bindels, 2021 Amsterdam

The copyright of articles that have been published has been transferred to the respective journals.

Layout and cover design: Rachel van Esschoten – www.divingduckdesign.nl

Printed by: Ipskamp Printing, Enschede - www.ipskampprinting.nl

DOING WELL, GETTING BETTER

*Facilitating physicians' reflection
On their professional performance*

DISSERTATION

to obtain the degree of Doctor at Maastricht University,
on the authority of the Rector Magnificus,
Prof. dr. Rianne M. Letschert
in accordance with the decision of the Board of Deans,
to be defended in public
on Monday, 22nd of February 2021, at 16:00 hours

by

ELISA BINDELS

Supervisors

Prof. dr. S. Heeneman

Prof. dr. M.J.M.H. Lombarts, *University of Amsterdam*

Prof. dr. A.J.J.A. Scherpbier

Assessment Committee

Prof. dr. W.N.K.A. van Mook (chair)

Prof. dr. E.W.M.T. ter Braak, *Utrecht University*

Prof. dr. P.L.P. Brand, *University of Groningen*

Dr. J.M. Frambach

Prof. dr. M.G.A. Oude Egbrink

TABLE OF CONTENTS

Chapter 1

General introduction	7
----------------------	---

Chapter 2

Reflection revisited: how physicians conceptualize and experience reflection in professional practice – a qualitative study	19
---	----

Chapter 3

Designing a system for performance appraisal: balancing medical specialists' accountability and professional development	41
--	----

Chapter 4

Sharing reflections on multisource feedback in a peer group setting: (how) does it stimulate physicians' professional performance and development?	65
--	----

Chapter 5

A new multisource feedback tool for evaluating the performance of specialty-specific physician groups: validity of the Group Monitor instrument	87
---	----

Chapter 6

General discussion	119
--------------------	-----

Appendices

Summary in English	136
Summary in Dutch	140
Valorization	144
Dissemination	150
Acknowledgements	152
About the author	155
SHE dissertation series	156

1

GENERAL INTRODUCTION

“Reflecting on action and acting on reflection”, as McKay notes¹, “is a professional imperative”. Indeed, reflection is considered to be a crucial component of physicians’ professional performance, i.e. what physicians are seen to do in practice.^{2,3} Most definitions of what it means to be a professional also include statements about reflection, lifelong learning and professional development. In general, reflection is understood as the process of learning through and from experience towards gaining new insights of self and/or practice.⁴⁻⁶ This often involves examining assumptions of everyday practice. It also tends to involve the individual practitioner in being self-aware and critically evaluating his/her own responses to practice situations. The point is to recapture practice experiences and mull them over critically in order to gain new understandings and so improve future practice. This is understood as part of the process of lifelong learning.⁷ Beyond these broad areas of agreement, however, contention and difficulty reign. There is debate about the extent to which physicians should focus on themselves as individuals rather than the larger social context. There are questions about how, when, where and why reflection should take place.^{8,9} For busy professionals short on time, reflection runs the risk to be applied in bland, mechanical, unthinking ways.¹⁰ Within this tangle of understandings and challenges, exactly how to facilitate reflection effectively to support physicians in their lifelong learning and professional development has become something of a conundrum. The ways in which this facilitation process currently takes shape for hospital-based physicians is the topic of this thesis.

The purpose of this introductory chapter is to set the stage for this thesis and its topic. First, I will discuss physicians’ professional performance in relation to the privilege of self-regulation of the medical profession. Second, I will elaborate on shifting notions of physicians’ continued learning and development in relation to the changing demands of modern healthcare. I will discuss how the increasing complexity of healthcare and associated calls for accountability require physicians to demonstrate that they are developing professionally and that their activities improve their practice. Third, I will formulate the main argument of this thesis, in which I will argue that reflection can only be a ‘meaningful ethic’ if it is not just approached in terms of what physicians are *supposed* to do, but as something that physicians *want* to do and feel they are *able* to do. Thereafter, I will formulate the main research question and continue to provide information about the requirements for professional re-registration of physicians in the Netherlands, which form the context of the four studies included in this thesis.

SELF-REGULATION OF THE MEDICAL PROFESSION

Society has granted physicians the privilege of self-regulation: the right to define, apply and enforce their own rules.^{11,12} This privilege boils down to the profession defining and monitoring its own quality standards, educating and training future peers, and regulating entrance to and removal from the profession.^{2,13} A leading role for physicians is crucial in order to protect the patient against too much market or government influence. Since on entering the medical profession physicians take an oath to always put patients' interests first, self-regulation is essentially the best guarantee for high-quality care. The altruistic manner of this oath implies a level of self-improvement that submits physicians to pursuing excellence, practicing compassionately, and demonstrating accountability for their performance.² From this felt responsibility, many quality improvement activities and systems have been developed, implemented and institutionalized, often later also legally established, such as guideline development, continuing medical education activities, peer visits and quality assessment.^{2,14} Clearly, quality arises not only at the hands of physicians, but is the result of the collective efforts and influences of many stakeholders involved in healthcare, such as healthcare organizations and governments as well as academic and quality management experts.^{2,13}

The privilege of self-regulation cannot be taken for granted - in fact, the medical profession continuously negotiates with society on how to meet existing expectations and concludes a so-called 'social contract' with society on how to make it happen. This means that the profession must respond whenever serious failings are reported, and the response must be effective and hence visible in individual and collective professional conduct.^{2,12} While self-regulation is a long-standing tenet of the medical profession, in recent years it has increasingly been a topic of debate both professionally and publicly.¹³ Reports about persistent gaps in healthcare quality and safety, increasing demands for transparency and involvement of many stakeholders in regulating patient care have changed and indeed put under pressure the self-regulating powers of physicians as a group.^{13,15} In response to this debate about self-regulation, the medical profession has put in place (additional) programs for maintenance of certification (USA),¹⁶ revalidation (Canada and UK),^{17,18} and re-registration (the Netherlands).¹⁹ These programs are intended to ensure that physicians maintain and enhance their competence and provide safe and high-quality care for the public. Participating in lifelong learning and professional development is a crucial component of these programs.

PHYSICIANS' CONTINUED LEARNING IN MODERN HEALTHCARE

Historically, physicians' continued learning was viewed rather narrowly as continuing medical education (CME). Learning was regarded as an "add-on" to daily practice, for which physicians needed to leave their workplace, for example, to attend a conference or lecture. This model of CME includes a number of educational strategies designed to keep physicians updated regarding new diagnostic approaches, the management of clinical conditions and the development of new technologies that are wholly separate from the clinical environment and healthcare system.²⁰ CME has been criticized for consistently limiting evaluation simply to satisfaction with the learning activities rather than any rigorous measurement of gains in knowledge or technical skills, improved performance and, more importantly, improved patient outcomes.^{21,22} Over the past decades, there has been an increased use of the concept of continuing professional development (CPD) to promote a new model of learning that engages physicians in study across a broad range of competencies beyond medical knowledge, such as communication, collaboration, leadership and professionalism.²⁰ Such competencies are considered to be more relevant and realistic in view of the range of skills required for working in today's complex healthcare systems and delivering high-quality care.¹⁵ The rationale for adoption of competencies has equally been based on the need to demonstrate greater accountability to the public.²³

A critical component of CPD is physicians' motivation and ability to self-direct their learning. Self-directed learning is a central concept in adult learning that has a long history prior to its adoption as one of the central elements of CPD.²⁴ Educational psychologists view self-directed learning as 'a complex cycle comprising intermingling elements such as psychological characteristics (e.g. self-efficacy, motivation, beliefs and learning style), personal choices (effort expended, learning strategies), judgments (self-assessment, attributions) and personal actions (e.g. goal setting, adjustment).'²⁵ Knowles identified five components or cognitive activities undertaken during a self-directed learning cycle: (1) diagnosing learning needs, (2) formulating learning goals, (3) identifying resources for learning, (4) choosing and implementing appropriate learning strategies, and (5) evaluating learning outcomes.²⁶ For CPD in medicine, Sargeant et al. suggest a conceptual framework of self-directed learning that incorporates two elements: one being the five cognitive activities identified by Knowles, and the other being the willingness of physicians to drive their own learning.²⁰ According to Sargeant et al., physicians' willingness to self-direct their learning is the key to lifelong learning and effective CPD.

Although self-directed learning enables physicians to undertake the learning process at their own pace and time, formulate their own goals and choose the learning strategies that accommodate their work-context pressures, several studies have raised questions about the effectiveness of self-directed learning. For example, there are potential risks for learner misunderstandings, overestimation of preparedness for practice and development of unchecked bad habits.^{27,28} Most importantly, self-directed learning heavily relies on accurate self-assessment, while several studies have found inaccuracies in physicians' self-assessments of their knowledge and skills. Physicians with the lowest external performance ratings and with less experience tend to overestimate their competence.²⁹ To generate a valid appraisal of their own performance and related learning needs, physicians need to 'inform' or 'guide' their self-assessment by using external data and feedback.³⁰⁻³² These data can take many forms, ranging from informal verbal feedback provided by colleagues, team members or patients, to formal performance assessments. For CPD to be effective, it will be necessary for the educational and developmental strategies to be closely linked to both clinical workplace needs and the needs of healthcare systems.

CPD can take place in traditional programs such as lectures and workshops, and in more informal and self-directed ways such as learning in practice from patients and critical incidents, in discussion with colleagues and coworkers and through reflection upon practice, relationships, professionalism and learning needs. It can also occur through a formal external performance or practice assessment such as chart audit, multisource feedback (MSF), peer practice visits or patient satisfaction surveys. Given the growing understanding of the importance of teamwork to healthcare delivery, opportunities are explored to focus on the shared work of a team of healthcare professionals, patients and their families in achieving health outcomes that cannot be achieved by individual healthcare professionals.³³ Within the workplace, the team, as the 'unit' of learning and assessment, is often the focus of organizational quality improvement strategies.³⁴ The team is also the subject of patient feedback and critical incident reviews to address and improve both individual and team performance.³⁵ It is agreed that successful implementation of CPD programs must be a collective responsibility of physicians, CPD provider organizations, regulators and the health system, if it is to meet expectations of the social contract between the public and the profession.³⁶

MAIN ARGUMENT AND RESEARCH QUESTION

As we learn more about the linkages between assessment, feedback, and continued learning and professional development, two themes are emerging. The first is the notion of assessment *for* learning, not just *of* learning, that is, using assessment data to guide learning and practice change.³⁷ The second is the realization that learning occurs within the context of one's practice, through the everyday contact with patients and colleagues and exploring the questions which arise in clinical practice.³⁸ Integral to this practice learning is physicians' ability to identify the need for learning and practice improvement, and also the willingness to respond to this need.²⁰ Nowadays, the use of internal and external sources of data as the foundation for developing a learning plan to continuously improve performance and practice is mandated by medical regulatory authorities in many western healthcare systems, as reflected in maintenance of certification, revalidation or re-registration requirements.¹⁶⁻¹⁹ However, the interlinking of CPD practices and requirements for maintenance of certification presents itself as a delicate matter: external accountability requirements can easily collide with more internally motivated notions of 'self-direction' and 'reflection'. Suspicion may arise that the documentation of reflection promotes skills of self-portrayal rather than other forms of learning. Furthermore, concerns have been expressed that the over-formalizing of professional activity may lead to a loss of personal investment and a disposition of compliance.^{9,10} In striving to make reflection visible and authentic, care must be taken to ensure that it is not enforced in an overly prescriptive manner, turning reflection into a tick-box exercise instead of an opportunity for learning and improving.^{9,10} Therefore, this thesis addresses the following research question:

How can reflection on physicians' professional performance be best facilitated to serve their developmental needs?

RESEARCH CONTEXT

In this thesis, the requirements for professional re-registration of hospital-based physicians in the Netherlands are taken as a starting point for the studies included in this thesis. The Dutch healthcare system can be characterized as a regulated market oriented system within the context of a universal insurance system.³⁹ The majority (approximately 12.500 of 17.000) of physicians in the Netherlands are independent entrepreneurs organized in 'medical specialist companies' through which specialist care is contracted by healthcare organizations. The remaining physicians are employed by healthcare organizations. Healthcare organizations and physicians legally share the responsibility for quality of patient care. In Dutch terminology, "medical specialist" is used to refer to physicians who have completed their specialty/residency training and are registered as a specialist with the Registration Committee Medical Specialists (in Dutch: Registratiecommissie Geneeskundig Specialisten, RGS). Registration is granted on the condition that physicians demonstrate that there is responsible functioning and adequate expertise. In order to maintain their registration as a medical specialist and thus their license to practice, physicians need to meet a number of re-registration requirements, which are assessed once every five years.³⁹ These include the following:

1. Working in clinical practice for a minimum of 16 hours per week;
2. Participating in accredited expertise-promoting activities equaling a minimum of 40 hours per year (CME activities);
3. Participating in the external peer review (called *visitatie*) program of a physician's professional association. This program focuses on physicians' group performance, assessed during peer site-visits in which attention is focused on work processes⁴⁰;
4. Participating in evaluation of individual performance. This evaluation procedure is referred to as the IFMS requirement (Individual Functioning of Medical Specialist) and includes the development and maintenance of a personal development plan (PDP; in Dutch: persoonlijk ontwikkelplan, POP) which allows the physician to work specifically on improvement actions resulting from the individual evaluation. This requirement has formally been in place as per 2020.³⁹

OVERVIEW OF STUDIES

This thesis comprises four empirical studies on (the facilitation of) reflection on physicians' professional performance in the clinical workplace. In accordance with Dutch healthcare policy on the monitoring of physicians' professional performance, a distinction is made between facilitation of reflection in the context of physicians' *individual* performance and physicians' *group* performance.

- **Chapter 2** comprises an open exploration of how hospital-based physicians describe their relation to reflection, by mapping how they conceptualize and experience reflective activity in their daily professional practice.
- **Chapter 3** reports on the design and implementation of a performance evaluation system, which enables physicians to collect MSF on their performance and meet with a coach to reflect on the feedback received and develop a plan to improve. The development and implementation of a tailored coaching approach was part of this study. The resulting system reflects the compromises made in dealing with the conflicting nature of accountability and development purposes.
- **Chapter 4** reports on physicians' experiences with participation in a professionally facilitated group session, in which physicians share and discuss the results of personal MSF with peers with whom they work on a regular basis. It is assumed that involving peers in a physician's professional performance reflections may provide a meaningful context to interpret the multisource feedback results and may promote feedback acceptance and follow up.
- **Chapter 5** reports on the development process of a new multisource feedback tool for evaluating the performance of specialty-specific physician groups, and the validity evidence to support its use. Three sources of validity evidence were examined: content validity, response process validity and internal structure validity.
- **Chapter 6** summarizes the main findings of the individual studies and discusses these findings in the light of the existing literature. In addition, methodological strengths and limitations are reflected on, implications for practice and recommendations for stakeholders are provided and an agenda for future research is presented.

Table 1.1 provides an overview of the individual studies, the research questions posed, the data sources, research approaches and study designs used and the setting in which the studies were conducted.

Table 1.1 *Overview of studies, research questions, data sources, designs and settings*

Chapter	Research question(s)	Data source	Approach/ Design	Setting
2	How do practicing physicians conceptualize reflection, and how do they experience reflective activity in their daily practice?	Semi-structured interviews	Interpretative phenomenological analysis approach	Dutch general hospitals and academic medical centers
3	How to design a performance appraisal system that will successfully comply with both revalidation requirements and the need for professional development?	Process analysis and semi-structured interviews	Design-based research approach	Dutch academic medical center
4	How do physicians experience participation in a professionally facilitated peer group session? How do they perceive the session to contribute to their professional development and performance improvement? Which factors do they perceive as hindering or helpful in this process?	Semi-structured interviews	Interpretative phenomenological analysis approach	Dutch general hospitals
5	What are the psychometric properties of the Group Monitor instrument (GM) for each of the four rater classes? Are there interpretation differences between rater classes? How many raters are necessary per class for reliable measurements?	Surveys among four rater classes: 1) physicians (self-assessment) 2) staff 3) peers 4) managers	Exploratory factor analyses, reliability and generalizability analyses	Dutch general hospitals

This dissertation is based on journal articles; some repetition of information across chapters cannot be prevented.

REFERENCES

1. McKay EA. Reflective practice: doing, being and becoming a reflective practitioner. Skills for Practice in Occupational Therapy. Oxford: Elsevier Ltd; 2008.
2. Lombarts K. Physicians' professional performance: between time and technology. 2019. Rotterdam: 2010 Publishers.
3. Sandars J. The use of reflection in medical education: AMEE Guide No. 44. *Med Teach*. 2009;31:685-695.
4. Boud D, & Fales A. Reflective learning: key to learning from experience. *J Humanist Psychol*. 1983;23(2):99-117.
5. Mezirow J. A critical theory of adult learning and education. *Adult Ed*. 1981;32(1):3-24.
6. Jarvis P. Reflective practice and nursing. *Nurse Educ Today*. 1992;12(3): 174-181.
7. Finlay, L. Reflecting on 'Reflective practice' PBPL paper 52. 2008. www.open.ac.uk/pbpl.
8. Reynolds. M & Vince R. (eds). Organizing reflection. 2004. Aldershot, Surrey: Ashgate.
9. Mann K, Gordon J, MacLeod A. Reflection and reflective practice in health professions education: a systematic review. *Adv Health Sci Educ Theory Pract*. 2009;14(4):595-621.
10. Ng SL, Kinsella EA, Friesen F, Hodges B. Reclaiming a theoretical orientation to reflection in medical education research: a critical narrative review. *Med Educ*. 2015;49(5):461-475.
11. Freidson E, Professionalism: the third logic. Oxford, UK: Blackwell Publishers; 2001.
12. Collier R. Professionalism: the privileges and burden of self-regulation. *Can Med Assoc J*. 2012;184(14):1559-1560.
13. Cruess S, Cruess R. The medical profession and self-regulation: a current challenge. *AMA J Ethics*. 2005;7(4):320-324.
14. Schulpen T, Lombarts K. Quality improvement of paediatric care in the Netherlands. *Arch Dis Child*. 2007;92(7):633-636.
15. Crossing the quality chasm: a new health system for the 21st century. The National Academies Press; 2001.
16. American Board of Medical Specialties. Standards for the ABMS Program for Maintenance of Certification (MOC): For implementation in January 2015.2014.www.abms.org/%20media/1109/%20standards-for-the-abms-program-for-moc-final.pdf. Accessed 16 April 2020.
17. Federation of Medical Regulatory Authorities of Canada Revalidation Working Group. Physician revalidation: maintaining competence and performance. Ottawa (ON): The Federation;2007.www.fmrac.ca /committees/documents/final_reval_position_eng.pdf. Accessed 16 April 2020.
18. General Medical Council. Ready for revalidation; the Good medical practice framework for appraisal and revalidation. Published March 2013, Manchester, United Kingdom.
19. College Geneeskundig Specialismen. KNMG. Besluit herregistratie specialisten. 2015. <https://www.demedischspecialist.nl/sites/default/files/Besluit%20herregistratie%2013%20juli%202016.pdf>. Accessed 16 April 2020.
20. Sargeant J, Bruce D, Campbell CM. Practicing physicians' needs for assessment and feedback as part of professional development. *J Contin Educ Health Prof*. 2013;33:554-62.
21. Forsetlund L, Bjorndal A, Rashidian A, Jamtvedt G, O'Brien MA, Wolf FM, Davis D, Odgaard-Jensen J, Oxman AD. Continuing education meetings and workshops: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews*. 2009;2:CD003030.

22. Mansouri M, Lockyer J. A meta-analysis of continuing medical education effectiveness. *J Contin Educ Health Prof.* 2007;27(1):6-15.
23. Boelen C, Dharamsi S, Gibbs T. The social accountability of medical schools and its indicators. *Education for Health (Abingdon).* 2012;25:180-194.
24. Brockett R, Hiemstra R. Bridging the theory-practice gap in self-directed learning. *New Dir Adult Con Ed.* 1985;25:31-40.
25. White CB et al. 2014. Self-regulated learning in medical education. In: *Understanding Medical Education: Evidence, Theory and Practice* (ed. Swanwick T). Chichester, West Sussex: Wiley.
26. Knowles M. 1975. *Self-directed learning: a guide for learners and teachers.* New York, NY: Association Press.
27. Brydges R, Dubrowski A, Regehr G. A new concept of unsupervised learning: directed self-guided learning in the health professions. *Acad Med.* 2010;85:S49-S55.
28. Violato C, Lockyer J. Self and peer assessment of pediatricians, psychiatrists and medicine specialists: implications for self-directed learning. *Adv Health Sci Educ Theory Pract.* 2006;11(3):235-244.
29. Davis D, Mazmanian P, Fordis M, Van Harrison R, Thorpe K, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA.* 2006;296(9):1094-1102.
30. Eva K, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. *Acad Med.* 2005;80(Suppl):S46-S54.
31. Sargeant J, Armson H, Chesluk B et al. The processes and dimensions of informed self-assessment: a conceptual model. *Acad Med.* 2010;85(7):1212-1220.
32. Lockyer J, Armson H, Chesluk B, Dornan T, Holmboe E, Loney E, Mann K, Sargeant J. Feedback data sources that inform physician self-assessment. *Med Teach.* 2011;33(2):e113-120.
33. Royal College of Physician and Surgeons of Canada. *CanMEDS 2015 Physician Competency Framework.* www.royalcollege.ca/rcsite/documents/canmeds/canmeds-full-framework-e.pdf.%20 . Accessed 16 April 2020.
34. Ferlie EB, Shortell SM. Improving the quality of healthcare in the United Kingdom and the United States: a framework for change. *Milbank Quarterly.* 2001;79(2):281-315.
35. Manser T. Teamwork and patient safety in dynamic domains of healthcare: a review of the literature. *Acta Anaesthesiol Scand.* 2009;53(2):143-151.
36. Cook DA, Blachman MJ, Price DW, West CP, Berger RA, Wittich CM. Professional development perceptions and practices among US physicians: a cross-specialty national survey. *Acad Med.* 2017;92(9):1335-1345.
37. Schuwirth L, van der Vleuten C. General overview of the theories used in assessment: AMEE Guide no. 57. *Med Teach.* 2011;33(10):783-797.
38. Horsley T, O'Neill J, Campbell C. The quality of questions and use of resources in self-directed learning: personal learning projects in the maintenance of certification. *J Contin Educ Health Prof.* 2009;29(2):91-97.
39. Seddon N. Is the future Dutch? *The Lancet.* 2008;372(9633):103-104.
40. Lombarts, MJMH. *Visitatie of medical specialists. Studies on its scope and impact.* PhD thesis; 2003.

2

REFLECTION REVISITED: HOW PHYSICIANS CONCEPTUALIZE AND EXPERIENCE REFLECTION IN PROFESSIONAL PRACTICE – A QUALITATIVE STUDY

Bindels, E
Verberg, C
Scherpbier, AJJA
Heeneman, S
Lombarts, MJMH

BMC Medical Education. 2018;18(1);105.

ABSTRACT

Introduction: For the purpose of continuous performance improvement, physicians are expected to reflect on their practice. While many reflection studies are theoretically oriented and often prescriptive in the sense that they conceptualize what reflection should look like, the current study starts with practicing physicians themselves and maps how these physicians conceptualize and experience reflection in daily professional practice.

Methods: We conducted a qualitative study using in-depth interviews with 13 hospital-based physicians from various specialties and institutions. The interviews were transcribed verbatim and were analyzed iteratively, following the interpretative phenomenological analysis approach.

Results: Data analysis resulted in the identification of three main topics: *fuzziness*, *domain specificity* and *dialogical dynamics* of reflection in professional practice. Reflection was conceptualized as a fuzzy process of contemplation and action, leading to change and hopefully improvement of personal performance and health care in general. Physicians' experiences with reflection were different for the patient domain and the team domain. Whereas experiences in the patient domain were recalled first and discussed in relatively clear terms, those in the team domain came second and were discussed in more ambiguous terms. In order to achieve improvement in daily practice, honest and open dialogues were perceived as necessary. These dialogues were regarded as the result of an interplay between an internal and an external dialogue. The internal dialogue required sensitivity and courage of the individual; the external dialogue required psychological safety and encouragement of the environment. Within the team domain however, handling the external dialogue effectively was not self-evident, underlining the importance of psychological safety.

Discussion: This study draws attention to the interdependence between the individual and the collective contributions to reflective activity in professional practice. Apart from its importance to physicians' individual medical performance, reflective activity is also important to the functioning of a team of physicians. To allow reflection to rise from an individual activity to a team activity, it is necessary to invest in a safe environment in which people are encouraged to think, act, and be engaged.

INTRODUCTION

In the last decades, the dynamics within health care have changed: nowadays, physicians are confronted with knowledgeable patients, ever larger teams and increasing medical knowledge. As a result, the desired attributes of a 'good doctor' have evolved significantly, bringing concomitant change in the goals and expectations of the medical education field. For most of the 20th century, the predominant focus of medical education across the professional continuum was the acquisition of and reflection upon medical knowledge and procedural skills. Professional practice today, however, requires skills that extend beyond mere medical knowledge, such as interpersonal communication, team building, management and leadership skills. This shift in attention toward generic skills has fueled the movement of competency-based medical education (CBME), in which all aforementioned skills have been clustered and translated into competencies.¹⁻³ It is widely acknowledged that an important mechanism for learning and improving this broad range of competencies and thus for becoming and staying a 'good doctor' is 'reflection'.⁴⁻⁷

Since reflection is recognized as a crucial component of lifelong learning, there has been thoroughly invested in its theoretical understanding.⁸⁻¹⁰ Reflection is often described as a cyclic process of analysis, asking questions and paraphrasing professional experiences or, as Finlay¹¹ (p. 1) states: "the point is to recapture practice experiences and mull them over critically in order to gain new understandings and improve practice accordingly". In the context of CBME and lifelong learning, the focus of reflection is on developing the self by expanding one's own knowledge base and interpretive capacity, and broadening one's understanding of certain issues. This inward-looking and individual focus is thought to contribute to self-insight and self-improvement through analysis and understanding of specific learning experiences that teach knowledge, skills and attitudes.¹²⁻¹⁸ Across the continuum of medical education and medical practice, reflection is documented in portfolios, checklists and protocols; also group reflection activities and one-on-one interaction between teachers/mentors and students/physicians are used to stimulate and assess reflection.¹⁹⁻²²

However, the way in which reflection has been taken up in medical education is problematic, as has been addressed in recent years, for instance by Ng et al.¹² They call for caution with regard to reductionist approaches that break down reflection and its complexity into discrete components of activity, or steps of a process. In this way, reflection is approached in an overtly prescriptive way, i.e. reflection is conceptualized 'as it should look like', thereby establishing the link between using reflection as a tool to develop oneself and using it to assess individuals. Several authors, among whom Sumsion²³ and Rolfe,²⁴ argue that reflection has been turned

into a 'tick box exercise'; the very essence of reflection may be compromised when it is enforced in an overly prescriptive manner, or when it is subjected to evaluation and reductive approaches.

Despite the situation as outlined above, within CBME literature, reflective capacity is discussed as an attribute of a competent professional and therefore physicians are expected to reflect on their performance. Most continuous professional development (CPD) programs and quality assurance policies now include procedures or regulations to show some evidence of reflective practice. Therefore, there is a growing need to develop new ways of thinking about reflection that recognize the complexities and relational qualities of today's professional practice, without reducing reflection to a 'product'.²⁵⁻²⁷ In order to inform reflective activities for practicing physicians, it is necessary to openly explore how physicians themselves describe their relation to reflection. In this study, we therefore addressed the following two interrelated research questions: (1) How do practicing physicians conceptualize reflection and (2) how do they experience reflective activity in their daily practice?

METHODS

Setting, participants and sampling strategy

We conducted this study in the Netherlands, where physicians are either employed by the hospital or organized in independent entrepreneur partnerships. We invited hospital-based physicians from various specialties and (non-)academic institutions to participate in an in-depth interview with the first author (EB). They were purposively sampled to provide a heterogeneous participant group in terms of medical specialty, age, and gender.

Study design

Being interested in understanding the phenomenon of reflection by physicians' lived experiences, we chose an interpretative phenomenological research approach (IPA).²⁸⁻²⁹ The IPA approach allowed us to investigate individual experiences and accounts of physicians while constructing an overall impression of their experiences with reflection in daily practice.³⁰⁻³¹ We drew from the epistemological stance that people are self-interpretative beings, therefore knowledge is subjective and there is no ultimate truth.³² According to this approach, it is only through a process of interpretation that meanings can be understood. It explicitly acknowledges and accepts the importance of the researchers' interpretation and embraces the assumption that pre-understanding and co-creation by the researcher and the participants are what makes interpretations meaningful.^{29,32} Since we asked physicians to reflect on their reflections, the research itself is in fact a reflective

activity. Because the IPA approach is one of active engagement in which researchers bring their own backgrounds to the analytical process, practicing reflexivity is critical. In this reflexive spirit, we provide the following contextual information: the lead author (EB) has a background in art history and clinical neuropsychology; her collaborators have significant experience in studying medical education, but their own disciplinary backgrounds include teacher education (CV), medicine and philosophy (AS), biomedical science and competency-based education (SH) and healthcare policy and management (KL). Since these backgrounds might influence the approach to the subject, all researchers have been involved in the process of analysis.

Data collection

Initially, we informed participants by email and telephone that the purpose of the interviews was to explore reflection. Subsequently, we invited them to consider participation. Upon acceptance of the invitation, we requested an individual consent. In the interviews we focused on in-depth exploration, allowing the participants to talk freely without interference from others. An open ended interview guide was constructed based on our research questions and tested out in a pilot interview. The interviewer (EB) imposed direction as little as possible, enabling participants to tell their own reflection story. She gave no specific introduction on reflection and the interview started with a so-called association question. Participants were invited to write down on post-its impressions, ideas and words coming to mind when thinking of 'reflection'. Drawing on these connotations, we covered more specific items like feelings, thoughts and reactions, as well as concrete examples ('Can you describe an actual situation in which you reflected? What was the start? How did the process look like? Where did it lead to?') To close the interview, we invited participants to formulate a metaphor or transcending image of reflection. We conducted a pilot interview with the purpose of gathering feedback on the content of the initial interview guide as well as the interview technique. The first author (EB) conducted all interviews between March 2016 and November 2016. Interviews lasted approximately one hour and were conducted in the privacy of the physician's office, whereas confidentiality was assured at the start of the interview. All interviews were audio-recorded and transcribed verbatim.

Data analysis

To deduce meaning from the data, the first author engaged in mind mapping as means to reflect on the participants' reflections on reflection. Mind mapping involves creating visual diagrams with thought bubbles for brainstorming themes and drawing connections between them in a visual and non-linear way.³³⁻³⁴ Following each interview, a mind map was created in order to begin to track preliminary emerging

themes, using the post-its as a starting point. First, the transcript was read to obtain a holistic view of the data as a whole and an naïve understanding was formulated. Second, the first author looked for 'parts' in the 'whole' by identifying passages, key words and phrases that elicited understanding about what the participant was saying about the concept and experience of reflection in daily practice. Third, these parts were categorized and - together with the author's memos - converted into a mind map, which was compared to the original post-its and transcripts. Through constant comparative analysis, individual mind maps were accommodated in overarching mind maps. Saturation was reached at the point where two subsequent interviews did not yield any new themes or input for recurring themes. In order to establish credibility in the interpretation of data, initial understandings of interviews were discussed within the research team and interview fragments were re-listened by CV, SH and KL. Transcripts were discussed among EB and CV. Individual mind maps and overarching mind maps were discussed within the research team; a final comprehensive understanding was formulated by the first author in consultation with the research team.

Ethical considerations

This study received approval from the Dutch Association for Medical Education (NVMO) on March 14th, 2016 under reference number 572. All participants consented to participate in this study by providing written informed consent.

RESULTS

Participants

In total, 19 invitations were sent and all were received positively. However, six invitees were not able to take part because of agenda related reasons. Participants were 8 men and 5 women, ranging in age from 31 to 65 years, representing 11 different specialties at 8 (academic and general) hospitals (Table 2.1).

Table 2.1 Participants: gender, age, specialty and setting

Participant #	Gender	Age	Specialty	Setting
1	Male	59	Gynecology and obstetrics	Academic
2	Female	40	Pediatrics	Academic
3	Female	64	Internal medicine	Academic
4	Male	39	Psychiatry	General
5	Male	44	Trauma surgery	Academic
6	Female	44	Neurology	Academic
7	Male	31	ENT surgery	Academic
8	Male	50	Clinical pathology	General
9	Female	40	Radiology	General
10	Female	46	Anesthesiology	General
11	Male	65	Gastro-enterology	Academic
12	Male	46	Surgical oncology	General
13	Male	59	Neurology	General

Overall results

Data analysis resulted in the identification of three main topics related to our two research questions: the *fuzziness*, *domain specificity* and *dialogical dynamics* of reflection in daily practice. In short, during the interviews reflection was conceptualized as an ongoing movement between contemplation and action orientation, leading to change and hopefully improvement of personal performance and health care in general. Physicians' experiences with reflection were different for the patient domain and the team domain. Whereas reflection experiences in the patient domain were recalled first and discussed in relatively clear terms, reflection experiences in the team domain came second and were discussed in more ambiguous terms. In order to achieve improvement in daily practice, honest and open dialogues were perceived as necessary. These dialogues were regarded as the result of an interplay between an 'internal dialogue' or private conversation with oneself, and an 'external dialogue' or the act of sharing experiences with others.

The internal dialogue required sensitivity and courage of the individual; the external dialogue required psychological safety and encouragement of the environment. We will now discuss these themes in detail below.

1. Reflection as a fuzzy process

'Reflection' elicited process-related verbs, such as 'becoming aware', 'holding still', 'looking back', 'observing and analyzing' and 'gaining insight'. Newly acquired insights were meant to 'determine direction' and led to outcomes such as 'balanced conduct', 'quality improvement in own performance and health care in general' and 'ongoing learning'. Although the concept elicited these clear process-related verbs, participants also pointed to the fuzziness of the concept, stating that it was something 'fuzzy', 'foggy' and 'difficult to grasp', as Interviewee 3 observed:

"I think most people do understand what is meant by it, but that this is a concept that encompasses a lot of things."

Participants referred to a dialectical movement between contemplation and action or action orientation within the reflection process, as Interviewee 4 stated:

"If reflection does not lead to action or action-orientation, then it is like there is insight without movement. Then reflection stays in the atmosphere of associations, but it has to clench together and become clear, so that you can apply it to something. Because when something is cloudy, it disturbs and can deter. Reflection is like a process of rising above the fog: you seek clarity to get a view of the right direction and when you see the path ahead of you, you also want to go that direction, because it suits you."

2. Domain specificity of reflection

Participants described a broad array of practice situations that triggered reflection. Reflection experiences that came to mind first were related to the patient domain; reflection experiences that seemed to be surrounded by more 'fog' were mostly related to the team domain and were reported in a later instance.

Patient domain

Overall, reflective activities that were mentioned first were closely tied to the medical domain, such as clinical reasoning, adaptive expertise and patient communication. Participants seemed to have little difficulty in finding the right words to describe how these reflection processes evolved. We interpreted this as a sign that participants felt most comfortable with reflection related to medical issues on the patient level. Reflection in these situations was associated with a state of

concentration, as a means to assist the participant to pinpoint what was meaningful and of fundamental importance. Reflection was described as 'slowing down', 'being on your guard', 'taking responsibility' and 'being open to other perspectives and possibilities yet standing your ground'. Reflection required effort, as other habits of mind might be more rewarding, such as speed or no-nonsense. All participants reported that the above mentioned elements lead them to ask themselves questions, such as 'Did I miss things?', 'Did I close off too soon?' or 'How can I best 'tune in' on this patient and establish or restore a sense of connection?' In the case of such questions, it was obvious to ask a colleague for help or discuss these issues during consultations with colleagues. As Interviewee 1 explained:

"In a consultation with peers, there is of course partial misunderstanding of each other's point of view, or people are not entirely well informed. It is exactly in that process where we analyze the whole case and sort out what happened exactly. Ultimately we reconstruct the story and then we can make a decision. Constructing this story is also very important in the sense that others can learn from it, because it will come to mind when they will come across a similar situation."

Team domain

Reflective activities that were only mentioned later on in the interviews were tied to communication issues with colleagues, team work and team dynamics, as participants started to reflect on aspects of reflection that are probably harder to describe. This cluster of situations did not only contain ad hoc situations that called for direct response, but also situations that seemed to develop gradually over time, such as an unequal division in tasks, responsibilities and participation within the team. With regard to these type of situations, participants seemed to experience more difficulty or 'fog' in describing how they navigated these situations and how reflection processes evolved. As Interviewee 4 explained:

"In your role toward the patient there are not that many barriers, you just do your work, it has to do with your medical competence. But there is also cooperation and communication with colleagues, those things. Some colleagues are very diplomatic in their communication, others are very direct. Well, it forces me to think about my own style, and in which situation I use which style."

The proportion of 'fog' in these reflection processes was tied to the degree of clarity about possible outcomes and the dependency on others. Interviewee 2 described her encounter with her new group of colleagues:

"In new situations as these you just need to test the waters and see how the wind blows. You ask yourself, where do I fit in and what will be my role?"

In describing a similar situation, Interviewee 9 looked at it from a team perspective:

"When you enter a new team, you need to deal with this whole interaction, this group process. Who makes and who breaks the policy? Is there policy anyway? Where do we stand and where are we going? Have we thought about it at all?"

This interviewee further widened her team perspective by relating her teamwork to the bigger picture of the organization:

"As a physician, you are a little cog in a very large whole and this can lead to a kind of indifference. I mean, you would like to change something but then you notice that the whole organization is inert and slow like syrup. Then you run the risk of developing this 'every man for himself' attitude and you just focus on your own little piece."

3. Dialogical dynamics

From the participants' reflection stories, reflection's alternating movement between contemplation and action orientation seemed to unfold itself in an interplay between an internal dialogue (or private conversation with oneself) and an external dialogue (or actual conversation with others). These dialogues were triggered by situations in daily practice in both the patient domain as well as the team domain.

Internal dialogue

All participants noticed that a certain 'emotional sensitivity', 'observational ability' and 'courage to have an honest look at oneself' was conditional to start any internal dialogue. Interviewee 3 described this internal dialogue in the context of a patient contact:

"When I get irritated by a patient, and that just happens sometimes, I feel it inside. It happens quite often when someone shows claiming behavior. That is the moment that I think 'hey, I need to pay attention now'. And yes, in that moment you should actually be able to take a step back and not give in to your feeling of irritation, although that may be your natural reaction. I know the moments in which I need to be extra alert and I also know what I can do to turn the tide."

The same interviewee also mentioned the importance of 'taking a critical look at your own share' in the case of a complication or medical error:

"You can say, 'well, it was the other person, it has nothing to do with me, I don't feel addressed or affected, I don't feel responsible..' But you can also think, what was my role in this whole story, that is what I mean with 'taking a critical

look at something'. Reflection just requires courage, because you also have to face your own imperfection and stupidity."

Interviewee 13 also mentioned that reflection was observing a situation, taking a step back and asking a question, instead of reacting based on your first impression. Ideally, the internal dialogue took place prior to the external dialogue:

"I think that over time, I try to understand things more by asking questions, instead of giving my opinion straight away. [Interviewer: How did you learn that?] Yes... through observations and by seeing how things happen, how people do things in conversations. My response also depends a lot on the moment and on the person: who is this person, do I know him or her better? Ideally you have some sort of repertoire at your disposal, like, you approach one colleague in this way and the other in that way.. [Interviewer: so before you start a conversation you have a conversation with yourself?] Well, ideally yes, unless I am so angry that I cannot control myself, then I cannot find enough space for this conversation with myself."

External dialogue

Participants described that when a situation related to the patient domain or the team domain was discussed collectively, an external dialogue was initiated. In the case of a (dissatisfying) patient contact or lacking medical knowledge, participants frequently stated that they could easily call for help from a colleague. In the case of a complication or medical error, external dialogue was considered to be of vital importance for patient care, as Interviewee 12 mentioned:

"Talking to others about this is a possibility to unwind. Because you talk to each other about it, you can sort out how to prevent it from happening next time. And it creates this possibility for other members to say 'you know, actually I experience the same thing'. So then you can see how you can improve practice as a group."

Other interviewees not only pointed to the importance of the external dialogue for direct patient care, but also for the team performance in the long run. As Interviewee 9 explained:

"If we as a group continue like this, we will get into trouble, eventually. Not so much in the sense that we don't do well as individual medical specialists, but more in the sense that working together is challenging. Because you will get frustrations, because it seems like, well, over time you develop some sort of unequal division in tasks, responsibilities and participation and in discussion people keep aloof. And in our group there are quite some differences in characters and we just have to make it work."

Several participants mentioned the challenge of team performance and described how team issues could get out of hand, for example in the case of an approaching merger. Interviewee 8 described how team discussions got stuck because of a conflict of interest and he described the experiences with an external counselor:

"Then this external person does the reflection in your place, so to say. He is a kind of father figure and he can simply guide the group in this whole reflection process. He spells out the consequences and what you can or cannot expect from each other. And what choices there are and also the limitations of these choices, so that you get a clear picture again."

Interviewee 4 described his experiences with an approaching merger and mentioned that he individually looked for sparring partners outside his team:

"Sometimes I meet with peers outside my team to discuss the uncertainties about the merger. Like, how do I position myself? Am I too careful? Then you get tips, advice, feedback from colleagues who work elsewhere. So that is very good, yes. And because you do not work with these people directly, it is certainly safer to speak your mind."

Interplay of internal and external dialogue

Participants described an interplay between an internal and an external dialogue in daily practice. In the patient domain, discussing patient cases with others was recognized as a way of reflecting collectively; the tipping point between the internal and external dialogue was relatively straightforward and the process of collective reflection was clearly framed. In the team domain, however, the tipping point between the internal and external dialogue was less straightforward and the process of collective reflection was less clear because of 'fuzzy' team dynamics, issues related to the division of tasks and responsibilities, the sorting out of a shared group vision and an overall (diffuse) sense of psychological safety. The importance of this sense of security was reported by all participants. From the participants' stories it became clear that the experience of psychological safety was a condition for the internal dialogue to be translated into an external dialogue. As Interviewee 5 explained:

"In my department, we take a critical stance toward each other and sometimes my colleagues ask me about how I handled things in a surgical procedure. I know that people can just wonder and be curious, but where I work, it is like, you have done something and colleagues take this as an opportunity to rise above you, to display their knowledge to residents and nurses at the expense of you, that is just something I feel."

The importance of psychological safety also came to the fore in the case of a medical error, a situation that started in the patient domain and underwent a follow-up in the team domain. Although some participants noted that team discussions on these matters were meant as a possibility to 'unwind', not everyone felt safe enough to even report an error, because of the possibility of 'becoming known as that person'. As Interviewee 8 mentioned:

"You just know that your own reaction to a medical error partly depends on how the other person reacts, to whom you must report. If you have some negative experiences with this, you are less inclined to report an error, while at the same time it is crucial for a good hospital setting."

From participants' stories, the need for commitment and reciprocity arose. As Interviewee 3 talked about the teamwork with her residents, she characterized her attitude toward building trust, working together and supervising them as follows:

"I tell them in advance about me being a bit neurotic.. in the sense that I find it difficult to sort of let them go.. I tell them that if they feel that I am looking over their shoulder too much that they should blow the whistle. And I always ask honest feedback from them. By doing so, I hope that they will feel that as a supervisor, I am not so much standing above them, but next to them. Working together simply cannot be a one-way street."

DISCUSSION

In this study it was shown that physicians conceptualized reflection as a fuzzy process in which elements of contemplation and action alternated. Participants described this process as an interplay between an internal dialogue and an external dialogue. The sequence in which reflection experiences in daily practice were mentioned provided information about a layering in the experience of reflective activity. Participants first referred to reflection experiences in the patient domain (the context of physician-patient contacts and medical performance), which they discussed in relatively straightforward terms. In the sequence of the interview, participants also referred to reflection experiences in the team domain (the context of working together with other professionals), which they discussed in more ambiguous terms. From this finding we concluded that reflection experiences in the patient domain were more readily available than reflection experiences in the team domain. In the following section, we will elaborate on (1) the conceptualization of reflection as a two-fold activity on both the individual level as well as the collective level; (2) the importance of reflection in the team domain to enable team development and foster improvement in the patient domain and (3) the need to create psychological safety in the team domain as a condition for effective reflection.

Our findings indicate that physicians do not view reflection solely as a clear or straightforward attribute of the competent individual professional. When openly exploring, physicians conceptualize reflection as a hard to frame 'fuzzy process' that takes place both in the individual atmosphere – by means of an internal dialogue – as well as in the collective atmosphere – by means of an external dialogue. This view on reflection is somewhat contradictory to the view on reflection as a solidified individual competency, as expressed in the CBME literature.^{3-5,12,19,25} In CBME literature, reflective capacity is portrayed as an individual possession, an ability to engage in reflective activity which can be acquired. This contrast between reflection viewed as an individual attribute and reflection viewed as something that commutes between the individual and his/her social environment resonates with the work of Lingard about individualist and collectivist discourses of medical competence. As described by Lingard, the individualist discourse views competence as a construct which individuals acquire and possess, which is context-free and which represents a state to be achieved. In the collectivist discourse, competence evolves from participation in authentic situations, is distributed across networks of persons and manifests itself in interconnected behaviors.³⁵⁻³⁶ If we consider the findings of the current study in the light of these individualist and collectivist discourses, then the notion of reflection requires expansion from 'self-reflection' to 'collective reflection'. This collective dimension of reflection has been described by several authors in communication and learning sciences, emphasizing the role of sharing experiences for the purpose of both individual learning as well as collective learning.³⁷⁻³⁹ As Boud⁴⁰ states, "reflection in collective settings cannot be an individual act if it is to influence work that takes place with others." Reflection on one's individual process or the team process can be facilitated either in a group of physicians or in a one-on-one interaction between a physician and peer or a facilitator. By explicating preliminary individual understandings and sharing individual experiences with others, a shared understanding about work practices and work-related experience is established and collective learning can emerge. Outcomes of this process of sharing and joint reflection, either during a one-on-one or a group session, can prompt both individual learning as well as collective learning.

Our analysis rendered two layers of reflection experiences, first related to the patient domain and second related to the team domain. Although distinguishable from each other, these two layers were closely connected to each other. For example, the situation of an approaching merger, which could cause disturbance within the team, could eventually impact the way care was organized and provided for in the patient domain. Collective reflection in the team domain could be a powerful method to ensure that team members learn what is necessary to remain motivated individually and as a team, to solve work-related problems collectively, to adequately discuss tensions among team members and to make group decisions that can count on sufficient support from all members. In communication and learning sciences this

is defined as collective learning: the joint creation of collective rules, procedures, norms and values that determine the behavior of the group of team members.⁴¹ Collective learning also implies that members regularly take time to reflect together and, if necessary, adjust the working method based on experiences and new insights. This ongoing process of reflecting together on common concerns means that the collaboration within the team is constantly improving, so that the team can maintain a high level of functioning when circumstances change.⁴¹⁻⁴² This calls for a broadening of the reflective focus of team members by concentrating on both their individual and collective part in team work. Since team work will always be a challenge, the question remains as to how collective reflection can be handled properly: by increasing openness, honesty and space for feedback. This brings us to our last point, the importance of psychological safety.

From the findings it became clear that reflecting together and sharing experiences with others is only possible if there is a certain degree of psychological safety. Only when team members feel safe, they will be willing to take interpersonal risks, be vulnerable and give feedback to colleagues without being afraid that this might have negative consequences for the relationship with team members.⁴³ When a psychologically safe environment is lacking, team members would prefer to remain silent, especially when it concerns learning from mistakes (as is the case with medical errors).⁴⁴⁻⁴⁵ In addition, a sense of joint responsibility and commitment and the extent to which someone identifies with the team is important.⁴³ Our findings suggest that reflecting together with team members on team issues is not self-evident; participants noticed that – in some matters – they preferred to speak with people outside their own ‘working bubble’. Participants also noticed that in team discussions colleagues might keep aloof and engage in abstract conversation as an alternative to action that meets the cultural norm. Possible consequences of this state of affairs may be mutual tension or development of an ego centric attitude, thereby possibly undermining performance in its broadest sense.⁴³ Creating a climate of psychological safety, therefore, is of central importance for fostering open dialogues, thereby fueling both individual and team learning and development.

Limitations

As with any qualitative study, a conducive partnership between the researcher and participant is necessary for rendering a thorough insight into the topic of interest. For this purpose, not only the interviewer’s listening sensitivity is important, but also participants’ abilities to verbalize their experiences. Although all participants’ accounts of their conceptualization and experience of reflective activity in daily practice are unique, IPA methodology requires that the researcher looks for themes that are shared by all participants.³⁰ Once identified, the themes may shape the direction of the analysis, rendering IPA vulnerable to researcher bias. As we were mindful of this issue,

we took steps to minimize such bias, primarily through ongoing discussions among the authors to verify our coding scheme, interpretations and conclusions. Apart from this methodological consideration, it should also be noted that all participants were hospital-based physicians working within the Dutch healthcare system. Therefore, it is not self-evident that the findings of the current study are transferable to physicians outside the Netherlands, or to physicians outside the hospital setting, such as general practitioners. Caution is also required when it comes to transferability of the findings to other groups of healthcare professionals, such as nurses.

Implications for practicing reflection

In the wake of previous studies, this study suggests to explore reflection as a collective activity affecting the larger health profession team or organization, rather than approaching reflection as a solitary, purely introspective activity. In the context of CPD programs, quality assurance policies and re-registration procedures, dedicated time should be provided for physicians to reflect on both their individual process as well as on common concerns within the team. Written reflective activities can be used, but based on the results of the current study, a reflection format in which an important role is reserved for the dialogical aspects of reflection could be more fitted to practice. Reflection could be facilitated in meetings following the completion of a multisource feedback (MSF) round within the team.⁴⁶ These meetings could either be one-on-one meetings with a facilitator or peer or team meetings led by a facilitator. In view of the importance of psychological safety, the use of skilled facilitators would be desirable. The team discussion with peers may also contribute to a sense of solidarity and mutual respect, since previous research has shown that sharing personal reflections with peers could improve the quality of collegial relationships and heighten the chance of performance improvement.⁴⁷

Recommendations for future research

Future research should continue to focus on reflection related to both the individual as well as the team functioning of physicians and the mutual influence of these two. As noted above, reflection on the physician's individual professional development or the team process can be facilitated in both one-on-one meetings between a physician and a peer or facilitator, or in group meetings led by a facilitator. Further research into the impact and perceived usefulness of different strategies to facilitate reflection in these meetings is warranted. It is known from previous research that guided reflection with the help of a facilitator helps to achieve performance improvement. However, it is unclear as to whether or how the presence of colleagues in reflection meetings to debrief individual performance feedback impacts on the process or its outcomes. Furthermore, investments must be made to develop and validate tools that can be used to evaluate the collective performance of groups of physicians.

CONCLUSIONS

The contribution of this reflection study is twofold: a majority of studies deals with the theoretical aspects of reflection and are prescriptive, conceptualizing what reflection should look like, whereas the current study starts with the practitioners themselves and maps how these practitioners describe their own relation to reflection. Secondly, it confirms some of the points that have been described in the reflection literature, for instance that reflection is an ambiguous concept which can have benefits even when it does not lead to improved practice. Physicians experienced reflection as an activity related to their medical performance, but also as an activity related to their performance within the team to which they belong. In realizing open dialogues with colleagues on the team level, psychological safety plays an important role. Both the individual physician as well as the team of physicians and other health professionals have a responsibility in the development of abilities such as listening, sensitivity and community spirit. We suggest that attention to these aspects will foster synergy between reflective physicians and will culminate in reflective teams.

REFERENCES

1. Institute of Medicine. Health Professions Education: a Bridge to Education, Washington, DC: National Academies Press; 2003.
2. Frenk J, Chen L, Bhutta ZA, Cohen J, Crisp N, Evans T, Fineber H, Garcia P, Ke Y, Kelley P, Kistnasamy B, Meleis A, Naylor D, Pablos-Mendez A, Reddy S, Scrimshaw S, Sepulveda J, Serwadda D, Zurayk H. Health professionals for a new century: transforming education to strengthen health systems in an interdependent World. *Lancet*. 2010; 376:1923-1958.
3. Carraccio CL, Englander R. From Flexner to competencies: reflections on a decade and the journey ahead. *Acad Med*. 2013;88(8):1067-1073.
4. Sandars J. The use of reflection in medical education: AMEE Guide No. 44. *Med Teach*. 2009;31(8):685-695.
5. Mann K, Gordon J, MacLeod A. Reflection and reflective practice in health professions education: a systematic review. *Adv Health Sci Educ*. 2009;14(4):595-621.
6. Epstein RM. Reflection, perception and the acquisition of wisdom. *Med Educ*. 2008;42(11):1048-1050.
7. Mamede S, Schmidt HG. The structure of reflective practice in medicine. *Med Educ*. 2004;38(12):1302-1308.
8. Kinsella EA. Professional knowledge and the epistemology of reflective practice. *Nurs Philos*. 2010;11(1):3-14.
9. Nguyen QD, Fernandez N, Karsenti T, Charlin B. What is reflection? A conceptual analysis of major definitions and a proposal of a five-component model. *Med Educ*. 2014;48(12):1176-1189.
10. Clarà M. What is reflection? Looking for clarity in an ambiguous notion. *J Teach Edu*. 2015;66(3):261-271.
11. Finlay L. Reflecting on reflective practice. PBPL paper 52. 2008. www.open.ac.uk/pbpl. Accessed 14 Dec 2017.
12. Ng SL, Kinsella EA, Friesen F, Hodges B. Reclaiming a theoretical orientation to reflection in medical education research: a critical narrative review. *Med Educ*. 2015;49(5):461-475.
13. Carr SE, Johnson PH. Does self-reflection and insight correlate with academic performance in medical students? *BMC Med Educ*. 2013;13(1), 113.
14. Stephens MB, Reamy BV, Anderson D, Olsen C, Hemmer PA, Durning SJ, Auster, S. Writing, self-reflection, and medical school performance: the human context of health care. *Mil Med*. 2012;177(9S):26-30.
15. Ballon BC, Skinner W. "Attitude is a little thing that makes a big difference": reflection techniques for addiction psychiatry training. *Acad Psychiatry*. 2008;32(3):218-224.
16. DasGupta S, Charon R. Personal illness narratives: using reflective writing to teach empathy. *Acad Med*. 2004;79(4):351-356.
17. Stark P, Roberts C, Newble D, Bax N. Discovering professionalism through guided reflection. *Med Teach*. 2006;28(1):e25-e31.
18. Nevalainen MK, Mantyranta T, Pitkala KH. Facing uncertainty as a medical student—a qualitative study of their reflective learning diaries and writings on specific themes during the first clinical year. *Patient Educ Couns*. 2010;78(2):218-223.

19. Hall P, Byszewski A, Sutherland S, Stodel EJ. Developing a sustainable electronic portfolio (ePortfolio) program that fosters reflective practice and incorporates CanMEDS competencies into the undergraduate medical curriculum. *Acad Med.* 2012;87(6):744-751.
20. Whitehead C, Sellegger V, Kreeke J, Hodges B. The 'missing person' in roles-based competency models: a historical, cross-national, contrastive case study. *Med Educ.* 2014;48(8):785-795.
21. Paul R, Beamish AJ, Suter VE, Ruprai CK, Al-Muzaffar L, Jiraporncharoen W. Assessing reflective practice. *Education for Primary Care.* 2013;24(5):388-390.
22. Wald HS, Borkan JM, Taylor JS, Anthony D, Reis SP. Fostering and evaluating reflective capacity in medical education: developing the REFLECT rubric for assessing reflective writing. *Acad Med.* 2012;87(1):41-50.
23. Sumsion J, Fleet A. Reflection: can we assess it? Should we assess it? *Assess Eval High Educ.* 1996;21(2):121-131.
24. Rolfe G. Rethinking reflective education: what would Dewey have done? *Nurse Educ Today.* 2014;34(8):1179-1183.
25. Batalden P, Leach D, Swing S, Dreyfus H, Dreyfus S. General competencies and accreditation in graduate medical education. *Health Aff.* 2002;21(5):103-111.
26. Holmboe ES, Foster TC, Ogrinc G. Co-Creating Quality in Health Care Through Learning and Dissemination. *J Contin Educ Health.* 2016;36:S16-S18.
27. Kotzee B. Private practice: Exploring the missing social dimension in 'reflective practice'. *Stud Cont Educ.* 2012;34(1):5-16.
28. King N, Carroll C, Newton P, Dornan T. "You can't Cure it so you have to Endure it": The Experience of Adaptation to Diabetic Renal Disease. *Qual Health Res.* 2002;12(3):329-346.
29. Finlay L, Ballinger C. Qualitative research for allied health professionals: Challenging choices. New Jersey: John Wiley & Sons; 2006.
30. Starks H, Brown Trinidad S. Choose your method: A comparison of phenomenology, discourse analysis, and grounded theory. *Qual Health Res.* 2007;17(10):1372-1380.
31. Bunniss S, Kelly DR. Research paradigms in medical education research. *Med Educ.* 2010;44(4):358-366.
32. Wojnar DM, Swanson KM. Phenomenology: An exploration. *J Holist Nurs.* 2007;25:172-180.
33. Buzan T, Buzan B. The mind map book: Radiant thinking – major evolution in human thought. London: BBC Books; 2003.
34. Tattersall C, Powell J, Stroud J, Pringle J. Mind mapping in qualitative research. *Nurs Times.* 2011;107(18):20-22.
35. Lingard L. Rethinking competence in the context of teamwork. In: Lingard L, Hodges B, editors. The question of competence: Reconsidering medical education in the twenty-first century. Ithaca NY: Cornell University Press; 2012.
36. Lingard L. Paradoxical truths and persistent myths: Reframing the team competence conversation. *J Contin Educ Health.* 2016;36:S19-S21.
37. Reynolds M, Vince R. Organizing reflection. London: Ashgate; 2004.
38. Senge P. The fifth discipline: The art and practice of the learning organization. New York: Doubleday; 2006.

39. Van Woerkom M, Croon M. Operationalising critically reflective work behavior. *Personnel Review*. 2008;37(3):317-331.
40. Boud D. Relocating reflection in the context of practice. In: Bradbury H, Frost N, Kilminster S, Zukas M, editors. *Beyond Reflective Practice. New approaches to professional lifelong learning*. New York: Routledge; 2010.
41. Edmonson A. Psychological safety, trust, and learning: A group-level lens. In: Kramer R, Cook K, editors. *Trust and distrust in organizations: Dilemmas and approaches*. New York: Russell Sage Foundation; 2004.
42. Gibb JR. Climate for trust formation. In: Leland LP, Gibb JR, Benne KD, editors. *T-group theory and laboratory method*. New York: Wiley; 1964.
43. Edmonson A. Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*. 1999;44(2):350-383.
44. Carmeli A, Hoffer Gittel J. High-quality relationships, psychological safety, and learning from failures in work organizations. *J Organ Behav*. 2009;30(6):709-729.
45. Edmonson A. Learning from failure in health care: frequent opportunities, pervasive barriers. *Qual Saf Health Care*. 2004;13. Suppl 2;ii3-ii9.
46. Sargeant J, Lockyer J, Mann K, Holmboe E, Silver I, Arson H, Driessen EW, MacLeod T, Yen W, Ross K, Power M. 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):1298-1706.
47. Overeen K, Driessen EW, Arah OA, Lombarts MJMH, Wollersheim HC, Grol RPTM. Peer mentoring in doctor performance assessment: strategies, obstacles and benefits. *Med Educ*. 2010;44(2):140-147.

3

DESIGNING A SYSTEM FOR PERFORMANCE APPRAISAL: BALANCING MEDICAL SPECIALISTS' ACCOUNTABILITY AND PROFESSIONAL DEVELOPMENT

Bindels, E
Boerebach, B
Scheepers, R
Nooteboom, A
Scherpbier, AJJA
Heeneman, S
Lombarts, MJMH

Under editorial review

ABSTRACT

Introduction: In many healthcare systems, medical specialists are accustomed to periodically participate in individual performance appraisals to guide their professional development. For the purpose of revalidation, or maintenance of certification, they need to demonstrate that they have engaged with the outcomes of these appraisals. The combination of taking ownership in professional development and meeting accountability requirements may cause undesirable interference of purposes. To support specialists in their professional development, new Dutch legislation requires that specialists discuss their performance data with a non-hierarchical (peer)coach and draft a personal development plan. In this study, we report on the design of this system for performance appraisal in a Dutch academic medical center.

Methods: Using a design-based research approach, a hospital-based research group had the lead in drafting and implementing a performance appraisal protocol, designed a multisource feedback tool, co-developed and piloted a coaching approach, implemented a planning tool, recruited peer coaches and facilitated their training and peer group debriefings.

Results: The system consisted of a two-hour peer-to-peer conversation based on the principles of appreciative inquiry and solution-focused coaching. Sessions were rated as highly motivating, development-oriented, concrete and valuable. Peer coaches were considered suitable, although occasionally specialists preferred a professional coach because of their expertise. The system honored both accountability and professional development purposes. By integrating the performance appraisal system with an already existing internal performance system, specialists were enabled to openly and safely discuss their professional development with a peer, while also being supported by their hierarchical leader in their self-defined developmental goals. Although the peer-to-peer conversation was mandatory and participation in the process was documented, it was up to the specialist whether or not to share its results with others, including their superior.

Discussion: In the context of mandatory revalidation, professional development can be supported when the appraisal process involves three characteristics: the appraisal process is appreciative and explores developmental opportunities; coaches are trustworthy and skilled; and the medical specialist has control over the disclosure of the appraisal output. Although the peer-to-peer conversations were positively evaluated, the effects on specialists' professional development have yet to be investigated in longitudinal research designs.

INTRODUCTION

For the purpose of continuing professional development (CPD), medical specialists (Netherlands) – or physicians (USA, Canada) or consultants (UK) - periodically participate in performance appraisals. Performance appraisal is an educational intervention, which involves a formative conversation between two professionals. Originally, the appraisal process was intended to provide a specialist with feedback on his or her performance, map a specialist's progress and identify areas for further development. Performance appraisal was not designed as an assessment of competence which a specialist would either pass or fail.¹ In response to society's call for transparency however, performance appraisal has been incorporated into procedures for revalidation (UK), recertification (USA) and re-registration (Netherlands). In this context, participation in performance appraisal is assumed to provide more objective assurance that a specialist is up to date and fit to practice medicine.²⁻⁴ The interlinking of appraisal and re-registration has presented itself as a dilemma, as this may cause undesirable interference of two types of goals, professional development versus accountability. The literature indicates that there may be negative backwash effects on professional development when performance appraisal is mandatory and will be used to prove fitness to perform. For instance, medical specialists' engagement in the appraisal process could be jeopardized, possibly turning the appraisal process into a tick-box-exercise.^{5,6}

In many healthcare systems, i.e. in the USA, Canada, the UK and the Netherlands, the appraisal process comprises the use of multisource feedback (MSF).⁷⁻⁹ Feedback from peers and colleagues is found to be necessary to inform specialists' self-assessment by providing a more realistic view of how they perform.¹⁰ Research on how specialists use MSF, however, has demonstrated that this feedback does not self-evidently find its way into performance change. Feedback can be perceived as disconfirming or disappointing, it may evoke an emotional reaction that can interfere with the ability to assimilate and learn from it, or the feedback may not be specific enough to catalyze certain performance changes.⁹ Also, one's self-efficacy and motivation and the collegial culture can influence the process of using feedback for learning and change. Facilitated reflective discussions guided by another person (a peer, coach or mentor) can foster feedback reconciliation with one's own self-assessment and lead to engagement and the preparation of realistic and useful development plans.^{11,12} The addition of 'coaching' to the feedback lexicon is significant, because it places feedback in a different light. Coaching is developmental, and philosophically it moves feedback away from its historic tie to assessment and moves it towards learning.¹² Coaching supports personal and professional development; coaching is facilitative and guides the recipient in identifying their own needs and goals in developing a realistic plan.^{11,12}

In the Netherlands, the requirements for re-registration have recently changed; new legislation concerning specialists' individual performance appraisal has been adopted. As per 2020, medical specialists are required to five-yearly collect performance feedback, discuss their feedback with a non-hierarchical facilitator and define and follow-up on a personal development plan (PDP).¹³ The aforementioned concern over the two types of potentially conflicting goals of re-registration and performance appraisal, however, may jeopardize this appraisal process. The needs for psychological safety and intrinsic motivation seem principally incompatible with mandatory accountability and transparency requirements. In this study, we report on the design of an individual performance appraisal system in a Dutch academic medical center, in which this incompatibility was taken into account. The design problem was formulated as follows: how to design a system that will successfully comply with both re-registration requirements and the need for professional development? By using a design-based research (DBR) approach, our aim was to define, and build a system around characteristics that are crucial for the facilitation of professional development in the context of mandatory re-registration. We aim for this design to serve as an inspiring example for other institutions or health care systems dealing with the same challenge.

METHODS

Dutch healthcare system

The Dutch healthcare system can be characterized as a regulated market oriented system within the context of a universal insurance system.¹⁴ The majority (approximately 12.500 of 17.000) of medical specialists in the Netherlands are independent entrepreneurs organized in 'medical specialist companies' through which specialist care is contracted by healthcare organizations. The remaining medical specialists are employed by healthcare organizations. Healthcare organizations and medical specialists – legally – share the responsibility for quality of patient care. Medical specialists are considered a self-regulating profession. After completion of medical specialty training, they are registered as medical specialists. In order to maintain their license, medical specialists need to meet a number of re-registration requirements, which is assessed once every five years. These involve (1) working in clinical practice for a minimum of 16 hours per week, (2) participating in accredited expertise-promoting activities equaling a minimum of 40 hours per year, and (3) participating in the external peer review (called *visitatie*) program of a specialist's professional body. The latter national program focuses on specialist group performance, assessed in highly protocolled peer site-visits. Participation in the *visitatie* program is mandatory since 1995. Since the performance of *individual* medical specialists is not the primary focus of the *visitatie* program, the

Dutch medical regulatory and registration authorities recently added a fourth re-registration requirement. As per 2020, specialists (4) must demonstrably work on their individual professional development.¹³ This new requirement is referred to as the IFMS (Individual Functioning of Medical Specialist) procedure.⁴ Meeting the requirements for re-registration is the responsibility of the individual specialist. The medical specialists in this study are specialists working in an academic medical center. Because academic specialists are employed by the hospital, the compliance is a shared responsibility of the medical staff and the Executive Board of the hospital. Therefore, they collaborate in the organization, implementation and safeguarding of these procedures.

IFMS system

The IFMS requirement consists of a five-yearly evaluation of individual performance and dictates the use of multisource feedback (MSF). The procedure entails three consecutive steps: (1) take part in the collection on MSF, (2) discuss the feedback with a trained facilitator with whom there is no hierarchical or otherwise dependent relationship, and (3) define and yearly evaluate/adjust a PDP. Every medical specialist must maintain a portfolio in which he or she shows evidence of his or her professional performance activities and outcomes covering all professional (medical and generic) competencies.¹⁵ The objective of the IFMS procedure is to maintain or enhance the quality of specialists' individual performance; its purpose is not to identify malfunctioning specialists. Other systems are available to manage (suspected) malfunctioning or poor performance.⁴

IFMS system in academic medical centers

Participation in IFMS is mandatory for all Dutch medical specialists, regardless of their healthcare organization. Because academic medical specialists are employed by the hospital, they are obliged to demonstrate accountability for their performance to the hospital. Within the hierarchical structure of academic medical centers, there is an existing procedure involving an annual review of the specialist's performance. The annual review is executed by the specialist's hierarchical superior, i.e. head of department (HoD), and provides a means to review the specialist's individual career development in alignment with the strategic objectives of the hospital. This annual review allows for customization of individual development, but does not take into account actual performance feedback from peers and coworkers.¹⁶ The new to be designed and implemented IFMS system must be embedded within the existing structure of annual reviews. Redundancy between the IFMS procedure and the procedure of annual reviews had to be prevented or at least minimized. On the one hand, the specialist needs to be able to safely discuss his or her performance with a non-hierarchical facilitator; on the other hand, it must be assured that the specialist

can be supported by his or her superior in intended developmental activities. The present study is about the design of the IFMS system in one specific hospital, the Amsterdam University Medical Centers, location Academic Medical Center (AMC).

Study design

In this study, we adopted a design-based research (DBR) approach. DBR is a fruitful approach for the (re)design of work-based learning environments and assessment programs. The five important characteristics of DBR are as follows: (1) it takes place in continuous cycles of design, evaluation and redesign; (2) it takes place in an authentic practice context; (3) it is aimed both at testing and refining theories and also advancing practice; (4) it is characterized by the use of mixed-methods; (5) designers, researchers and practitioners with different expertise interact frequently and share their ideas.¹⁷ In the following two sections, we will elaborate on the role and composition of the project team, the stakeholders involved and the different steps in the design, evaluation and redesign of the IFMS system.

Project team and stakeholders

The Executive Board of the hospital assigned the design of the IFMS system, from protocol development through evaluation and implementation, to a hospital-based research group. This research group, named Professional Performance and Compassionate Care (PP&CC), has extensive experience in specialists' performance evaluation and improvement (www.professionalperformance-amsterdam.nl). Four members of PP&CC (KL, RS, BB and EB) formed the project team, all also co-authoring this paper. The other two authors (SH, AS) have extensive experience in educational research and development and acted as a sounding board in the conception and drafting of this article. In 2016 the project team sought the opinions of both internal and external experts on medical leadership, hospital administration and representatives of the Dutch medical regulatory and registration authorities. This was done to assure a good position of the IFMS system within the organizational context of the hospital. Expertise from a psychologist (AN) was sought to develop an approach to discuss the performance feedback, which would facilitate specialists' professional development. The hospital's specialist group of (17) neurologists was found willing to pilot this approach. Involvement of various stakeholders was essential to tailor the new system to the specific hospital context. For example, coordination with department heads was needed to let the IFMS system fit in with the existing procedure of annual reviews.

Design, evaluation and redesign

A cyclic process of defining underpinning design principles, evaluation activities and refinement was the basis for the development of the IFMS procedure in the AMC. The process consisted of five steps:

1. Development of an IFMS protocol
2. Selection of an MSF tool;
3. Development, piloting and evaluation of a coaching approach;
4. Recruitment, training and evaluation of coaches
5. Hospital-wide implementation

At the beginning of the design process, the project team together with the hospital board, set out a number of design principles or criteria to be taken into account in the system's design:

- First, the IFMS protocol would have to be developed in consultation with the various stakeholders and should align with the ambitions of the AMC regarding the integration of the IFMS procedure with the existing procedure of annual reviews.
- Second, the MSF tool should cover the broad range of competencies as described in the competency framework used by the Dutch Central College of Medical Specialists (CCMS).¹⁸ Also, the tool should be scientifically sound, easy to use and avoid administratively burdening the specialists being evaluated as well as their assessors. Other feedback tools known to the hospital's specialists were to be included as well.
- Third, the coaching approach to discuss the feedback report had to optimally facilitate specialists' professional development. In practice, this meant: facilitating interpretation of feedback data, supporting specialists in setting developmental goals in the context of their own practice, and develop a route to improvement and change.
- Forth, the coaches needed to be skilled, trustworthy and have close affinity with the medical workplace. Also, as a means to increase specialists' involvement in the IFMS system, the system needed to be based on the principle 'for specialists, by specialists'. For these reasons, coaches should be recruited among the medical specialist workforce (peer coaches).
- Fifth, the redesign process would be subjected to evaluation and research. It was thus decided to pilot-test the IFMS procedure in one department, using several qualitative methods. Evaluation results would serve as input for redesign or adjustment of the procedure. After piloting the procedure and recruiting and training peer coaches, the procedure would be evaluated in two other departments at the start of hospital-wide implementation. Thereafter, the implementation process would be scaled up to include all medical specialists. Last, the handling of the IFMS procedure would ultimately have to be transferred from the (temporary) project group to the organization.

RESULTS

In this section, we will subsequently describe the phases of the design and development as indicated in the methods section.

1. IFMS protocol

In consultation with the various stakeholders and based on incremental evaluation (see further), the system was written down in detail in the IFMS protocol. Since the AMC was the first Dutch academic medical center to present its solution to the new IFMS requirement, the AMC offered its protocol for review to the (national) medical registration authority (RGS). The RGS sanctioned the protocol. Next, the protocol was approved by the hospital's highest medical leadership and the Works Council, and finally ratified by the hospital's Executive Board.

The protocol describes the process of collecting MSF data, the (planning of the) coaching session, the documentation of the session, drafting the personal development plan (PDP), and the reporting and discussion of the PDP with the (hierarchical) peer superior (see Figure 3.1). It contains specific articles reflecting 'the middle ground' between what is mandated by law and what is preferred in terms of prioritizing and promoting professional growth. Notably, the protocol contains the 'negotiated compromises' related to confidentiality of the coaching session and (non-)disclosure of its findings. Specifically, the protocol determines the face-to-face (peer) coaching session as strictly confidential and the medical specialist is invited to bring any (performance) topic to the table for discussion. Also, the protocol prescribes that the resulting navigation document – a preliminary PDP - (see Figure 3.1) remains confidential; it is up to the specialist whether or not to share this document with others, including his or her superior. Following the protocol, the only thing the coach is allowed to report to the medical specialist's superior is that the coaching session has taken place. Specialists may craft their navigation document into a final PDP. In contrast to the navigation document, the final PDP which is modified and approved by the specialist, is sent to the superior for discussion during the annual review. The superior will confirm and note in the specialist's HR file that the specialist in question has met the IFMS requirement; this documentation serves as prove of compliance to be forwarded to the authorities, the RGS.

Some articles in the protocol show how the AMC chose for exceeding the minimum legal requirements in favor of facilitating professional development:

- First, instead of once every five years as required by law, it was decided to facilitate participation in the IFMS procedure twice every five years.
- Second, for intervening years, the protocol recommends to additionally

discuss the progress of the PDP with peers during a peer group meeting, under supervision of an external coach.

- Third, in addition to the required MSF data (see further) all specialists are expected to also submit other available systematically collected performance information. In particular, medical specialists should submit their feedback reports on their supervisory and teaching skills as collected through the well-validated System for Evaluation of Teaching Qualities (SETQ).¹⁹

Figure 3.1 Timeline IFMS trajectory in the Amsterdam Academic Medical Center

2. MSF tool: INCEPT

The INCEPT (INviting Co-workers to Evaluate Physicians Tool) questionnaire was selected to collect multisource feedback.²⁰ The INCEPT had been previously developed by the research group PP&CC and had been investigated for its psychometric qualities and feasibility. The INCEPT covers three domains of professional performance: professional attitude, patient-centeredness and organization and (self-)management. It is a 21 item questionnaire and facilitates

respondents' narrative comments. Collection of feedback takes place in a web-based environment, through which specialists invite 8 peers, 8 co-workers and 8 residents to evaluate their performance and also fill in a self-evaluation. Once the feedback is gathered, on average after four weeks, an individualized MSF report is generated. Specialists can download this report through their personal online profile and send the report to the coach. They are encouraged to also include other sources of feedback such as residents' evaluations regarding training activities. Optionally, specialists can fill in a reflection form in the same web-based environment in preparation for the coaching session.

3. Coaching approach: DANA

Development of the DANA

A coaching approach was developed by a psychologist (AN), who had approximately 20 years of experience in coaching medical specialists. The newly developed approach was based on Appreciative Inquiry practices and solution-focused coaching and was named the Developmental Appreciative Navigational Approach (DANA). The aim of DANA was to optimize the use of each individual specialist's talent, to identify visible and hidden personal qualities and to formulate an overarching professional development goal. To enable the specialist to expand on intrinsic motivation and individual qualities and to prevent the coaching session from becoming a tick-box-exercise in which the coach would skim through the feedback, the duration of the coaching session was set at two hours. The session was preferably held in the privacy of the specialist's office, where confidentiality was assured at the start of the session. The session resulted in the formulation of a (short) navigation document summarizing the specialist's qualities and ambitions, a description of one or more identified developmental goals and multiple concrete actions, preferable set in time and place to achieve those goals. An overview of the substantive elements of the DANA is displayed in Table 3.1.

Evaluation of the DANA - methods

The newly developed DANA was pilot-tested in the department of neurology; all neurologists participated in an individual coaching session with the experienced psychologist. Within two weeks after the session, a researcher (KL) conducted semi-structured interviews to evaluate how they had experienced the DANA session in the light of their professional development. The interview contained questions about how the session had contributed to their insight into personal qualities and developmental opportunities. In addition, several grading scales were used, such as "I have experienced this session as 'empowering' vs. 'discouraging'/'appreciative' vs. 'judgmental'/'safe' vs. 'unsafe'. Furthermore, specialists were asked how they had experienced the role of the coach and whether they had any objection to discuss the documentation of the coaching session with their superior during their

Table 3.1 *Steps of the Developmental Appreciative Navigational Approach (DANA)*

Step	Content
1. Introduction	Build a foundation for a trusting relationship.
2. Confidentiality	Discuss the purpose of the coaching session and explain the procedure.
3. 'In your element'	Explore the ideal work situation, when the medical specialist is able to work in a way that gives energy and brings out the best in him/her.
4. Desired future development	Explore ambitions and wishes for the future and satisfaction with the current range of duties.
5. Performance feedback data	Recognize a pattern in the personal feedback data.
6. Personal qualities	Name qualities and connect with feedback data.
7. Improvement goal	Identify and develop a plan for improvement; use scale questions to gain insight into motivation and self-efficacy.
8. Document	Jointly reflect on the session and provide a written summary of the content of the session and the development plan discussed.

annual review. Lastly, they were asked about the appropriateness of the five-yearly frequency of this coaching session. Interviews lasted for 30-50 minutes and were audio-taped. The content of the interviews was summarized and used as input for adjustment of the procedure. The interview protocol is displayed in Appendix 3.1.

Evaluation of the DANA - results

The pilot test in the neurology department took place from January till March 2017. All 17 neurologists rated the session as highly motivating, appreciative, development-oriented, concrete and valuable. They characterized the session as open and safe. For most of them, the session did not reveal new insights about their performance. Rather, it confirmed existing ideas, or was instrumental in rendering, accepting or internalizing the feedback received. The session did however deepen their insights in their developmental opportunities and improvement goals. For all neurologists, at least one improvement goal was formulated. They reported to be highly committed to and motivated for this goal. The role of the coach was greatly appreciated and unanimously rated with a minimum score of 8 out of 10. Some expressed to be interested in additional coaching sessions to follow up on the defined development goals. Most neurologists reported to consider peer coaches to potentially also be good coaches, although some expressed their preference for a professional coach. This preference was related to both the confidentiality of the coaching session and the expertise of a professional coach. Criticism was expressed in terms of the disclosure of the coaching document to their superior, i.e. the Head of Department, as this could hamper the openness during the coaching session. The neurologists supported the proposal to schedule a coach session every two years,

instead of once every five years as required by law. Based on the evaluation of the pilot test the protocol was adjusted.

4. Coaches

Recruitment and training of coaches

Given the importance for coaches to be trustworthy, and bearing in mind the requirement of budget neutrality, coaches were recruited among the hospital's own medical specialist workforce. The Executive Board sent out a letter to all senior medical leadership – i.e. chairs of the organizational divisions- in which they were requested to nominate a number of potential coaches in proportion to the number of medical specialist staff working in the division. Qualities such as good communication and empathic skills were emphasized as crucial selection criteria. The coaching job would take approximately 30 - 40 hours per year for eight to ten coaching sessions, including preparation and drafting the navigation report. At least once a year, coaches would participate in a two-hour meeting with other coaches to exchange experiences and improve skills. The appointment as a coach was set for the duration of three years. Participation in a training guided by the experienced psychologist was mandatory. The training took eight hours and consisted of information transfer, direct teaching, experiential personal development exercises and coaching skills practice, using role play. For participation in the training, eight accreditation points were awarded.

Evaluation of coaches

Peer coaches were recruited among the medical specialist workforce in July 2017. In October 2017, ten peer coaches were trained by the psychologist (AN). Subsequently, the project team planned coaching sessions for them within two departments, neonatology and rehabilitation medicine. Within two weeks after the coaching sessions in these departments, two researchers (BB, RS) conducted semi-structured interviews with specialists to evaluate the session and the role of the coach, using a shortened version of the pilot interview protocol. Interviews lasted for 15-20 minutes and were audio-taped. The content of the interviews was again summarized and used as input for adjustment of the procedure.

Fifteen specialists were interviewed; their experiences were comparable to those of the neurologists in the pilot phase. However, a number of specialists noted that the emphasis in the coaching session had been very much on the appreciation of positive qualities, leaving less room and attention for the formulation of developmental goals and drafting a plan for attaining these goals. In response to this, more attention was paid to strengthening the developmental orientation of the coaching conversation in the DANA training and the meetings with the coaches, by paying special consideration to the factors which may both impede and enable progress

and success, setting timelines and considering how to measure success. Once again, the evaluation results showed that specialists sometimes had a preference for a professional non-peer coach. Based on these results, the inclusion criteria for the recruitment of additional coaches were extended to also include psychologists and professionals who otherwise had insight into the medical specialist workplace. A list of external coaches was compiled to be able to refer specialists to a professional coach for additional sessions if desirable or necessary, which could be funded from their personal budget.

5. Hospital-wide implementation

Gradual scale-up

After (initial) implementation in the departments of neurology (pilot) and neonatology and rehabilitation medicine in 2017, the IFMS procedure was implemented in another six departments in 2018. The project team contacted the heads of these departments and informed the medical staff about the IFMS procedure. Practical matters such as the management of MSF through the web-based environment were entrusted to secretary staff, who were provided with manuals and templates. Upon request, heads of department were provided with a training on how to follow-up on the coaching session during the annual review in a positive and development-oriented manner. In October 2018, the group of coaches was expanded with nine coaches. In the same period, an online planning tool for efficient scheduling of sessions was introduced. Coaches registered their availability and provided a profile text about themselves, so that medical specialists could make an informed decision for a coach. To support coaches in their development as a coach, coach meetings were organized which were supervised by the psychologist. By the end of 2018, 96 specialists had taken part in the IFMS procedure.

Transfer to HR department

In May 2019, the design, evaluation and implementation of the IFMS system was completed and the procedure could be embedded into the organization. The operational responsibility of the IFMS procedure was transferred from the research group PP&CC to the hospital's Human Resources (HR) department. The research group continued to support and advise the HR staff on content matters, and also continues investigations into the follow-up on coaching sessions. The HR department supported continuous recruitment and training of coaches, as well as the scheduling of coach meetings and the monitoring of the degree of participation and lead time on department level. In 2019, another 12 departments took part in the new system and the group of coaches was brought to final strength with the addition of 11 coaches, resulting in a total number of 30 coaches. By the end of 2019, 162 out of 624 specialists in the hospital had completed the IFMS procedure.

DISCUSSION

This study reports on a design-based research (DBR) development of a new performance appraisal system that serves both specialists' professional development needs and mandatory accountability, or re-registration requirements. The new system reflects the compromises made in dealing with the conflicting nature of accountability and development purposes. Characteristics that appeared crucial and non-negotiable for the facilitation of professional development were an appreciative and development-oriented performance appraisal process, the trustworthiness and skills of the peer coaches and a non-disclosed appraisal output. The DBR approach contributed to the successful implementation and positive reception by specialists, organizational leadership and regulators.

The operational heart of the new system is the newly developed coaching approach to guide peer-to-peer conversations. The two hour sessions were designed to maximize meaningfulness of the conversation for the specialists. Peer coaches were trained to create a positive and safe conversation by stressing the professional and personal strengths of the specialist, offering guidance in specific and realistic goal setting and in the development of a concrete action plan. The evaluation results of the pilot and the initial implementation showed that specialists highly appreciated the session and the coaches. The person of the coach and the coach-coachee relationship was found to be important. Peer coaches were considered trustworthy given their familiarity with the working context. The choice for peer coaches was supported by the principle of budget neutrality. Nevertheless, some specialists expressed a preference for an external professional coach, for two reasons. First, specialists valued the specific skills and ample experience of a professional coach. Second, specialists attached great importance to a coach who is not part of "the system", this was particularly true for the more senior specialists in leadership positions. Both expertise and confidentiality were considered of special importance because of the potential sensitivity of the topics that were discussed during the coaching session. Although we did not systematically collect these data, anecdotal evidence showed that specialists struggled with high workloads, organizational changes, wellbeing problems or issues related to interpersonal dynamics with colleagues and managers. This is in line with research on the relationships among specialists' work load, social support and well-being.²¹⁻²³ To build coaching expertise, it was found important to ensure that the peer coaches conduct sufficient coaching sessions per year. Therefore coaches are asked to commit to eight to ten sessions. Although formally they are granted dedicated time for this, coaches' availability remains a challenge in combination with busy clinical practice. In the second round of recruiting coaches, a few non-specialists (i.e. from the HR department) were also selected and trained to ensure a larger and a

more diverse pool of coaches. This combination of peer coaches, non-peer coaches and external coaches may very well be the ideal pool to work with.

One of the tougher challenges in the design of the system was to effectively combine the new non-hierarchical, formative coaching session and the existing hierarchical annual review with the peer superior. The initial proposal of offering specialists a personal coaching session - to facilitate in-depth reflection and a tailor-made PDP - which was then to be presented to their superior for the purpose of supporting the implementation of the PDP in practice, was not deemed feasible. Medical specialists brought to the table that sharing the results of the coaching session would potentially inhibit the agenda setting, openness and depth of the conversation. In practice, this could mean that areas in need of improvement, or personal ambitions that conflict with the strategic objectives of the organization, would remain underexposed in the coaching session when this information would be forwarded to management. In other words, disclosing results would seriously weaken the formative function of the coaching session. Moreover, the specialists claimed that sharing the outcomes was *not* mandated for complying with the re-registration requirements; therefore some suggested it could be considered an (illegitimate) intrusion of their autonomy. Instead, specialists opted for the separation of the two procedures; in their view autonomy, safety and confidentiality should take precedence over involvement of organizational leadership. This view may indicate a lack of trust in or experience of the organization in the personal development of medical specialists. Although in the literature it has been pointed out that the traditional professional – manager dichotomy is no longer a valid reflection of contemporary professions and organizations,²⁴⁻²⁶ the experience of such a dichotomy may still be present. The compromise found was to give the specialist ultimate say in (not) sharing parts or all of the personal development plan with their peer superior. Although the future will have to show how this turns out in practice, we expect that most specialists will decide to share their coaching results. Nevertheless, we consider the negotiated compromise an important success factor for meaningful conversations.

In times of bureaucratic overload of medical specialists,²⁷ it was crucial that the new system would not add too much administrative burden to the medical specialists and the organization in general. This was largely accomplished by using the existing –for specialists familiar–platform for the mandatory MSF data collection and by generously supporting the specialists in every step of the process, i.e. by planning and organizing the data collection and by having the coaches draft the specialist's professional development plan. In particular, the timing of the appraisal process within a department, to make it fit with other ongoing projects, showed to be important in specialists' commitment to providing meaningful feedback to their peer specialist being evaluated. After the system's implementation, the HR department continues facilitation of the appraisal procedure, recruitment and training of

coaches, and communication with departmental and executive leadership. The organization's investments in HR support also shows its commitment to and felt responsibility for a high performing medical staff.

Strengths and limitations

In accordance with the principles of DBR, the study was carried out as a cyclic process of design, evaluation and redesign, in an authentic real-life setting, in which the researchers, designers and practitioners interacted frequently to share their ideas. The DBR approach allowed us to propose, test and amend solutions for the challenging goal of satisfactorily serving the different purposes of accountability and professional development. The resulting system, as reported in this paper, reflects the various compromises that were reached by extensive deliberations. For the success of the system, during the design process we constantly kept in mind the importance of the quality of the feedback conversation. We did this by developing a new coaching approach and ensuring a thorough evaluation, first when the conversation was conducted by a psychologist and later again when the conversation was conducted by a peer coach. Since this study took place in the local setting of one Dutch academic medical center, its findings may not generalize to other contexts. However, the fact that other (academic) medical centers have copied our protocol or used it as a basis to build their own, suggests that the design is at least to some extent transferrable to other settings. Nevertheless, we did undertake steps to stimulate a broader relevance than the local situation by including a detailed description of the context of the study and formulating clear design principles. Similar to many other DBR studies, the roles of researcher and designer were fulfilled by the same persons. In order to prevent the findings of the study to be influenced by the researchers' biases, two researchers from another institution (AS and SH) were involved in the research process.

Future research

For future evaluation purposes, it is crucial to invest in a permanent evaluation program to enable longitudinal research into the effects of the new appraisal process on specialists' performance and professional development. By using mixed methods, different sources of information can be triangulated that explain why the process is successful under which circumstances and which context. Continued close collaboration and interaction between IFMS researchers and the HR department is crucial. Given the nature of the topics discussed during a coaching session, it is important to use person-centered measures of effectiveness focused on work engagement, wellbeing and interpersonal communication. These self-reported measures should be complemented with more distal outcomes of coaching related to the wider organization, such as impact on team climate and overall hospital

culture. In this regard, the proposed group meetings to discuss the progress of the PDP with peers, led by an external coach, may offer additional opportunity for gaining insight into how the new system can contribute to performance change. In the event that the appraisal procedure and the annual review procedure remain separated, hospitals need to be – in some way – informed about the content or results of the coaching sessions. This is imperative, as hospitals and medical specialists bear a joint responsibility for the quality of care.

CONCLUSIONS

In this study, we designed a new performance appraisal procedure for medical specialists undergoing re-registration that would successfully comply with both accountability requirements as well as fulfill the need for professional development. The study took place in the local context of a large academic medical center in the Netherlands and the findings can be translated into design guidelines broadly applicable in other hospital settings, and even healthcare systems. In the context of mandatory re-registration, professional performance and development can be supported when the appraisal process is appreciative and development-oriented, when the coach is skilled and trustworthy and when the medical specialist has control over the disclosure of the output of the process to others.

REFERENCES

1. Brennan N, Bryce M, Pearson M, Wong G, Cooper C, Archer J. Understanding how appraisal of doctors produces its effects: a realist review protocol. *BMJ Open*. 2014;4(6):e005466.
2. General Medical Council. Ready for revalidation; the Good medical practice framework for appraisal and revalidation. Published March 2013, Manchester, United Kingdom.
3. American Board of Medical Specialties. Standards for the ABMS Program for Maintenance of Certification (MOC): For implementation in January 2015. 2014. www.abms.org/media/1109/standards-for-the-abms-program-for-moc-final.pdf. Accessed 16 April 2020.
4. Orde van Medisch Specialisten. Leidraad IFMS. 2014. https://www.demedischspecialist.nl/sites/default/files/Leidraad%20IFMS_definitief.pdf. Accessed 16 April 2020.
5. McGivern G, Ferlie E. Playing tick-box games: Interrelating defences in professional appraisal. *Hum Relat*. 2007;60(9):1361-1385.
6. Chamberlain JM. Portfolio-based performance appraisal for doctors: a case of paperwork compliance. *Sociol Res Online*. 2010;15(1):68-76.
7. Holmboe ES, Ross K. Commentary: realizing the formative potential of multisource feedback in regulatory-based assessment programs. *Acad Med*. 2012;87(12):1657-1659.
8. Lockyer J, Horsley T, Zeiter J, Campbell C. Role for assessment in maintenance of certification: physician perceptions of assessment. *J Contin Educ Health Prof*. 2015;35(1):11-17.
9. Sargeant J, Bruce D, Campbell CM. Practicing physicians' needs for assessment and feedback as part of professional development. *J Contin Educ Health Prof*. 2013;33(S1):S54-S62.
10. Sargeant J, Armson H, Chesluk B, Dornan T, Eva K, Holmboe E, van der Vleuten C. The processes and dimensions of informed self-assessment: a conceptual model. *Acad Med*. 2010;85(7):1212-1220.
11. Sargeant J, Lockyer J, Mann K, Holmboe E, Silver I, Armson H, Power M. 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):1698-1706.
12. Sargeant J, Lockyer JM, Mann K, Armson H, Warren A, Zetkovic M, Holmboe E. The R2C2 model in residency education: how does it foster coaching and promote feedback use?. *Acad Med*. 2018; 93(7):1055-1063.
13. College Geneeskundige Specialisten. Besluit herregistratie specialisten. 2015. www.demedischspecialist.nl/sites/default/files/Besluit%20herregistratie%2013%20juli%202016.pdf. Accessed 16 April 2020.
14. Seddon N. Is the future Dutch?. *The Lancet*. 2008;372(9633):103-104.
15. Federatie Medisch Specialisten. Visiedocument Medisch Specialist 2025. 2017. <https://www.demedischspecialist.nl/sites/default/files/Visiedocument%20Medisch%20Specialist%2025-DEF.pdf>. Accessed 16 April 2020.
16. Sehlbach C, Govaerts MJ, Mitchell S, Rohde GG, Smeenk FW, Driessen EW. Box-ticking and Olympic high jumping—physicians' perceptions and acceptance of national physician validation systems. *Med Teach*. 2018;40(9):886-891.
17. Dolmans DH, Tigelaar D. Building bridges between theory and practice in medical education using a design-based research approach: AMEE Guide No. 60. *Med Teach*. 2012;34(1):1-10.

18. Centraal College Medisch Specialisten. Kaderbesluit CCMS. 2009. Available on www.knmg.nl.
19. Boerebach BC, Lombarts KM, Arah OA. Confirmatory factor analysis of the system for evaluation of teaching qualities (SETQ) in graduate medical training. *Eval Health Prof.* 2016;39(1):21-32.
20. Van der Meulen MW, Boerebach BCM, Smirnova A, Heeneman S, oude Egbrink MGA, van der Vleuten CPM, Arah OA, Lombarts MJMH. Validation of the INCEPT: A Multisource Feedback Tool for Capturing Different Perspectives on Physicians' Professional Performance. *J Contin Educ Health Prof.* 2017;37:9-18.
21. Aalto AM, Heponiemi T, Josefsson K, Arffman M, Elovainio M. Social relationships in physicians' work moderate relationship between workload and wellbeing—9-year follow-up study. *Eur J Public Health.* 2018;28(5):798-804.
22. Heponiemi T, Kouvonen A, Vänskä J, Halila H, Sinervo T, Kivimäki M, Elovainio M. Effects of active on-call hours on physicians' turnover intentions and well-being. *Scand J Work Environ Health.* 2008;34(5):356-363.
23. Dewa CS, Loong D, Bonato S, Thanh NX, Jacobs P. How does burnout affect physician productivity? A systematic literature review. *BMC Health Serv Res.* 2014;14(1):325.
24. Waring J, Currie G. Managing expert knowledge: organizational challenges and managerial futures for the UK medical profession. *Organ Stud.* 2009;30(7):755-778.
25. Noordegraaf M. Hybrid professionalism and beyond:(new) forms of public professionalism in changing organizational and societal contexts. *Journal of Professions and Organization.* 2015;2(2):187-206.
26. Noordegraaf M. Risky business: How professionals and professional fields (must) deal with organizational issues. *Organ Stud.* 2011;32(10):1349-1371.
27. Sinsky C, Colligan L, Li L, Prgomet M, Reynolds S, Goeders L, Blike G. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. *Ann Intern Med.* 2016;165(11):753-760.

APPENDIX 3.1.

Interview protocol for the evaluation of the DANA / coaching session

The purpose of this interview is to evaluate your recent coaching session. Based on the results, an advice will be formulated regarding the implementation of coaching sessions for all medical specialists. Central to the evaluation is your experience with the Developmental Appreciative Navigational Approach (DANA) used in the coaching session.

1. Looking back on the session, which three words best describe your experience?
.....
2. a. Did the coaching session provide you with a (better) insight into your qualities as a medical specialist?
surely / slightly / hardly / not at all
b. Can you indicate what contributed to (not) gaining this insight?
.....
3. a. Did the coaching session provide you with a (better) insight into your areas for improvement as a medical specialist?
surely / slightly / hardly / not at all
b. Can you indicate what contributed to (not) gaining this insight?
.....
4. Can you characterize the coaching session as you experienced it on the following scales(put an X on the line):

Discouraging _____ Empowering

Normative _____ Educational

Energy-consuming _____ Energizing

Unsafe _____ Safe

Judging _____ Appreciative

Focus on weaknesses | | | Focus on strengths
| | | |

Closed | | | Open
| | | |

De-motivating | | | Motivating
| | | |

Worthless | | | Valuable
| | | |

Superficial | | | Profound
| | | |

Tense | | | Relaxed
| | | |

Little development-oriented | | | Highly development-oriented
| | | |

Abstract | | | Concrete
| | | |

Theoretical | | | Practical
| | | |

5. a. Have you been able to draft a preliminary personal development plan (PDP) during the coaching session?
Yes / No

b. Does the preliminary PDP contain the most important issues related to your performance and development?
Yes / No

c. On a scale from 1 to 10, to what extent do you feel committed to realizing/ implementing this PDP?

1 = I don't feel committed at all

10 = I feel highly committed

My grade:

d. Expressed in percentages, how do you estimate the chance that you will have realized your PDP in one year?

..... %

e. Are there any objections for you to discuss this PDP with your superior during the annual review?

Yes / No

If yes, can you explain your objections

.....

6. The coaching session is in principle a one-time session, with a follow-up during the annual review (discussion of the PDP). How do you feel about this as a starting point?
-

7. Do you consider the IFMS procedure (multisource feedback + coaching session + PDP) relevant for re-registration?

Yes / No

Any explanation?

.....

8. If you could summarize the coaching session in an image or a metaphor, which image or metaphor would you choose?
-

9. Based on your experience with the coaching session, what would be your advice on coaching sessions in the Amsterdam Academic Medical Center (AMC)?
-

Thank you for your cooperation. This data will be processed anonymously. Results will be reported in such a way that they cannot be traced to individuals.

4

SHARING REFLECTIONS ON MULTISOURCE FEEDBACK IN A PEER GROUP SETTING: (HOW) DOES IT STIMULATE PHYSICIANS' PROFESSIONAL PERFORMANCE AND DEVELOPMENT?

Bindels, E
Van den Goor, M
Scherpbier, AJJA
Lombarts, MJMH
Heeneman, S

Under editorial review

ABSTRACT

Introduction: Reflecting on and using feedback is important for physicians' continuous professional development (CPD). A common format is the discussion of multisource feedback (MSF) in a one-on-one session with a trusted peer or coach. A new approach is to discuss MSF in a collegial group setting. This qualitative study explored how physicians experience participation in a professionally facilitated peer group session in the context of their CPD.

Methods: Between March and July 2018, 26 physicians were interviewed, representing 13 mono-specialty physician groups from 5 Dutch general hospitals. Each group participated in a session, led by a professional facilitator, in which they reflected on their personal MSF data. Interviews were transcribed verbatim and analyzed iteratively, following the interpretative phenomenological approach.

Results: Participation was experienced as a process of disclosing and sharing personal reflections with others while striking a balance between interpersonal proximity and distance to peers. Sharing reflections with others rendered the feedback more meaningful, deepened collegial relationships and created a sense of urgency for improvement. Improvement goals were mostly related to relational fine-tuning in collaboration; goals related to individual career management remained in the background. Influential factors for the effectiveness of the group sessions were related to the facilitator's expertise, group size, continuity and quality of collegial relationships, personal vulnerabilities and the context of CPD policy.

Discussion: Sharing personal reflections on MSF in a peer group setting may contribute to physicians' CPD through a reinforcement of collegial relationships, generating opportunities for performance improvement. Conditions for success are an experienced facilitator, longitudinal trusting relationships and a limited group size.

INTRODUCTION

Inherent in being and remaining a 'good doctor' is the motivation for ongoing learning and the ability to engage in reflection on one's professional performance.^{1,2} 'Professional performance' extends beyond the ongoing acquisition of and reflection upon medical knowledge and procedural skills; it also encompasses ongoing development of skills such as interpersonal communication, collaboration and leadership and aspects of personal development and well-being.³⁻⁶ In the context of requirements for both continuous professional development (CPD) and maintenance of certification (MoC) or revalidation, physicians are expected to periodically reflect on their professional performance.^{4,6} In many health care systems, i.e. in the USA, Canada, the UK and the Netherlands, it is common to include the results of multisource feedback (MSF) in this reflective activity.^{7,8} Research on how physicians use MSF, however, has demonstrated that MSF does not self-evidently find its way into physicians' performance change; multiple studies underscore that *facilitation* of reflection is required to achieve actual improvement.⁹⁻¹²

Indeed, physicians who have been assessed using MSF approaches have identified that a facilitator is helpful in reflecting on feedback. It can promote a deeper understanding of and insights into one's own behaviors and opinions about factors that could enhance the impact of the feedback.¹²⁻¹⁴ One-on-one sessions have shown to result in clear goal setting and intention to change; the effectiveness of facilitated reflection is dependent on the development of rapport, trust and mutual respect between the physician and the facilitator.^{15,16} In addition to one-on-one sessions, also *group* sessions are being explored. During a group session, which is moderated by a professional facilitator, physicians are invited to reflect on their personal MSF data with same-specialty peers, with whom they work together on a daily basis. The mechanisms through which a group session can impact CPD still need further investigation.

There are obvious pragmatic reasons for choosing group sessions instead of individual sessions, i.e. group sessions are more time and cost effective. Nevertheless, reasons of principle may also underpin such a choice. Although group sessions have not yet been studied in the specific context of MSF, reflecting on MSF in a peer group setting may provide opportunities for deep engagement in the learning process, as physicians are learning through their relationships with peers. Furthermore, a peer group setting can create a space for self and peer assessment of a formative kind by providing a context for comparing oneself to others.¹⁷ Several authors have indicated that peers can be viewed as instructional resources for one another, sharing learning intentions and criteria for success.¹⁸⁻²¹ Also, sharing personal reflections with colleagues could improve the quality of collegial relationships and increase the chance of real performance improvement.¹⁵

The question arises whether and how group sessions indeed are helpful for physicians to reflect on their professional performance and contribute to their development. In this study, we therefore explored three research questions: in the context of (mandatory) CPD in which MSF is used, (1) how do physicians experience participation in group sessions? (2) how do they perceive the session to contribute to their professional performance and development? and (3) which factors do they perceive as hindering or helpful in this process?

METHODS

Setting and participants

Within the context of CPD, Dutch medical specialists periodically take part in performance appraisal. This process is characterized by maintaining a portfolio and acquiring feedback, which together form the basis for a personal conversation with a trained facilitator. As of 2020, participation in performance appraisal is mandatory for revalidation. This revalidation requirement entails three consecutive steps: (1) take part in the collection of MSF every five years, (2) discuss feedback with a trained facilitator with whom there is no hierarchical or otherwise dependent relationship, and (3) define and yearly evaluate a PDP.²² Initially only one-on-one sessions with a facilitator were used, but gradually also group sessions are being used. For our study, we conducted individual, semi-structured interviews with physicians from Dutch general hospitals who had participated in a group session. In our study, MSF data was gathered through an online system collecting anonymous feedback from three groups of respondents: peers (same-specialty peers and peers from adjacent specialties), residents, and supporting staff (e.g. nurses, pharmacists, therapists and secretary staff).²³ The feedback results were collated into a personal report. In our study, participants were invited to discuss their reports with their peers. These peers had provided feedback through the online system and they belonged to the same medical specialty in the same hospital, with whom the physician in question shared the same patients and worked together as team members. For most participants, it was the first encounter with a group session in the context of revalidation; the choice for a group session was made by the medical staff boards of the hospitals in question.

Group sessions

A professional facilitators would lead the group sessions; in case the group would count more than six participants, there would be two facilitators. The facilitators, ten in total, were all independent consultants from one private firm with experience in the medical specialist workplace. Physicians received an invitation from the

facilitator for a 2-2,5 hour group session. The facilitator communicated the objectives of the session as follows: (1) sharing personal performance strengths and areas for improvement; (2) gaining (better) insight into each other's performance strengths and exploring how the group can learn from this insight; (3) providing opportunities to support each other or explore ways to support each other and (4) determining which topics need concrete action. At the start of the session, the facilitator invited every physician to share one or more points from his or her feedback report and elaborate on aspects that they found to be remarkable, recognizable, confronting or amusing. Hereafter, the facilitator invited peers to share their views on what was shared by the physician. Before the start of the session, the facilitator had taken note of all feedback reports; in the event that a physician would remain silent about important parts of his/her feedback, the facilitator would bring this up in the discussion, albeit in a subtle way. In the invitation to the session it was explicitly stated that the session was confidential; no information was disclosed to others inside or outside the hospital.

Sampling strategy and procedure

We used purposive and convenience sampling to compose a diverse sample of participants, representing physicians from various hospitals and specialties, belonging to physician groups that varied in size and team membership. The facilitator informed the physicians taking part in the group session about the current study via email. In this email, they were asked whether they agreed with the presence of the first author (EB) during the group session. It was emphasized that she had not taken notion of the content of the feedback reports. The purpose of her presence was twofold: on the one hand, it was intended to familiarize herself with the course of events during a group session and to obtain an impression of the group to which participants belonged; on the other hand, it was intended to lower the threshold for physicians to participate in an interview with her after the session. At the end of each session, the facilitator asked whether physicians were willing to participate in an interview. We aimed to conduct two interviews per physician group within two weeks after the group session.

Study design and data collection

Being interested in physicians' experiences with reflection during a group session and their perceptions of its contribution to their performance and development, we chose an interpretative phenomenological approach (IPA).²⁴⁻²⁶ In the literature, the process of IPA is described as engaging in a double hermeneutic, "whereby the researcher is trying to make sense of the participant trying to make sense of what is happening to him (p. 10)."²⁶ Since we asked physicians to reflect on the reflective activity in the group session, the research itself is, in fact, a reflective activity. IPA is

concerned with the balance of convergence and divergence within the sample, not only by searching for shared themes but also pointing to the particular way in which these themes play out for individuals. We collected data by using semi-structured individual interviews. The first author (EB) conducted the interviews between March 2018 and July 2018. Interviews lasted 45-60 minutes and were held in the privacy of the physician's office, where confidentiality was assured at the start of the interview. Based on discussions within the research team, an interview guide with five open-ended questions was designed to structure the interviews. The questions related to (1) the physician's first impression of the group session, (2) the topics that had been shared by the physician and the reactions he/she had received, (3) the physician's experience of the presence of peers/facilitator and the interpersonal dynamics during the session, (4) the perceived contribution of the session to his/her performance and development, and (5) factors that were perceived as helpful or hindering. All interviews were audio-recorded and transcribed verbatim.

Data analysis

Common to IPA, we adopted an iterative approach: we performed data collection and data analysis simultaneously, whereby insights from previous interviews shaped subsequent data collection. As IPA prescribes that the analysis should point to both convergence and divergence, we adopted the strategy of identifying stories of participants who experienced the session as particularly valuable and participants who did not hold this experience. We sought to track down underlying processes by revising interview questions in the next interviews. Two researchers (EB, MvdG) independently read and open coded the first transcripts and compared the naming and interpretation of codes, resolving differences through consensus. During discussions with the research team, emerging themes were identified and subsequent transcripts were coded with these themes in mind. This coding cycle was repeated several times. Data saturation was reached at the point where further data analysis no longer resulted in new analytical themes. A comprehensive understanding was formulated by the first author in consultation with the research team.

Reflexivity

Since IPA is an approach in which researchers bring their own backgrounds, researcher roles and assumptions to the analytical process, practicing reflexivity is critical.²⁵ In this reflexive spirit, we provide the following contextual information: the first author (EB) has a background in arts, humanities and psychology; other authors have significant experience in studying medical education, and their own disciplinary backgrounds include medicine, management and consultancy (MvdG), medicine and philosophy (AS), biomedical sciences and competency-based education (SH)

and health services research (KL). The second author (MvdG) fulfilled a dual role, given her role as a researcher and professional facilitator of two of the 13 group sessions included in this study.

Ethical considerations

This study was exempt from Institutional Board Review under Dutch law. A waiver of ethical approval was provided by the medical ethics review committee of the University Medical Center in Amsterdam, waiver number W18_o89. All participants consented to participate in this study by providing written informed consent.

RESULTS

Participants

In total, 26 physicians from 13 mono-specialty physician groups in five Dutch hospitals participated in the interviews. Table 4.1 shows the details of the participants regarding gender, age, specialty and group size.

Table 4.1 Participants: gender, age, specialty and group

#	Gender and age	Specialty	Group	Group size (3-11)
1	Male in his 40s	Emergency medicine	A	3-5 group members
2	Female in her 30s			
3	Male in his 50s	Geriatrics	B	3-5 group members
4	Female in her 50s			
5	Male in his 40s	Anesthesiology	C	6-8 group members
6	Male in his 40s			
7	Male in his 60s	Pulmonology	D	≥ 9 group members
8	Male in his 60s			
9	Male in his 60s	Internal medicine	E	6-8 group members
10	Female in her 30s			
11	Male in his 40s			
12	Female in her 30s	Emergency medicine	F	6-8 group members
13	Male in his 30s			
14	Female in her 40s	Geriatrics	G	3-5 group members
15	Female in her 50s			
16	Male in his 50s	Orthopedics	H	≥ 9 group members
17	Male in his 40s			
18	Female in her 40s			
19	Female in her 30s	Geriatrics	I	3-5 group members
20	Male in his 60s			
21	Female in her 50s	Medical rehabilitation	J	3-5 group members
22	Male in his 40s			
23	Male in his 40s	Pediatrics	K	≥ 9 group members
24	Male in his 40s	Rheumatology	L	3-5 group members
25	Female in her 40s	Gynecology	M	3-5 group members
26	Female in her 30s			

Overall results

We identified four themes:

1. Disclosing and sharing

Participation was experienced as a process of disclosing and sharing personal reflection with others - potentially eliciting feelings of vulnerability - while striking a balance between interpersonal proximity and distance with peers.

2. Understanding feedback in a broader perspective

The session contributed to insight into own performance, although the content of the feedback was mostly not surprising. Listening to colleagues' explanation and nuance of feedback led to perceptions of increased meaningfulness of the feedback, prompting participants to open up to the perspectives of others and to feel more connected to them.

3. Creating urgency for improvement

The process of sharing reflections and generating interpersonal closeness fueled the urgency for improvement. Improvement goals were mostly related to making adjustments in collaborative practice; to a lesser extent, these goals were related to individual career management.

4. Influential factors

Experiences with participation and perceptions of value for CPD were influenced by factors related to the facilitator, the peer group, the individual physician and the context of CPD policy.

We describe these aspects below. To exemplify our findings, we provide quotes identified by a numerical code assigned to the participant and a letter code assigned to the group session (e.g., participant 5, group session C).

1. Disclosing and sharing

In all cases, except for one, participants had not taken notion of each other's feedback reports before the start of the group session. Participants noted that they had experienced sharing information from their feedback report as something unfamiliar. Most of them reported to have felt a little uncomfortable; some of them even felt somewhat vulnerable and 'exposed':

"I noticed that I needed to cross a threshold and I was glad to notice the same in my colleagues. As they are all quite 'strong personalities', it was reassuring to see them a bit vulnerable too." (13F)

The display of vulnerability led others to feel encouraged to engage with their peer's story:

"Now that I knew about the bottlenecks and uncertainties of my colleague, I felt a bit more free to just say something about that, while I might not do that normally." (14G)

Due to the self-disclosure of peers, participants were not only inclined to respond more to their peer's story, but they were also inclined to let their guard down and engage in self-disclosure themselves:

"You hear about what goes on in others; it was a slightly different atmosphere than I had expected. I opened up more and revealed more of myself and my own struggles." (25M)

Parallel to opening up to each other, participants also emphasized their need to be in control of what they shared and how openly it was discussed. They reported the need to strike a balance between proximity and professional distance. Some participants considered this process of balancing to be fairly straightforward, as one participant stated rather explicitly:

"My colleagues are not my friends; a group of friends at work, I think that's deadly. I need some distance to be able to collaborate with my colleagues in a professional way." (20I)

Another participant articulated this need to strike a balance between proximity and professional distance as a "need for self-monitoring": she noted that the high level of self-disclosure of others could be experienced as an implicit expectation to "return the favor" and share personal information as well. Instead, she chose to only share reflections to the extent that she felt comfortable enough:

"For many of my colleagues, the conversation became very personal. I struggled with that because I'm just new to the group and I didn't want to feel too vulnerable at this point. I think this session should not be a therapy group; you should only share personal things that are relevant for your functioning as a physician.. but that is of course a grey area." (26M)

After all, engaging in self-disclosure in reaction to peers' self-disclosure was a personal consideration. However, as more peers remained on the sideline, this could translate into superficial talk, missing a critical undertone and undermining honest and constructive feedback:

"It was mostly chatting safely about each other's qualities. It was obvious that some of us preferred not to play off the back foot." (11E)

2. Understanding feedback in a broader perspective

Although participants indicated that the feedback in their personal reports was not very surprising, participants did indicate that the session helped them to understand their feedback in a broader perspective. By listening to peers' explanation and nuance of the feedback, participants perceived the feedback as more meaningful and – as a result - were more inclined to open up to their peers' perspectives:

"For me, it was insightful to hear how my colleagues perceive my working style and I too got a better understanding of how their styles developed over time. I noticed that by discussing this directly, but with the presence of a facilitator, I was more open to their side of the story and I became milder in my feedback towards them." (24L)

Participants referred to the group session as an opportunity to re-experience a personal connection, to take time to be "in real contact", by asking and being asked: "How are you doing, but *really*?" (15 G). The aforementioned process of mutual exchange contributed to an atmosphere in which (hidden) emotions were allowed to emerge. In this way, it was possible to discuss matters that were not enclosed under normal working circumstances, such as feelings of insecurity, overload or misjudgment:

"I shared more about personal issues than I thought I would. I think my colleagues already felt these insecurities of mine beneath the surface. Because of their reaction I felt appreciated for who I am. During the session they explored with me what I need to stay on my feet." (25M)

Overall, many participants mentioned the session's beneficial emotional effects, related to feelings of empathy, recognition, appreciation, solidarity and support. Participants underlined the importance of bonding and feeling connected:

"I did not expect to get emotional during the session, but it happened anyway.. In my colleagues' reaction, I felt genuine interest, concern and empathy. I mean, patient contact is very important, but so is working with a group of colleagues you feel comfortable and safe with.. that makes up for three quarter of your job satisfaction." (15G)

3. Creating urgency for improvement

Sharing reflections with peers deepened collegial relationships and created a sense of urgency for improvement. During the session, participants formulated improvement goals, related to personal communication habits and professional pitfalls, such as "leaving more room for nuance during conversations", "involving

others in referral of patients in an earlier stage”, “delegating more/not keeping information to yourself too long”, “letting things take their course” and “having more confidence in own judgment”. Overall, most improvement goals were related to relational fine-tuning in collaboration:

“I learned about the small irritations of my colleagues towards my way of supervising residents. For me, this does not necessarily mean that I am going to change my way of doing things, but perhaps I can better explain why I do things the way I do them. In the long run, this could help to create a mutual understanding of each other’s methods and to open up possibilities for more alignment.” (13F)

Also, participants talked with their peers about their personal and professional challenges, such as work/life balance and wellbeing struggles, and about character traits that are difficult to change but do influence long-term performance, such as perfectionism and fear of failure. In this context, several participants emphasized the importance of being in good touch with one’s colleagues in order to support both individual and group performance:

“Now that I am back after a period of illness, it is important to discuss with my colleagues how much I can handle. By sharing how you feel in your work, you can – as a group – better choose a strategy in the division of tasks. The setting in which you share these things helps, because in daily practice we hardly take a moment to do so or have dedicated time for this important thing.” (14G)

Apart from improvement goals related to collaborative practice, participants also reported improvement goals related to individual career management, albeit to a much lesser extent. Participants formulated goals such as “additional specialization”, “updating certain medical knowledge”, “expansion of/profiling within an administrative role” or “focus on the management of an outpatient clinic”. Implementation of these intended changes in practice were not discussed in detail. Sharing these plans mainly served as a way to inform one’s peers, in order to deal with arrangement issues later.

“During a session like this, there is opportunity to express personal wishes and needs for professional development. However, shifting to other activities or investing in individual development is something that needs to fit in with the organization and philosophy of the group.” (9K)

4. Influential factors

Participants' experiences with the group session and their perceptions of its contribution to their performance and development were influenced by a number of factors. These factors were related to the facilitator, the peer group, the individual physician and the context of CPD policy/revalidation.

Facilitator

Overall, participants spoke highly appreciative about the facilitator's role as an intermediary. The facilitator's explicit invitation to share experiences or add nuance to a peer's story was noted in particular:

"In a normal work setting, I often tend to get defensive, but with the facilitator, things were discussed quietly. She structured the conversation in a pleasant way without being overly present, thereby creating a safe atmosphere." (24L)

Participants also noted the facilitator's power to let the participants validate each other in their qualities. Some participants highlighted the facilitator's ability to elicit information in a friendly and playful way or to ask in-depth questions.

Peer group

Participants agreed that the context of the group had an impact on the (effects of the) interactional process. Group context factors entailed factors such as size, history and composition of the group. In a group that stood out in homogeneity in terms of age, career stage and years of experience, one participant noted the lack of diversity or seniority within the group to be a factor in acquiring insight into own performance:

"We are all young emergency doctors and sometimes I feel like we have to reinvent the wheel. It would be nice if there were some more experienced doctors in our group, to hear different perspectives, but also to provide us with an overview and offer us some reassurance." (12F)

In case the group was rather large (seven people or more), participants reported that there was not enough quality time to thoroughly discuss individual matters. When the group was rather 'young', for example, because of organizational changes such as a merger, participants indicated that they needed time to get to know each other better:

"There was too little time for so many people. And besides that, it felt rather superficial because we don't know each other that well yet. It is also a matter of time that people want to show more of themselves." (16H)

Individual physician

Personal circumstances were recognized as a factor in determining the experiences with participation in the session. Many participants explicitly mentioned their level of experience and career stage to be of influence:

"As a young specialist, I still need to get used to this new status and gain more confidence. Thanks to my colleagues' reactions and the way they received me, I felt very much appreciated and respected." (19I)

Also, participants differed from each other in terms of overall vulnerability. In the extreme case of a participant who acknowledged that her well-being was under severe pressure, the group session was a complete overstimulation. This participant experienced the use of a group format as an illustration of the "pursuit of efficiency in healthcare", undermining an honest and meaningful conversation about performance and development:

"The session had a big impact. Discussing the numbers in my feedback report with my colleagues, the whole thing just did not feel genuine, especially since it was about that theme of professional functioning. There should be more attention to topics such as sustainable employability and career development, but in my opinion, a group session like this is not the way." (4B)

Context of CPD policy

Many participants were enthusiastic about the group session; some stated that sessions of this kind should take place more often. Although they considered the session valuable, however, they did note that it cost energy and effort and that it should not be 'overdone'. One participant also underlined that physicians should experience autonomy in using group sessions as a tool for development and performance improvement. In this context of professional (self-) regulation, others indicated that they had only participated because it was part of CPD policy and mandatory for revalidation purposes. As one of them cynically remarked:

"If the profession requires participation in these sessions, I will of course respond to it. However, if these sessions were abolished, I would not slip into a deep depression." (20I)

Participants also mentioned that these group sessions took place only once every two years. They wondered how the follow-up on improvement goals could be jointly monitored:

"How are we going to ensure that this is not just a formality but that something is really going to change? How are we going to notice or measure any changes?" (14G)

DISCUSSION

Our study provides insight into the importance and (im)possibilities of interpersonal relationships for learning from feedback during a group session. This study revealed that sharing reflections with others rendered feedback more meaningful, deepened collegial relationships and created a sense of urgency for improvement. Conditions for success were an experienced facilitator, longitudinal trusting relationships and a limited group size. In the following section, we will discuss our findings by elaborating on (1) the interdependence of 'process' and 'outcome' and the role of interpersonal trust; (2) the use of communication strategies that influence the quality of a group session and (3) the role of these group sessions in realizing performance improvement and development.

Although in our research questions we separated 'process' and 'outcomes', these were difficult to distinguish in our results. Physicians reporting high levels of interpersonal trust and openness during the process also tended to report beneficial (emotional) outcomes. Interpersonal trust and openness seemed to be enacted by the set-up of the session as an invitation to co-construct the conversation. Physicians reported that by hearing their peers' personal reflections, they were inclined to respond and share personal reflections as well. This cascade of processes is reminiscent of a mechanism called 'the dyadic effect'. The notion of the dyadic effect was coined by the psychologist Jourard (1971), which states that if self-disclosure by one person increases (in this case the sharing of personal information or reflections), so does that by the other.²⁷ Jourard regarded communication as an interpersonal transaction and conceptualized self-disclosure as a deliberate invitation to the other to know and share his/her experiences.²⁷ Although self-disclosure is a relatively infrequent form of communication, it is claimed to be important in a variety of contexts.²⁸⁻³³ Several authors have identified individuals' ability and willingness to self-disclose as determinants of their personal health and satisfaction, success in being understood and working competently with others, and the ability to provide communication experiences that others find satisfying and therapeutic.^{28,29} In communication science, considerable attention has been given toward identifying those characteristics of positive relationships most closely associated with self-disclosure.^{28,29} Not surprisingly, and in accordance with our findings, several authors refer to interpersonal trust as a key factor for self-disclosure.^{28,29,34,35} Our findings show that in many cases the exchange of personal reflections went naturally. One physician noticed an implicit norm of reciprocity: the perception that sharing should be equal among group members. In the light of this experience, physicians emphasized that the session took place within a professional working context, requiring attention to the relationship between 'professional' and 'personal'.

The findings of our study suggest that the quality of a group session is affected by communication strategies that may or may not be deliberately used by participants. Group members differ in their willingness to self-disclose and provide others with constructive, honest feedback, creating an atmosphere of 'positive yet superficial talk'. This state of affairs points to the use of so-called politeness strategies.^{36,37} Brown and Levinson's Politeness theory states that in social interactions, positive and negative 'face' play a role. There is positive face when individuals only emphasize positive things during feedback conversations to avoid damaging relationships and the self-esteem of the feedback receiver. There is negative face when constructive feedback is provided with the intention to support an individual's longitudinal growth, but which may be a breach of the norms of expected politeness. A polite or face-saving learning culture may thus have a negative impact on feedback conversations.³⁸ Based on the findings of our study, sharing personal reflections in a group setting more often could be a way to foster openness and meaningful conversation within a physician group.

As our findings illustrate, physicians reported that the group sessions were felt to be supportive in terms of reflecting on performance and prompted them to compare their views on their own professional performance with the perspectives of others. Physicians noted that sharing reflections could increase a sense of urgency for improvement and foster a community spirit, which might help in implementing intended changes. However, the application of coaching principles, such as drawing up a detailed plan of action and investing in (the monitoring) of a follow-up process³⁹, appeared to be a challenge in a group setting, especially in larger groups as we found in this study. Nevertheless, it may be premature to conclude that group sessions fall short in their objective to achieve performance improvement and development. Performance improvement and development may have a longer lead time, and the effects of the session may be more pervasive as it has taken the group context as the starting point.^{34,40} It was striking that improvement goals tended to center around relational fine-tuning in collaborative practice, whereas goals related to individual career management were discussed less extensively. It might be argued that multisource feedback and reflecting on it together with peers predisposes to improvement of collaborative practice and group performance, which also affects individual performance.

Strengths and limitations

The strength of our study is its exploratory character. Since this feedback group initiative is gaining more acceptance in CPD practices, it is important to delve into the experiences of the main stakeholder group, physicians themselves. The involvement of the second author (MvdG) as both a researcher and group facilitator provided a valuable insider perspective. As with any qualitative research, however,

we were mindful of the issue – or challenge – of different researcher perspectives. We took steps to carefully manage these differences, primarily through ongoing discussions among the authors to verify our coding scheme and interpretations and by letting the first author (EB) be present at the group sessions to assist her in forming an understanding of its context. As an additional way to triangulate findings across participants, we performed at least two interviews for each group session. The principal limitation of this study relates to transferability. Our findings represent the collected information of 26 individual interviews over the course of 13 group sessions in five Dutch hospitals. It should be acknowledged that this study was performed in a general hospital setting. We recognize that there are unique national policies, organizational factors and institutional cultures, which may not be transferable to other settings, such as academic settings or hospital settings outside the Netherlands.

Areas for further research

The findings of this study draw attention to several research and pedagogical challenges. Of particular importance is the development of techniques for medical educators and facilitators to increase physicians' awareness of their actual and potential levels of disclosure which characterize their interactions with others. Since high levels of disclosure facilitate group effectiveness, mutual understanding and personal satisfaction, knowledge of self-disclosure should be included in the repertoire of professional competence to improve communication behavior during group sessions. An important avenue for future research is that of studying the interactional nature of group sessions in various contexts and exploring which interactions may be beneficially affected by a (professional) facilitator or change agent. The question whether and how the emotional effects of the group session are retained and whether and how these effects contribute to a climate of openness and psychological safety requires longitudinal research. Also, more research is needed to explore how the institutional culture can influence the quality and impact of feedback, feedback-seeking, acceptance and performance improvement and development. Understanding socio-cultural factors in various work environments is essential before designing initiatives to promote meaningful feedback exchanges and enhance impact on behavior change and professional development. Finally, future research should look at the effects of group sessions on actual behavior change, by using a variety of information sources, including - but certainly not limited to - self-reported changes in practice.

Implications for practice

The findings of our study suggest that when facilitating performance feedback in a peer group setting, there are a number of factors that need to be taken into account: the expertise of the facilitator, the size of the group (should not be too large), the

presence of longitudinal trusting relationships and the absence of (profound) organizational disturbances or (excessive) personal struggles. It should be recognized that the facilitator's task is challenging in terms of navigating group dynamics and tailoring the session to the needs of the individual. For young physicians, a group session can be helpful in becoming better aligned with the group, investing in interpersonal relationships and gaining insight in both themselves as well as others. For vulnerable individuals, a group session may generate too much exposure and/or stress. Finally, it is important to provide clarity about the purpose and the design of group sessions in light of CPD policy and its connection to revalidation requirements. Mandatory participation in group sessions could undermine physicians' motivation, as evidenced by the findings of this study as well as literature reporting on 'reflection fatigue' and cynicism about reflection as a 'tick box exercise'.^{41,42} To optimize the facilitation of performance feedback uptake, it is important to keep in mind that it is not about – as expressed by Ramani et al. – following recipes, but about investing in relationships in order to provide an environment in which people are encouraged to disclose, discuss and learn from feedback.⁴³

CONCLUSIONS

Reflecting on personal MSF data during a peer group session offered the possibility to discuss and compare self and others' perceptions and assisted in gaining a nuanced insight into one's professional performance. Sharing reflections with others deepened collegial relationships and created a sense of urgency for improvement, being mostly related to interpersonal communication and collaboration issues. Sharing reflections with others was experienced as a source of social support and was perceived as helpful in realizing actual change. Factors influencing experiences with participation and perceptions of its contribution to CPD were related to the expertise of the facilitator, the continuity and quality of collegial relationships, personal vulnerabilities and the context of CPD policy and its connection to revalidation.

REFERENCES

1. Epstein RM. Mindful practice. *JAMA*. 1999;282(9):833-839.
2. Guest CB, Regehr G, Tiberius RG. The lifelong challenge of expertise. *Med Educ*. 2011;35(1):78-81.
3. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA*. 2002;287:226-35.
4. Bindels E, Verberg C, Scherpbier A, Heeneman S, Lombarts K. Reflection revisited: how physicians conceptualize and experience reflection in professional practice – a qualitative study. *BMC Med Educ*. 2018;1:105.
5. Ericsson KK. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med*. 2004;79:70-81.
6. Mann K, Gordon J, MacLeod A. Reflection and reflective practice in health professions education: A systematic review. *Adv Health Sci Educ Theory Pract*. 2009;14:595-621.
7. Sargeant J, Bruce D, Campbell CM. Practicing physicians' needs for assessment and feedback as part of professional development. *JCEHP*. 2013;33:S54-62.
8. Pelgrim EAM, Kramer AWM, Mekkink HGA, Van der Vleuten CPM. Reflection as a component of formative assessment appears to be instrumental in promoting the use of feedback; an observational study. *Med Teach*. 2013;35(9):772-8.
9. Sargeant J, Mann K, Sinclair D, Van der Vleuten C, Metsemakers J. Understanding the influence of emotions and reflection upon multi-source feedback acceptance and use. *Adv Health Sci Educ*. 2008;13(3):275-88.
10. Eva KW, Armson H, Holmboe E, et al. Factors influencing responsiveness to feedback: On the interplay between fear, confidence, and reasoning processes. *Adv Health Sci Educ Theory Pract*. 2012;17:15-26.
11. Sargeant JM, Mann KV, Van der Vleuten CP, Metsemakers JF. Reflection: A link between receiving and using assessment feedback. *Adv Health Sci Educ Theory Pract*. 2009;14:399-410.
12. Sargeant J, Lockyer J, Mann K, Holmboe E, Silver I, Arson H, Driessen E, MacLeod T, Yen W, Ross K, Power M. 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.
13. Telio S, Ajjawi R, Regehr G. The "educational alliance" as a framework for reconceptualizing feedback in medical education. *Acad Med*. 2015;90:609-14.
14. Overeem K, Driessen EW, Arah OA, Lombarts KMJM, Wollersheim HC, Grol RPTM. Peer mentoring in doctor performance assessment: strategies, obstacles and benefits. *Med Educ*. 2010;44(2):140-7.
15. Overeem K, Wollersheim H, Driessen E, Lombarts K, van de Ven G, Grol R, Arah O. Doctors' perceptions of why 360-degree feedback does (not) work: a qualitative study. *Med Educ*. 2009;43(9):874-882.
16. Brehaut JC, Colquhoun HL, Eva KW, Carroll K, Sales A, Michie S et al. Practice feedback interventions: 15 suggestions for optimizing effectiveness. *Ann Intern Med*. 2016;164:435-41.
17. Boud D. Sustainable assessment: rethinking assessment for the learning society. *Studies in Contin Educ*. 2000;22(2):151-167.

18. Black P, William D. Developing the theory of formative assessment. *Assess Eval Acc.* 2009;21:5-31.
19. Tigelaar DE, Dolmans DH, Meijer PC, de Grave WS, van der Vleuten CP. Teachers' interactions and their collaborative reflection processes during peer meetings. *Adv in Health Sci Educ.* 2008;13:289-308.
20. Boerboom TB, Jaarsma D, Dolmans DH, Scherpbier AJ, Mastenbroek NJ, van Breukelen P. Peer group reflection helps clinical teachers to critically reflect on their teaching. *Med Teach.* 2011;33:615-623.
21. Van Lierop M, de Jonge L, Metsemakers J, Dolmans D. Peer group reflection on student ratings stimulates clinical teachers to generate plans to improve their teaching. *Med Teach.* 2018;40:302-309.
22. Federatie Medisch Specialisten. Besluit Herregistratie specialisten. <https://www.demedischspecialist.nl/onderwerp/herregistratie>. Accessed on 14 July 2019.
23. Van der Meulen MW, Boerbach BC, Smirnova A, Heeneman S, oudeEgbrink MG, van der Vleuten CP, Lombarts K. 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.
24. Brocki JM, Wearden AJ. A critical evaluation of the use of interpretative phenomenological analysis (IPA) in health psychology. *Psychology and Health.* 2006;21(1):87-108.
25. Smith JA, Flowers P, Larkin M. Interpretative phenomenological analysis: Theory, method, research. London: Sage. 2009.
26. Smith JA. Evaluating the contribution of interpretative phenomenological analysis. *Health Psychology Review.* 2011;5(1):9-27.
27. Jourard SM. Self-disclosure: An experimental analysis of the transparent self. New York, Wiley- Interscience, 1971.
28. Egan G. Encounter: Group processes for interpersonal growth. Belmont, California: Wadsworth, 1970.
29. Beebe SA, Beebe SJ, Redmond MV. Interpersonal communication. Scarborough, Ont: Prentice-Hall Canada, 2000.
30. Wheelless LR, Grotz J. The measurement of trust and its relationship to self-disclosure. *Hum. Commun. Res.* 1977;3:250-7.
31. Lannutti PJ, Strauman EC. Classroom communication: The influence of instructor self-disclosure on student evaluations. *Communication Quarterly.* 2006;54:89-99.
32. Ledbetter AM, Mazer JP, De Groot JM, Meyer KR, Mao Y, Swafford B. Attitudes toward online social connection and self-disclosure as predictors of Facebook communication and relational closeness. *Commun. Res.* 2011;38:27-53.
33. Schrodtt P, Phillips KE. Self-disclosure and relational uncertainty as mediators of family communication patterns and relational outcomes in sibling relationships. *Commun. Monogr.* 2016;83:486-504.
34. Argyris C. Interpersonal competence and organizational effectiveness. Homewood, Ill: Dorsey, 1962.
35. Ramani S, Konings K, Ginsburg S, van der Vleuten CP. Relationships as the backbone of feedback: exploring preceptor and resident perceptions of their behaviors during feedback conversations. *Acad Med.* 2019; published ahead of print.

36. Brown P, Levinson S. Universals in language usage: Politeness phenomena. Cambridge: Cambridge University Press. 1978.
37. Ramani S, Konings S, Ginsburg S, van der Vleuten CP. Twelve tips to promote a feedback culture with a growth mindset: swinging the pendulum from recipes to relationships. *Med Teach*. 2018;1-7.
38. Ramani S, Post SE, Konings K, Mann L, Katz JT, van der Vleuten CP. "It's just not the culture": A qualitative study exploring residents' perceptions of the impact of institutional culture on feedback. *Teach Learn Med*. 2017;29:153-161.
39. Ramani S, Konings K, Ginsburg S, van der Vleuten CP. Feedback redefined: Principles and practice. *J Gen Intern Med*. 2019;34(5):744-749.
40. Brennan N, Bryce M, Pearson M, Wong G, Cooper C, Archer J. Understanding how appraisal of doctors produces its effects: a realist review protocol. *BMJ Open*. 2014;4:e005466.
41. Rolfe G. Rethinking reflective education: what would Dewey have done? *Nurse Educ Today*. 2014;34(8):1179-83.
42. Ng SL, Kinsella EA, Friesen F, Hodges B. Reclaiming a theoretical orientation to reflection in medical education research: a critical narrative review. *Med Educ*. 2015;49(5):461-75.
43. Ramani S, McMahon GT, Armstrong EG. Continuing professional development to foster behavior change: from principles to practice in health professions education. *Med Teach*. 2019:1-8.

5

A NEW MULTISOURCE FEEDBACK TOOL FOR EVALUATING THE PERFORMANCE OF SPECIALTY-SPECIFIC PHYSICIAN GROUPS: VALIDITY OF THE GROUP MONITOR INSTRUMENT

Bindels, E
Boerebach, B
Van der Meulen, MW
Donkers, J
Van den Goor, M
Scherpbier, AJJA
Lombarts, MJMH
Heeneman, S

ABSTRACT

Introduction: Since clinical practice is a group oriented process, it is crucial to evaluate performance on the group level. The Group Monitor (GM) is a multisource feedback tool that evaluates the performance of specialty specific physician groups in hospital settings, as perceived by four different rater classes. In this study we explored the validity of this tool.

Methods: We explored three sources of validity evidence: (1) content, (2) response process and (3) internal structure. Participants were 254 physicians, 407 staff, 621 peers and 282 managers of 57 physician groups (in total 479 physicians) from 11 hospitals.

Results: Content was supported by the fact that the items of the existing tool were closely related to daily practice. Pilot rounds resulted in reformulation and reduction of items. Four subscales were identified for all rater classes: *Medical practice*, *Organizational involvement*, *Professionalism* and *Coordination*. 'Physicians' and 'staff' had an extra subscale, *Communication*. However, the results of the generalizability analyses showed that variance in GM scores could mainly be explained by the specific hospital context and the physician group specialty. Optimization studies showed that for reliable GM scores, 3 to 15 evaluations were needed, depending on rater class, hospital context and specialty.

Discussion: The GM provides valid and reliable feedback on the performance of specialty specific physician groups. When interpreting feedback, physician groups should be aware that rater classes' perceptions of their group performance are colored by the hospitals' professional culture and/or the specialty.

INTRODUCTION

In the hospital setting, patient care is provided by physicians working together in specialty specific groups (e.g. cardiologists, gynecologists, neurologists). In order to perform at an adequate level as a physician group, mutual relationships are essential, both between physician group members as well as with other partners in the hospital organization, e.g. physicians from other medical specialties, nurses, pharmacists, therapists, secretary staff and managers. Since clinical practice is becoming more group oriented, it is important to recognize that performance improvement is more effectively accomplished at a group level.¹ Therefore, the evaluation and monitoring of the performance of specialty specific physician groups is increasingly being recognized as an integral part of quality requirements.^{2,3,4}

Up until now, investments have been made to develop tools for the evaluation of healthcare teams which have high degrees of multidisciplinary collaboration, such as palliative care teams or teams in the operating room. These tools use both self-assessment measures as well as assessment measures from outside observers.^{5,6,7,8} However, tools to particularly evaluate the performance of *specialty specific physician groups* are scarce and mainly use self-assessment measures.^{9,10} This situation is not optimal for two reasons. First, to meet higher standards of quality of care, it is necessary to evaluate physician performance at the level of the physician group. Second, to meet higher standards of transparency, it is essential to not only consider self-assessment measures but also take into account the perspectives of collaborative partners.^{11,12,13}

We aimed to fill this gap by developing and testing a *group* multisource feedback tool, the Group Monitor (GM). The GM was designed in collaboration with hospital-based physicians and is currently used as a tool for quality improvement by hospital-based physician groups in Dutch healthcare. The GM evaluates the performance of specialty specific physician groups as perceived by four rater classes: (1) physician group members themselves (self-assessment), (2) supporting staff (e.g. nurses, pharmacists, therapists, secretary staff), (3) peers (physicians from other medical specialties in the same institution) and (4) hospital managers, see Figure 5.1 for a schematic overview. The tool consists of 35 items concerning specific components of group performance (e.g. communication, collaboration and organization) and two global items. Uniform items have been developed for each of the four rater classes, so that the questionnaire can be used across a wide range of clinical settings and that the perceptions of group performance can both be examined per rater class as well as be compared among the different rater classes.

This article describes the development process for the GM and the validity evidence we collected to support its use, using a validity framework developed by the

American Psychological Association (APA), the American Educational Research Association (EARNA), and the National Council on Measurement in Education (NCME), as described by Downing.¹⁴ The “APA framework”, with its five domains of validity evidence, has been described as “the current standard of assessment validation.”¹⁵ We also provide the psychometric data we measured to document its reliability, based on Generalizability Theory (GT).^{16,17}

Figure 5.1 Group Monitor instrument - schematic image of the four rater classes: physicians, staff, peers and managers

METHODS

Context

This study was conducted in 11 general hospitals in the Netherlands, where physicians are organized in specialty specific groups. The evaluation and monitoring of the performance of these physician groups is an integral part of quality requirements as set out by the medical profession itself.^{2,3,4} Every five years, physician groups are required to take part in a formal external quality review carried out by their own specialty society.^{9,10} With this quality inspection, which is of a formative rather than a summative nature, an overview of the performance level is obtained. Based on this information, which is gathered from multiple sources, the group's quality policy, standards and instruments (e.g. guidelines, protocols, agreements between scientific societies, training, continuing education and patient education) are adjusted if necessary. Participation in this quality review is mandatory for re-registration as a physician. In 2013, the Quality Assurance Advisory Committee introduced the guideline 'Appraisal system for quality inspections', which describes the general standards that can be tested during the quality review.¹⁸ The basic standards included in the guideline are based on four quality domains: (1) evaluation of patient care, (2) professional development, (3) specialist group functioning and (4) the patient's perspective. Although the Group Monitor was not explicitly developed for use in the five-yearly quality review, its use may contribute to the evaluation of the third quality domain.

Development of the Group Monitor

- In 2012, the development of the GM was initiated by the GM project group, consisting of seven people: a chairperson of the medical staff of a general teaching hospital, a policy coordinator of that same hospital, an expert in the field of instrument development and four healthcare management consultants. The instrument was designed in collaboration with an advisory group of physicians, human resource managers, representatives of scientific boards and experts in the field of instrument development.
- The items of the GM are based on a review of an existing instrument for self-assessment of group performance, the Quick Scan.^{10,19} This instrument was developed in the late nineties as part of the first initiatives for quality reviews in Dutch healthcare.¹⁰ Since then, the instrument has been used – and has proved its usefulness – in practice by mapping the strengths and weaknesses of specialty specific physician groups.¹⁹ In the Quick Scan, five subscales were investigated, each subscale containing 15 items: (1) *Policies and procedures*, (2) *Group culture and structure*, (3) *Decision making*, (4) *Reputation* and (5) *Results*. For the development of the GM, these five subscales were reworked into six pre-

defined subscales of the GM: (1) *Medical practice*, (2) *Innovation and development*, (3) *Professionalism*, (4) *Teamwork*, (5) *Organization* and (6) *Communication*. The phrasing of the subscale headings was guided by the competency framework of CanMEDS²⁰, as the content and terminology of this framework has already found recognition and acceptance with the target group.

- In 2013 and 2014, the GM was tested in two pilot rounds with in total 16 physicians (self-assessment), 14 staff, 28 peers and 17 managers. Items were tested on applicability, comprehensibility and prioritization, leading to reformulation of items, a reduction from 43 to 36 items and the addition of one global rating ("I would recommend this group to others"). In the first pilot round, items could be rated on a 6-point scale (1= totally disagree, 2= strongly disagree, 3= disagree, 4= agree, 5= strongly agree, 6 = totally agree) with an additional 'cannot judge' option. However, the score distribution indicated that the distance between label 3 and 4 was experienced as too big, so that respondents were more inclined to opt for label 4. Therefore, in the second pilot round, the answer scale was adjusted to a 5-point scale (1= bad, 2= weak, 3= needs attention, 4 = sufficient, 5= good). Label 3 ('needs attention') was deliberately chosen instead of a phrasing that tends towards the safe middle position, in order to obtain the greatest possible differentiation. Both pilot rounds were evaluated among the staff boards, the Board of Directors and the Board of Governors.
- In 2016, the GM was incorporated into the Professional Performance Online (PPO) platform, an online platform aimed at facilitating physicians' professional development (www.professionalperformanceonline.nl). This internet-based environment was used to facilitate the data collection for the current study on a broad scale. When incorporated into the PPO platform, the items of the GM were checked for clarity once more. One item originally belonging to the subscale 'Professionalism' (the item "This is a physician group that I trust") was labeled as a global item, resulting in the final questionnaire consisting of 35 items on six pre-defined subscales and two global items. Once the questionnaires were completed, on average after one month, the evaluated groups received a feedback report, which was automatically generated by the web-based system.

Format and content of the Group Monitor

The GM was developed as one uniform questionnaire for four different rater classes: (1) members of the physician group themselves (self-assessment), (2) supporting staff (e.g. nurses, secretary staff, therapists, pharmacists), (3) peers (physicians from other medical specialties) and (4) hospital managers.

- The GM taps into behavioral aspects of group performance, such as communication, collaboration, use of all contributors' expertise, effort, shared decision making and sharing workload. The questionnaire consists of six pre-defined subscales, containing 35 items in total: (1) M - *Medical practice*, (2) I

- *Innovation and development*, (3) *P - Professionalism*, (4) *T - Teamwork*, (5) *O - Organization*, and (6) *C - Communication*.
- Items can be rated on a 5-point scale (1 = bad, 2 = weak, 3 = needs attention, 4 = sufficient, 5 = good) with an additional 'cannot judge' option. Furthermore, the questionnaire contains two global ratings: (1) "This physician group is a group that I trust", using a 5-point scale (1= bad, 2 = weak, 3= needs attention, 4= sufficient, 5= good) and (2) "I would recommend this group to others" to be rated on a 10-point scale (ranging from 1= definitely not to 10 = definitely yes). In addition, raters are encouraged to complement their responses for each subscale with narrative comments, as previous studies indicate that narrative comments are valuable and informative data sources in addition to numerical feedback.^{21,22}
- To ensure that raters evaluate the entire physician group and not just one individual belonging to this group, at the start of each subscale raters are explicitly asked to answer the items for the entire physician group by means of the following statement: "We would like you to answer the following items for the entire physician group".

Data collection

To distribute the GM on a broad scale, we used the PPO platform. From September 2014 to July 2017, 57 physician groups (18 surgical, 39 non-surgical) from 11 general hospitals in the Netherlands participated in the study. Physician groups were actively recruited using the network of the GM project group. Data collection was not evenly spread over hospitals and specialties: one hospital counted 26 physician groups, one hospital counted 16 physician groups and the other nine hospitals counted 1-3 physician groups. Members of the participating physician groups self-evaluated their own group performance and were asked to invite at least 12 peers from other physician groups in the same hospital, 8 supporting staff and 8 managers to fill out the questionnaire. Although the pilot rounds had shown that raters provided narrative comments, we did not use this qualitative information for the current study. The dataset for this study consists of quantitative scores only.

Data analysis: demonstrations of validity

The APA framework was used as a validity framework, consisting of five categories of validity evidence: (1) content, (2) response process, (3) internal structure, (4) relationship with other variables and (5) consequences.¹⁴

1. The content validity is concerned with ensuring that the content of the test is sufficiently similar to, and representative of, the task that it is intending to measure. We described the instrument elaboration, structure and content in the preceding section.

2. The response process validity is concerned with ensuring that all sources of error associated with the administration of the test are recognized and limited to the full extent possible. We described the evaluation of the response process and the subsequent adjustments in the GM questionnaire in the preceding section.
3. The internal structure validity is concerned with the statistical and psychometric characteristics of the questions or prompts and the psychometric properties of the model used to score/scale the assessment. This aspect of validity is involved in determining the generalizability and reproducibility of the assessment. This source of validity evidence is explored by using explorative factor analyses, item-total correlations and inter-scale correlations (see further, section Exploration of psychometric properties). Also, this source of validity evidence is investigated by using Generalizability Theory (GT), performing G studies and D studies (see further, section *Generalizability Theory*).^{16,17}
4. The relationship to other variables type of validity evidence is concerned with the correlational or relationship of assessment results with other previous or existing measures of performance. This source of validity evidence was beyond the scope of this study.
5. The consequences or consequential validity is concerned with the impact that the assessment has on both the examinees, as well as on the health service, patients, and wider society. This source of validity evidence was beyond the scope of this study.

Exploration of psychometric properties

- Before starting to investigate the internal structure of the GM, evaluations with more than 50% missing data were excluded from further analysis. Evaluations with less than 50% missing data were imputed using expectation-maximization (EM) technique as the data were assumed to be missing at random. Since recent perspectives on rater cognition state that it is unreasonable to expect different rater classes to interpret group performance identically,^{21,23,24} psychometric analyses were performed for each rater class separately.
- Component structure was investigated by conducting principal component analysis (PCA) on the 35 items. The promax rotation method was used to extract the components, since the data showed that components were correlated.²⁵ The Kaiser-Guttman criterion (eigenvalue above 1.0) was used to determine the number of components to extract and also the scree plot was checked. Interpretation of the components was guided by statistical results (factor loadings above 0.30) and whether items clustered logically.
- Internal consistency was determined by computing Cronbach's alphas for all subscales; a coefficient of 0.70 or higher was considered acceptable.²⁶ As an additional measure of the consistency of the subscales, homogeneity of each subscale was assessed by item-total correlations, which should be above

0.40.²⁷ The overlap between the subscales was investigated using inter-scale correlations. Ideally, inter-scale correlations are below 0.70 (which corresponds to an overlap of less than 50%).²⁷

- Construct validity was investigated by examining correlations of the GM subscales with the two global items: (1) "This physician group is a group that I trust" and (2) "I would recommend this physician group to others". We hypothesized that physician groups that scored high on the subscales would also score high on being trusted and recommended to others. We expected these correlations to fall within the range of 0.40 to 0.80 for an indication of good construct validity.²⁸
- To not only explore the internal structure of the scores for each rater class separately, but to also compare the scores between the different rater classes, the overall mean scores were aggregated on the level of the rater class. The overlap between the overall mean scores of the different rater classes was investigated by studying the correlation matrix. We hypothesized that correlations above 0.70 indicate that the rater classes are measuring the same aspects of performance and that correlations below 0.70 are an indication that the rater classes are measuring different aspects of performance.²⁷

Generalizability Theory (GT)

Generalizability theory (GT) can be used to generalize a physician group's score to the average score of that physician group under all possible and acceptable conditions of the taking of the questionnaire. Through generalizability studies (G studies), the effects of specific sources of variance ("facets") on the score are isolated and those effects that introduce an error (bias) in a score are identified. Through design studies (D studies), we can use the results of a G study to determine the minimal number of raters per class needed to obtain a sufficient level of reliability.^{16,27}

- In the G studies for the GM, five facets were considered: (1) the physician groups (G), which were the object of measurement. The other facets, which could introduce biases, were (2) the raters (R), (3) the specialty of the physician group (Spec), (4) the hospital in which the physician group works (Hosp) and (5) the subscales of the instrument (S). Since the data collection was not evenly spread over hospitals and specialties, it was not possible to combine the facets Spec and Hosp in one G Study. We therefore performed two separate G studies per rater class. The first model included the facet Spec and the second model included the facet Hosp. Every rater provided one rating on one physician group on one occasion. Raters were thus specific to physician groups: the R facet was *nested* within the G facet. In turn, the R and G facets together were *nested* within the Spec or the Hosp facet respectively. Since all raters (in the same class) used the same subscales, the S facet was crossed with all other facets. In all models, the number of subscales was considered as fixed.

- In summary, we performed two G studies for each of the rater classes separately. The first model consisted of Specialty, Group, Rater and Subscale, with Rater *nested* within Group (Group:Rater) and Group and Rater *nested* within Specialty (Spec:Group:Rater). The second model consisted of Hospital, Group, Rater and Subscale, with Rater nested within Group (Group:Rater) and Group and Rater nested within Hospital (Hosp:Group:Rater).
- Following G studies, D studies determine how many ratings will be needed to ensure reliability of the GM scores (i.e. reliability coefficient above 0.7).³⁷ As a measure of reliability, we calculated the G coefficient (G) and standard error of measurement (SEM) for varying numbers of ratings. A G of 0.60 was considered acceptable but indicative of a need for improvement and 0.80 was considered very reasonable.²⁹ The SEM can be interpreted on the original scoring scale, in this case a 5-point scale. Similar to validation studies by Boor et al³⁰, Boerboom et al³¹, Silkens et al³² and Van der Meulen et al²¹, we decided to accept a maximum “noise level” of 1.0 on the scale. In the case of a 5-point scale, we considered a $SEM \leq 0.26$ ($1.96 \times 0.26 \times 2 \approx 1.0$) as adequate for a 95% confidence interval interpretation. For the generalizability analyses we used Lme4 package in R software with restricted maximum likelihood (REML) estimation³³; for all other analyses we used SPSS version 20 (SPSS Inc., Chicago, IL).

Ethical considerations

This study was exempt from Institutional Board Review under Dutch law. A waiver of ethical approval was provided by the Institutional Review Board of the Academic Medical Center of Amsterdam, Amsterdam, the Netherlands, waiver number W18_095.

RESULTS

Participants

We collected data of 57 physician groups from 11 general hospitals from 2014 to 2017. In total, 254 evaluations were completed by physician group members themselves (self-assessment). Physician groups received evaluations from 407 supporting staff, 621 peers and 282 managers (Table 5.1). In total, 19 evaluations were excluded because more than 50% of the data were missing. Response rate was 53% for physician group members, 66% for staff, 66% for peers and 72% for managers.

Psychometric properties

- Principal component analysis revealed a 4-component model for the 'peers' and the 'managers', explaining 67.7% and 65.7% of the variance respectively. These components or subscales were labeled as *Medical practice*, *Organizational involvement*, *Professionalism* and *Coordination*. For 'self' and 'staff', the analysis showed a model with an additional 5th subscale, labeled as *Communication*. Total explained variance was 59.4% and 67.3% respectively. The factor loadings of the items belonging to each subscale, with variations among the four rater classes, are displayed in Table 5.2.
- Internal consistency analyses showed that Cronbach's alphas for subscales ranged from 0.83 to 0.89 for 'self', from 0.86 to 0.93 for 'staff', from 0.85 to 0.95 for 'peers' and from 0.85 to 0.94 for 'managers'. Corrected item-total correlations were all higher than 0.48 for all rater classes. The inter-scale correlations for each of the four rater classes ranged from 0.60 to 0.74 for 'self', from 0.53 to 0.84 for 'staff', from 0.68 to 0.86 for 'peers' and from 0.66 to 0.84 for 'managers' (Appendix 5.1).
- Construct validity analyses revealed that the correlations between each subscale and the two global items ranged from 0.57 to 0.66 for 'self', from 0.52 to 0.79 for 'staff', from 0.62 to 0.82 for 'peers' and from 0.59 to 0.80 for 'managers' (Appendix 5.2).
- Comparison between the overall mean scores of the four rater classes showed that correlations ranged from 0.54 to 0.65, indicating an overlap of less than 50% (Appendix 5.3).

G studies and D studies

- The sources of variance in the GM scores for the rater classes in both models are given in Table 5.3. The model for the rater class 'self' could not be statistically identified in the Hospital G Study, due to the uneven distribution of these raters over the hospitals. Therefore, a smaller facet model was used for this case. Table 5.4 shows the results of the D studies. It reports the number of evaluations (raters) needed from each rater class for a G coefficient above 0.60 and a SEM below 0.26. For the rater class 'self', for both the hospital model as well as the specialty model, 3 raters were needed to obtain reliable measurements. For staff, peers and managers, 7-9 raters, 8-15 raters and 6-7 raters were needed respectively.

Table 5.1 Group Monitor instrument - characteristics of the rater classes from 2014 to 2017

	Physicians (self- assess- ment)	Staff	Peers	Managers	Total
No. of raters (% of those invited)	254 (53)	407 (66)	621 (66)	282 (72)	1564 (65)
Minimum-maximum raters per physician group	4-21	2-21	1-41	3-13	
Mean number of evaluations per physician group (SD)	8.43 (3.75)	7.92 (3.79)	10.89 (6.3)	5.91 (2.05)	
No. of hospitals					11
No. of physician groups					57*
Total mean score (SD)	4.23 (0.44)	4.10 (0.52)	4.08 (0.57)	4.06 (0.54)	
Mean scale scores (SD)					
Medical practice	4.44 (0.41)	4.39 (0.53)	4.18 (0.59)	4.22 (0.52)	-
Organizational involvement	4.04 (0.62)	4.19 (0.56)	3.98 (0.66)	4.05 (0.63)	-
Professionalism	4.22 (0.51)	4.12 (0.61)	4.16 (0.61)	4.09 (0.62)	-
Coordination	4.26 (0.46)	4.02 (0.58)	3.99 (0.65)	3.88 (0.64)	-
Communication	4.18 (0.57)	3.77 (0.69)	-	-	-

*Specialties include: surgery, gynecology, ENT, ophthalmology, orthopedics, urology, plastic surgery, oral surgery, anesthesiology, cardiology, pediatrics, gastroenterology, neurology, radiology, psychiatry, dermatology, medical microbiology, geriatrics, rheumatology, pathology, internal medicine, intensive care, emergency aid, clinical chemistry, lung diseases, allergology, medical psychology, mental health, rehabilitation medicine (29 specialties in total).

Table 5.2 *Group Monitor instrument – items and factor loadings on the identified subscales for all four rater classes*

Item	Factor loadings				
	This physician group	Medical practice	Organizational involvement	Professionalism	Coordination
M1..has the right knowledge and skills		0.76 (self) 0.78 (staff) <u>0.80 (peers)</u> 0.64 (managers)			
M2..provides the right care in an acute situation		0.55 (self) 0.75 (staff) <u>0.68 (peers)</u> 0.79 (managers)			
M3..shows consistency in their action		0.39 (self) <u>0.42 (peers)</u> 0.66 (managers)			0.58 (staff)
M4..provides good (patient) care		0.70 (self) 0.69 (staff) <u>0.60 (peers)</u> 0.76 (managers)			
P1..puts the patient's interest first		0.65 (self) 0.44 (staff) <u>0.44 (peers)</u> 0.70 (managers)			
P2..holds a good balance in proximity and professional distance		0.41 (self) 0.49 (staff) 0.68 (managers)		0.62 (peers)	
P3..discusses unprofessional behavior				0.61 (self) 0.62 (staff) <u>0.75 (peers)</u> 0.28 (managers)	
P4..acts with integrity				0.59 (self) 0.57 (staff) <u>0.75 (peers)</u> 0.53 (managers)	
P5..shows respect to people with whom they work together				0.80 (self) 0.74 (staff) <u>0.82 (peers)</u> 0.71 (managers)	
P6..contributes to an open discussion culture				0.79 (self) 0.68 (staff) <u>0.88 (peers)</u> 0.55 (managers)	
P7..is open to criticism				0.66 (self) 0.58 (staff) <u>0.78 (peers)</u> 0.52 (managers)	

Table 5.2 Continued

Item	This physician group	Medical practice	Organizational involvement	Professionalism	Coordination	Communication
I1..is open to new techniques/ technologies/ innovation			0.41 (self) 0.71 (staff) <u>0.50 (peers)</u> 0.51 (managers)			
I2..invests in research			0.77 (self) 0.93 (staff) <u>0.81 (peers)</u> 0.91 (managers)			
I3..is active in committees/ administrative activities			0.41 (self) 0.76 (staff) <u>0.82 (peers)</u> 0.75 (managers)			
I4..plays an active role in training and education		0.52 (self)	0.68 (staff) <u>0.60 (peers)</u> 0.73 (managers)			
I5..deals carefully with (consequences of) reports (complaints)			0.40 (staff)	0.36 (self) <u>0.55 (peers)</u> 0.44 (managers)		
O1..clarifies who I can contact for which case					0.69 (self) 0.57 (staff) <u>0.54 (peers)</u> 0.61 (managers)	
O2..is accessible to discuss something				0.50 (staff) 0.64 (managers)	0.62 (self) <u>0.53 (peers)</u>	
O3..is easily accessible					0.76 (self) 0.59 (staff) <u>0.74 (peers)</u> 0.56 (managers)	
O4..works effectively with other colleagues					0.60 (self) 0.56 (staff) <u>0.63 (peers)</u> 0.61 (managers)	
O5..is well organized			0.43 (self)		0.49 (staff) <u>0.65 (peers)</u> 0.76 (managers)	
O6..acts as an entity			0.44 (self)		<u>0.65 (peers)</u> 0.81 (managers)	0.48 (staff)
O7..presents itself clearly			0.50 (self) <u>0.53 (peers)</u> 0.63 (managers)		0.36 (staff)	

Table 5.2 Continued

Item This physician group	Medical practice	Organizational involvement	Professionalism	Coordination	Communi- cation
T1..provides a clear answer in a consultation situation	<u>0.86 (peers)</u> 0.74 (managers)			0.54 (self) 0.58 (staff)	
T2..complies with agreements made	0.58 (peers)			0.33 (self) 0.77 (staff) 0.68 (managers)	
T3..performs adequately in a consultation situation	0.62 (peers)			0.45 (self) 0.62 (staff) 0.57 (managers)	
T4..engages others in a timely manner to deliver optimal care	<u>0.57 (peers)</u>			0.75 (self) 0.75 (staff) 0.47 (managers)	
T5..uses available knowledge/facilities effectively	0.45 (peers)			0.78 (self) 0.69 (staff) 0.39 (managers)	
T6..holds a good relationship with management/board		<u>0.58 (peers)</u> 0.53 (managers)		0.39 (self) 0.58 (staff)	
C1..provides clear information				<u>0.38 (peers)</u> 0.74 (managers)	0.72 (self) 0.52 (staff)
C2..listens carefully to my input			0.60 (staff) <u>0.48 (peers)</u>	0.51 (managers)	0.80 (self)
C3..approaches others empathically and respectfully			0.85 (staff) <u>0.70 (peers)</u> 0.79 (managers)		0.67 (self)
C4..acts in a professional way			0.71 (staff) <u>0.71 (peers)</u> 0.69 (managers)		0.56 (self)
C5..seeks contact timely in case of a problem			0.39 (self) <u>0.37 (peers)</u>	0.71 (managers)	0.42 (staff)
C6..listens carefully to each other				<u>0.52 (peers)</u> 0.69 (managers)	0.66 (self) 0.67 (staff)

Table 5.3 Group Monitor instrument - G study: variance by source for two models

Model 1: Hospital				
Variance component (%)	Self*	Staff	Peers	Managers
Hosp:Group:Rater		0.116 (29)	0.163 (39)	0.121 (29)
Group:Rater	0.103 (35)	0.076 (19)	0.070 (17)	0.073 (17)
Group	0.076 (26)	0.048 (12)	0.027 (7)	0.052 (12)
Hosp		0.001 (0)	0.050 (12)	0.057 (14)
Subscales	0.026 (9)	0.051 (13)	0.011 (3)	0.019 (4)
Residual	0.086 (30)	0.108 (27)	0.089 (22)	0.103 (24)
Model 2: Specialty				
Variance component (%)	Self	Staff	Peers	Managers
Spec:Group:Rater	0.061 (21)	0.116 (29)	0.104 (26)	0.117 (30)
Group:Rater	0.042 (14)	0.077 (19)	0.129 (32)	0.076 (20)
Group	0.064 (22)	0.038 (9)	0.052 (13)	0.064 (16)
Spec	0.013 (4)	0.010 (3)	0.011 (3)	0.009 (2)
Subscales	0.026 (9)	0.051 (13)	0.011 (3)	0.019 (5)
Residual	0.086 (30)	0.108 (27)	0.090 (23)	0.104 (27)

*Due to uneven distribution of raters over the hospitals, this model could not be statistically identified. Therefore, a smaller facet model is used for this case.

Table 5.4 Group Monitor instrument - D study: the G-coefficient and SEM as a function of the number of raters (N) of the four rater classes

	Model 1: Hospital			Model 2: Specialty		
	N ^a	G ^b	SEM ^c	N ^a	G ^b	SEM ^c
Self	3	0.64	0.20	3	0.60	0.20
Staff	7	0.61	0.17	9	0.61	0.15
Peers	15	0.61	0.13	8	0.62	0.18
Managers	7	0.62	0.18	6	0.63	0.19

a N = number of raters needed for reliable measurements. Each rater equals one evaluation.

b a G-coefficient of ≥ 0.60 is considered acceptable.

c a SEM of ≤ 0.26 is considered acceptable.

DISCUSSION

The purpose of this study was to investigate whether the GM instrument, a uniform questionnaire developed for four different rater classes (physicians/self-assessment, staff, peers and managers), provides reliable and valid information about the performance of specialty specific physician groups. The GM taps into behavioral aspects of group performance, such as communication, collaboration, use of all contributors' expertise and sharing workload. Using the APA framework³⁴, three categories of validity evidence were tested: (1) content, (2) response process and (3) internal structure.

- The content was supported by the fact that the instrument was developed in close collaboration with the target group (specialty specific physician groups) and the fact that the items were based on a review of an existing instrument.
- The response process was investigated in two pilot rounds, leading to reformulation and reduction of items, addition of global items and adjustment of the answer scale.
- The internal structure was examined in two ways, by studying the psychometric properties of the GM for each rater class separately, and performing G studies for each rater class separately. The principal component analysis identified an underlying structure of four performance subscales for all rater classes: (1) *Medical practice*, (2) *Organizational involvement*, (3) *Professionalism* and (4) *Coordination*. The rater class 'self' and 'staff' appeared to have an additional 5th performance scale, *Communication* (Table 5.2). Since we found that items clustered differently for the four rater classes, we further examined the interpretation differences between the rater classes. Data was aggregated on rater class level and overall mean scores were compared among the classes. The overall mean scores did correlate, but not very strongly. The strongest correlation was the correlation between 'self' and 'peers' (0.65), which is logical, given the comparability of (the nature of) their professional activities.

Given that items clustered differently for the rater classes and correlations between the overall mean scores were not particularly high, we may infer that different rater classes indeed perceive and measure aspects of group performance differently. The use of a uniform questionnaire for different rater classes thus comes with limitations in terms of comparability of scores. It could be reasoned that these mutual differences are an expression of legitimate, experience-based interpretations of group performance. In this context, Crossley et al²³ state that rater classes differ in their cognitive structures or frames of reference, also called *designations*. These designations stem from different standards and thus different experiences with the physician group that was evaluated. The results of the G studies, however, showed that these designations are not only an attribute of the specific rater class, but that they are also colored by aspects of the working environment.

For both G models, 70-88% of the variance in GM scores was explained by the facets, which means that the models did reflect the main sources of bias. The hospital and specialty of the rater, together with the specific composition of the rater panels (facets Hosp:Group:Rater, Spec:Group:Rater and Group:Rater) explained 50% or more of the variance. The variance explained by the several subscales was small (although a bit higher for 'self' and 'staff' due to the extra 5th subscale), which would indicate that the identified constructs in the subscales only had a limited contribution. The percentage of variance that was true variance of the group score (facet 'Group') was relatively low for the rater classes staff, peers and managers (7-16%), indicating that there was a lot of bias. This percentage was much higher for the rater class self (22-26%), pointing to relatively less bias.

Based on these findings, we may infer that physicians themselves could most reliably rate their group performance. This was also reflected by the results of the D studies: for 'self' only 3 ratings were needed for reliable scores, compared to 6-15 ratings for staff, peers and managers. Staff and managers were comparable to each other in terms of the number of ratings needed. Peers showed the largest spread in terms of the number of ratings needed: when the hospital was taken into account, twice as many ratings were needed than when specialty was considered. This finding may be explained by the uneven distribution of groups across hospitals in combination with the large spread in the number of peer-raters per physician group (see Table 5.1).

The finding that the rater class 'self' apparently reached agreement about their group performance the fastest may be an expression of the phenomenon of *in-group favoritism*.³⁴ Social psychology research shows that under certain circumstances, people prefer and have affinity for one's in-group, which can be expressed in one's evaluation of others (e.g. colleagues with whom one works on a daily basis).^{35,36} The finding that of all rater classes, the rater class 'self' had the highest total GM score (and highest subscale scores, although subscales did not consist of exactly the same items for each rater class, see Table 5.1) could also point to this phenomenon.

Strengths and limitations

To the best of our knowledge, this study was the first to examine validity evidence of a multisource feedback instrument for the evaluation of the performance of specialty specific physician groups. It is also one of the few studies that explore the validity of the same instrument with a uniform questionnaire for different rater classes.

- In terms of methodology, it should be noted that principal component analyses were performed on a robust sample, thereby contributing to the stability of the findings.³⁷ Follow-up research is needed to determine whether the proposed structure remains intact following a confirmatory factor analysis in a larger

- sample. Nevertheless, it must be kept in mind that the distinction of subscales only explained a relatively small percentage of variance in GM scores.
- Furthermore, it must be noted that – consistent with other MSF tools – self, staff, peer and manager ratings were skewed toward favorable impressions of group performance.^{38,39} In research into MSF for *individual* performance, an explanation for these positive ratings could be the individual's self-selection of raters, which may result in selecting only positive-minded raters. In the case of MSF for *group* performance, such as the Group Monitor instrument, this explanation would be less obvious from a practical point of view.
 - A last point of attention is the fact that the GM was developed to be used in the Dutch hospital setting. Therefore, caution is required when this instrument is used in other healthcare systems or other types of settings, for example the primary care setting. Items may need to be changed to reflect these differences and then further validated.

Implications for practice and future research

The inclusion of different perspectives in evaluation procedures adds value to any evaluation. In the case of the GM, the outcomes per rater class can inform subsequent improvement of group performance by identifying qualities and areas for improvement per rater class. As mentioned before, the interpretation of the subscales differed for the four rater classes. When interpreting the GM results, physician groups should be aware of the fact that it is not possible to compare the evaluations of the rater classes *per subscale*. Instead, the evaluations of the different respondent groups should be treated as separate sources of feedback about the group performance, providing a diverse palette of information.⁴⁰ In addition, physician groups should keep in mind that the GM results should be interpreted within the specialty/hospital specific context. For reliable measurements, it is crucial to invite sufficient raters, especially for the rater class 'peers' – depending on whether this is feasible for the physician group in question. Discussing the GM results within the physician group could support the group in implementing improvement plans and could therefore be used for quality assurance purposes, for example in the context of an external quality review.^{41,42,43}

To conclude, the importance should be underscored of viewing validation as an ongoing process. Validity evidence needs to be updated over time for the instrument's continued relevant and appropriate use in various contexts and groups. In this study we have looked into the first three categories of validity evidence (content, response process and internal structure); further research should explore the last two categories (relation to other variables and consequences).¹⁴ GM scores should be examined in relation to other measures of group performance and to measures pertaining to the

other quality domains within the aforementioned five-yearly quality review. Also, investigating the effects of the GM on practice would be a useful step in considering the effects of the measurement, thereby contributing to the discussion on how to improve physician group performance as part of quality assurance programs.

CONCLUSIONS

The Group Monitor instrument is a uniform questionnaire for various rater classes, providing valid and reliable formative feedback on the performance of specialty specific physician groups. It should be noted that different rater classes perceive or experience aspects of group performance differently. Therefore, GM results of different rater classes should not be compared on the level of performance subscales, but should be treated as separate valuable information sources about group performance, to be considered in relation to each other. Also, when interpreting the GM results, it should be kept in mind that rater classes perceptions of group performance are colored by the professional culture within the hospital and/or the specialty of the physician group in question. Future research should investigate whether and how physician groups use this type of feedback in their ongoing pursuit of group performance and quality improvement.

LESSONS FOR PRACTICE

- The Group Monitor (GM) instrument is a uniform MSF tool which provides feedback on the performance of specialty specific physician groups from four rater classes: (1) physicians (self-assessment), (2) staff, (3) peers and (4) managers.
- Four subscales are distinguished: *Medical practice*, *Organizational involvement*, *Professionalism* and *Coordination*; the rater classes 'physicians' and 'staff' have an additional subscale, *Communication*.
- Rater classes perceive items differently; therefore physician groups should treat the GM results of the rater classes as separate information sources about their group performance.
- It should be kept in mind that rater classes' perceptions of group performance are colored by the hospital culture and the physician group's specialty.

REFERENCES

1. Hodges B. Assessment in the post-psychometric era: Learning to love the subjective and collective. *Med Teach*. 2013;35:564-568.
2. KNMG 'Kwaliteitskader medische zorg: Staan voor kwaliteit', 2012. www.knmg.nl/web/file?uuid=5807beb4-60b5-4443-a09d-f1bb592a36b3&owner=5c945405-d6ca-4deb-aa16-7af2088aa173&contentid=698. Accessed 13 Feb 2018.
3. Federatie Medisch Specialisten. Vision document 'De medisch specialist 2015'. 2012. www.demedischspecialist.nl/sites/default/files/Visiedocument%20web.pdf. Accessed 13 Feb 2018.
4. Federatie Medisch Specialisten. Vision document 'De medisch specialist 2025'. 2017. www.demedischspecialist.nl/sites/default/files/Visiedocument%20Medisch%20Specialist%202025-DEF.pdf. Accessed 13 Feb 2018.
5. Jeffcott SA, Mackenzie CF. Measuring team performance in healthcare: Review of research and implications for patient safety. *J Crit Care*. 2008;23, 188-196.
6. Rosen MA, Bedwell WL, Wildman JL, Fritzsche BA, Salas E, Burke, CS. Managing adaptive performance in teams: Guiding principles and behavioral markers for measurement. *HRMR*. 2011;21(2),107-122.
7. Valentine MA, Nembhard IM, Edmondson AC. Measuring teamwork in health care settings: a review of survey instruments. *Medical Care*. 2015;53(4), e16-e30.
8. Marlow S, Bisbey T, Lacerenza C, Salas E. Performance measures for health care teams: a review. *Small Group Research*. 2018;49:306-356.
9. Schulpen TW, Lombarts KM. Quality improvement of paediatric care in the Netherlands. *Arch Dis Child*, 2007;92:633-636.
10. Fossen JA, Hagemeyer JW, de Koning JS, van Logtestijn SL, Lombarts MJMH. Kwaliteitsvisite nieuwe stijl. Handboek voor wetenschappelijke verenigingen. Alphen aan den Rijn: Van Zuiden Communications, 2005.
11. Sargeant J, Bruce D, Campbell CM. Practicing physicians' needs for assessment and feedback as part of professional development. *J Contin Educ Health*. 2013;33:554-562.
12. Donnon T, Al Ansari A, Al Alawi S, et al. The reliability, validity, and feasibility of multisource feedback physician assessment: a systematic review. *Acad Med*. 2014;89:511-516.
13. Al Ansari A, Donnon T, Al Khalifa K, et al. The construct and criterion validity of the multisource feedback process to assess physician performance: a meta-analysis. *Adv Med Educ Pract*. 2014;5:39-51.
14. Downing SM. Validity: on the meaningful interpretation of assessment data. *Med Educ*. 2003;37:830-837.
15. Cook DA, Zendejas B, Hamstra SJ, et al. What counts as validity evidence? Examples and prevalence of a systematic review of simulation-based assessment. *Adv Health Sci Educ*. 2014;19:233-250.
16. Crossley J, Russell, J, Jolly B, et al. 'I'm pickin' up good regressions': the governance of generalisability analyses. *Med Educ*. 2007;41:926-934.
17. Brennan RL. Generalizability theory. New York: Springer. doi.org/10.1007/978-1-4757-3456-0%20(2001). Accessed 26 April 2019).
18. Federatie Medisch Specialisten. Guideline 'Leidraad Waarderingsstelsel voor de Kwaliteitsvisites'. www.demedischspecialist.nl/onderwerp/kwaliteitsvisites. Accessed 18 Feb 2019.

19. Federatie Medisch Specialisten. Quick Scan Manual 'Quick Scan Handleiding'. www.demedischspecialist.nl/onderwerp/kwaliteitsvisitaties. Accessed 18 Feb 2019.
20. Frank JR, Danoff D. The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. *Med Teach*. 2007;29: 642-647.
21. Van der Meulen MW, Boerebach BC, Smirnova A, Heeneman S, Oude Egbrink MG, van der Vleuten CP, Arah OA, Lombarts, KM. Validation of the INCEPT: a multisource feedback tool for capturing different perspectives on physicians' professional performance. *J Contin Educ Health*. 2017;37:9-18.
22. Overeem K, Lombarts MJ, Arah OA, Klazinga NS, Grol, RPTM, Wollersheim HC. Three methods of multi-source feedback compared: a plea for narrative comments and coworkers' perspectives. *Med Teach*. 2010;32(2), 1441-147.
23. Crossley J, Jolly, B. Making sense of work-based assessment: ask the right questions, in the right way, about the right things, of the right people. *Med Educ*. 2012;46:28-37.
24. Gingerich A, Kogan J, Yeates P, Govaerts M, Holmboe E. Seeing the 'black box' differently: assessor cognition from three research perspectives. *Med Educ*. 2014;48:1055-1068.
25. Costello AB, Osborne JW. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *PARE*. 2005;10(7), 1-9.
26. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16,297-334.
27. Arah OA, Hoekstra JB, Bos AP. New tools for systematic evaluation of teaching qualities of medical faculty: results of an ongoing multi-center survey. *PLoS One*. 2011;6, e25983.
28. Streiner DL, Norman GR. Health measurement scales: A practical guide to their development and use. Oxford UK: Oxford University Press, 2008.
29. Gronlund NE. How to construct achievement tests. Englewood Cliffs, NJ: Prentice-Hall, 1988.
30. Boor K, Van der Vleuten C, Teunissen T, Scherpbier A, Scheele F. Development and analysis of D-RECT, an instrument measuring residents' learning climate. *Med Teach*. 2011;33:820-827.
31. Boerboom TB, Dolmans DHJM, Jaarsma ADC, Muijtjens AMM, Van Beukelen P, Scherpbier AJJA. Exploring the validity and reliability of a questionnaire for evaluating veterinary clinical teachers' supervisory skills during clinical rotations. *Med Teach*. 2011;33:e84-e91.
32. Silkens ME, Smirnova A, Stalmeijer RE, Arah OA, Scherpbier AJ, Van der Vleuten CP, Lombarts, KM. Revisiting the D-RECT tool: validation of an instrument measuring residents' learning climate perceptions. *Med Teach*. 2016;38:476-481.
33. Zhehan, J. Using the Linear Mixed-Effect Model Framework to Estimate Generalizability Variance Components in R: A lme4 Package Application. *Methodology*;14:133-142.
34. Tropp LR, Wright SC. Ingroup identification as the inclusion of ingroup in the self. *Pers Soc Psychol Bull*. 2001;27:585-600.
35. Chen Y, Li SX. Group identity and social preferences. *Am Econ Rev*. 2009;99:431-457.
36. Turner JC, Reynolds KJ. "Self-categorization theory," in: *Handbook of Theories in Social Psychology*. UK: Sage Publications, 2011.
37. Wetzel AP. Factor analysis methods and validity evidence: A review of instrument development across the medical education continuum. *Acad Med*. 2012;87:1060-1069.
38. Boerebach BC, Arah OA, Heineman MJ. Embracing the complexity of valid assessments of clinicians' performance: A call for in-depth examination of methodological and statistical contexts that affect the measurement of change. *Acad Med*. 2016;91:215-220.

39. Beckman TJ, Ghosh AK, Cook DA, et al. How reliable are assessments of clinical teaching? A review of the published instruments. *J Gen Intern Med.* 2004;19:971-977.
40. 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: the effects of multiple occasions and assessor groups. *Acad Med.* 2015;90:1093-1099.
41. Sargeant JM, Mann KV, Van der Vleuten CP, Metsemakers JF. Reflection: a link between receiving and using assessment feedback. *Adv Health Sci Educ.* 2009;14:399-410.
42. DeNisi AS, Kluger AN. Feedback effectiveness: Can 360-degree appraisals be improved? *Acad Manag Perspect.* 2000;14:129-139.
43. Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, O'Brien MA, Johansen M, Grinshaw J, Oxman AD. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev.* 2012;6.

APPENDIX 5.1

Group Monitor instrument – corrected item-total correlations and Cronbach's α for the subscales of each of the four rater classes

Item This physician group...	Corrected item- total correlation	Cronbach's α
Physicians (self-assessment)		
<i>Subscale 1: Medical Practice</i>		<i>0.83</i>
M1 ... has the right knowledge and skills	0.63	
M4 ... provides good (patient) care	0.68	
P1 ... puts the patient's interest first	0.59	
M2 ... provides the right care in an acute situation	0.62	
I4 ... plays an active role in training and education	0.48	
P2 ... holds a good balance in proximity and professional distance	0.54	
M3 ... shows consistency in their action	0.56	
<i>Subscale 2: Organizational involvement</i>		<i>0.83</i>
I2 ... invests in research	0.48	
O7 ... presents itself clearly	0.69	
O6 ... acts as an entity	0.74	
O5 ... is well organized	0.69	
I1 ... is open to new techniques / technologies / innovation	0.48	
I3 ... is active in committees / administrative activities	0.54	
<i>Subscale 3: Professionalism</i>		<i>0.87</i>
P5 ... shows respect to people with whom they work together	0.65	
P6 ... contributes to an open discussion culture	0.74	
P7 ... is open to criticism	0.70	
P3 ... discusses unprofessional behavior	0.70	
P4 ... acts with integrity	0.62	
C5 ... seeks contact timely in case of a problem	0.61	
I5 ... deals carefully with (consequences of) reports (complaints)	0.56	
<i>Subscale 4: Coordination</i>		<i>0.89</i>
T5 ... uses available knowledge/facilities effectively	0.64	
O3 ... is easily accessible	0.55	
T4 ... engages others in a timely manner to deliver optimal care	0.70	
O1 ... clarifies who I can contact for which case	0.58	
O2 ... is accessible to discuss something	0.66	
O4 ... works effectively with other colleagues	0.74	
T1 ... provides a clear answer in a consultation situation	0.61	
T3 ... performs adequately in a consultation situation	0.67	
T6 ... holds a good relationship with management/board	0.57	
T2 ... complies with agreements made	0.61	

Item This physician group...	Corrected item- total correlation	Cronbach's α
<i>Subscale 5: Communication</i>		0.87
C2 ... listens carefully to my input	0.77	
C1 ... provides clear information	0.63	
C3 ... approaches others empathically and respectfully	0.69	
C6 ... listens carefully to each other	0.74	
C4 ... acts in a professional way	0.65	
Staff		
<i>Subscale 1: Medical Practice</i>		0.88
M1 ...has the right knowledge and skills	0.73	
M2 ...provides the right care in an acute situation	0.71	
M4 ...provides good (patient) care	0.77	
P2 ... holds a good balance in proximity and professional distance	0.71	
P1 ... puts the patient's interest first	0.66	
<i>Subscale 2: Organizational involvement</i>		0.85
I2 ... invests in research	0.67	
I3 ... is active in committees / administrative activities	0.75	
I1 ... is open to new techniques / technologies / innovation	0.64	
I4 ... plays an active role in training and education	0.73	
I5 ... deals carefully with (consequences of) reports (complaints)	0.57	
<i>Subscale 3: Professionalism</i>		0.93
C3 ... approaches others empathically and respectfully	0.75	
P5 ... shows respect to people with whom they work together	0.76	
C4 ... acts in a professional way	0.77	
P6 ... contributes to an open discussion culture	0.82	
P3 ... discusses unprofessional behavior	0.72	
C2 ... listens carefully to my input	0.71	
P7 ... is open to criticism	0.80	
P4 ... acts with integrity	0.69	
O2 ... is accessible to discuss something	0.69	
<i>Subscale 4: Coordination</i>		0.93
T2 ... complies with agreements made	0.73	
T4 ... engages others in a timely manner to deliver optimal care	0.76	
T5 ... uses available knowledge/facilities effectively	0.69	
T3 ... performs adequately in a consultation situation	0.78	
O3 ... is easily accessible	0.62	
T1 ... provides a clear answer in a consultation situation	0.69	
T6 ... holds a good relationship with management/board	0.69	
O1 ... clarifies who I can contact for which case	0.66	
O4 ... works effectively with other colleagues	0.78	
O5 ... is well organized	0.78	
O7 ... presents itself clearly	0.67	

Item	Corrected item- total correlation	Cronbach's α
This physician group...		
<i>Subscale 5: Communication</i>		<i>0.86</i>
C6 ... listens carefully to each other	0.76	
M3... shows consistency in their action	0.61	
C1 ... provides clear information	0.70	
O6 ... acts as an entity	0.70	
C5 ... seeks contact timely in case of a problem	0.65	
Peers		
<i>Subscale 1: Medical Practice</i>		<i>0.94</i>
M2...provides the right care in an acute situation	0.76	
T3 ... performs adequately in a consultation situation	0.80	
T4 ... provides good (patient) care	0.80	
T2 ... complies with agreements made	0.74	
T4 ... engages others in a timely manner to deliver optimal care	0.79	
T5 ... uses available knowledge/facilities effectively	0.76	
P1 ... puts the patient's interest first	0.72	
M3...shows consistency in their action	0.72	
<i>Subscale 2: Organizational involvement</i>		<i>0.85</i>
I3 ... is active in committees / administrative activities	0.66	
I2 ... invests in research	0.72	
I4 ... plays an active role in training and education	0.71	
O7 ... presents itself clearly	0.64	
I1 ... is open to new techniques / technologies / innovation	0.64	
<i>Subscale 3: Professionalism</i>		<i>0.95</i>
P6 ... contributes to an open discussion culture	0.84	
P5 ... shows respect to people with whom they work together	0.83	
P7 ... is open to criticism	0.80	
P4 ... acts with integrity	0.79	
P3 ... discusses unprofessional behavior	0.77	
C4 ... acts in a professional way	0.79	
C3 ... approaches others empathically and respectfully	0.79	
P2 ... holds a good balance in proximity and professional distance	0.78	
T6 ... holds a good relationship with management/board	0.68	
I5 ... deals carefully with (consequences of) reports (complaints)	0.69	
C2 ... listens carefully to my input	0.75	
C5 ... seeks contact timely in case of a problem	0.74	
<i>Subscale 4: Coordination</i>		<i>0.91</i>
O3 ... is easily accessible	0.68	
O5 ... is well organized	0.78	
O6 ... acts as an entity	0.75	
O4 ... works effectively with other colleagues	0.79	
O1 ... clarifies who I can contact for which case	0.63	
O2 ... is accessible to discuss something	0.74	
C6 ... listens carefully to each other	0.77	
C1 ... provides clear information	0.66	

Item This physician group...	Corrected item- total correlation	Cronbach's α
Managers		
<i>Subscale 1: Medical Practice</i>		<i>0.90</i>
M2 ...provides the right care in an acute situation	0.79	
M4 ...provides good (patient) care	0.76	
T1 ... provides a clear answer in a consultation situation	0.70	
P1 ... puts the patient's interest first	0.72	
P2 ... holds a good balance in proximity and professional distance	0.74	
M3... shows consistency in their action	0.65	
M1 ...has the right knowledge and skills	0.65	
<i>Subscale 2: Organizational involvement</i>		<i>0.85</i>
I2 ... invests in research	0.65	
I3 ... is active in committees / administrative activities	0.66	
I4 ... plays an active role in training and education	0.76	
O7 ... presents itself clearly	0.65	
I1 ... is open to new techniques / technologies / innovation	0.60	
<i>Subscale 3: Professionalism</i>		<i>0.93</i>
C3 ... approaches others empathically and respectfully	0.76	
P5 ... shows respect to people with whom they work together	0.75	
C4 ... acts in a professional way	0.74	
O2 ... is accessible to discuss something	0.68	
P6 ... contributes to an open discussion culture	0.76	
T6 ... holds a good relationship with management/board	0.76	
P4 ... acts with integrity	0.69	
P7 ... is open to criticism	0.79	
I5 ... deals carefully with (consequences of) reports (complaints)	0.69	
P3 ... discusses unprofessional behavior	0.73	
<i>Subscale 4: Coordination</i>		<i>0.94</i>
O6 ... acts as an entity	0.79	
O5 ... is well organized	0.81	
C1 ... provides clear information	0.74	
C5 ... seeks contact timely in case of a problem	0.73	
C6 ... listens carefully to each other	0.79	
T2 ... complies with agreements made	0.75	
O1... clarifies who I can contact for which case	0.64	
O4 ... works effectively with other colleagues	0.80	
T3 ... performs adequately in a consultation situation	0.74	
O3 ... is easily accessible	0.68	
C2 ... listens carefully to my input	0.72	
T4 ... engages others in a timely manner to deliver optimal care	0.69	
T5 ... uses available knowledge/facilities effectively	0.66	

APPENDIX 5.2

Group Monitor instrument – inter-scale correlations and scale-global correlations for each of the four rater classes

Physicians (self-assessment)					
	Medical practice	Organizational involvement	Professionalism	Coordination	Communication
Medical practice	1	-	-	-	-
Organizational involvement	0.65	1	-	-	-
Professionalism	0.65	0.65	1	-	-
Coordination	0.69	0.73	0.74	1	-
Communication	0.60	0.65	0.71	0.68	1
Global item 1: This is a physician group that I trust	0.63	0.64	0.63	0.66	0.60
Global item 2: This is a physician group that I recommend	0.63	0.66	0.57	0.60	0.60
Staff					
	Medical practice	Organizational involvement	Professionalism	Coordination	Communication
Medical practice	1	-	-	-	-
Organizational involvement	0.66	1	-	-	-
Professionalism	0.71	0.59	1	-	-
Coordination	0.75	0.64	0.84	1	-
Communication	0.67	0.53	0.80	0.82	1
Global item 1: This is a physician group that I trust	0.67	0.56	0.79	0.73	0.70
Global item 2: This is a physician group that I recommend	0.64	0.52	0.72	0.71	0.67

Peers				
	Medical practice	Organizational involvement	Professionalism	Coordination
Medical practice	1	-	-	-
Organizational involvement	0.71	1	-	-
Professionalism	0.86	0.68	1	-
Coordination	0.83	0.71	0.86	1
Global item 1: This is a physician group that I trust	0.82	0.67	0.82	0.78
Global item 2: This is a physician group that I recommend	0.73	0.62	0.71	0.72
Managers				
	Medical practice	Organizational involvement	Professionalism	Coordination
Medical practice	1	-	-	-
Organizational involvement	0.66	1	-	-
Professionalism	0.76	0.68	1	-
Coordination	0.73	0.66	0.84	1
Global item 1: This is a physician group that I trust	0.66	0.59	0.80	0.76
Global item 2: This is a physician group that I recommend	0.66	0.67	0.77	0.78

All correlations had $p < 0.01$

APPENDIX 5.3

Group Monitor instrument - comparisons between the four rater classes: correlations between overall mean scores, aggregated on the level of the rater class

		Overall mean score on item 1-35			
		Self	Staff	Peers	Managers
Overall mean score on item 1-35	Self	1	-	-	-
	Staff	0.54	1	-	-
	Peers	0.65	0.55	1	-
	Managers	0.59	0.54	0.56	1

All correlations had $p < 0.01$

6

GENERAL DISCUSSION

Inherent to being a physician is the willingness to provide the best possible patient care. In order to remain a good physician, physicians need to engage in lifelong learning and performance improvement. Currently, our understandings of physician professional learning are in transition. While traditionally viewed rather narrowly as continuing medical education (CME), it is now viewed more broadly as continuing professional development (CPD).^{1,2} In the past decade, participation in continuing professional development (CPD) programs has become increasingly interlinked with procedures for professional re-registration.³⁻⁶ This thesis was written with the understanding of this state of affairs as an attempt to stimulate physicians' efforts in line with their responsibility for their performance. CPD programs aim to address physicians' intrinsic motivation for ongoing learning and development and encourage them to improve collaboration with other healthcare professionals. To guide learning and practice improvement, CPD requires physicians to use assessment and feedback data from the workplace. This approach aims to raise the quality of physicians' performance to a higher level, in line with the adage "doing well, getting better". In this thesis, the following research question was addressed:

How can reflection on physicians' professional performance be best facilitated to serve their developmental needs?

In this last chapter, we will pinpoint the main findings of the studies and discuss these findings in relation to existing literature. The main findings will be discussed under the header of two themes: (1) the facilitation of reflection and the process of fostering physicians' needs for competence, autonomy and relatedness, and (2) physicians' collective competence as complementary to their individual competence. Thereafter, we pay attention to strengths and limitations, and consider implications for practice and directions for future research.

MAIN FINDINGS

Performance appraisal systems – It is possible to design, develop and implement a performance appraisal system that complies with accountability requirements and stimulates professional development. To facilitate reflection on professional performance, these systems dictate the use of multisource feedback, discussion of the feedback results with a trained facilitator and design of a personal development plan (chapter 3 and 4).

Multisource feedback tools - Multisource feedback can be collected at the level of the individual physician, but also at the level of the physician group. There is now a tool (GM) available to feasibly and reliably collect feedback data on physician group performance. The tool captures five performance domains of a physician group's

performance, such as medical practice, professionalism and communication. Various respondent groups (multiple sources) view the physician groups differently (chapter 5).

Facilitation of reflection – ‘Reflection’ can be characterized as a ‘fuzzy’, relational and interactional process, which takes place both ‘within people’ (internal dialogue) as well as ‘between people’ (external dialogue). Reflection’s fuzziness appears to be closely linked to the complexity of practice and the power dynamics inherent to working with other professionals in a complex healthcare organization (chapter 2). Prerequisites for facilitating reflection on professional performance are a challenging yet psychologically safe environment, careful attention to the process of reflection and the opportunity to (privately) speak with a skilled and trustworthy coach (chapter 2, 3 and 4).

Conversation formats – Conversation formats to guide physicians’ reflection on their professional performance include a one-on-one conversation between the individual physician and a coach (chapter 3), and a group-based conversation led by a coach (chapter 4). Both formats can contribute to physicians’ individual professional development, albeit in a different way. The one-on-one format offers protected quality time and assistance with the development of an improvement plan that matches individual learning needs and interests. The group-based format offers high levels of interactivity and efficiency and established a clear link between the individual physician and his or her work environment. The characteristics of the two formats, their perceived contribution to professional development and (contra) indications for each are summarized in Table 6.1.

Realizing change - Physicians who were individually coached reported that they felt safe to discuss personal matters. The one-on-one conversation did not reveal new insights about their performance but was instrumental in rendering, accepting or internalizing feedback that they received. The conversation did deepen their insights in their developmental opportunities; physicians were highly committed to and motivated for attaining the improvement goal that they formulated during the session (chapter 3). Physicians who participated in a group-based conversation reported that a group-based format could potentially elicit feelings of vulnerability. However, they also reported that listening to colleagues’ additional explanation and nuance could increase the meaningfulness of feedback, create urgency for improvement, deepen collegial relationships and strengthen feelings of relatedness. Due to the nature of the conversation format, improvement goals were mostly related to relational fine-tuning in collaborative practice (chapter 4).

Table 6.1 Characteristics, perceived contributions and (contra)indications for one-on-one and group-based coaching, based on the findings in chapter 3 and 4.

	One-on-one coaching	Group-based coaching
Characteristics	<ul style="list-style-type: none"> ▪ Duration: 2 hours per medical specialist ▪ Resource intensive ▪ Maximum psychological safety ▪ Based on principles of Appreciative Inquiry and solution-focused coaching ▪ Assistance with development of improvement plan 	<ul style="list-style-type: none"> ▪ Duration: 2,5 hours for the whole group ▪ Efficient ▪ High level of interactivity ▪ Direct link between individual physician and work environment
Perceived contribution to professional development	<ul style="list-style-type: none"> ▪ Accepting and internalizing feedback ▪ Deepening insight into developmental opportunities ▪ Enhancing motivation and commitment to change 	<ul style="list-style-type: none"> ▪ Increasing the meaningfulness of feedback ▪ Creating urgency for improvement ▪ Deepening collegial relationships
Indications	<ul style="list-style-type: none"> ▪ Increased personal vulnerability 	<ul style="list-style-type: none"> ▪ Presence of longitudinal trusting relationships ▪ Group size not more than approx. 7 participants
Contra-indications	<ul style="list-style-type: none"> ▪ None 	<ul style="list-style-type: none"> ▪ Increased personal vulnerability ▪ Heavy organizational disturbances

FOSTERING PHYSICIANS' NEEDS FOR COMPETENCE, AUTONOMY AND RELATEDNESS

CPD activities, such as participation in performance appraisal and discussing feedback, serve to support physicians in their capacity to be self-directed and function capably as team members, to make decisions on their own and know when to ask for help, and to maintain their own health and sense of well-being in the presence of multiple system and other demands.⁷ A theory that can contribute to our understanding of the role of performance appraisal and discussing feedback play in promoting professional growth is the self-determination theory (SDT). This is a theory of personal motivation and development proposing that humans have a natural tendency toward autonomous regulation of behavior and are intrinsically motivated to learn and take on challenges.^{8,9} To do so requires three basic psychological needs: a sense of competence, autonomy and relatedness. The findings of this thesis can be meaningfully discussed in light of these three needs and offer insight in the way in which one-on-one conversations and group conversations can foster each of these needs.

Competence refers to “the desire to feel effective in whatever actions one pursues and performs”.^{8,9} It is about feeling effective in continuous interaction with the social environment and experiencing confidence in one’s own abilities. The findings in chapter 3 show that the newly developed coaching approach (Developmental

Appreciative Navigational Approach, DANA) was received positively because physicians felt confirmed in their abilities and skills as a physician. During these one-on-one conversations, physicians not only discussed their MSF results regarding their role as clinician (INCEPT), but could also bring residents' feedback regarding their role as supervisor (SETQ) into the conversation. The findings in chapter 4 showed that a group conversation could provide a positive boost to the sense of competence of young physicians who were still struggling with their newly acquired status and position within the physician group. By receiving reassurance, appreciation and positive feedback – both through the MSF system as well as during the conversation -, their confidence was strengthened.

Autonomy refers to “the desire to be one’s own origin or source of behavior”.^{8,9} This does not mean acting entirely on one’s own. It means having a sense of free will in one’s actions which carries with it the responsibility to recognize when to ask for help or seek external resources. Since autonomy, competence and insight go hand in hand in relation to one’s professional performance, tensions and sensitivities may arise when in coaching conversations the boundaries of one’s competence come into the picture. In chapter 3, the individual physician was invited by the coach to self-direct his or her learning process while taking into account external feedback. Identifying personal development goals and drawing up an improvement plan for which the physician was intrinsically motivated may well have enhanced the experience of autonomy. The findings in chapter 4 showed that a group session offered the opportunity to discuss differences in working style, so that understanding of each other’s viewpoint was created and more mutual alignment could be realized. In this process, the coach played an important role by steering the conversation in the right direction while ensuring a safe and inclusive atmosphere.

Relatedness refers to “the desire to feel connected with others, to having a sense of belongingness, both with significant individuals as well as with a significant community”.^{8,9} It is not just being a member of that community, but also knowing that one is accepted and valued by its members. The findings of chapter 4 showed that a group session offered an excellent opportunity to foster the need for relatedness: by taking time to reflect on each other’s qualities, trust and interpersonal closeness could arise and physicians could experience more relatedness. For example, physicians took a sincere interest in letting their peers know that “they were there for them”. This finding is consistent with recent educational research that demonstrates that meaningful conversations are grounded in an alliance between feedback provider and feedback recipient.^{10,11} Due to the social and interactive nature of the group format, group sessions may lend themselves better to fostering this basic need than one-on-one sessions. This finding is important, since in the practice of a demanding clinical setting, addressing the need for physicians to feel related in meaningful ways to their team members is often overlooked.¹² However,

it should be kept in mind that internalizing positive feedback during a one-on-one session could also enhance the sense of belonging and relatedness. Regarding the group sessions, it should be taken into account that physicians did not always feel maximally at ease because of the aspect of self-disclosure in a social context. This was especially the case when the group was large and/or consisted of members who did not yet know each other well. In some cases, it was also more difficult to deepen the reflection process and ensure that blind spots were opened up for discussion; participants could use politeness strategies, i.e. giving superficial or non-critical feedback in order to not put pressure on professional relationships.

PHYSICIANS' COLLECTIVE COMPETENCE AS COMPLEMENTARY TO THEIR INDIVIDUAL COMPETENCE

In keeping with the complexity of modern healthcare, both individual and team-based approaches to quality improvement are necessary. These approaches need not only to enhance physicians' individual competence, but also their collective competence. The studies in this thesis have addressed both. Although the individual competence of physicians is affected by contextual factors, it is assumed that during their training physicians develop a level of competence that is more or less permanent. With this in mind, individual competence can be conceptualized as a rather stable possession that, once acquired, holds across contexts.¹³ Collective competence, instead, is conceptualized as a distributed capacity of a system, an evolving relational phenomenon that emerges from the resources and constraints of particular contexts.^{13,14} Thus, delivery of high quality patient care is dynamic and strongly influenced by relationships among team members, which are shaped by perceptions of trust, power and respect, and by structural or material factors like schedules and staffing. Collective competence can vary with the situation, just as individuals may be members of multiple teams and their membership may change.^{13,15} Although the concept of collective competence might be clear, difficulty arises when trying to clarify what the concept would mean translated into physicians' practice. There is still a long way to go to translate the idea of collective competence into competency frameworks, regulatory frameworks, assessment and accreditation strategies. Lingard insists on the importance of not seeing collective competence as antithetical to individual competence, but as complementary: it involves seeing individuals in relation to each other.¹³ Understanding how the personalities and individual competencies in a team intersect, conflict and relate - which can arise during a team-based conversation (chapter 4) - is relevant to collective competence.

It is suggested that CPD will continue to become an increasingly team-based, interactive activity.¹⁶ Team-based CPD can be defined as an activity 'undertaken

after initial qualification when members of two or more health and/or social care professions learn with, from, and about each other to improve collaboration and the quality of care'.^{17,p.143} In the CPD literature, there is an increasing interest in inter-professional teams. Of course, CPD can also be delivered to intra-professional teams, involving learners from the same profession or specialty. In chapter 5 we reported on the development and validation of the Group Monitor instrument, a multisource feedback tool to provide feedback on the performance of intra-professional teams, i.e. specialty-specific physician groups. At this point, a comment regarding the wording 'team' and 'group' is appropriate. Although the words 'team' and 'group' are often used interchangeably, sometimes a distinction is made. A group can be characterized as a collection of individuals with separate goals who are brought together by common interests and experience. Even though everyone shares information and resources with other group members, each member is responsible for their own work.¹⁸ A team can be characterized as an interdependent group of individuals who share responsibility and are focused on a common goal. People in a team have a mutual understanding with other members. By working together, they tend to maximize each other's strengths and minimize weaknesses. Unlike a group, where each member is expected to contribute separately, the most important characteristic of a team is synergy: the whole is greater than the sum of its parts.¹⁸ Although the new feedback tool in chapter 5 has been called *Group Monitor*, it goes without saying that a physician group ideally behaves like a team. After all, providing the best possible patient care is every physician's goal and can only be achieved through collaboration. By using the Group Monitor instrument, physician groups received input from four rater classes: (1) members of the physician group itself (self-assessment), (2) support staff, (3) hospital managers, and (4) peers from other medical specialties. An important finding was that rater classes perceived aspects of group performance differently, probably as a result of different standards and thus different experience-based interpretations of group performance. Group performance ratings were not only influenced by the membership of a particular rater class; they were also colored by the hospitals' professional culture and/or the medical specialty. This latter finding could be considered an indication of many things: the context-specific, fluid and dynamic character of collective competence, the assessment of physicians' individual competencies and many intangible factors, including a long-established reputation of physician groups.

STRENGTHS AND LIMITATIONS

In order to formulate implications for practice and recommendations for future research, first the strengths and limitations of the studies in this thesis will be discussed that should be taken into account. This thesis is among the first attempts to address the professional development of practicing physicians as a continuation

of the learning path that stretches from undergraduate and postgraduate medical education into lifelong learning. Starting from an open exploration of the perspective of physicians themselves is a major strength of this thesis. In addition, the use of various research designs and research methods, the active involvement of relevant stakeholders, the multispecialty and multicenter study design and the collaboration in a multidisciplinary team of researchers have contributed to the quality of this thesis.

Limitations should be considered when interpreting our findings to ensure a nuanced view of the implications for practice and research. In addressing the limitations of our research, the specific limitations of the various studies were discussed in the individual chapters. The first overarching point of attention is the fact that although the ultimate goal of CPD is to enhance the quality and safety of patient care and to enhance health outcomes, it was not investigated how reflective feedback conversations impacted physicians' long-term behavior in practice. Options to address this major limitation are discussed under the header of 'directions for future research'. A second overarching point of attention is the fact that the studies included in this thesis have been conducted in the Dutch hospital setting. It should be acknowledged that there are unique national policies, organizational factors and institutional cultures, which may not be transferable to other settings and healthcare systems outside the Netherlands, or other groups of healthcare professionals, such as nurses. A pleasant observation, however, was that some of the findings in this thesis did appear to be transferable to other settings – at least to some extent. The design and the protocol for implementing the new performance appraisal system in one academic medical center (chapter 3) has been adopted or used as a basis for building the same type of system in other (academic) medical centers.

IMPLICATIONS FOR PRACTICE

The results of these studies provide a number of recommendations for the facilitation of reflection on physicians' professional performance. In the section below, the recommendations are discussed under the header of three themes: (1) supporting physicians' professional development, (2) combining one-on-one and group based coaching, and (3) allocating resources for performance appraisal and coaching.

Supporting physicians' professional development

To support practicing physicians' CPD, whether or not in the context of mandatory re-registration, the use of coaching conversations is an emerging trend.¹⁹⁻²¹ Coaching conversations include reviewing assessment and feedback data and engaging

physicians in a conversation about those which are most meaningful to them and which they see as providing the most compelling opportunity for improving their performance. Attending to the physician's own assessment of his or her performance and personal learning needs and interests is essential to enhance the person's sense of autonomy. The coaching conversation then moves to guiding the physician to set specific and realistic improvement goals and to developing a plan for attaining them. Planning involves paying special attention to the factors which may both impede and enable progress and success, setting timelines, and considering how success will be measured or noticed.²² An important function of the coach is to aid physicians in identifying resources that they can draw upon or approaches they might use; in other words, to help them think differently about how they might accomplish their goal. Helping physicians realize that just "working harder" or "doing more" is often not the best approach to making sustainable practice change, is an important facilitation message.²⁰ In the academic medical center in which the performance appraisal system with coaching session was implemented (chapter 3), some physicians indicated their need for follow-up coaching sessions. It is recommended that hospital management and regulatory authorities ensure that the performance appraisal and coaching systems accommodate physicians' need for follow-up sessions.

Combining one-on-one and group-based coaching

In this thesis, two coaching formats have been investigated, a one-on-one conversation and a peer group conversation guided by a coach (chapter 3 and 4). Both formats can contribute to physicians' individual professional development, albeit in a different way (Table 6.1). When conducting coaching conversations, it should always be kept in mind that physicians' individual performance cannot be separated from the practice context. Practice is necessarily contextualized, and does not exist outside of a particular setting, within a cohort of professionals, a community of activity or a set of social and professional engagements. Therefore, it is recommended to use a meaningful combination of one-on-one and group-based coaching conversations. In group-based conversations, discussing the working climate within the department can be helpful in addressing individual differences, roles and culture and strengthen the physician group cohesion. In particular, group-based conversations may be appropriate after the arrival of new colleagues. For young physicians, a group-based conversation may be helpful in becoming better aligned with the peer group and strengthening in interpersonal relationships.

In contrast to discussing physicians' *individual* performance (whether in a one-on-one or group-based conversation), also physicians' *group* performance must be discussed on a regular basis. Such a conversation necessarily requires a group format and can take place among the members of the physician group or with partners

within the healthcare organization, such as nurses, managers or peers of adjacent medical specialties. Discussing the results of the Group Monitor instrument (chapter 5) may serve as input for these conversations and may assist physician groups in identifying rater class-specific areas for improvement. With a view on developing performance assessment *programs*, a connection must be established between performance assessment initiatives on the individual level and on the group level. Individual development plans can be included within the objectives of the physician group and can contribute to a more thorough 'department profile'. In this way, implementing the goals set in development plans becomes a shared responsibility of the individual physician *and* the physician group/department.

Allocating resources for performance appraisal and coaching

To support physicians' professional development, healthcare institutions need to invest in two types of resources. First, resources are needed to implement systems that support easy access to performance assessment and feedback data. In times of bureaucratic overload of physicians, performance assessment systems should not add too much administrative burden to physicians and the healthcare organization in general. Second, resources are needed for training and supervision of coaches. Selection criteria for coaches are qualities such as good communication and empathic skills. In relation to the guiding of effective group-based conversations, it should be recognized that simultaneously navigating group dynamics and tailoring the conversation to the needs of individual physicians is challenging. Crucial attributes of coaches are confidence in working with groups and creating a non-threatening environment, flexibility, approachability and a sense of humor. To build coaching expertise, it is important that coaches meet each other on a regular basis to exchange practice experiences and discuss areas for improvement. Within healthcare institutions, a combination of peer coaches, non-peer coaches and external coaches may be the ideal pool to work with. Collaboration among all relevant stakeholders involved (physicians, professional and organizational leadership within healthcare institutions and hospital administration) is necessary to design and implement performance appraisal systems that enable physicians to comply with external accountability requirements while serving their needs for professional development. In order to make a long-term statement about the actual contributions of coaching to physicians' professional development, it is imperative that regulatory bodies carefully monitor physicians' experiences with one-on-one and group-based coaching conversations.

DIRECTIONS FOR FUTURE RESEARCH

Following the results of the studies in this thesis, several directions for future research can be provided. Although the requirement for CPD is clear, it is important to ensure that CPD activities are meaningful and that they contribute to improving physicians' performance and the quality of care. The goal of CPD is to support physicians in their work across various settings, disciplines and professional challenges. In considering this goals, Allen et al. point to the need to expand beyond traditional outcomes of CPD activities (what physicians are able to know, do or perform) by including the recording of less tangible impacts on professional relationships in the areas of networking, collaboration, career development and organizational change.²³ Communities of practice may offer an important theory to conceptualize or explain the impact of CPD in further studies.²³ In order to investigate less tangible impacts of CPD on professional relationships, it seems effective to opt for (longitudinal) studies with combinations of various research methods (such as action research, process research, case study, mixed methods and qualitative methods).

The studies in this thesis have shed light on physicians' experiences with formal performance appraisal and coaching in the context of professional re-registration. Across healthcare institutions, however, there may be considerable variation among coaching practices. Future studies need to address this variation in terms of the value of coaching as part of formal performance appraisal for physicians' professional development and quality of care. In particular, the role of patient feedback should be considered: what does the implementation of formal performance appraisal systems mean for their trust in their physician and in the medical profession in general? How could the involvement and input of patients in formal performance appraisal enhance this sense of trust? Furthermore, research is needed to gain insight into the range of topics that are discussed during coaching conversations in order to adjust coaching methodologies. These topics may include work engagement and vitality, occupational well-being, career development and sustainable employability, and inter-professional teamwork and collaboration.

Finally, an important direction for future research is examining the contribution of MSF *in combination with* coaching conversations to physicians' long term behavior change. Over the years, multiple studies have reported performance improvement after receiving MSF. For example, Boerebach et al.²⁴ found that over time supervisors received higher teaching performance scores from their residents, and Van der Meulen et al.²⁵ likewise reported an increase in clinical performance scores in subsequent years. However, the specific role of coaching in achieving these improvements was not investigated in these studies. In the Netherlands, Overeem et al.²⁶ were the first ones to evaluate the combination of MSF and guided reflection in the context of a voluntary IFMS procedure. In 2012, they examined physicians'

self-reported changes in the six months after they had received MSF and meeting with a coach. A small majority of physicians indicated that they had improved one or more aspects of their performance. In addition to self-reported changes, future studies also need to look into behavior change as perceived by peers and coworkers. In doing so, attention should also be paid to changes in the working climate. Furthermore, it is important to consider the contribution of coaching conversations to the functioning of healthcare organizations, by identifying hospital-wide problems and areas for improvement.

A FINAL REFLECTION

As a final reflection of the research presented in this thesis, I believe that reflection has the potential to support meaningful professional development that, in turn, may contribute to the improvement of patient care. However, it needs to be considered that these intended outcomes will only be achieved if there is careful attention to the process of reflection. The main focus areas need to relate to challenging physicians in a supportive environment and recognizing that effective reflection requires personal and/or group motivation for reflection, development of personal meta-cognitive “thinking about thinking” skills (including noticing, processing and action planning), the opportunity to meet with a trained facilitator and a wider organizational system that recognizes and supports these focus areas.

REFERENCES

1. Davis DA, Barnes BE, Fox RD, eds. *The Continuing Professional Development of Physicians: from Research to Practice*. 1st ed. Chicago, IL: American Medical Association; 2003.
2. Davis D, Bordage G, Moores LK, et al. The science of continuing medical education: terms, tools, and goals: effectiveness of continuing medical education: American college of chest physicians evidence-based educational guidelines. *Chest*. 2009;135(suppl):8S-16S.
3. American Board of Medical Specialties. Standards for the ABMS Program for Maintenance of Certification (MOC): For implementation in January 2015.2014.www.abms.org/%20media/1109/%20standards-for-the-abms-program-for-moc-final.pdf. Accessed 16 April 2020.
4. Federation of Medical Regulatory Authorities of Canada Revalidation Working Group. Physician revalidation: maintaining competence and performance. Ottawa (ON): The Federation;2007.www.fmrac.ca/committees/documents/final_reval_position_eng.pdf. Accessed 16 April 2020.
5. General Medical Council. Ready for revalidation; the Good medical practice framework for appraisal and revalidation. Published March 2013, Manchester, United Kingdom.
6. College Geneeskundig Specialisten. KNMG. Besluit herregistratie specialisten. 2015. www.demedischspecialist.nl/sites/default/files/Besluit%20herregistratie%2013%20juli%202016.pdf. Accessed 16 April 2020.
7. Mann K, van der Vleuten C, Eva K, et al. Tensions in informed self-assessment: How the desire for feedback and reticence to collect and use it can conflict. *Acad Med*. 2011;86:1120-1127.
8. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol*. 2000;55:68-78.
9. Ten Cate TJ, Kusurkar RA, Williams GC. How self-determination theory can assist our understanding of the teaching and learning processes in medical education. *AMEE guide no. 59. Med Teach*. 2011;33:961-973.
10. Ajjawi R, Regehr G. When I say... feedback. *Med Educ*. 2019;53:652-654.
11. Ramani S, Könings KD, Mann KV, Pisarski EE, van der Vleuten CPM. About politeness, face and feedback: Exploring resident and faculty perceptions of how institutional feedback culture influences feedback practices. *Acad Med*. 2018;93:1348-1358.
12. Bernabeo EC, Holtman MC, Ginsburg S, Rosenbaum JR, Holmboe ES. Lost in transition: The experience and impact of frequent changes in the inpatient learning environment. *Acad Med*. 2011;86:591-598.
13. Lingard L. Paradoxical truths and persistent myths: Reframing the team competence conversation. *J Contin Educ Health Prof*. 2016;36:S19-S21.
14. Boreham N. A theory of collective competence: Challenging the neo-liberal individualization of performance at work. *Br J Educ*. 2004;52(1):5-17.
15. Langlois S. Collective competence: Moving from individual to collaborative expertise. *Perspect Med Educ*. 2020;9:71-73.
16. Olson CA. Twenty predictions for the future of CPD: implications of the shift from the update model to improving clinical practice. *J Contin Educ Health Prof*. 2012;32:151-152.
17. Reeves S. An overview of continuing interprofessional education. *J Contin Educ Health Prof*. 2009;29:142-146.

18. Thompson LL. Making the team: A guide for managers. 4th ed. Upper Saddle River, NJ: Prentice Hall, 2011.
19. Lovell B. What do we know about coaching in medical education? A literature review. *Med Educ*. 2018;52:376-390.
20. Sargeant J, Bruce D, Campbell CM. Practicing physicians' needs for assessment and feedback as part of professional development. *J Contin Educ Health Prof*. 2013;33:S54-62.
21. Sargeant J. Future research in feedback: How to use feedback and coaching conversations in a way that supports development of the individual as a self-directed learner and resilient professional. *Acad Med*. 2019;94:S9-S10.
22. Reed VA, Schifferdecker KE, Turco MG. Motivating learning and assessing outcomes in continuing medical education using a personal learning plan. *J Contin Educ Health Prof*. 2012;32:287-294.
23. Allen LM, Palermo C, Armstrong E, Hay M. Categorising the broad impacts of continuing professional development: a scoping review. *Med Educ*. 2019;53:1087-1099.
24. Boerebach BC, Arah OA, Heineman MJ, Busch OR, Lombarts KM. The impact of resident and self-evaluations on surgeon's subsequent teaching performance. *World J Surg*. 2014;38(11):2761-2769.
25. Van der Meulen M, Heeneman S, Oude-Egbrink M, Van der Vleuten C, Arah O, Lombarts K. When feedback backfires: influences of negative discrepancies between physicians' self and assessors' scores on their subsequent multisource feedback ratings. 2020. In preparation.
26. Overeem K, Wollersheim HC, Arah OA, Crujlsberg JK, Grol RP, Lombarts KM. Factors predicting doctors' reporting of performance change in response to multisource feedback. *BMC Med Educ*. 2012;12(1):52.

Appendices

SUMMARY IN ENGLISH
SUMMARY IN DUTCH
VALORIZATION
DISSEMINATION
ACKNOWLEDGEMENTS
ABOUT THE AUTHOR
SHE DISSERTATION SERIES

SUMMARY IN ENGLISH

'Reflection' has become established as a key principle underpinning maintenance of standards within professional education and practice. Reflection is considered to be the driving force of lifelong learning and performance improvement. As described in **chapter 1**, the ongoing transition from continuing medical education (CME) to continuing professional development (CPD) signifies that physicians' learning is no longer limited to the updating of medical knowledge, but that physicians also regularly reflect on the broad spectrum of competencies that are necessary for optimal performance in modern healthcare. Currently, there is great emphasis on ensuring that physicians can demonstrate that they are engaging in the process of reflection on professional practice; CPD activities are increasingly made mandatory for professional re-registration. Physicians are required to use assessment and feedback data from their workplace to guide their learning and practice improvement. This approach aims to raise the quality of physicians' performance to a higher level, in line with the adage "doing well, getting better". However, adopting an instrumental approach to reflection may result in the creation of rituals more than any meaningful insight. When reflection is experienced in an overly prescriptive manner and subject to formal evaluation, the very essence and purpose of reflection may be compromised. Therefore, in this thesis, the following research question was addressed:

How can reflection on physicians' professional performance be best facilitated to serve their developmental needs?

In the four studies included in this thesis, the concept of reflection was explored by taking the perceptions of physicians themselves as a starting point. It was investigated how, in the context of mandatory re-registration, physicians' professional development could be stimulated and performance appraisal conversations could be rendered meaningful, for example by investing in in-depth coaching and drafting realistic personal development plans or by involving peers in these conversations. Finally, the focus shifted from individual physicians in the context of their peer group to the physician group itself. A new multisource feedback tool for evaluating the performance of physician groups was developed and tested.

In **chapter 2**, we turned to the context of daily practice to shed light on physicians' developmental needs and to find clues for meaningfully facilitating reflection to support physicians' professional development. We investigated how physicians themselves describe their relation to reflection in professional practice. In the perception of physicians, reflection does not only play a role in maintaining medical expertise, but also in maintaining relationships, both with others (patients, peers, collaborative partners in the healthcare organization) and with oneself.

Reflection was understood as an interplay between an internal dialogue (a private conversation with oneself) and an external dialogue (a conversation with others). Physicians highlighted the fuzziness of the reflective process, comparing it to a process of 'rising above the fog' in order to determine strategy and take informed action. The fuzziness was closely linked to the complexity of practice and the power dynamics inherent to working with others in a complex healthcare organization. A psychologically safe environment appeared as a prerequisite for facilitating reflection; occasionally speaking to someone outside of one's professional 'bubble' was deemed helpful.

In **chapter 3**, we report on the design and implementation of a performance appraisal system in an academic medical center. Making reflection and professional development mandatory for professional re-registration presents itself as a dilemma, since mandatory accountability and transparency requirements seem principally incompatible with physicians' needs for psychological safety and intrinsic motivation. The performance appraisal system was designed in such a way that the mandatory nature of participating in the system did not detract from satisfying physicians' developmental needs. The heart of the system consisted of a two-hour peer-to-peer coaching session, in which coaches guided physicians' reflection upon their multisource feedback results and assisted them in drafting a personal development plan for which the physician was intrinsically motivated. Important factors for rendering the coaching session meaningful included (1) an appreciative and development-oriented coaching process, (2) a skilled and trustworthy coach, and (3) separation of the formative coaching session from the hierarchical annual review with the physician's superior/head of department. The successful implementation and positive reception of the system was due to a collaborative approach to designing the system and the involvement of physicians, organizational leadership and regulators.

In **chapter 4**, we investigated an alternative format for a one-on-one coaching session, i.e. a peer group session. During a peer group session, led by a professional facilitator, physicians shared their reflection on their multisource feedback results with the peers of their specialty-specific group with whom they worked together on a daily basis. This group format is increasingly used as an addition or alternative to one-on-one coaching sessions, as it establishes a clear link between physicians and their social context. Overall, sharing reflections was experienced as a source of social support and a means for deepening collegial relations. Physicians perceived the group session as helpful in realizing actual change in practice by creating a sense of urgency for improvement. Important factors for rendering the group session meaningful included (1) an experienced facilitator capable of navigating group dynamics, (2) longitudinal trusting relationships, (3) a limited group size, and (4) absence of profound organizational disturbances or excessive personal struggles.

In **chapter 5**, we shifted the focus from individual physicians in the context of their peer group to the physician group itself in the context of the wider healthcare organization. In this chapter, we report on the development and validation of the Group Monitor instrument, a multisource feedback tool to provide feedback on physicians' group performance. Physician group received input from four rater classes: (1) members of the physician group itself (self-assessment), (2) support staff, (3) hospital managers, and (4) peers from other medical specialties. Four subscales were distinguished: Medical practice, Organizational involvement, Professionalism, and Coordination. The rater classes "physicians (self-assessment)" and "support staff" had an additional subscale, Communication. An important finding was that rater classes perceived items differently; the fuzziness in these perceptions appeared to be attributed to contextual factors, such as the medical specialty of the physician group and the hospital in which the physician group worked.

In **chapter 6**, the main findings of the individual studies were summarized and considered in relation to physicians' needs for competence, autonomy and relatedness and literature regarding physicians' individual and collective competence.

Implications for practice include the following:

- Healthcare institutions need to allocate sufficient resources to implement systems that support easy access to performance assessment feedback data, while adding not too much administrative burden to physicians and the healthcare organization in general. Resources should also be allocated to train and supervise coaches.
- Hospital management and regulatory authorities need to ensure that performance appraisal and coaching systems accommodate physicians' need for follow-up sessions. Physicians' experiences with one-on-one and groups-based coaching conversations should be carefully monitored.
- In addition to reflection on individual performance, attention must be paid to reflection on the performance of physician groups. In order to develop a performance assessment *program*, a link must be established between performance assessment initiatives at the individual and group level.

Directions for future research include the following:

- Future studies should look at the contributions of one-on-one and group-based coaching conversations to physicians' long-term behavior change, as perceived by physicians themselves (self-report), but also by colleagues and patients.
- Variation among coaching practices across healthcare institutions deserves attention, as does the added value of patient feedback to the performance appraisal process.

- In order to adjust coaching methodologies, the range of topics that are discussed during coaching sessions needs to be explored. These topics may include work engagement, vitality and occupational well-being, career development and sustainable employability, and inter-professional teamwork and collaboration.

SUMMARY IN DUTCH

'Reflecteren op het eigen functioneren' is een belangrijk uitgangspunt geworden voor het handhaven van normen binnen het beroepsonderwijs en de klinische praktijk. Reflectie wordt beschouwd als de drijvende kracht achter levenslang leren en verbeteren. Zoals beschreven in **hoofdstuk 1** betekent de geleidelijke shift van *continuing medical education* (CME) naar *continuing professional development* (CPD) dat het leren van artsen niet langer beperkt blijft tot het up-to-date houden van medisch inhoudelijke kennis. Artsen dienen ook regelmatig te reflecteren op het brede spectrum van competenties die nodig zijn voor optimaal functioneren in de moderne gezondheidszorg. Momenteel ligt er grote druk op artsen om te laten zien dat zij 'lerend' zijn en reflecteren op hun functioneren; om deze reden worden CPD-activiteiten steeds vaker verplicht gesteld voor professionele herregistratie. Artsen worden geacht om beoordelings- en feedbackgegevens van hun werkplek te gebruiken als input voor hun leerproces. Deze aanpak heeft tot doel artsen zo optimaal mogelijk te laten functioneren, in lijn met het adagium "goed zijn en beter worden". Het verplicht stellen van reflecteren en het benaderen van reflecteren op een instrumentele manier kan echter sociaal wenselijk gedrag in de hand werken, in plaats van dat het reflecteren daadwerkelijk bijdraagt aan inzicht in eigen functioneren. Om deze reden staat in dit proefschrift de volgende onderzoeksvraag centraal:

Hoe kan reflectie op het professioneel functioneren van artsen het beste worden gefaciliteerd om tegemoet te komen aan hun leer- en ontwikkelbehoeften?

In de vier studies die in dit proefschrift zijn opgenomen, is het concept van reflectie onderzocht door de percepties van artsen zelf als uitgangspunt te nemen. Onderzocht werd hoe in het kader van verplichte herregistratie de professionele ontwikkeling van artsen gestimuleerd kon worden en functioneringsgesprekken betekenisvol konden worden ingericht, bijvoorbeeld door te investeren in diepgaande coaching en het opstellen van realistische persoonlijke ontwikkelplannen of door collega's te betrekken bij deze gesprekken. Ten slotte verschoof de focus van individuele artsen in de context van hun collega-groep naar de vakgroep zelf. Er is een nieuwe multisource feedbacktool ontwikkeld en getest om het functioneren van vakgroepen te evalueren.

In **hoofdstuk 2** hebben we ons gericht op de context van de dagelijkse praktijk om inzicht te krijgen in de leer- en ontwikkelbehoeften van artsen en om aanwijzingen te vinden voor het zinvol faciliteren van reflectie ter ondersteuning van hun professionele ontwikkeling. We onderzochten hoe artsen zelf hun relatie tot reflectie in de beroepspraktijk ervaren. In de beleving van artsen speelt reflectie niet alleen een rol bij het up-to-date houden van medisch-inhoudelijke kennis,

maar ook bij het onderhouden van relaties, zowel met anderen (patiënten, peers, samenwerkingspartners in de zorgorganisatie) als met zichzelf. Reflectie werd opgevat als een samenspel tussen een interne dialoog (een privégesprek met jezelf) en een externe dialoog (een gesprek met anderen). Artsen karakteriseerden het reflectieve proces als 'fuzzy' en vergeleken het met een proces van 'boven de mist uitstijgen' om strategie te bepalen en om actie te ondernemen op basis van vergaarde informatie. Deze neveligheid was nauw verbonden met de complexiteit van de praktijk en de machtsdynamiek die inherent was aan het samenwerken met anderen in een complexe zorgorganisatie. Een psychologisch veilige omgeving bleek een voorwaarde te zijn voor zinvolle reflectie; af en toe van gedachten wisselen met iemand buiten de eigen professionele 'bubbel' werd nuttig geacht.

In **hoofdstuk 3** hebben we gerapporteerd over het ontwerp en de implementatie van een systeem voor functioneringsgesprekken in een academisch ziekenhuis. Het verplicht stellen van reflectie en professionele ontwikkeling in het kader van herregistratie vormt een dilemma, aangezien verantwoordings- en transparantie-eisen haaks lijken te staan op de intrinsieke motivatie van artsen en hun behoefte aan psychologische veiligheid. Het systeem voor functioneringsgesprekken werd zo ontworpen dat het verplichte karakter van deelname aan het systeem geen afbreuk zou doen aan de ontwikkelbehoeften van artsen. Het hart van het systeem bestond uit een twee uur durend collegiaal gesprek. Zogenaamde peer-coaches begeleidten artsen bij het reflecteren op hun multisource feedbackresultaten en ondersteunden hen bij het opstellen van een persoonlijk ontwikkelplan (POP) waarvoor de arts in kwestie intrinsiek gemotiveerd was. Belangrijke factoren om het gesprek zinvol te maken waren onder meer (1) een waarderende en ontwikkelingsgerichte coachingsmethodiek, (2) een bekwame en betrouwbare coach, en (3) scheiding tussen het formatieve functionerings-gesprek en het hiërarchische jaargesprek met het afdelingshoofd. De succesvolle implementatie en positieve ontvangst van het systeem was te danken aan een gezamenlijke aanpak bij het ontwerpen van het systeem en de betrokkenheid van artsen, afdelingshoofden, ziekenhuisbestuurders en toezichthouders.

In **hoofdstuk 4** hebben we groepsgesprekken onderzocht als alternatief voor de één-op-één collegiale gesprekken. Tijdens een groepsgesprek, (be)geleid door een professionele facilitator, deelden artsen hun persoonlijke reflecties op hun feedbackresultaten met de collega's van hun vakgroep met wie zij dagelijks samenwerkten. Deze gespreksvorm wordt steeds vaker gebruikt als aanvulling op of alternatief voor één-op-één gesprekken, omdat het een duidelijke link legt tussen de individuele arts en zijn/haar sociale context. Over het algemeen werd het delen van reflecties met collega's ervaren als een bron van sociale steun en een middel om collegiale relaties te verdiepen. Artsen achtten het groepsgesprek nuttig voor het op gang brengen van verandering in de praktijk omdat het gesprek een gevoel

van urgentie voor verandering creëerde. Belangrijke factoren om de groepssessie zinvol te maken waren onder meer (1) een ervaren facilitator die in staat was op een goede manier om te gaan met groepsdynamiek, (2) reeds langer bestaande vertrouwensrelaties, (3) een beperkte groepsgrootte en (4) de afwezigheid van ernstige organisatorische ongeregelheden of persoonlijke problemen.

In **hoofdstuk 5** hebben we de focus verlegd van de individuele arts in de context van zijn/haar collega-groep naar de vakgroep zelf in de context van de bredere zorgorganisatie. We ontwikkelden en valideerden een nieuwe multisource feedbacktool om de performance van vakgroepen te evalueren, de Groepsmonitor. De vakgroep ontving input van vier respondentgroepen: (1) leden van de vakgroep zelf (zelfevaluatie), (2) ondersteunend personeel, (3) ziekenhuismanagers en (4) collega's uit aanpalende medische specialismen. Er werden vier subschalen onderscheiden: Medische praktijkvoering, Betrokkenheid bij de organisatie, Professionaliteit en Coördinatie. De respondentgroepen 'vakgroepleden (zelfevaluatie)' en 'ondersteunend personeel' hadden een aanvullende subschaal, Communicatie. Een belangrijke bevinding was dat respondentgroepen de items verschillend interpreteerden; de 'fuzziness' in hun percepties bleek toegeschreven te kunnen worden aan contextfactoren, zoals het medisch specialisme van de vakgroep en het ziekenhuis waarin de vakgroep werkzaam was.

In **hoofdstuk 6** werden de belangrijkste bevindingen van de individuele studies samengevat en in verband gebracht met de behoeften van artsen aan competentie, autonomie en verbondenheid. Ook werden de bevindingen besproken in relatie tot literatuur over de individuele en collectieve competentie van artsen.

Implicaties voor de praktijk zijn onder meer:

- Zorginstellingen dienen voldoende middelen vrij te maken om systemen te implementeren die eenvoudige toegang tot feedbackgegevens mogelijk maken en administratieve rompslomp voor de arts en de organisatie beperken. Daarnaast dienen er middelen te worden vrijgemaakt om coaches op te leiden en te ondersteunen.
- Ziekenhuismanagers, bestuurders en toezichthouders dienen erop toe te zien dat systemen voor functioneringsgesprekken voorzien in de mogelijkheid voor vervolgesprekken. De ervaringen van artsen met collegiale één-op-één gesprekken en groepsgesprekken moeten zorgvuldig worden gemonitord.
- Naast reflectie op het functioneren van individuele artsen dient er aandacht te worden besteed aan reflectie op het functioneren van vakgroepen. Om een solide programma van functioneringsgesprekken te ontwikkelen dient er een link te worden gelegd tussen functioneringsgesprekken op individueel en groepsniveau.

Aanbevelingen voor toekomstig onderzoek zijn onder meer:

- Toekomstige studies dienen zich te richten op de bijdragen van één-op-één en groeps gesprekken aan gedragsverandering van artsen op de lange termijn. Hierbij is het van belang de zelfrapportage van artsen in ogenschouw te nemen, maar zeker ook de rapportage van collega's en patiënten.
- Variatie in coachingspraktijken tussen zorginstellingen verdient aandacht, evenals de toegevoegde waarde van patiëntfeedback.
- Voor een verdere verfijning van coachingsmethodieken dienen de verschillende onderwerpen die tijdens gesprekken de revue passeren nader te worden verkend. Deze onderwerpen kunnen betrekking hebben op bevlogenheid, vitaliteit en welzijn, loopbaanontwikkeling en duurzame inzetbaarheid, en interprofessionele samenwerking.

VALORIZATION

In this addendum we summarize and reflect on how the knowledge created by the research presented in this doctoral thesis can be made suitable for use outside the academic field. We reflect on the social relevance of the research and we indicate how the results of this thesis are relevant for a number of target groups. Furthermore, we discuss activities that are undertaken and products that are developed to bring the results of the research to the attention of a wider audience. Lastly, we focus on the innovation emanating from the research and we discuss a schedule for the implementation of research findings.

Social relevance

Quality improvement has always been one of the core responsibilities of physicians and one of the primary tasks of their professional associations. In the process of forming and formalizing the medical profession, as we know it today, many developments were explicitly initiated to assure and improve the quality and safety of patient care. In the Netherlands, the accreditation program of the medical professional associations (in Dutch: *visitatie*) can be considered the core of professional quality improvement practice: it brings together all the efforts and results of professional development, the implementation of (evidence based) guidelines, the structure, culture and functioning of the specialist group, and the use of complication registration with an interest in patient safety. In addition to the group focused accreditations, in 2008 the evaluation of the individual performance of physicians was introduced in Dutch hospitals. The need for individual performance evaluation fits the time frame of the past two decades: the authority of physicians is no longer taken for granted, public pressure on the medical profession grows stronger and the demands for transparency are increasing. The social trend is that individual physicians must be held accountable for (the results of) their professional actions.

The evaluative system to (further) improve the quality of physicians' individual performance is referred to as the IFMS system (Individual Functioning Medical Specialist).¹ It includes the collection of multisource feedback, discussion of the feedback with a trained coach or facilitator and the development and adjustment of a personal development plan. Now that participation in IFMS has recently been included as a requirement for professional re-registration, and this obligation applies from 2020 onwards,² it is important to carefully monitor the implementation of and physicians' experiences with IFMS. Attention needs to be paid to designing and implementing the system in such a way that physicians are enabled to comply with accountability requirements, while also providing a psychologically safe environment and fostering their intrinsic motivation. The research presented in this

doctoral thesis provides insight into how physicians experience their relationship to reflective activity in professional practice (chapter 2) and how they feel about discussing multisource feedback on their performance with a coach, be it in a one-on-one setting (chapter 3) or in a collegial group setting (chapter 4). In addition to providing insight into physicians' experiences with feedback and coaching, the research in this thesis generates knowledge about effective strategies for the development and implementation of the IFMS system in healthcare institutions, in particular academic medical centers (chapter 3). To contribute to the optimization of the group-based accreditation practices, a new multisource feedback questionnaire for the evaluation of the group performance of specialty-specific physician groups was developed and tested (chapter 5).

Target groups

The results of this thesis are particularly relevant to individual physicians and physician groups/medical departments, and organizational leadership and regulatory bodies. Since participation in IFMS is relatively new for physicians in academic medical centers, it is important for physicians to be informed about the different conversation formats (one-on-one format or group-based format) and about the experiences of physicians with these two formats. The results in chapter 3 and 4 showed that participation in the IFMS procedure promotes awareness of strengths and weaknesses. The peer group may play an important role in providing social support and appreciation, creating urgency for improvement and promoting mutual alignment. IFMS can be an important means to make quality measurable and discussable, to gain insight into mutual differences, identifying roles and culture and address possible impaired functioning at an early stage. The advice is therefore to make use of IFMS on a regular basis. Participation in IFMS and other quality systems, such as accreditation, need to be carefully aligned so that results can reinforce each other. The coordination of the various quality systems promotes continuous attention to both the individual and group performance of physicians.

In addition to the responsibility of the individual physician and physician groups/medical departments, the hospital's Board of Directors and the staff board also have a joint responsibility to continuously improve and safeguard the quality and safety in the healthcare institution. The results in chapter 3 show that healthcare institutions need to allocate resources to implement systems that support easy access to performance assessment and feedback data. In times of bureaucratic overload of physicians, performance assessment systems should not add too much administrative burden to physicians and the healthcare organization in general. Resources should also be allocated to train and supervise coaches; careful coordination with quality experts, program developers and researchers is necessary. Organizational leadership and regulatory bodies need to ensure that performance

appraisal and coaching systems accommodate physicians' need for follow-up sessions. In order to make a long-term statement about the actual contributions of coaching to physicians' professional development, it is imperative that physicians' experiences with performance appraisal and coaching conversations are carefully monitored and that coaching methodologies remain the subject of research.

Activities and products

The results of this thesis have been published in academic journals and have been disseminated to the scientific community by presentations at various national and international conferences. Chapter 2 has been published in *BMC Medical Education* and chapter 4 has been published in the *Journal of Continuing Education in the Health Professions*, potentially reaching a high number of researchers, educational scientists, physicians and quality managers. Sharing the knowledge resulting from this thesis with a broad audience has been done by presenting the work at conferences on medical education in the Netherlands, Switzerland, Spain, Norway, Malaysia and the United Arab Emirates (see Dissemination section).

Of particular note, the (preliminary) results of the studies reported in chapter 3 and 4 have been presented at a symposium hosted by the research group Professional Performance & Compassionate Care (PP&CC) of the Amsterdam University Medical Centers. During this symposium, which took place on 30 November 2018, the development and implementation of the IFMS system in the Amsterdam University Medical Centers was brought to the attention of physicians and policy advisors on quality improvement. Special attention was paid to the development of the DANA (Developmental Appreciative Navigation Approach). The purpose of the DANA is to optimize the use of physicians' unique qualities to initiate a growth process that optimally supports its developmental goal. It can also be used to transform teams and medical departments. The DANA approach has been developed within the medical specialist setting by Annemiek Nooteboom, a psychologist with 20 years of experience in coaching physicians. The approach is based on training techniques from Appreciative Inquiry practices and solution-focused coaching, both belonging to the field of Positive Psychology. DANA is a structured and validated coaching method on the basis of which formative performance appraisal conversations can be conducted. Performance appraisal conversations are conducted by facilitators who have been trained as a Professional Development Coach using the DANA method. Prior to the conversation, physicians are required to collect multisource feedback on their individual professional performance and discuss this with the trained Professional Development Coach. The coach is a fellow physician from a division other than the physician's own. For the collection of multisource feedback, the validated INCEPT (INViting Coworkers to Evaluate Physicians Tool) questionnaire is used.³ This questionnaire has been developed by the aforementioned research

group and is available through the online platform www.professionalperformance-amsterdam.com. Using this platform, physicians can invite colleagues to fill out the INCEPT questionnaire to provide feedback on their professional performance. Responses are summarized in a feedback report, in which the scores and narratives are presented in such a way that it reveals areas for improvement. The purpose of the performance appraisal conversation is to help physicians to get the best out of themselves, for example by making conscious and unconscious competences visible. In addition, development goals are formulated in a solution-oriented manner and linked to the intrinsic motivation of the physician so that transfer to daily practice can be realized. The result of the conversation is a Navigation document that is then used by the head of department to draw up a Personal Development Plan during the annual appraisal. Department heads are also trained in the DANA methodology so that they can monitor development goals in a positive and empowering manner. In addition, the head of department can use the detailed information to optimally match the physician's range of duties with his/her skills and motivation.

Apart from designing and implementing an IFMS system which focuses on the evaluation of physicians' *individual* professional performance, this thesis has also produced an instrument for the evaluation of physicians' *group* performance. Chapter 5 reports on the development of the Group Monitor instrument, a multisource feedback tool for the evaluation of the performance of specialty-specific physician groups. This instrument is available through the platform www.professionalperformance-amsterdam.com. Although the Group Monitor was not explicitly developed for use in accreditation programs of the medical professional associations, its use may well fit this context.

Innovation

In the past two decades, there has been a remarkable rise in research on reflection in medical education. Many of these studies are theoretically oriented and are often prescriptive in the sense that they conceptualize what reflection should look like. However, several authors have identified flaws in the way reflection has been operationalized: medical education has translated the age-old concept into a teachable and measurable construct, putting reflection at risk of becoming a tick-box exercise.⁴ In this thesis, care has been taken to avoid reductionist approaches to reflection that break down reflection and its complexity into discrete components of activity, or steps of a process. This thesis has an innovative character in the sense that it starts from an open exploration of the perspective of physicians themselves and focuses on the *practice* of reflection.

The practical aspect of reflection is closely related to learning from feedback and developing a coaching approach to feedback conversations. As described in the

previous section, a coaching approach (DANA) was developed that was tailored to the needs of practicing physicians. Furthermore, innovative approaches were used to study strategies to facilitate physicians' reflection on their professional performance. Of particular note is the design-based research (DBR) approach in chapter 3, which was used for the design and implementation of the IFMS system in the Amsterdam University Medical Centers. By using this approach, the aim was to define and build a system around characteristics that are crucial for the facilitation of professional development in the context of mandatory re-registration. The DBR approach allowed us to propose, test and amend solutions for the challenging goal of satisfactorily serving the different purposes of accountability and professional development. The resulting system reflects the various compromises that were reached by extensive deliberations. Other (academic) medical centers have copied this protocol or used it as a basis to build their own. The DBR approach has the potential to improve the quality of similar studies in the future.

Lastly, this thesis has made a contribution to the field of questionnaire development by developing and testing a new multisource feedback tool, the Group Monitor (chapter 5). The Group Monitor is a 35-item uniform questionnaire developed for four different rater classes (physicians/self-assessment, supporting staff, peers and managers) to rate the performance of specialty-specific physician groups. The Group Monitor taps into behavioral aspects of group performance, such as communication, collaboration, use of all contributors' expertise and sharing workload. The items are based on a review of an existing instrument of group performance, the Quick Scan.⁵ This existing instrument, however, only considered self-assessment measures, whereas the Group Monitor also takes into account the perspectives of collaborative partners in the healthcare organization. To the best of our knowledge, this instrument is the first multisource feedback instrument for the evaluation of physicians' *group* performance.

Schedule and implementation

Results of the conducted studies will be incorporated into future workshops on quality improvement of medical specialist care given by the PP&CC research group. To help disseminate the research findings, printed copies of the thesis will be sent to members of key stakeholders in Dutch hospital care, including the Dutch regulatory bodies, such as the Federation of Medical Specialists (Federatie van Medisch Specialisten) and the Royal Dutch Medical Association (Koninklijke Nederlandse Maatschappij tot bevordering der Geneeskunst), quality improvement departments, hospital-wide education committees, department heads, program directors and clinical supervisors, as well as international researchers in the same field.

Despite the fact that by 2020 every hospital is obliged to have a well-functioning IFMS system in place, refinement and adaptation of the procedure and operational processes will of course continue. In the Amsterdam University Medical Centers, the operational responsibility of the IFMS system has been transferred to the HR department. The PP&CC research group will remain involved in substantive matters. Currently, this research group is looking for opportunities to collaborate with research groups in other hospitals. By joining forces, variation among coaching practices across hospitals can be addressed and topics that deserve more attention can be explored. These topics may include work engagement, vitality and occupational well-being, career development and sustainable employability and inter-professional teamwork and collaboration.

References

1. Federatie Medisch Specialisten. Leidraad Individueel Functioneren Medisch Specialisten (IFMS). (In Dutch). Utrecht: Orde van Medisch Specialisten; 2014. https://www.demedischspecialist.nl/sites/default/files/Leidraad%20IFMS_definitief.pdf. Published September 2014. Accessed November 27, 2019.
2. College Geneeskundig Specialismen. KNMG. Besluit herregistratie specialisten. 2015. <https://www.demedischspecialist.nl/sites/default/files/Besluit%20herregistratie%2013%20juli%202016.pdf>. Accessed 16 April 2020.
3. Van der Meulen MW, Boerebach BCM, Smirnova A, Heeneman S, oude Egbrink MGA, van der Vleuten CPM, Arah OA, Lombarts MJMH. Validation of the INCEPT: A Multisource Feedback Tool for Capturing Different Perspectives on Physicians' Professional Performance. *J Contin Educ Health Prof.* 2017;37:9-18.
4. Ng SL, Kinsella EA, Friesen F, Hodges B. Reclaiming a theoretical orientation to reflection in medical education research: a critical narrative review. *Med Educ.* 2015;49(5):461-475.
5. Federatie Medisch Specialisten. Quick Scan Manual 'Quick Scan Handleiding'. <https://www.demedischspecialist.nl/onderwerp/kwaliteitsvisitaties>. Accessed 18 Feb 2019.

DISSEMINATION

This thesis

- **Bindels E**, Verberg C, Scherpbier A, Heeneman S, Lombarts K. Reflection revisited: how physicians conceptualize and experience reflection in professional practice - a qualitative study. BMC Med Educ. 2018;18(1):105
- **Bindels E**, Boerebach B, van der Meulen M, Donkers J, van den Goor M, Scherpbier A, Lombarts K, Heeneman S. A New Multisource Feedback Tool for Evaluating the Performance of Specialty-Specific Physician Groups: Validity of the Group Monitor Instrument. JCEHP. 2019;39(3):168-177.
- **Bindels E**, van den Goor M, Scherpbier A, Lombarts K, Heeneman S. Sharing reflections on multisource feedback in a peer group setting: (how) does it stimulate physicians' professional performance and development? Under editorial review.
- **Bindels E**, Boerebach B, Scheepers S, Nooteboom A, Scherpbier A, Heeneman S, Lombarts K. Designing a system for performance appraisal: balancing medical specialists' accountability and professional development. Under editorial review.

Presentations

- Rogano Meeting 2016
Oral presentation: Reflection revisited
- Ottawa Conference 2018
Poster presentation: Reflection revisited: how physicians conceptualize and experience reflection in professional practice - a qualitative study
- AMEE Conference 2018
Poster presentation: Reflection revisited: how physicians conceptualize and experience reflection in professional practice - a qualitative study
- NVMO Congress 2018
Oral presentation: Validatie van de Groepsmonitor, een 360 graden feedback-instrument voor de professional performance van medisch specialistische vakgroepen: verschillende respondentgroepen, verschillende percepties
- EAPH Conference 2019
Oral presentation: Feedback conversations in a peer group setting: Physicians' experiences with guided reflection on their professional performance in the context of CPD
- Rogano Meeting 2019
Oral presentation: The proof of the pudding is in the eating: implementation and evaluation of a CPD coaching system for physicians in an academic hospital
- Ottawa Conference 2020
Oral presentation: Development and implementation of a CPD coaching system for physicians at a university hospital in the Netherlands

Workshops

- Professional Performance: 2010-2020, symposium 26/02/2016
Workshop title: Feedback, en dan?
Bindels, E., van den Goor, M., van der Leeuw, R.M.
- Professional Performance: 2010-2020, pre-conference NVMO Congress 16/11/2016
Workshop title: Van feedback naar performance verbetering
Bindels, E., van der Meulen, M.W.
- Liever leren dan (her)registreren, symposium 30/11/2018
Workshop title: Groepsreflectie en persoonlijke ontwikkeling
Bindels, E., van den Goor, M.

ACKNOWLEDGEMENTS

Een aantal mensen wil ik graag in het bijzonder bedanken voor hun bijdrage aan dit proefschrift.

Promotoren, Kiki, Sylvia en Albert, dank voor jullie begeleiding de afgelopen jaren. Jullie vormden een driemanschap met een bonte mix van expertise, kwaliteiten en perspectieven. Hoewel dit bij mij nogal eens leidde tot momenten van 'lost in translation' zijn we elkaar allengs gelukkig steeds beter gaan verstaan. Ik ben trots dat we het als team tot zo'n goed einde hebben gebracht. Kiki, ik heb bewondering voor je drijvende kracht en toewijding aan het onderzoek naar de professional performance van artsen, je bent een 'woman with a mission'! Dankjewel voor je betrokkenheid, kritische zin en het grote vertrouwen dat je in mij hebt gesteld. Sylvia, ook al was Maastricht niet om de hoek, je voelde toch altijd dichtbij. Je had een nuchtere kijk op mijn studies en onderwierp mijn schrijfsels altijd aan een kritisch taalkundige check. Dankjewel voor je open ontvangst, je zorgvuldigheid en je zin voor structuur. Albert, ik was jouw laatste promovendus en in aanloop naar je pensioen stelde je jezelf wat meer op de achtergrond. Desalniettemin was je altijd een snelle reageerder en had je een duidelijk pragmatische inbreng. Dankjewel voor je aanmoediging en luchtige toets.

Artsen die hebben deelgenomen aan de afzonderlijke studies middels een interview en/of hun data ter beschikking hebben gesteld voor dit onderzoek, veel dank voor jullie bijdrage. Zonder jullie tijd en moeite was dit proefschrift er nooit gekomen.

Mede-auteurs, Christel, Myra, Jeroen, Benjamin, Renée S, Annemiek en Mirja, dank voor jullie meedenken en meeschrijven aan de afzonderlijke studies. In het bijzonder wil ik Christel, Myra en Jeroen noemen. Christel, wat een geluk dat ik je voor je vertrek bij het AMC heb ontmoet. Je hebt me onder je hoede genomen en me wegwijs gemaakt in kwalitatieve onderzoeksmethoden. Tijdens de eerste studie heb je sloten geduld met me gehad en je had altijd een luisterend oor; je voelt voor mij echt als een officieuze extra promotor. Myra, dankjewel voor je positieve energie en je *can do*-mentaliteit. Ik kon met jou op ieder gewenst moment sparren; je bracht me op nieuwe ideeën en je enthousiasme werkte aanstekelijk. Wat fijn dat ik via jou meer voeling met de praktijk kon ontwikkelen en dat de Q3-club mij zo warm heeft verwelkomd. Jeroen, toen ik in mijn maag zat met een ingewikkelde statistische kwestie was je mijn reddende engel. Dankjewel voor de reeks consultaties waarin je samen met mij het probleem hebt uitgedroogd en de uitkomsten daarvan in begrijpelijke taal aan me hebt uitgelegd.

Collega's van de Professional Performance & Compassionate Care onderzoeksgroep, door de jaren heen zijn er veel onderzoekers uitgevlogen en zijn er ook weer nieuwe bijgekomen. Irene, Renée vdL, Renée S, Benjamin, Joost, Lenny, Milou, Alina, Kirsten, Maarten, Iris, Renée V en Rosa, dank voor jullie steun en support, jullie kritische feedback op mijn stukken en alle gezellige momenten op de afdeling en tijdens de Heusden-weken en congressen. Dank ook Onyi voor de gezellige Heusden-conversaties. Guusje, dankjewel voor de leuke en leerzame beginjaren, je hebt me nieuwsgierig gemaakt naar medisch antropologisch onderzoek! Mirja, wij zijn ons PhD-avontuur ongeveer tegelijk gestart en waren in de laatste fase kamergenootjes. We konden samen heel serieus werken, maar er was gelukkig ook gekkigheid. Tijdens congressen kwam ik erachter dat jij niet alleen een superfijne collega was maar ook een heel welkome partner-in-crime! Dankjewel voor je opgewektheid, je scherpe analyses, je gestructureerde en nauwkeurige werkwijze, je gezelligheid en onze wandelingetjes rondom het oh-zo inspirerende AMC-bastion.

De niet-onderzoekers van het Center for Evidence Based Education, waar de onderzoeksgroep aanvankelijk was gehuisvest, wil ik ook graag noemen, in het bijzonder Manon, Floor en Coby: dank voor jullie interesse, support, de praatjes bij de koffie-automaat en de wijntjes op het NVMO-congres. Dank ook aan de onderzoekers van de afdeling Medische Psychologie; nadat de onderzoeksgroep naar deze afdeling was overgestapt hebben jullie ons met open armen ontvangen. Naomi en Maartje, het was leuk om jullie te treffen in de pauzes of bij een borreltje; samen met Maarten, Iris en Rosa waren jullie een gezellig clubje daar op die kamer aan het einde van de gang!

Maastrichtse collega's van de School of Health Professions Education (SHE), dank dat jullie zo'n warm collega-nest op afstand waren. Ondanks dat ik niet zo heel vaak in het Zonnige Zuiden was, voelde ik me wel altijd heel erg welkom. Maastricht was voor mij een plek om inspiratie op te doen en (buitenlandse) PhD'ers te ontmoeten in een sfeer die uitnodigt tot uitwisseling en constructieve feedback. Carolin, Sanne, Stephanie en Tiuri, bedankt voor jullie leuke gezelschap tijdens borrels en congressen. Dank ook aan het secretariaat, in het bijzonder Lilian: je was een lichtend baken voor eigenlijk alle (administratieve) aangelegenheden in Maastricht, dankjewel voor alle hulp en ondersteuning!

Leden van de Journal Club, dank jullie wel voor de inspirerende feedback- en kill-your-darlings-sessies! Bedankt voor alle reflecties en inzichten, maar ook de gezelligheid op borrels en congressen.

Team van Q3, in het bijzonder Guido, Anja, Gabrielle en Rob: heel veel dank voor jullie open ontvangst en jullie oprechte nieuwsgierigheid naar mijn observaties van de groepsgesprekken. Ik heb ervan genoten jullie 'in actie' te zien! Zonder de mogelijkheid om de groepsgesprekken daadwerkelijk bij te wonen en met jullie na te bespreken zou ik veel contextinformatie hebben gemist en zou de studie in hoofdstuk 4 lang niet zo goed zijn geworden.

Marit, wat fijn dat ik er in de eindsprint van de PhD zo'n coole AMC-collega van de HR-afdeling bij kreeg. En wat een geluk dat ik erachter kwam dat jij een goed feestje niet schuwt en de schone kunst verstaat van het 'dingen over de schutting gooien'!

Prof. dr. Bert Felling, u bent vanaf het begin van mijn PhD-traject op de achtergrond betrokken geweest; dank voor uw interesse, aanmoediging en adviezen.

Lieve vrienden, in het bijzonder Eline, Maeike, Kim, Angelica, Aukje en Manouk: ik ben blij met jullie! Dank jullie wel voor jullie luisterend oor en opbeurende woorden, jullie gezelligheid en onze gezamenlijke biertjes, wijntjes en dansjes. Proost!

Tot slot wil ik mijn thuisfront bedanken. Lieve Ruben en Simone, jullie zijn het liefste broertje en de liefste zus die ik me kan voorstellen, dank jullie wel dat jullie als paranimfen naast mij staan. Dankjewel ook Anatol en Sofietje, ook al ben je nog zo ieniemienie... Lieve pappa en mamma, dank voor jullie onvoorwaardelijke liefde. Ik kon altijd bij jullie schuilen tijdens dit bewogen PhD-avontuur; jullie hebben me op meer manieren gesteund dan ik kan zeggen. I love jullie dat 't verrekt!

ABOUT THE AUTHOR

Elisa Bindels was born in Boxmeer on January 11, 1986. She attended the Gymnasium at Scholengemeenschap in Cuijk, graduating cum laude in 2004. Elisa started studying Ancient History at the VU University in Amsterdam and later switched to Art History, of which she completed the bachelor's program cum laude in 2010. During her studies, Elisa followed internships at the Dutch Art History institutes in Rome and Florence and spent a semester in Bologna as part of the European Erasmus program. Also, she was editor for some time at the magazine *Kunstlicht*, a magazine for visual arts, culture and architecture, affiliated with the VU University.

Driven by a wide range of interests, Elisa then decided to change course and study Psychology at the University of Amsterdam. She was selected to participate in the interdisciplinary Honours-program and completed the bachelor's program cum laude in 2013. Elisa stayed at the university of Amsterdam to continue the master program in Psychology, specializing in Clinical Neuropsychology. She completed the program with merit in 2014. During the master program, Elisa followed an extensive internship at the Onze Lieve Vrouwe Gasthuis (OLVG) in Amsterdam, at the department of Medical Psychology and Psychiatry. Supervised by psychologists, psychiatrists and medical residents, she got acquainted with the setting of clinical teaching practice. During her studies, Elisa worked as a clinical interviewer for a mental health care institution. After graduation from the master program, she deepened her knowledge and skills by following a post-graduate course in cognitive behavioral therapy. Also, she fulfilled multiple teaching roles at the University of Amsterdam and acted as a mentor for first-year Psychology students.

In 2015, Elisa started her PhD at the School of Health Professions Education (SHE) at Maastricht University in collaboration with the Professional Performance and Compassionate Care research group of the Amsterdam University Medical Center. During her PhD she supervised communication practicals for medical students at the University of Amsterdam and obtained a University Teaching Qualification. As part of one of her PhD studies she fulfilled the role of program coordinator of a large-scale project at the Amsterdam University Medical Center. During this project, a system was developed and implemented to evaluate the individual performance of medical specialists. Currently, Elisa works as a researcher at the VU University in Amsterdam at the department of Clinical Psychology, on a project into autonomy-enhancing interventions for patients with anxiety and depression. In the future, it is her ambition to combine patient care, education and research.

SHE DISSERTATION SERIES

The SHE Dissertation Series publishes dissertations of PhD candidates from the School of Health Professions Education (SHE) who defended their PhD theses at Maastricht University. The most recent ones are listed below, For more information go to: <https://she.mumc.maastrichtuniversity.nl>

Oudkerk Pool, A. (06-11-2020) Competency-based portfolio assessment: unraveling stakeholder perspectives and assessment practices

Stammen, L. (16-10-2020) Pursuing high-value, cost-conscious care: the role of medical education

Van der Meulen, M. (15-10-2020) Assessment of physicians' professional performance using questionnaire-based tools

Matsuyama, Y. (05-10-2020) Contextual attributes fostering self-regulated learning in a teacher-centered culture: learners' professional identity formation is a trigger

Rovers, S. (16-09-2020) Growing knowledge: supporting students' self-regulation in problem-based learning

Guiliani, M. (19-05-2020) A critical review of global curriculum development, content and implementation in oncology

Schreurs, S. (20-03-2020) Selection for medical school: the quest for validity

Schumacher, D. (19-03-2020) Resident Sensitive Quality Measures: defining the future of patient-focused assessment

Sehlbach, C. (21-02-2020) To be continued. Supporting physicians' lifelong learning

Kikukawa, M. (17-12-2019) The situated nature of validity: exploring the cultural dependency of evaluating clinical teachers in Japan

Kelly, M. (10-12-2019) Body of knowledge: an interpretive inquiry into touch in medical education

Klein, D. (06-11-2019) The performance of medical record review as an instrument for measuring and improving patient safety

Bollen, J. (01-11-2019) Organ donation after euthanasia: medical, legal and ethical considerations

Wagner-Menghin, M. (25-09-2019) Self-regulated learning of history-taking: looking for predictive cues

Wilby, K. (02-07-2019) When numbers become words: assessors' processing of performance data within OSCEs

Szulewski, A. (20-06-2019) Through the eyes of the physician: expertise development in resuscitation medicine

McGill, D. (29-05-2019) Supervisor competence as an assessor of medical trainees. Evaluating the validity and quality of supervisor assessments

Van Rossum, T. (28-02-2019) Walking the tightrope of training and clinical service: the implementation of time variable medical training

- Amalba, A.** (20-12-2018) Influences of problem-based learning combined with community-based education and service as an integral part of the undergraduate curriculum on specialty and rural workplace choices
- Melo, B.** (12-12-2018) Simulation design matters: improving obstetrics training outcomes
- Olmos-Vega, F.** (07-12-2018) Workplace learning through interaction: using socio-cultural theory to study residency training
- Chew, K.** (06-12-2018) Evaluation of a metacognitive mnemonic to mitigate cognitive errors
- Sukhera, J.** (29-11-2018) Bias in the mirror: exploring implicit bias in health professions education
- Mogre, V.** (07-11-2018) Nutrition care and its education: medical students' and doctors' perspectives
- Ramani, S.** (31-10-2018) Swinging the pendulum from recipes to relationships: enhancing impact of feedback through transformation of institutional culture
- Winslade, N.** (23-10-2018) Community pharmacists' quality-of-care metrics: a prescription for improvement
- Eppich, W.** (10-10-2018) Learning through talk: the role of discourse in medical education
- Wenrich, M.** (12-09-2018) Guided bedside teaching for early learners: benefits and impact for students and clinical teachers
- Marei, H.** (07-09-2018) Application of virtual patients in undergraduate dental education
- Waterval, D.** (26-04-2018) Copy but not paste: an exploration of cross-border medical curriculum partnerships
- Smirnova, A.** (04-04-2018) Unpacking quality in residency training and healthcare delivery