McMaster Modular Assessment Program (McMAP) Through the Years: Residents' Experience With an Evolving Feedback Culture Over a 3-year Period

Shelly-Anne Li, MSc, PhD (candidate), Jonathan Sherbino, MD, MEd, and Teresa M. Chan, MD, MHPE

ABSTRACT

Background: Assessing resident competency in emergency department settings requires observing a substantial number of work-based skills and tasks. The McMaster Modular Assessment Program (McMAP) is a novel, workplace-based assessment (WBA) system that uses task-specific and global low-stakes assessments of resident performance. We describe the evaluation of a WBA program 3 years after implementation.

Methods: We used a qualitative approach, conducting focus groups with resident physicians in all 5 postgraduate years ($n = 26$) who used McMAP as part of McMaster University's emergency medicine residency program. Responses were triangulated using a follow-up written survey. Data were analyzed using theory-based thematic analysis. An audit trail was reviewed to ensure that all themes were captured.

Results: Findings were organized at the level of the learner (residents), faculty, and system. Residents identified elements of McMAP that were perceived as supporting or inhibiting learning. Residents shared their opinions on the feasibility of completing daily WBAs, perceptions and utilization of rating scales, and the value of structured feedback (written and verbal) from faculty. Residents also commented extensively on the evolving and improving feedback culture that has been created within our system.

Conclusion: The study describes an evolving culture of feedback that promotes the process of informed self-assessment. A programmatic approach to WBAs can foster opportunities for feedback although barriers must still be overcome to fully realize the potential of a continuous WBA system. A professional culture change is required to implement and encourage the routine use of WBAs. Barriers, such as familiarity with assessment system logistics, faculty member discomfort with providing feedback, and empowering residents to ask faculty for direct observations and assessments must be addressed to realize the potential of a programmatic WBA system. Findings may inform future research in identifying key components of successful implementation of a programmatic workplace-based assessment system.

Competency-based medical education (CBME) has driven the development of new methods of assessment of clinical competence.[1,2] The central tenets of the CBME paradigm describe physician

competence as a multifaceted, dynamic, contextual, and ever-changing process.[2–4] CBME emphasizes outcomes, progression of ability, and learner-centeredness. Building on the premise that combining multiple samples from multiple raters over time can provide an appropriate representation of a learner's professional development, a programmatic assessment system allows faculty to make numerous low-stakes assessments that can be aggregated into a high-stakes judgment about global performance.[5–9] Programmatic assessment is a systematic method to collate data across the learner's training from multiple domains of physician competence, allowing educators to create tailored future instruction to assist with remediation or acceleration of learning. An essential element of a programmatic assessment system is a workplace-based assessment (WBA).[8]

A WBA is a direct observation in the authentic clinical environment of a specific element of many interconnected competencies that a professional performs.[10] They can be structured (e.g., the MiniCEX) or unstructured and may utilize both qualitative and quantitative measures to describe performance.[10–13] WBAs hold the promise to capture authentic performance of trainees in the work environment. In CBME, WBAs are typically brief, criterion-based, and low-stakes, prompting an immediate feedback encounter.[14] Given the shift toward CBME, WBAs are frequently used as a method of performance evaluation. These regular, direct observations may enable a culture of ongoing feedback and formative assessment, which, in turn, can potentially foster a culture of professional educational support (i.e., coaching).[14–17] Currently, there is limited literature exploring medical trainees' reactions and perceptions of receiving feedback from WBA in postgraduate medical education beyond recent experiences in the United Kingdom.[18] These British studies were, however, immediately completed after a short period of implementation, rather than studies of a system after a period of continuous quality improvement and cultural change.

The McMaster Modular Assessment Program (McMAP) is a programmatic assessment system that collects and combines data from 58 WBA instruments based on emergency medicine (EM) clinical tasks.[19–22] McMAP has been in operation since 2012. The instruments are divided into three levels (junior, intermediate, senior) and comprehensively mapped to the CanMEDS physician competency framework.[23] Each day, trainees are assessed by faculty members who directly observe them performing a representative task specific to EM, and then they are also rated along a global rating scale (GRS) tailored to one of three levels.[19] A shared mental model among assessors is facilitated via checklists and behavioral anchors included in the rating instruments. At the end of a rotation, 16 task-specific assessments and 16 daily global ratings are aggregated by a program administrator and given to a rotation supervisor to generate a narrative end-of-rotation report using a standardized template.

Recent literature has linked the importance of medical students' receptivity and perceptions around feedback to the assessment culture within their medical schools.[24] Programmatic WBA systems, such as McMAP, require a substantive commitment by trainees to achieve the number of assessments required. In this era of CBME, exploring trainees' experience with direct observations and determining their response to a programmatic WBA system will help to improve and inform subsequent iterations of assessment systems and policies.

The purpose of this study was to explore the resident experience of a programmatic WBA assessment system and to understand the benefits and drawbacks that emerge several years after full implementation. To our knowledge, this is the first study to explore resident perceptions about a CBME programmatic assessment system postimplementation.

METHODS

Study Design

This was a qualitative study that used semistructured, moderated, focus group discussions to explore resident physicians' perceptions about a programmatic assessment system (i.e., McMAP). Inspired by previous program evaluations at the undergraduate medical education level,[25] we used a realist evaluation framework.[26] This evaluation framework asks us to consider what mechanisms are at play within a specific context to generate a particular outcome. For the analysis, we used an interpretive descriptive approach[27] to report on the residents' perceptions about McMAP. Interpretative description was used because it allows researchers to examine a phenomenon by identifying themes and patterns among subjective perspectives, while being mindful of individual variations of the topic under study.[28] The product of interpretive description

aims to have application potential; the findings are intended to inform reasoning and to act as a platform for assessment, planning, and intervention.[27]

We used focus groups to explore potentially disparate views of the residents' perspectives on McMAP and to generate point–counterpoint discussions. Recognizing that group dynamics are powerful in the discovery process, participants were asked to complete an anonymous survey (See Appendix 1) after attending the focus group session to ensure that all participants had a chance to voice their opinions on how McMAP may or may not have influenced their performance as a trainee. The survey comprised 10 open-ended questions that supplemented the focus group questions. Because these questions prompted self-reflection and may be sensitive (i.e., questions pertaining to residents' views of their own performance, initial reactions of receiving their assessment), the survey was distributed to each resident after focus group discussions so that a protected setting permitted more open responses.

The survey responses were analyzed together with the focus groups transcriptions. We actively compared the findings between the survey and the focus groups to take note of any similarities or differences.

Study Setting and Population

Our setting was a single academic EM program, with four teaching hospitals affiliated with McMaster University, Hamilton, Ontario, Canada. All 37 current resident physicians were invited to participate through e-mail. The invitation letter included a summary of the study, duration of participation, and participant expectations. The participant expectations included: voluntary participation, that selective response was allowed, and that discontinuation of participation could happen at any time for any reason. A reminder e-mail was sent to all participants 1 week after the initial invitation. Separate focus group sessions for each postgraduate year (PGY) 1 to 5 were conducted to compare findings across PGYs.

Study Protocol

The interview guide (See Appendix 2) was adapted from Heeneman et al.[25] based on input from content experts (TC, JS), one of which is presently the Canadian Specialty Committee's chair (JS). We pilot tested the interview guide on nonparticipatory (graduated) residents to ensure that questions were posed clearly and appropriately. No changes were made after piloting. From July to August 2015, focus groups (which lasted between 35 and 60 minutes) were conducted. Identities of each resident and names or affiliations mentioned during the interviews were deidentified. This project was reviewed by the institutional review board chair at McMaster University and was granted an exemption.

Data Management and Analysis

Interviews were audio-recorded and then transcribed verbatim by a transcriptionist independent of the project. The non-MD interviewer (SAL) had a background in medical research but was not affiliated with the hospital to reduce respondent bias and conflicts of interest. The interviews were collected and analyzed concurrently using interpretive description[27] and independently reviewed by two investigators (SAL and TC). The investigators held iterative discussions to populate the list of codes. Once sufficiency was reached and no new codes were generated, the complete set of interviews was coded by a single investigator using the finalized list of codes. Codes that had similar concepts were grouped together in categories. Themes emerged from codes and categories. Several strategies were used to maintain study rigor, including: 1) employing intercoder agreement (transcripts were coded separately by two coders and in duplicate, resolving any discrepancies), 2) use of an audit trail (record of key decisions during the study), 3) engaging in reflexivity (investigator reflects on how her background and assumptions may impact research findings), and 4) actively searching for deviant quotes that did not agree with the main themes. To increase the rigor of our analysis, a third investigator (JS) reviewed the transcripts in full, performing a formal audit of our analysis. All suggestions from the third investigator were merged into the final code via a consensus process.

RESULTS

Twenty-six (70.3%) resident physicians participated in the study (Table 1). All residents in the study had completed at least 3 months of residency and were enrolled in McMAP.

The focus group transcripts were read in totality, with few new additional themes found in the third focus group transcript. The additional source triangulation surveys confirmed our sufficiency point, as they yielded no new codes or themes. Table 2 shows the codes and categories organized into four themes.

Table 1
Demographics of Participants

	Focus Group	Survey
PGY		
1	7 (26.9)	2 (13.3)
2	4 (15.4)	3 (20.0)
3	3 (11.5)	1 (6.7)
4	5 (19.2)	5 (33.3)
5	7 (26.9)	2 (13.3)
Unknown	0	1 (6.7)
Gender		
Male	18 (69.2)	11 (73.3)
Female	7 (26.9)	4 (26.7)
Total	26 (100)	15 (100)

Data are reported as *n* (%).

Fifteen (57.7%) of 26 participants completed the post-session anonymous survey (Table 2). All results were guided by the realist evaluation framework, identifying the "[w]hat works, for whom, and under which contexts"[26] in McMAP for resident physicians. Codes for themes have been organized into two subcategories (benefits and drawbacks). Because two of four themes expanded beyond the learner level, these findings were also organized at levels of the faculty and system.

Clinical Tasks

Residents appreciated the design and preselection of clinical tasks as a way to explore skill deficits that were previously unknown to them and to help target specific skills to improve. For example,

... Discharge instructions is something that I thought was good actually because I tended to speak a lot of jargon like from before like early on I didn't "talk" down, so those pieces of feedback were helpful to learn how to explain things carefully and clearly and to write things down. (PGY2)

Some residents felt that the tasks were helpful in promoting informed self-assessment. One resident shared:

I think much of the true value of the tasks winds up being self assessment, but it is hard to know where you actually are at compared to your expected level as we don't tend to work with each other much, so I find it useful to have an idea of how I am actually doing overall. (PGY4)

During a rotation residents had to develop a plan to facilitate the completion of specific WBAs. The list of available clinical tasks also provided a structured and formal way to track their progress in completing all competencies. Residents also spoke about feeling "pressured" to complete tasks to progress in their residency training. There was a tendency to avoid tasks perceived as overly cumbersome or minimally valuable to their professional development. With highly valued tasks, however, residents proactively completed the task. Finally, residents had difficulty recognizing which tasks were suitable for observation in the clinical environment.

At the systems level, the progression from junior to senior tasks pushed residents into learning and practicing new skills.

... [T]he ones that I [liked] are the ones that we didn't really get exposure to as juniors, so things like ED flow and time management as well as quality assurance are things that we will have to get better at and be more comfortable with in our senior years instead of just taking care of one patient at a time and knowing that patient well. (PGY3)

Residents also discussed "what didn't work" at the systems level, which prevented completion of all required tasks. The unpredictable nature of the types of patients and encounters within the emergency department (ED) was considered as a barrier. Residents also shared that the dynamic, quick-paced ED setting made the completion of several tasks nonfeasible. One resident reflected,

... The role modeling health promotion which basically requests that in front of a junior, counsel a patient on some sort of health promotion measure such as quitting smoking and be observed by your staff. So that involves all three of them being in the room, which will never happen. (PGY5)

Specifically, residents stated that specific tasks impacted the workflow of other ED practitioners (e.g., feedback from ED nurses) or that there are competing patient care needs that demand attention from faculty.

Observation and Feedback

Residents felt that feedback encounters allowed for informed self-assessment and identified the encounter as an opportunity to learn firsthand from a variety of faculty with various strengths. When asked: "If at all, how does a WBA system help you understand your

Table 2
Themes Discussed by Residents About Our Programmatic WBA System

Clinical tasks
  Systems level
    Benefits   Allows for defining skills required of EM practice
                Promotes feedback and learning culture in clinical environment
                Progression to senior tasks pushes residents into new skill sets
                Helps faculty in exploring learner deficits for teaching
    Drawbacks   Workflow and incorporating tasks is difficult in the clinical environment
                Criteria/expectations on how to complete a task
                Relation between end-of-rotation report and rotation lead minimize impact of reflection
                Fluency with tasks for faculty and residents
                Feasibility with tasks
  Learner level
    Benefits   Helps improve practice
                Fosters learning experience
                Helps residents explore deficits unknown to self
                Finds focal deficiencies and provides opportunities to address
                Provides structured and formal way to track their progress for all competencies
    Drawbacks   Difficulty in planning for tasks
                Performance anxiety
                Avoiding tasks in areas of weakness
                Avoiding overly cumbersome tasks
                Avoiding tasks that they did not feel were important (i.e., health advocacy tasks)
                Residents feeling "pressured" to complete tasks to progress
                Fluency with tasks—familiarity with tasks allows for better integration in workflow
Observation and feedback
  Learner level
    Benefits   Feedback encounters allow for opportunity to grow
                Feedback creates opportunity to learn from faculty specialized in specific areas
    Drawbacks   Day-to-day variations in resident's desire for feedback
                Individual variations in seeking feedback (laid-back vs. active seekers)
                Junior residents perceive that they do not have power to ask for assessments (but became more empowered over time)
                Context influences variability in feedback even if task is the same
  Faculty level
    Drawbacks   Variable in uptake, usage, skills in teaching and assessing, and engagement
                Competing demands in ED for faculty time
                Response process issues (faculty gaming system, social desirability bias, too focused on the numbers)
                Residents perceive an inter-rater reliability difference
                Faculty unfamiliar with tasks (fluency)
  Systems Level
    Benefits   System empowers residents to ask faculty members for feedback
                System creates opportunity for frequent and daily feedback
                Feedback allows for residents to learn and grow
                Promotes undocumented feedback
                Spillover effect of increasing feedback frequency is building a feedback culture
    Drawbacks   Verbal feedback not always created when form filled
                Different education cultures between sites
                McMAP may inhibit feedback on performance not directly related to task
                Linking assessment data gaps to progression for training/graduation not clear
                Culture shift from UGME to PGME; culture shift between sites
Global assessments
  Rating scales
    Benefits   Multiple assessments creates trend for resident's own perceived progress
                Helps identify weak competencies easily (e.g., through outlier scores)
                Helpful tool for precipitating reflection for improvement
    Drawbacks   Global ratings perceived not as meaningful as feedback
                Perceived variance in GRS (interstaff, context, complexity of patients)
                Skepticism of whether faculty members understand the ratings
                Using a criterion scale does not allow for normative comparison, which has been a traditional experience
  Logistics   Difficulty accessing records/unaware of all capabilities of the software
                Bugs in system linking tasks to modify GRS
                Length of written (typed) feedback may be influenced by faculty proficiency with technology
                Design/layout has influence on scoring

GRS = global rating scale; PGME = postgraduate medical education; UGME = undergraduate medical education.

weaknesses and strengths?" all except for one resident who responded to the survey identified McMAP as a conduit for seeking feedback from faculty during a rotation. A resident shared, "McMAP is useful as a springboard to consider areas of performance I might not traditionally ask for feedback on my own. It is useful in that a time for focused feedback is built into the shift." (PGY4).

Quantity of feedback depended on the day-to-day variations in the residents' desire for feedback,

personal motivation for seeking feedback (laid back vs. active seekers), and PGY (i.e., seniors were more likely to actively seek feedback than juniors). Residents also valued real-time feedback as they completed specific clinical tasks:

> I do think there is value to having as I mentioned before somebody stand at the bedside with you while you give discharge instructions or explain a care plan and give feedback on what went well and what didn't go well. And I feel like if we didn't have this McMAP program in place you would probably never get that. So it kind of facilitates that type of feedback. (PGY2)

At the faculty level, residents described heterogeneity in faculty members' involvement in McMAP, motivation for teaching learners, comfort level and skills in assessment and providing feedback, and familiarity with the learners' tasks. Residents understood that inter-rater reliability of the feedback may be difficult to achieve, based on faculty members' varying skill sets, comfort, motivations, engagement, and the competing administrative and patient care demands.

At the systems level, McMAP empowered residents to ask faculty members for feedback and created the culture for frequent and daily feedback, which allowed residents to learn and grow in their professional expertise. One senior resident commented:

> I feel that the benefits [sic] of McMAP is that it facilitates you asking that question without you having to be explicit about it. So perhaps if you are a little bit more nervous about asking for that, it provides you a conduit by which you can ask it. And it does ensure that you will receive feedback of some quality. (PGY3)

Residents identified that they received more feedback on clinical tasks outside of the McMAP curriculum. For example, the McMAP prompted feedback in areas not previously discussed between the faculty-learner dyad. One junior resident reflected,

> We do get feedback that isn't included in the McMAP. And I think that is one of the good things and the bad things of the McMAP is it causes the staff to give us feedback, but it doesn't record all of the feedback that we get. (PGY1).

Most residents preferred informal, verbal feedback versus structured written feedback. Verbal feedback was deemed more substantive, demonstrative, and timely than written feedback. Residents also perceived that some faculty may be reluctant to document negative feedback in their record. Further, residents experienced anxiety about written negative feedback impacting their end-of-rotation assessment.

At the systems level, McMAP propelled the feedback culture by creating a platform for faculty to conduct direct observation and to provide timely feedback to residents performing the clinical tasks. One senior resident elaborated:

> I don't think that McMAP is a tool for facilitating learning. I think that McMAP is a tool for facilitating feedback. It might help you learn things about yourself but I don't think it is the source of telling you where you need to go and learn more. Or changing how you learn before a shift. It tells you I need to be more assertive in resuscitation or I'm not as good at teaching a junior as I thought. (PGY5, female)

Inherently, direct observation and feedback were inextricably linked in the minds of the learners. So too was the impact of the busy clinical environment on their performance during direct observation tasks. One learner remarked:

> I find that sometimes it is not even an accurate reflection of what I do normally because if I know my staff is looming outside to get back to patients I find for the ones that are standing by the bedside waiting for me I feel rushed and I don't feel like I have my normal flow and I feel like it is not actually an accurate representation of what I would do . . .. (PGY2)

Global Assessments

Overall, residents appreciated the inclusion of the criterion-based GRS because the multiple assessments on performance facilitated the documentation of progress over time. It was considered an efficient tool to identify areas of dyscompetence. For example, residents learned to scan for "outliers" (low numbers on the 7-point scale) and to concentrate on improving weaker abilities. The GRS also promoted an environment for ongoing self-assessment. There was a tendency for residents to use the aggregate GRS scores provided on the

end-of-rotation assessment report to benchmark their progress against peers in the same PGY and against more senior residents. A senior resident reflected:

> I usually go through them (global rating scale), I just like to remind myself from the whole picture to look like what stood out either good or bad and then I just like to see kind of roughly if there anything that I just have to refresh in my head for the next time. (PGY3)

Even though residents used the GRS as a source of feedback, these scores were perceived as less meaningful than verbal feedback. Whereas verbal feedback from faculty members provided a further understanding of the context and elaboration on their performance, GRS scores provided very limited guidance on how to improve. The numbers on the rating scale did not provide the depth of feedback (despite the behavioral anchors) when compared to the verbal feedback also received. Residents were also skeptical about faculty comprehension on rating. Further, they remarked that the ratings varied by faculty member, complexity of the patient and task, context and postgraduate year.

> I have had staff specifically say, oh it's the start of the year I should put you somewhere in the middle right, so there is room to improve. And I was like I guess if that is what you think or you could go by what are the different levels and see where I fit in, not just what you think you should put as a number . . .. (PGY1)

Logistics

There were a number of logistic concerns expressed by the residents. The bulk of these were related to technical aspects of the digital platform that hosted McMAP. Specifically, the electronic interface was the main source of problems. Residents hypothesized that design and layout elements might influence scoring and a faculty member's ability to input data.

DISCUSSION

Using a realist evaluation framework that asks how a mechanism acts within a context to generate an outcome, our study reveals that a system with frequent task-specific and global daily assessments (McMAP) has created a *context* that foster a feedback culture (the

mechanism) that promotes informed self-assessment by our residents (the *outcome*).

McMAP's use of PGY-specific clinical tasks for performance tracking, the inclusion of formal observation and feedback encounters, and structured assessments using GRS appear to foster a culture of feedback, wherein the residents note that they receive more feedback than is captured by the system. This feedback culture seems to be promoting the process of informed self-assessment.[29] Many of the residents gave evidence of seeking feedback to improve their practice. Couched within social psychology theories, this set of processes describes learners' access, integration, and analysis of internal (cognitive) and external inputs from concrete and reliable sources such as comments from faculty to generate an appraisal of their own performance. Self-assessment is considered a multifaceted construct that is composed of many discrete activities including the selection of external data and standards, awareness of one's internal state, and critical reflection of one's own performance.[30] In our analysis, the residents repeatedly expressed that each component of the WBA system (clinical tasks, observation and feedback, formal assessment) cultivated an opportunity to receive feedback, one of the major external sources of self-assessment. The preselected clinical tasks reflected relevant areas of physician competence, in which completing these tasks were reported to improve practice. The observation and feedback encounters were considered as opportunities to receive guidance and suggestions from faculty and were often noted as valuable experiences that strengthened their knowledge in different medical situations.

Finally, the GRS alerted residents to areas of dyscompetence and provided a snapshot of their overall progress. McMAP also served as a launching pad for residents to actively solicit feedback from faculty; residents felt empowered and encouraged to ask for feedback, particularly among junior residents. This finding is consistent with a previous study that observed that learners are more apt in seeking feedback when feedback is included in training on a daily basis[31] and that well-implemented WBA programs facilitate feedback.[24,32]

Some of the drawbacks of using McMAP included the challenges with incorporating clinical tasks in their daily workflow, especially in a busy and dynamic ED. Residents perceived several tasks as cumbersome, nonfeasible, or of little value to their overall learning goals, which can help streamline the present system. Residents also observed heterogeneity in the quantity and

quality of feedback from faculty, as well as the level of engagement in their teaching.

LIMITATIONS

Our study has some key limitations. First, convenience sampling may not be inclusive of all possible viewpoints to ensure thematic sufficiency. Second, the study was intended to explore the initial perceptions and experiences of resident physicians on a WBA system after full implementation using a realist evaluation framework in a major teaching hospital in an urban Canadian city. Due to the context specificity of the ED setting (e.g., uncapped patient loads, short-term and episodic nature of patient cases, the multidisciplinary nature of ED healthcare teams) situated in a large, metropolitan, academic teaching hospital the transferability of the findings to other clinical settings (e.g., non ED environments, nonacademic teaching hospitals) may be inappropriate.

Finally, even though we identified "what works and what doesn't" in the ED context for resident physicians at the levels of learner, faculty, and system, the scope of the interview guide did not allow us to explore the "how's" and "why's" of the benefits and drawbacks. A deeper exploration of the residents' perspectives may generate a fuller understanding of the cognitive processes underlying informed self-assessment.

FUTURE DIRECTIONS

This study serves as a basis for our local program's continuous quality improvement cycle, but it also spurs us to consider how we might use less complicated procedures to replicate similar feedback encounters. As McMAP has recently been adopted by other residency training programs in Canada, it will be interesting to see if similar phenomena occur in these other locations. Does a robust, programmatic, WBA system facilitate the formation of a feedback culture? Involving multiple sites in a staggered case method study may allow us to determine what elements are easily transferrable and what are contextually bound.

CONCLUSIONS

Our study provides an example of how a programmatic approach to workplace-based assessments can enable feedback to serve as a meaningful guide for learners. To our knowledge, this is the first study to explore resident perspectives on a programmatic workplace-based assessment system postimplementation. The study describes an evolving culture of feedback that promotes informed self-assessment. The findings can inform medical educators and administrators about the potential challenges and successes that emerge longitudinally after full implementation, particularly how it influences a culture of feedback. Finally, the findings can guide future work in identifying key components of successful implementation of a programmatic workplace-based assessment system.

References

1. Frank JR, Snell LS, Cate OT, et al. Competency-based medical education: theory to practice. Med Teach 2010;32:638–45.
2. Iobst WF, Sherbino J, Cate OT, et al. Competency-based medical education in postgraduate medical education. Med Teach 2010;32:651–6.
3. Koens F, Mann KV, Custers EJ, Cate OT. Analysing the concept of context in medical education. Med Educ 2005;39:1243–9.
4. Frank JR, Mungroo R, Ahmad Y, Wang M, De Rossi S, Horsley T. Toward a definition of competency-based education in medicine: a systematic review of published definitions. Med Teach 2010;32:631–7.
5. van der Vleuten CP, Schuwirth LW, Driessen EW, et al. A model for programmatic assessment fit for purpose. Med Teach 2012;34:205–14.
6. Schuwirth LW, Van der Vleuten CP. Programmatic assessment: from assessment of learning to assessment for learning. Med Teach 2011;33:478–85.
7. Van Der Vleuten CP, Schuwirth LW, Driessen EW, Govaerts MJ, Heeneman S. Twelve tips for programmatic assessment. Med Teach 2015;37:641–6.
8. Bok HG, Teunissen PW, Favier RP, et al. Programmatic assessment of competency-based workplace learning: when theory meets practice. BMC Med Educ 2013;13:123.
9. Schuwirth LW, van der Vleuten CP. Programmatic assessment and Kane's validity perspective. Med Educ 2012;46:38–48.
10. Kogan JR, Holmboe ES, Hauer KE. Tools for direct observation and assessment of clinical skills of medical trainees: a systematic review. JAMA 2009;302:1316–26.
11. Hauer KE, Holmboe ES, Kogan JR. Twelve tips for implementing tools for direct observation of medical trainees' clinical skills during patient encounters. Med Teach 2011;33:27–33.
12. Govaerts MJ, Van Der Vleuten CP, Schuwirth LW, Muijtjens AM. Broadening perspectives on clinical performance

assessment: rethinking the nature of in-training assessment. Adv Health Sci Educ Theory Pract 2007;12:239–60.

13. Pelgrim EA, Kramer AW, Mokkink HG, van den Elsen L, Grol RP, van der Vleuten CP. In-training assessment using direct observation of single-patient encounters: a literature review. Adv Health Sci Educ Theory Pract 2011;16:131–42.

14. Hamdy H, Guidance A. Workplace based assessment: a guide for implementation. Med Teach 2010;31:59–60.

15. Watling C, Driessen E, van der Vleuten CP, Lingard L. Learning from clinical work: the roles of learning cues and credibility judgements. Med Educ 2012;46:192–200.

16. Watling C, Driessen E, van der Vleuten CP, Lingard L. Learning culture and feedback: an international study of medical athletes and musicians. Med Educ 2014;48:713–23.

17. Watling C, Driessen E, van der Vleuten CP, Vanstone M, Lingard L. Music lessons: revealing medicine's learning culture through a comparison with that of music. Med Educ 2013;47:842–50.

18. Massie J, Ali JM. Workplace-based assessment: a review of user perceptions and strategies to address the identified shortcomings. Adv Health Sci Educ Theory Pract 2015;21:455–73.

19. Chan T, Sherbino J. The McMaster Modular Assessment Program (McMAP). Acad Med 2015;90:900–5.

20. Chan TM, Sherbino J, editors. McMaster Modular Assessment Program: Junior Edition. 1st ed. San Francisco: Academic Life in Emergency Medicine, 2015. doi:10.13140/RG.2.1.3452.8168.

21. Chan TM, Sherbino J, editors. McMaster Modular Assessment Program: Intermediate Edition. 1st ed. San Francisco: Academic Life in Emergency Medicine, 2015.

22. Chan TM, Sherbino J, editors. McMaster Modular Assessment Program: Senior Edition. 1st ed. San Francisco: Academic Life in Emergency Medicine, 2015.

23. Frank JR, Danoff D. The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. Med Teach 2007;29:642–7.

24. Harrison CJ, Könings KD, Dannefer EF, Schuwirth LW, Wass V, van der Vleuten CP. Factors influencing students' receptivity to formative feedback emerging from different assessment cultures. 2016;5:276–84.

25. Heeneman S, Oudkerk Pool A, Schuwirth LW, van der Vleuten CP, Driessen EW. The impact of programmatic assessment on student learning: theory versus practice. Med Educ 2015;49:487–98.

26. Pawson R, Tilley N. Realistic Evaluation. Thousand Oaks, California: SAGE Publications Ltd., 1997.

27. Thorne S, Reimer Kirkham S, O'Flynn-Magee K. The analytic challenge in interpretive description. Int J Qual Methods 2004;3:1–11.

28. Hunt MR. Strengths and challenges in the use of interpretive description: reflections arising from a study of the moral experience of health professionals in humanitarian work. Qual Health Res 2009;19:1284–92.

29. Sargeant J, Armson H, Chesluk B, et al. The processes and dimensions of informed self-assessment: a conceptual model. Acad Med 2010;85:1212–20.

30. Epstein R, Siegel D, Silberman J. Self montiroing in clinical practice: a challenge for medical educators. J Contin Educ Health Prof 2008;28:5–13. Available at: http://on linelibrary.wiley.com/doi/10.1002/chp.149/abstract. Accessed Aug 29, 2013.

31. Delva D, Sargeant J, Miller S, et al. Encouraging residents to seek feedback. Med Teach 2013;35:e1625–31.

32. Saedon H, Salleh S, Balakrishnan A, Imray CH, Saedon M. The role of feedback in improving the effectiveness of workplace based assessments: a systematic review. BMC Med Educ 2012;12:25.

Appendix 1

McMAP post-focus group survey

The written component asks you to reflect on your experience with McMAP. Please provide answers to the questions below. The information you provide will remain confidential and anonymous. Please do not provide your name.

Gender: Male Female

PGY: 1 2 3 4 5

1. How do you feel when you know that you're being assessed on your performance?

2. Do you feel you behave any different from the usual (when you're not assessed)?

3. If at all, how does a programmatic assessment system help you understand your strengths and weaknesses?

4. How do the results of your assessment affect the perceptions of your performance?

5. In your own words, describe how frequent assessment makes you feel? (e.g. Do the results motivate you? Reassure you? Challenge you? Discourage you?) Why?

6. How did you react when you first saw your report of your block assessments? What about your first block exam results?

7. What did you think about the format and the way that the report was structured?

8. Did it help or prevent you from understanding more about your performance?

9. What motivates you towards excellence?

Does McMAP contribute to your growth and development as a physician in any way? Can you explain your answer?

Appendix 2

Interview guide

General questions

1. How long have you been assessed using the McMAP system? What were your initial reactions to it?
2. I'd like you to compare your expectations to when you were first introduced to it and now, after having had some time working with the system. Was it any different from what you expected?
3. I invite you to think about your medical school or off-service rotations when you are not programmatically assessed. How do those evaluations or assessments compare to the McMAP/programmatic assessment?

    a. Prompt: Any differences or similarities?
    b. Prompt: Which one do you prefer? Why?
4. Please take a look at the list of tasks in front of you. Which assessment or observed task was really informative for your learning in the past year?
5. How do you work on/prepare for the daily clinical assessments?
6. What information do you get from the clinical program of assessment we have for your clinical work (e.g. McMAP)? Do you find the information useful? Why or why not?
7. Which part of the McMAP motivates you to learn more or continue to learn? Which part does not motivate you? Explain.

8. What is your perception of longitudinal data (i.e. global rating scale trends) reported to you each month? How do you use these data, if at all? Explain why or why not.

Value of McMAP

9. I want to get your thoughts on the value of this program. How do you determine the value of the assessment information and feedback?
10. What determines whether you use this information and feedback in the end-of-block reflective activities? How do you use this information?
11. How do you value the written feedback and/or coaching of your mentor?
12. Is their written feedback/coaching ever different from their verbal feedback/coaching?

    a. Prompt: Why do you think this is?
13. What is your experience with self-directing your learning for the various assessments?
14. If you were to make a McMAP version 4.0, what do you want to include in there?

    a. Prompt: What will you want to take out or change?
    b. Prompt: Why would you implement that change?
15. Do you have any questions for me? Are there topics that we've yet to cover?