

Resident Perceptions of Assessment and Feedback in Competency-Based Medical Education: A Focus Group Study of One Internal Medicine Residency Program

Leora Branfield Day, MD, Amy Miles, MD, Shiphra Ginsburg, MD, PhD, and Lindsay Melvin, MD, MHPE

Abstract

Purpose

As key participants in the assessment dyad, residents must be engaged with the process. However, residents' experiences with competency-based medical education (CBME), and specifically with entrustable professional activity (EPA)-based assessments, have not been well studied. The authors explored junior residents' perceptions regarding the implementation of EPA assessment and feedback initiatives in an internal medicine program.

Method

From May to November 2018, 5 focus groups were conducted with 28 first-year internal medicine residents from the University of Toronto, exploring their

experiences with facilitators and barriers to EPA-based assessments in the first years of the CBME initiative. Residents were exposed to EPA-based feedback tools from early in residency. Themes were identified using constructivist grounded theory to develop a framework to understand the resident perception of EPA assessment and feedback initiatives.

Results

Residents' discussions reflected a growth mindset orientation, as they valued the idea of meaningful feedback through multiple low-stakes assessments. However, in practice, feedback seeking was onerous. While the quantity of feedback had increased,

the quality had not; some residents felt it had worsened, by reducing it to a form-filling exercise. The assessments were felt to have increased daily workload with consequent disrupted workflow and to have blurred the lines between formative and summative assessment.

Conclusions

Residents embraced the driving principles behind CBME, but their experience suggested that changes are needed for CBME in the study site program to meet its goals. Efforts may be needed to reconcile the tension between assessment and feedback and to effectively embed meaningful feedback into CBME learning environments.

Assessment in competency-based medical education (CBME) is reliant on frequent direct observations of trainee performance, coupled with regular and context-specific feedback. Observations and formative feedback based on performance on entrustable professional activities (EPAs) and milestones contribute to summative decisions regarding progression in training. Previous work has demonstrated that trainees must be engaged in the process of discussing, reviewing, and reflecting upon feedback for it to be seen as authentic and used effectively to drive learning.^{1,2} Resident engagement in assessment and feedback processes is

therefore paramount to the success of this new educational paradigm. However, the experiences of residents in CBME have not been well studied.

Faculty and trainees are receptive to milestone-based assessment, but common concerns by both include the increased time required for frequent assessment and the need for high-quality, contextually specific feedback.³ A recent study of internal medicine (IM) residents showed that the perceived quality of feedback had not improved after implementation of milestone-based assessment in a CBME framework.⁴ To achieve sustained and effective assessment activities in CBME, timely clinical assessments based on direct observation must be coupled with meaningful formative feedback. Currently, little is known about the key barriers and facilitators from the resident perspective to achieving this aim. An understanding of the resident perspective on completing EPA-based assessments is essential to optimize engagement from frontline learners and facilitate authentic assessment activities as medical education moves forward with CBME.^{5,6}

We aimed to address this gap in the literature through conducting semistructured focus groups to explore the experiences of first-year IM residents with EPA-based assessment and feedback initiatives in the clinical setting. Through understanding facilitators and barriers to assessment, programmatic initiatives can be targeted to address these challenges and thus facilitate improved trainee-supervisor encounters.

Method

We used a constructivist grounded theory (CGT) approach to explore learners' perceptions of EPA-based assessment during their residency training. In CGT, researchers' prior experiences in the field and knowledge of the literature are used to inform research questions, shape data exploration, and provide depth of interpretation.⁷⁻⁹ Accordingly, our research team consisted of 2 senior IM residents (A.M. and L.B.D.) and 2 faculty educators (L.M. and S.G.) who are also clinical attendings in IM. All researchers are involved in the development and implementation of CBME in the IM program. Researchers A.M. and L.B.D.

Please see the end of this article for information about the authors.

Correspondence should be addressed to Lindsay Melvin, Department of Medicine, 399 Bathurst St. 8E-425, Toronto, ON, Canada M5T 2S8; email: Lindsay.melvin@uhn.ca; Twitter: @LmelvinMD.

Acad Med. 2020;95:1712-1717.

First published online March 17, 2020

doi: 10.1097/ACM.00000000000003315

Copyright © 2020 by the Association of American Medical Colleges

Supplemental digital content for this article is available at <http://links.lww.com/ACADMED/A844>.

frequently receive EPA-based feedback in their clinical work, and researchers L.M. and S.G. assess trainees using the EPA format. Our perspectives represent both sides of the assessment dyad, which helped to foster reflexive dialogue and approaches to data interpretation. We obtained ethical approval for this study from the University of Toronto Research Ethics Board.

Setting

In July 2017, the University of Toronto IM training program piloted the Royal College of Physicians and Surgeons of Canada's brand of CBME: Competency by Design (CBD). Assessment is designed around stage-specific EPAs: key clinical tasks of a specialty that a resident can be trusted to perform independently once sufficient competence has been demonstrated.¹⁰ Each EPA contains multiple milestones: observable markers of trainee ability needed to perform the clinical task.¹¹ Our assessment forms use numeric scales and narrative comments to document overall entrustment.¹²

In our program, residents are encouraged to initiate EPA-based assessment encounters twice per week. Residents and faculty are told that each assessment encounter is intended to be low stakes, with an emphasis on coaching and feedback to guide learning.¹²

Data collected contribute to discussions regarding residents' performance at the IM program competence committee, where summative decisions are made about whether residents can progress to the next stage of their training. Residents also receive monthly in-training evaluation reports (ITERS). We provided residents and faculty with information about the EPA-based assessment and feedback, the electronic submission platform, and how data would be used once the rollout had begun through local and academic rounds.

Sample

By email, we recruited a convenience sample from the group of IM residents beginning their training at the University of Toronto coinciding with the beginning of the CBD pilot (the cohort entering IM residency in July 2017).⁹ All residents had completed medical school in non-CBME models of medical education. Two of us (L.B.D. and A.M.) conducted 3 separate

1-hour focus groups between May and June 2018, as the 2017 cohort was nearing the completion of their first training year. In November 2018, we conducted 2 additional focus groups to further explore and expand on ideas raised in the first 3 groups; this time, we recruited from the cohort beginning their residency training in July 2018, to maintain the first-year resident perspective and also to explore perceptions at an earlier time point in the academic year. Each cohort contained 70–80 residents, and in total, 28 residents participated in 5 focus groups consisting of 4–7 residents each.

During the focus groups, the interviewer (L.B.D. or A.M.) asked open-ended questions inviting participants to describe the process of initiating and completing EPA-based assessments and to reflect on the impact of these assessments on their workflow, supervisor–trainee relationships, and learning (see Supplemental Digital Appendix 1, at <http://links.lww.com/ACADMED/A844>). As the interviewers were residents themselves, they were poised to understand and elucidate the subtleties of EPA-based assessment and feedback encounters. The use of near peers as facilitators also served to minimize any power differentials. All interviews were audiorecorded, transcribed verbatim, and all data were deidentified before data analysis. We refined the interview guide throughout the iterative analysis process, according to CGT methodology.⁹

Data collection and analysis

We conducted data collection and analysis iteratively using line-by-line coding and constant comparative analysis.^{7–9} All 4 authors independently coded the first 3 transcripts to identify commonly occurring initial codes. We all then met to compare and discuss similarities in their codes and to identify initial themes. After the 2 additional focus groups in 2018, we met again to refine and finalize the coding structure, comparing it with earlier interviews. Data collection ceased at theoretical saturation, when we determined that we had sufficient depth and understanding of the data to develop a framework to describe learners' perceptions and experiences with EPA-based assessment.^{7,13,14} We used NVivo statistical software, version 12.2.0 (QSR International, Doncaster, Victoria, Australia) for data management.

Quotations below are identified with group (G) and participant (P) number.

Results

Trainees described overall positive attitudes toward the prospect of receiving meaningful feedback through multiple low-stakes assessments and coaching. They embraced the potential for CBME to create more opportunities for formative feedback and direct observation:

I want it to be something similar to coaching in sports, or musical instruction where somebody who is an expert at a given skill actually observes ... and can identify ways that [you] can improve doing that particular skill. (G1P1)

Residents indicated that they desired feedback that was “timely, specific, [and] actionable” (G5P6) to guide their learning. However, in practice, feedback seeking was perceived as onerous, anxiety provoking, and led to subtle tension between supervisors and trainees. EPA-based assessments were felt to have reduced feedback to a form-filling exercise, leading to learners feeling disengaged and sensing the same in supervisors. While learners reported that the quantity of feedback had increased, they described the perception that feedback quality had actually suffered. EPA-based assessments were also felt to be at odds with the workflow and culture of IM, as they increased cognitive load and workload, altered the dynamics of trainee–supervisor relationships, and diminished the distinction between formative and summative assessment.

Feedback seeking as onerous

Residents described the new process of initiating feedback encounters as distinctly challenging:

But practically speaking, as we all know, the time restrictions and the fact that we have to initiate [EPA-based assessments], I think is a big deal, a big barrier. (G2P1)

Residents expressed significant social anxiety associated with feedback seeking, reinforcing their hesitancy to initiate these encounters:

It's another thing on our plate. And it's another thing I need to figure out how to socially navigate, like is it appropriate for me to ask one more time? Or should I just drop it?... I'll just drop it because I don't want to ask them again. (G5P4)

Due to competing time constraints and patient care responsibilities, many learners described feeling “guilty” (G4P1) for “annoying” (G1P5), “burdening” (G2P3), or “nagging” (G5P4) their preceptors for assessments:

There’s this total inconvenience factor, especially if you are the only resident on... It’s extremely uncomfortable to be like, “Oh, by the way, on top of all of our dying patients, can you do an EPA [-based assessment] for me?” (G2P3)

Although some supervisors seemed enthusiastic about completing assessments, many others were perceived as being critical or “annoyed by doing it” (G2P1). Residents described that preceptors’ disinterest could be seen explicitly as expressed through negative comments overheard or stated directly to learners, or deduced from behaviors such as repeatedly deferring requests for assessment:

They go super, super quickly through the form and then don’t actually give any feedback, so you kind of get that sense that they don’t want to be doing it. (G2P1)

Consequently, some residents described a strain on their relationship and rapport with clinical supervisors, at least in the short term:

It affects my relationship with my staff ... makes it more transactional.... I don’t feel that I’ve developed the same level of rapport I may have [otherwise]. (G5P6)

Feedback reduced to form filling

Although residents reported that EPA-based assessments at times led to increased volume of feedback, they also described feeling that the quality of feedback had, on the whole, worsened. Trainees perceived that assessors often became preoccupied by “form filling” (G1P1) and “checking boxes” (G4P3) instead of providing meaningful advice and “actual coaching” (G1P3) that otherwise occurred when receiving formative feedback in the absence of the EPA-based assessment. As one resident explained:

Feedback from staff when discussing a case might be very fluid, very constructive, but the moment you move that feedback onto a [EPA-based assessment] form, their entire framework of how they give you feedback is changed. All of a sudden, all of their comments

become more generic, because the questions are very fixed in specifically what they want the staff to talk about. So, I find that instead of talking about actual performances related to a certain case, they just click through. (G1P2)

Trainees described a tendency for evaluators to provide “nonspecific,” “generic” narrative feedback (G1P4) on the EPA-based assessment forms. They expressed frustration with vague comments with insufficient suggestions for improvement:

I think that what has been lacking from my feedback is precision. There’s not necessarily been something that is identifiable or at least clear that needs to be improved upon. And so, without that, there’s no good advice on how to improve, and then there’s also no way for me to measure or reflect on whether or not I’ve improved. (G1P1)

Residents desired more than just an impression of their performance—they wanted actionable, specific feedback leading to a teaching moment. For residents, there was perceived greater learning value from feedback derived from coaching and “organic conversations,” compared with “running through a checklist” with EPA-based assessments (G3P4).

Consequently, some residents circumvented the form-filling process. They advised their supervisors to avoid documentation and to instead use the chosen time for coaching and a discussion:

I say, “You don’t have to type anything. I would prefer on this case; can we just talk?”... And then I do actually get constructive criticism. Like I’ve gotten in the past... If they were typing it out, I don’t know that [the constructive criticism] would have come out. (G5P5)

EPA-based assessments perceived as summative

Learners recognized that EPA-based assessments were intended to be formative and to provide opportunities for specific, contextually bound feedback. However, in practice, it was perceived that these assessments diminished the distinction between formative and summative assessments for supervisors:

I think the EPA [-based assessment]s are meant to be a formative assessment. But [they] are being used as summative assessments ... as a mini overall

evaluation, whereas it should be coaching, which are just 2 different things. It just feels like I’m being assessed, not coached. (G3P4)

EPA-based assessments were perceived as having a summative intent as trainees recognized that all individual data points eventually contributed to summative judgments. This perception was reinforced by aspects of the assessments that mimicked the process of ITER completion, including a tendency for assessors to provide broad, generic feedback and due to delays between patient encounters and form completion:

Ultimately many staff treat them similarly to ITERs and the comments will be, “Great student, great job! No areas to improve. Keep reading around cases.” And that doesn’t actually reach the objective that the EPA [-based assessment]s are meant for. (G5P2)

Furthermore, some supervisors were perceived to be interacting with the form as though an individual assessment might lead to high-stakes summative judgments, seemingly avoiding writing down constructive feedback for fear it would “hold back someone in training” (G4P1). Accordingly, residents were preoccupied with concerns about their performance. They inferred from these observed behaviors that any less-than-flawless performance documented on a single EPA-based assessment could have implications for their permanent record and even their future career prospects:

We’re just constantly under scrutiny and constantly being evaluated, and there’s just always this presence on our mind that whatever we do plays into our evaluation and our career choices down the line. (G3P4)

As a result, participants expressed difficulty engaging with the assessments as truly formative learning opportunities. They described performance anxiety and changed their feedback-seeking strategies to choose straightforward cases or “cases that you think you did well in” to avoid a critical evaluation (G2P1). Residents explained that they were less likely to ask for observation, clarification, and constructive criticism to avoid exposing themselves to vulnerability and risk. Reflecting on his reluctance to ask for observation and feedback, one participant stated, “I’m sort of scared to do it because I don’t want to be labeled as a poor clinical examiner [sic]” (G4P1).

Residents acknowledged that these behaviors diminished the learning value of the assessment process:

An important aspect of education and learning is recognizing, acknowledging limitations.... If you do that you worry, then that that will be reflected negatively in an EPA [-based assessment].... And so you kill the educational process, the learning cycle, by fixating on this evaluation component. (G4P2)

Tension with the culture of IM

EPA-based assessment was felt to be at odds with the culture and workflow of IM. Many residents described experiencing these assessments not as short, succinct feedback episodes but as more imposing activities that did not fit in to the daily routine of IM. Significant effort was required to reorganize a busy day in order to accommodate the assessment process in between patient care responsibilities:

The day is so broken up, and you're already so short for time for actual patient care ... it's very difficult to book this in to a separate time during the day. (G2P1)

EPA-based assessments were perceived to add inefficiency and disrupt workflow, as trainees were required to shift their concentration from patient care to receiving feedback and back again:

We have to partition those tasks away from our clinical duties and take ourselves out of the sort of clinical way of thinking. (G2P2)

Residents described that direct observation, upon which these assessments are intended to be based, also occurred very infrequently, if at all. Residents desired more frequent direct observation but conceded that it, too, was impracticable within the workflow of IM. Comparisons were often drawn to procedural-based or surgical disciplines where a preceptor may often be present to supervise and directly observe trainee performance. In contrast to these specialties, in IM, trainees are expected to function independently:

None of my EPA [-based assessment]s are direct observation, it's just not feasible in [internal] medicine. Either I'm presenting a case overnight that I've seen to the staff or it's something I've done that the staff later on comes in and fills in the form. It's just not feasible, I think, in any subspecialty unless it's a procedure, for direct observation. (G1P5)

Concerns were raised regarding the ability of the new assessment process to adequately capture many of the higher-order skills required of residents in IM. Key competencies such as communication skills or clinical reasoning were felt to be subjective, and their components complex and difficult to operationalize and to distill down to "checkboxes" (G3P1) on a form. For residents, EPA-based assessments were more appropriate for assessing objective performance outcomes such as procedural capabilities. As one participant remarked:

[Unlike with] procedural stuff and physical exams ... many of the higher function tasks, for instance, aspects of communication with patients are a lot trickier to get meaningful feedback on through an EPA [-based assessment] because it's not a standardizable [sic] task.... I don't think you could have a universal rubric. (G4P2)

The tensions between EPA-based assessment and the culture of IM contributed to a lack of trainee buy-in and perceived difficulty in implementing this new assessment paradigm.

Longitudinal relationships contribute to EPA-based assessment completion

Despite the challenges identified related to assessment in the program, there were several situations where learners derived significant learning value. Relationship building was valued, and longitudinal relationships with facilitated rapport and residents' comfort in asking for feedback:

But once you get to know somebody, you're just more comfortable reviewing cases with them, regardless of whether you're being evaluated. I feel like things just get a lot more relaxed once you get to know a staff. (G2P2)

Residents also appreciated the opportunity to demonstrate improvement over time with the same assessor:

Some of the times where it's gone better is where there's been opportunity to reconnect with that same person and they've given you feedback on the stuff you worked on [in the interim]. (G3P2)

Discussion

In CBME, EPA-based assessment serves a dual purpose: as feedback to guide learning and as low-stakes assessments

of competence to support summative evaluations and promotion. However, residents clearly described that frequent assessment in the new CBME framework does not necessarily equate to meaningful feedback. Residents indicated that the process of completing EPA-based assessments had made feedback-seeking onerous due to multiple environmental and social factors. The introduction of EPA-based assessments was perceived as having distorted feedback to a "checkbox" exercise. Trainees' feedback seeking became limited by their perception that EPA-based assessments have a primarily summative, not formative, intent. Cultural barriers and workflow constraints in IM further limited trainee engagement with assessment and feedback processes. Our findings suggest that our implementation of EPA assessment and feedback initiatives may have paradoxically diminished the likelihood of seeking and receiving constructive feedback, a significant unintended consequence given the importance of frequent formative feedback in CBME.^{12,15}

Residents clearly expressed a growth mindset orientation.¹⁶ They were excited by the prospect of receiving frequent coaching and formative feedback. This finding is consistent with studies demonstrating that trainees strongly desire meaningful feedback to identify knowledge gaps and highlight weaknesses, with the goal of developing clinical proficiency.¹⁷⁻¹⁹ Our results expand on work by Angus and colleagues,²⁰ who found that IM residents are receptive to feedback provided in milestone-based assessment. However, we found that resident expectations were not met with reality when EPA-based assessments were implemented in practice. Despite their receptivity to the idea of formative feedback, in practice, there were multiple barriers to resident engagement in the process, chief of which was the burdensome nature of initiating the assessment encounters.

One critical finding of our study was trainees' sense that the quality and utility of feedback did not meet their expectations and, in some cases, seemed worse compared with other formative feedback they received in the clinical setting. Consistent with previous literature, residents voiced their desire for actionable feedback from EPA-based

assessments in the form of written comments and verbal coaching.¹⁸ Our study also echoes past studies reporting that supervisors can seem preoccupied with form completion, reducing feedback to an exercise in “checking boxes” with diminished educational value.^{4,21–23} Formalizing the process of receiving formative “in the moment” feedback with the need to now complete frequent forms may undermine its quality. This may be an unintended and unexpected consequence of assessment in CBME, as high-quality, contextually specific feedback is essential to support trainees’ learning progression and to inform competence committee decisions.^{21,24}

Unexpectedly, we found that some residents actively encouraged faculty to *avoid* form completion, as a work-around to redirect the focus to verbal feedback. Form filling was viewed as disruptive, an “either-or” task in conflict with the delivery of feedback. This is another example of an unintended consequence of the new system, one that undermines the pressing need for assessment data. Now more than ever, the lines between assessment and feedback have blurred, yet a disproportionate focus on assessment does not support learners’ need for coaching and feedback to facilitate growth and development. As Watling and Ginsburg caution,²⁵ reconciling the need for robust data to support competence decisions with learners’ need for meaningful feedback and coaching may be challenging. Finding a balance between increasing frequency of assessment while maintaining, or ideally improving, meaningful feedback must be a key goal of future implementation strategies.

An additional barrier to trainee engagement in this new assessment system is that low-stakes feedback episodes are perceived as high-stakes assessment. Participants’ feedback-seeking behavior was hindered by concern of summative evaluation, affecting EPA selection, similar to findings among surgical trainees.^{26,27} Our findings complement work by Eva and colleagues²⁸ that it is not the intent of assessment that matters but rather the perception of the intent that affects how learners interact with assessment. In addition to practical limitations to the provision of meaningful feedback, the sociocultural dynamics of feedback delivery cannot be overlooked.^{29–31}

A prevailing culture of summative assessment has been thought to be a negative influence on trainees’ use of feedback.³² As Watling describes, while a reliance on summative assessment is necessary in medicine to ensure patient safety, the dominance of this summative learning culture may have unintended consequences for the acceptance of feedback.³³ Perhaps we should question the expectation that initiation of observations and assessments in CBME should be largely driven by trainees.

Finally, practical challenges with the process of engaging in EPA-based assessments detracted from resident experiences. Completing assessments did not fit in well with the culture and workflow of IM, as previously suggested by Hatala and colleagues.³⁴ Trainee engagement was hindered by their perception of these assessments as cognitive disruptions in a clinical environment already fraught with frequent interruption. Although the assessments were supposed to be completed quickly, in reality, they were not, and the process of engaging in them added cognitive load that was “the straw that broke the camel’s back” (G4P1). Further, learners felt that the emphasis on independent work in IM limited opportunities for direct observation and led to frequent indirect assessment of competence. We suspect that other nonprocedural specialties could suffer from similar challenges. We recognize that our findings reflect the early implementation of the CBME system at our training site and that these perceptions may change over time. However, there appears to be a need to evaluate and mitigate any negative impact of this new assessment process on frontline trainees and on patient care activities, while being mindful that a solution will not be a one-size-fits-all approach across specialty programs. Finding strategies to facilitate frontline implementation in each specialty will therefore be essential in successful implementation.

Limitations

There are several limitations to our study. First, these data were collected during the first 2 years of the pilot phase of CBME implementation in our program. This presents a unique opportunity to capture the resident perspective during

initial implementation efforts. However, it will be important for future studies to evaluate the evolution of trainees’ perceptions to the implementation of EPA-based assessment and feedback processes over time. We also focused on postgraduate trainees in IM in one residency program. Our findings have significant specialty-specific implications that may not be transferable to other programs of varying size, orientation, and culture, or to procedurally based disciplines. Future research is needed to develop a broader understanding of the experiences of residents in other settings and disciplines. Our use of first-year residents necessitated that they had not formally experienced the prior assessment program as residents for comparison. Future work should explore contextual factors that inhibit or support a culture of high-quality formative feedback.

Conclusions

For residents in our IM training program, our implemented program of assessment in CBME did not equate to meaningful feedback. To optimize the implementation of EPA-based assessment in clinical practice, efforts will be needed to reconcile the tension between assessment and feedback, and this may require a significant culture shift from a focus on assessment for performance to a focus on assessment for learning.

Acknowledgments: The study team would like to thank those who participated in the study for their time and insights.

Funding/Support: This research was supported by a grant from The HoPingKong Centre of Excellence for Education and Practice at the University of Toronto.

Other disclosures: None reported.

Ethical approval: Ethical approval was obtained from the Research Ethics Board at the University of Toronto, April 13, 2018; protocol no. 35776.

Previous presentations: Earlier versions of this work have been presented at the Canadian Conference on Medical Education on April 15, 2019, in Niagara Falls, Ontario, Canada.

L. Branfield Day is a fourth-year chief medical resident, internal medicine training program, Department of Medicine, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada.

A. Miles is a fourth-year resident, geriatric medicine subspecialty training program, Department of Medicine, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada.

S. Ginsburg is staff physician, Division of Respiriology, Mount Sinai Hospital and University Health Network, professor, Department of Medicine, Faculty of Medicine, University of Toronto, and scientist, Wilson Centre for Research in Education, University of Toronto, Toronto, Ontario, Canada.

L. Melvin is assistant professor, Department of Medicine, Faculty of Medicine, University of Toronto, and staff physician, Division of General Internal Medicine, University Health Network, Toronto, Ontario, Canada.

References

- Govaerts M. Workplace-based assessment and assessment for learning: Threats to validity. *J Grad Med Educ.* 2015;7:265–267.
- Watling CJ, Kenyon CF, Zibrowski EM, et al. Rules of engagement: Residents' perceptions of the in-training evaluation process. *Acad Med.* 2008;83(10 suppl):S97–S100.
- Aagaard E, Kane GC, Conforti L, et al. Early feedback on the use of the internal medicine reporting milestones in assessment of resident performance. *J Grad Med Educ.* 2013;5:433–438.
- Raum SE, Lappe K, Colbert-Getz JM, Milne CK. Milestone implementation's impact on narrative comments and perception of feedback for internal medicine residents: A mixed methods study. *J Gen Intern Med.* 2019;34:929–935.
- Harris P, Bhanji F, Topps M, et al. Evolving concepts of assessment in a competency-based world. *Med Teach.* 2017;39:603–608.
- Tekian A, Watling CJ, Roberts TE, Steinert Y, Norcini J. Qualitative and quantitative feedback in the context of competency-based education. *Med Teach.* 2017;39:1245–1249.
- Charmaz K, Belgrave L, Gubrium JF, Holstein JA, Marvasti AB, McKinney KD. Qualitative interviewing and grounded theory analysis. In: *The SAGE Handbook of Interview Research: The Complexity of the Craft.* Thousand Oaks, CA: Sage; 2012.
- Charmaz K. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis.* London, UK: Sage Publications Ltd; 2006.
- Watling CJ, Lingard L. Grounded theory in medical education research: AMEE guide no. 70. *Med Teach.* 2012;34:850–861.
- ten Cate O. Entrustability of professional activities and competency-based training. *Med Educ.* 2005;39:1176–1177.
- Englander R, Frank JR, Carraccio C, Sherbino J, Ross S, Snell L. Toward a shared language for competency-based medical education. *Med Teach.* 2017;39:582–587.
- Gofton W, Dudek N, Barton G, Bhanji F. *Workplace-Based Assessment Implementation Guide: Formative Tips for Medical Teaching Practice.* 1st ed. Ottawa, ON, Canada: The Royal College of Physicians and Surgeons of Canada; 2017:1–12. <http://www.royalcollege.ca/rcsite/documents/cbd/work-based-assessment-practical-implications-implementation-guide-e.pdf>. Accessed February 29, 2020.
- Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: How many interviews are enough? *Qual Health Res.* 2017;27:591–608.
- Morse JM. The significance of saturation. *Qual Health Res.* 1995;5:147–149.
- Holmboe ES, Edgar L, Hamstra S. *The Milestones Guidebook, Version 2016.* <http://www.acgme.org/Portals/0/MilestonesGuidebook.pdf>. Accessed February 29, 2020.
- Dweck C, Walton GM, Cohen GL. *Academic Tenacity: Mindsets and Skills That Promote Long-Term Learning.* Seattle, WA: Bill & Melinda Gates Foundation; 2014.
- Bing-You RG, Paterson J, Lewne MA. Feedback falling on deaf ears: Residents' receptivity to feedback tempered by sender credibility. *Med Teach.* 1997;19:40–44.
- Duijn CCMA, Welink LS, Mandoki M, Ten Cate OTJ, Kremer WDJ, Bok HGJ. Am I ready for it? Students' perceptions of meaningful feedback on entrustable professional activities. *Perspect Med Educ.* 2017;6:256–264.
- Watling C, Driessen E, van der Vleuten CP, Vanstone M, Lingard L. Beyond individualism: Professional culture and its influence on feedback. *Med Educ.* 2013;47:585–594.
- Angus S, Moriarty J, Nardino RJ, Chmielewski A, Rosenblum MJ. Internal medicine residents' perspectives on receiving feedback in milestone format. *J Grad Med Educ.* 2015;7:220–224.
- Cho SP, Parry D, Wade W. Lessons learnt from a pilot of assessment for learning. *Clin Med (Lond).* 2014;14:577–584.
- Bindal T, Wall D, Goodyear HM. Trainee doctors' views on workplace-based assessments: Are they just a tick box exercise? *Med Teach.* 2011;33:919–927.
- Tomiak A, Braund H, Egan R, et al. Exploring how the new entrustable professional activity assessment tools affect the quality of feedback given to medical oncology residents. *J Cancer Educ.* 2020;35:165–177.
- Holmboe ES, Sherbino J, Long DM, Swing SR, Frank JR. The role of assessment in competency-based medical education. *Med Teach.* 2010;32:676–682.
- Watling CJ, Ginsburg S. Assessment, feedback and the alchemy of learning. *Med Educ.* 2019;53:76–85.
- 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.
- Gaunt A, Patel A, Ruscus V, Royle TJ, Markham DH, Pawlikowska T. 'Playing the game': How do surgical trainees seek feedback using workplace-based assessment? *Med Educ.* 2017;51:953–962.
- Eva KW, Munoz J, Hanson MD, Walsh A, Wakefield J. Which factors, personal or external, most influence students' generation of learning goals? *Acad Med.* 2010;85(10 suppl):S102–S105.
- Ramani S, Könings KD, Mann KV, Pisarski EE, van der Vleuten CPM. About politeness, face, and feedback. *Acad Med.* 2018;93:1348–1358.
- Ramani S, Post SE, Könings K, Mann K, Katz JT, van der Vleuten C. "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.
- Schut S, Driessen E, van Tartwijk J, van der Vleuten C, Heeneman S. Stakes in the eye of the beholder: An international study of learners' perceptions within programmatic assessment. *Med Educ.* 2018;52:654–663.
- Harrison CJ, Könings KD, Schuwirth L, Wass V, van der Vleuten C. Barriers to the uptake and use of feedback in the context of summative assessment. *Adv Heal Sci Educ.* 2015;20:229–245.
- Watling C. The uneasy alliance of assessment and feedback. *Perspect Med Educ.* 2016;5:262–264.
- Hatala R, Ginsburg S, Hauer KE, Gingerich A. Entrustment ratings in internal medicine training: Capturing meaningful supervision decisions or just another rating? *J Gen Intern Med.* 2019;34:740–743.