Pediatric Critical Care Fellow Perception of Learning through Virtual Reality Bronchoscopy

Liron Talmi[1], Sabine Nabecker[2], Dominique Piquette[3,4]*, and Briseida Mema[1,3]*

[1]Department of Critical Care Medicine, Hospital for Sick Children, Toronto, Ontario, Canada; [2]Department of Anesthesiology and Pain Medicine, Sinai Health System, Toronto, Ontario, Canada; [3]Interdepartmental Division of Critical Care Medicine, University of Toronto, Toronto, Ontario, Canada; and [4]Department of Critical Care Medicine, Sunnybrook Health Sciences Center, Toronto, Ontario, Canada

ORCID IDs: 0000-0001-9524-0821 (S.N.); 0000-0001-9622-0304 (B.M.)

ABSTRACT

Background: Virtual reality (VR) simulators have revolutionized training in bronchoscopy, offering unrestricted availability in a low-stakes learning environment and frequent assessments represented by automatic scoring. The VR assessments can be used to monitor and support learners' progression. How trainees perceive these assessments needs to be clarified.

Objective: The objective of this study was to examine what assessments learners select to document and receive feedback on and what influences their decisions.

Methods: We used a sequential explanatory mixed methods strategy. All participants were pediatric critical care medicine trainees requiring competency in bronchoscopy skills. During independent simulation practice, we collected the number of learning-focused practice attempts (scores not recorded), assessment-focused practice (scores recorded and reviewed by the instructor for feedback), and the amount of time each attempt lasted. After simulation training, we conducted interviews to explore learners' perceptions of assessment.

(Received in original form August 4, 2023; accepted in final form November 28, 2023)

*These authors contributed equally as joint senior authors.

Supported by the Department of Critical Care Medicine, Hospital for Sick Children, Toronto, Ontario, Canada.

Author Contributions: L.T.: literature review, data analysis, and writing of manuscript. S.N.: statistical analysis and review of the manuscript. D.P.: literature review, conceptualization of the question, and review of manuscript. B.M.: literature review, the conceptualization of the questions, data collection and review, and manuscript writing.

ATS Scholar Vol 5, Iss 1, pp 174–183, 2024

DOI: 10.34197/ats-scholar.2023-0097IN

Results: There was no significant difference in the number of attempts for each practice type. The average time per learning-focused attempt was almost three times longer than the assessment-focused attempt (mean [standard deviation] 16 ± 1 min vs. 6 ± 3 min, respectively; $P < 0.05$). Learners perceived documentation of their scores as high stakes and only recorded their better scores. Learners felt safer experimenting if their assessments were not recorded.

Conclusion: During independent practice, learners took advantage of automatic assessments generated by the VR simulator to monitor their progression. However, the recording of scores from the simulation program to document learners' trajectory to a set goal was perceived as high stakes, discouraging learners from seeking supervisor feedback.

Keywords:
virtual reality simulator; bronchoscopy training; assessment; feedback; set standard

Fiberoptic bronchoscopes became commercially available in 1968, a year known as the "second revolution," and, since then, the bronchoscope has been used widely in many specialties (1–3). Twenty years ago, another "revolution" in procedural training was described: a training curriculum using a virtual reality (VR) simulator showing that novices have similar scores to experts after having been practicing for approximately 6 hours (4). What followed these "revolutions" were decades of dedication to training in bronchoscopy (5). Simulation has long been established as an effective instructional tool in medicine, especially for procedural teaching (6–9). Reviews dedicated to examining the role of simulation in bronchoscopy training confirm its effectiveness while recognizing the need for more work to understand the best curriculum design features (10–13).

The VR simulator provides immediate feedback on performance with automatic scoring of clinically relevant variables (4, 12). The automatic scores alone or when combined with a rater's judgment about knowledge of the anatomy have demonstrated good validity evidence (14, 15). Both learners and supervisors can use these scores to support and document learners' progression to the desired outcome. Outcome-based curricula are being introduced in education as critical care programs across North America have transitioned to competency-based medical education (CBME), an outcome-based model that deemphasizes time spent training and focuses on learning outcomes and learner-centeredness (16–18). One of the core principles of CBME is the use of multiple assessments that support and document the development and acquisition of competence (19, 20).

The foundational principle of mastery learning is analogous to CBME in that competence is expected and can be achieved by all learners (21). Evidence

Correspondence and requests for reprints should be addressed to Briseida Mema, M.D., M.H.P.E., Department of Critical Care Medicine, Hospital for Sick Children, 555 University Avenue, Toronto, ON, M5G 1X8 Canada. E-mail: briseida.mema@sickkids.ca.

suggests that mastery learning in simulation is superior to nonmastery instruction and improves the skills of advanced critical care providers (22, 23). We implemented a mastery model of simulation training for our pediatric critical care medicine (PCCM) trainees for fiberoptic bronchoscopy using a VR simulator.

This study aimed to explore how trainees perceived their bronchoscopy simulation assessments in an outcome-based training model. More specifically, we wanted to examine what assessments learners chose to document and receive feedback on and what influenced their decisions.

METHODS

Study Design

After institutional research ethics board approval was received, a sequential explanatory mixed methods strategy (Figure 1) was used to address our study objective (22). All participants were PCCM trainees enrolled in a 2-year program in which bronchoscopy skills are a required competency. This study occurred in 2018 while transitioning from a traditional, time-based post-graduate training program to CBME.

Simulation Curriculum

We took advantage of the established mastery model of simulation-based training in bronchoscopy (Figure 1) to explore our questions in a controlled setting. The VR bronchoscopy simulator used was ORSIM (Airway Simulation Limited) (24), selected on the basis of published evidence of its benefit in training (25) and its user-friendly interface, size, and portability, making it accessible to learners. The VR simulator was purchased in 2017 for $30,000 CDN. The automatic scoring included clinically relevant variables such as accuracy (number and identity of segmental bronchi successfully entered), speed (the time taken to complete the study), and dexterity (number of collisions between the bronchoscope tip and the bronchial wall).

The curriculum consisted of *1*) a large-group, instructor-led, 3-hour didactic session that provided an overview of the procedure, anatomy, and key normal

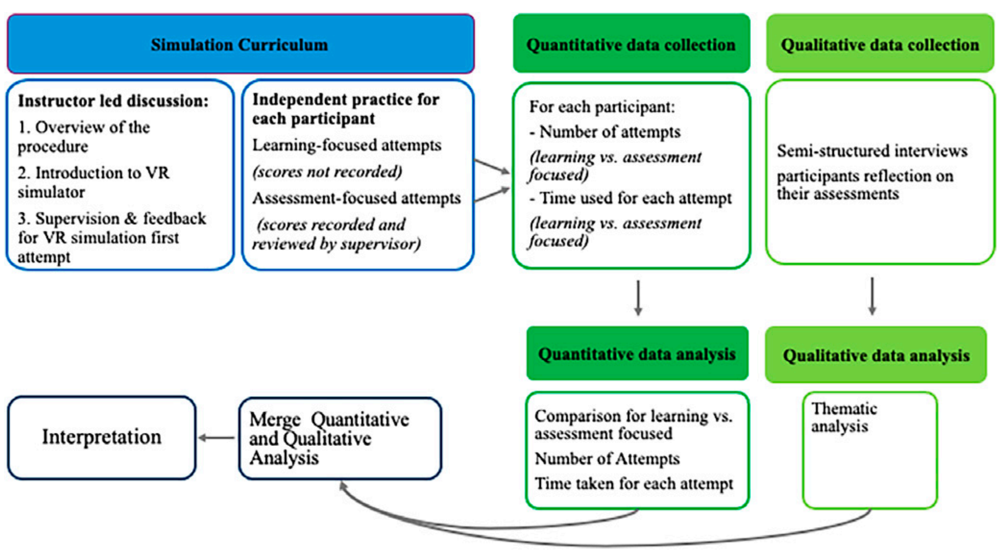

Figure 1. Overview of bronchoscopy simulation curriculum and study design. VR = virtual reality.

and abnormal findings through lectures and demonstrations; *2*) a small-group, 30-minute introduction to the VR simulator equipment; and *3*) a one-to-one supervised session in which the trainee received individualized instructor feedback and guidance during their first VR simulated bronchoscopy attempt, as well as explanations on the automatic scores generated by the simulator. This supervised practice was followed by independent practice on the VR simulator to achieve target scores established by the program. Although a specific simulation score that predicts success in the clinical context has not been established, evidence shows that learners perform better clinically when they have optimized simulation learning (26). Given the above, the target goal was determined by agreement from five experts in bronchoscopy and education. Their decision was based on performance and published average scores of expert physicians using the VR simulator (4). Learners were asked to compete a full bronchoscopy examination in less than 3 minutes and to recognize all 19 segments of the airway tree.

During the independent practice, learners had no requirements or limitations in the number of sessions, with the ability to decide their practice length provided that they reached the target goal, which was a prerequisite for the transition to supervised clinical practice. Although the instructor-led curriculum was completed within 1 day, learners were given 2 weeks and protected time to complete the independent component. The VR simulator was conveniently located in the trainees' office, allowing multiple practice opportunities. During independent practice, learners decided which attempts would count as learning-focused practice (scores not recorded) and assessment-focused practice (scores recorded and reviewed by the instructor for feedback).

Data Collection and Analysis

Quantitative data. During learners' independent practice with the VR simulator, for each practice type (learning focused vs. assessment focused), we collected the number of attempts in each practice type and the time taken for each attempt. Quantitative data were collected and analyzed using Stata version 16.1 (StataCorp). We calculated the average time for both learning-focused and assessment-focused practices and the total time used for simulation for each practice type. We applied *t* tests to compare each of the practice types (learning focused vs. assessment focused) for *1*) the total time taken for practice, *2*) the average time for attempt, and *3*) the number of attempts (Table 1). We used double-sided tests for two groups with similar variability. These comparisons allowed us to examine what assessments are most likely to be recorded by learners.

Qualitative data. The interviews with learners were conducted after the simulation training and collection and analysis of the quantitative data, which informed our semistructured interviews. Interviews targeted the explanation of documentation practices, learners' perceptions of the assessment data they received from the VR simulator, their learning curves, and their attitude toward recording scores (*see* Appendix E1 in the data supplement). The interviews were transcribed verbatim, and the transcripts then analyzed and coded inductively by a faculty member (B.M.) and a trainee (L.T.). Thematic analysis was conducted with open coding, iterative analysis, and constant comparison when coding, searching, reviewing, defining, and

Table 1. Total time, number of attempts, and mean time (in minutes) per attempt for each practice type (learning focused vs. assessment focused)

	Data	*P* Value*
Time		
Total time simulation, min	220 ± 73 (185–254)	
Total time learning-focused attempts, min	161 ± 53 (136–186)	<0.001
Total time assessment-focused attempts, min	59 ± 29 (45–72)	
Number of attempts		
Total number of attempts	21 ± 7 (18–25)	
Number of learning-focused attempts	10 ± 4 (8–12)	0.241
Number of assessment-focused attempts	11 ± 5 (9–14)	
Mean time		
Per attempt, min	11 ± 3 (9–12)	
Per learning-focused attempt, min	16 ± 1 (16–17)	<0.001
Per assessment-focused attempts, min	6 ± 3 (4–7)	

Numbers represent mean ± standard deviation (95% confidence interval).
*The *t* test was applied.

naming themes (27, 28), using dialogue and reexamination of the transcripts to resolve differences in interpretation. In line with a sequential explanatory design, the qualitative results were used to inform and enrich our interpretations of the quantitative data (29).

RESULTS

Quantitative Data

Twenty of 25 eligible learners (11 female, 9 male) consented to be included in the study. All the participants had completed a pediatric residency before enrolling in PCCM training. The participants were true novices because none had prior training or experience using fiberoptic bronchoscopy or VR simulators.

All learners practiced to the curriculum's target goal. There was no significant difference in the number of learning-focused versus assessment-focused practice attempts, mean (standard deviation [SD]), 11 ± 5 versus 10 ± 4, respectively ($P = 0.241$). The total time for learning-focused practice sessions was lengthier than the assessment-focused practice sessions: mean (SD), 161 ± 53 versus 59 ± 29 minutes, respectively ($P < 0.05$). Learning sessions were longer than assessment sessions, not because trainees made more attempts during practice sessions but because each attempt took longer. The average time per each learning-focused attempt was almost three times longer than for assessment-focused attempts: mean (SD), 16 ± 1 versus 6 ± 3 minutes, respectively ($P < 0.05$) (Table 1).

Qualitative Data

Within 1 week of completing their simulation training, 30-minute interviews with all 20 learners were conducted by B.M., who has experience in qualitative methodology. Our analysis explored trainees' perceptions of their scores generated from the VR simulator, documentation of their attempts, and their decision making on learning-focused versus assessment-focused practice. Whether the scores automatically generated by the VR simulator (measuring the time to complete a bronchoscopy) were perceived as learning-focused feedback or performance assessment depended on individual characteristics, contextual factors, and score representation. These perceptions influenced simulation practice behaviors where only the significantly shorter attempts (a better score) were documented as assessment focused and reviewed with the instructor. We describe three main themes below.

Theme 1: Individual Factors Affecting Score Perception

The usefulness of the automatic scores to guide independent practice toward the target goal was underscored by trainees. Some also commented on the dual aspect of scores and how assessment and learning are inextricably linked.

> *You can't possibly separate them: judgment/assessment or feedback. It's not as if you have one set of glasses on which only support learning and the other set of glasses on which only measure performance. (L5)*

Although assessment and learning were understood as inseparable and "two sides of the same coin," the program's recording of their attempts and scores caused apprehension. Documenting the scores transformed the practice in simulation into a high-stakes process. Some trainees were concerned the recording of their performance might be used for comparisons with others and as a measure of their learning speed.

> *The moment that the score is recorded it is "high stakes." You used the expression "for learning." Then why does the program need to record my scores? You can help me learn without it. (L3)*

Theme 2: Contextual Factors that Impacted Learners' Perception of Scores

Many learners alluded to the competitive culture of medicine and how recording their scores promoted a culture that encouraged them to show and prove their performance. Hence, the learners chose to treat assessment-focused practice as high stakes and document their best attempts.

> *The culture in medicine is about showing performance and more so now with having to record your growth curve (as we say in peds) to a set performance standard for so many competencies. (L16)*

Learners were concerned that substandard scores or a lack of score improvement during simulation training would result in fewer clinical opportunities to do a bronchoscopy. These opportunities were already limited by the relatively uncommon occurrence and inherent risks for patients in the intensive care unit.

> *I am reassured here that the assessment data is not used for any decision, but I can't help thinking that it would be if this wasn't a study. Many times, who does the procedure at bedside will be decided by what the supervisors know about their skills. If it's a complex, urgent situation, the best [physician] has to do it. (L11)*

Theme 3: Data Representation

Learners reported that the numerical representation of their performance automatically generated by the simulator was only sometimes informative. However, they tried to make sense of the scores and to understand how to apply the feedback to the clinical context.

> *It would be nice to make it more obvious for other aspects of the procedure to look more like feedback, because here you only get numbers. I don't think I would take away any of the other information. (L2)*

DISCUSSION

We examined learners' attitudes toward frequent assessments in an outcome-based bronchoscopy simulation training. Furthermore, we explored what assessments learners tended to document and review with the instructor while practicing the target goal and why. Given a choice, learners recorded their better assessments (scores), even in a simulation training designed as a low-stakes educational activity.

The relationship between assessment and learning is much more complex than the expression "assessment drives learning" (30, 31). Improvement in technology has revolutionized training in bronchoscopy with the use of VR simulators that provide realistic environments and immediate feedback through automatic scoring of variables that are easy to measure and clinically relevant (12, 14, 15). In the context of an outcome-based curriculum taking advantage of VR simulators, our learners used the assessments (scores) for learning, and all reached the set goal of the simulation curriculum. The findings, however, highlighted some limitations of this outcome-based simulation curriculum: Automatic feedback was not necessarily more informative or clearer than feedback provided by supervisors, learners felt safer experimenting if their assessments (scores) were not recorded, and, last, contextual factors such as the culture of medicine affect simulation-based assessments in the same way as clinically based assessments.

Learners had reservations about scores being recorded unless they were optimal. As soon as data were recorded, the assessment became high stakes and learners developed behaviors to maintain the appearance of continuous progression and competence. When assessments were perceived as suboptimal, trainees resorted to not documenting these data and not asking for external feedback from supervisors. They engaged with what they perceived as suboptimal scores by analyzing them independently. Anxiety with recorded assessments has been reported in the literature (32–34). Promoting a growth mindset has been suggested; however, in medicine, where performance is closely linked to patient outcome and the existing culture is competitive, trainees strive to look competent and demonstrate an optimal learning curve.

Similar to other studies showing that learners choose tasks in which they are likely to succeed or assessors who would rate them favorably, our study found that learners chose to record and receive feedback on their better scores (35, 36), supporting the notion that learners' views on assessments persist even in simulation, which is thought to provide a safe ground where learners can experiment and take risks (37).

One study found that direct observation, combined with assessment, created high-stakes environments where the perceived expectation of residents was to "perform" rather than to "practice" (38). Other authors explored the sociocultural influences of direct observation on medical training and found that learners valued the ability to work independently (39). These authors concluded that a lack of observation could support learning. Future work should explore how learners' behaviors and perceptions change when education programs use scores to make decisions about the group of learners rather than individuals and collect scores anonymously rather than connected to an individual learner.

In summary, the learners used the VR simulators and automatic scores to guide their learning. However, the goal of learning may be swayed by the difference between the intention of having assessment as feedback versus assessment to document performance, creating tensions in learners' minds. Our findings showed that even in a low-stakes environment, such as data in simulation training for a procedural skill, documenting assessments leads to learners focusing on performance rather than seeking expert feedback. There is significant evidence for the value of feedback in education (40, 41) as one of the most critical factors that support learning. The effects of using automatic rather than expert feedback could be further investigated in the context of bronchoscopy simulation with advanced subspeciality learners.

Our findings are important for educators designing simulation programs to train and support learners. It is possible that CBME will change the culture and the perception of assessments, but that is a topic that requires future investigation.

Author disclosures are available with the text of this article at www.atsjournals.org.

REFERENCES

1. Panchabhai TS, Mehta AC. Historical perspectives of bronchoscopy. Connecting the dots. *Ann Am Thorac Soc* 2015;12:631–641.
2. Efrati O, Sadeh-Gornik U, Modan-Moses D, Barak A, Szeinberg A, Vardi A, *et al.* Flexible bronchoscopy and bronchoalveolar lavage in pediatric patients with lung disease. *Pediatr Crit Care Med* 2009;10:80–84.
3. Nussbaum E. Pediatric fiberoptic bronchoscopy: clinical experience with 2,836 bronchoscopies. *Pediatr Crit Care Med* 2002;3:171–176.
4. Colt HG, Crawford SW, Galbraith O III. Virtual reality bronchoscopy simulation: a revolution in procedural training. *Chest* 2001;120:1333–1339.
5. Nilsson PM, Naur TMH, Clementsen PF , Konge L. Simulation in bronchoscopy: current and future perspectives. *Adv Med Educ Pract* 2017;8:755–760.
6. Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, *et al.* Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *JAMA* 2011;306: 978–988.
7. Reznick RK, MacRae H. Teaching surgical skills—changes in the wind. *N Engl J Med* 2006;355: 2664–2669.
8. Henricksen JW, Troy L, Siefkes H. Pediatric critical care medicine fellowship simulation use survey. *Pediatr Crit Care Med* 2020;21:e908–e914.
9. Naik VN, Matsumoto ED, Houston PL, Hamstra SJ, Yeung RY, Mallon JS, *et al.* Fiberoptic orotracheal intubation on anesthetized patients: do manipulation skills learned on a simple model transfer into the operating room? *Anesthesiology* 2001;95:343–348.
10. Chandra DB, Savoldelli GL, Joo HS, Weiss ID, Naik VN. Fiberoptic oral intubation: the effect of model fidelity on training for transfer to patient care. *Anesthesiology* 2008;109:1007–1013.
11. Marsland C, Larsen P, Segal R, Hunter S, Morris J, Mezzavia P, *et al.* Proficient manipulation of fibreoptic bronchoscope to carina by novices on first clinical attempt after specialized bench practice. *Br J Anaesth* 2010;104:375–381.
12. Davoudi M, Colt HG. Bronchoscopy simulation: a brief review. *Adv Health Sci Educ Theory Pract* 2009;14:287–296.

13. Kennedy CC, Maldonado F, Cook DA. Simulation-based bronchoscopy training: systematic review and meta-analysis. *Chest* 2013;144:183–192.

14. Moorthy K, Smith S, Brown T, Bann S, Darzi A. Evaluation of virtual reality bronchoscopy as a learning and assessment tool. *Respiration* 2003;70:195–199.

15. Ilgen JS, Ma IW, Hatala R, Cook DA. A systematic review of validity evidence for checklists versus global rating scales in simulation-based assessment. *Med Educ* 2015;49:161–173.

16. Accreditation Council for Graduate Medical Education. The milestones guidebook [accessed 2023 Aug 4]. Available from: https://www.acgme.org/globalassets/milestonesguidebook.pdf.

17. Frank JR, Snell LS, Cate OT, Holmboe ES, Carraccio C, Swing SR, *et al.* Competency-based medical education: theory to practice. *Med Teach* 2010;32:638–645.

18. Misra S, Iobst WF, Hauer KE, Holmboe ES. The importance of competency-based programmatic assessment in graduate medical education. *J Grad Med Educ* 2021;13(2 Suppl):113–119.

19. Van Melle E, Frank JR, Holmboe ES, Dagnone D, Stockley D, Sherbino J; International Competency-based Medical Education Collaborators. A core components framework for evaluating implementation of competency-based medical education programs. *Acad Med* 2019;94:1002–1009.

20. Konopasek L, Norcini J, Krupat E. Focusing on the formative: building an assessment system aimed on student growth and development. *Acad Med* 2016;91:1492–1497.

21. McGaghie WC. Mastery learning: it is time for medical education to join the 21st century. *Acad Med* 2015;90:1438–1441.

22. Cook DA, Brydges R, Zendejas B, Hamstra SJ, Hatala R. Mastery learning for health professionals using technology-enhanced simulation: a systematic review and meta-analysis. *Acad Med* 2013;88: 1178–1186.

23. Vitale KM, Barsuk JH, Cohen ER, Wayne DB, Hansen RN, Williams LM, *et al.* Simulation-based mastery learning improves critical care skills of advanced practice providers. *ATS Scholar* 2023;4: 48–60.

24. ORSIM. Homepage [accessed 2023 Aug 4]. Available from: https://www.orsim.com/.

25. Baker PA, Weller JM, Baker MJ, Hounsell GL, Scott J, Gardiner PJ, *et al.* Evaluating the ORSIM simulator for assessment of anaesthetists' skills in flexible bronchoscopy: aspects of validity and reliability. *Br J Anaesth* 2016;117(Suppl 1):i87–i91.

26. Issenberg SB, McGaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med Teach* 2005;27:10–28.

27. Creswell JW, Plano Clark VL. Designing and conducting mixed methods research. Thousand Oaks, CA: Sage Publications; 2007.

28. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;3:77–101.

29. Kiger ME, Varpio L. Thematic analysis of qualitative data: AMEE Guide No. 131. *Med Teach* 2020;42:846–854.

30. Scott IM. Beyond 'driving': the relationship between assessment, performance and learning. *Med Educ* 2020;54:54–59.

31. Bursac I, Mema B. Assessment in simulation versus clinical context: a different lens for different moments. *ATS Scholar* 2022;4:12–19.

32. Branfield Day L, Miles A, Ginsburg S, Melvin L. Resident perceptions of assessment and feedback in competency-based medical education: a focus group study of one internal medicine residency program. *Acad Med* 2020;95:1712–1717.

33. Martin L, Sibbald M, Brandt Vegas D, Russell D, Govaerts M. The impact of entrustment assessments on feedback and learning: trainee perspectives. *Med Educ* 2020;54:328–336.

34. 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.

35. Acai A, Li SA, Sherbino J, Chan TM. Attending emergency physicians' perceptions of a programmatic workplace-based assessment system: the McMaster Modular Assessment Program (McMAP). *Teach Learn Med* 2019;31:434–444.

36. Gaunt A, Patel A, Rusius 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.

37. Rudolph JW, Simon R, Dufresne RL, Raemer DB. There's no such thing as "nonjudgmental" debriefing: a theory and method for debriefing with good judgment. *Simul Healthc* 2006;1:49–55.

38. LaDonna KA, Hatala R, Lingard L, Voyer S, Watling C. Staging a performance: learners' perceptions about direct observation during residency. *Med Educ* 2017;51:498–510.

39. Watling C, LaDonna KA, Lingard L, Voyer S, Hatala R. 'Sometimes the work just needs to be done': socio-cultural influences on direct observation in medical training. *Med Educ* 2016;50: 1054–1064.

40. Ende J. Feedback in clinical medical education. *JAMA* 1983;250:777–781.

41. Veloski J, Boex JR, Grasberger MJ, Evans A, Wolfson DB. Systematic review of the literature on assessment, feedback and physicians' clinical performance: BEME Guide No. 7. *Med Teach* 2006; 28:117–128.