Faculty and Resident Perspectives on Using Entrustment Anchors for Workplace-Based Assessment

Nancy Dudek, MD, MEd
Wade Gofton, MD, MEd
Janelle Rekman, MD, MAEd
Allan McDougall, MA, PhD

ABSTRACT

Background Research suggests that workplace-based assessment (WBA) tools using entrustment anchors provide more reliable assessments than those using traditional anchors. There is a lack of evidence describing how and why entrustment anchors work.

Objective The purpose of this study is to better understand the experience of residents and faculty with respect to traditional and entrustment anchors.

Methods We used constructivist grounded theory to guide data collection and analysis (March–December 2017) and semistructured interviews to gather reflections on anchors. Phase 1 involved residents and faculty (n = 12) who had only used assessment tools with traditional anchors. Phase 2 involved participants who had used tools with entrustment anchors (n = 10). Data were analyzed iteratively.

Results Participants expressed that the pragmatic language of entrustment anchors made WBA (1) *concrete* and justifiable; (2) *transparent* as they explicitly link clinical assessment and learning progress; and (3) *align with training outcomes*, enabling better feedback. Participants with no prior experience using entrustment anchors outlined contextual concerns regarding their use. Participants with experience described how they addressed these concerns. Participants expressed that entrustment anchors leave a gap in assessment information because they do not provide normative data.

Conclusions Insights from this analysis contribute to a theoretical framework of benefits and challenges related to the adoption of entrustment anchors. This richer understanding of faculty and resident perspectives of entrustment anchors may assist WBA developers in creating more acceptable tools and inform the necessary faculty development initiatives that must accompany the use of these new WBA tools.

Introduction

The medical education system is transitioning from a time-based, linear curriculum to a competency-based medical education (CBME) curriculum. The Accreditation Council for Graduate Medical Education (ACGME) Milestones and the Royal College of Physicians and Surgeons of Canada Competence by Design are examples of this. The goal of this change is to ensure that trainees progress after demonstrating competence in a given area instead of simply completing a rotation. This represents a significant shift in pedagogical values and leaves educators searching for better ways to assess trainees within revised learning frameworks. Research suggests that workplace-based assessment (WBA) is an optimal method of assessing professional competence.[1] Consequently, there is a meaningful emphasis on this type of assessment within CBME curricula.[2] With the growth of CBME comes the need to develop and deploy WBA tools that accurately reflect the workplace performance of medical students and residents.[3,4]

There are numerous tools available for WBA, including the mini-clinical evaluation exercise (mini-CEX) and direct observation of practical skills.[5] These tools typically have a list of items to be assessed on an anchored or partially anchored rating scale and a space for narrative comments. Traditional anchors for these scales include rating trainees according to the stated expectations of performance (1, rarely meets expectations, to 5, consistently exceeds expectations) or the quality of the performance (1, poor, to 5, excellent).

The reliability of WBAs that use these traditional rating scale anchors is a concern.[6–8] Reliability issues have been attributed to several sources of error, including distributional rater errors such as the leniency/severity error ("dove/hawk") and range restriction error (central tendency), which is a failure to use the whole scale.[9,10] Recently, various WBA assessment tools have been published that use the

DOI: http://dx.doi.org/10.4300/JGME-D-18-01003.1

Editor's Note: The online version of this article contains the interview guides used in the study.

standard of competence, or independent performance, as their "top score" for their rating scale anchors.[11–18] Assessment tools with these anchors have demonstrated increased reliability when compared with traditional anchors.[11–18] These anchors have various names in the literature, including enstrustability, entrustment, and independence anchors. Although the specific wording of the behaviorally descriptive anchors vary, they are all conceptually based around the ordinal assessment of a progression to competence.[19] In this article we will refer to them as entrustment anchors.

The Ottawa Surgical Competency Operating Room Evaluation (O-SCORE) was developed to assess the performance of a trainee on a single surgical procedure.[14] The rating scale developed for this tool uses entrustment anchors. Multiple sources of validity evidence, including high reliability, were demonstrated for this tool.[14,15] Based on these results, the O-SCORE entrustment anchors were used with 2 other WBA tools: the Ottawa Clinic Assessment Tool (OCAT) and the Ontario Bronchoscopy Assessment Tool (OBAT).[20–22] Given the success of these entrustment anchors, national assessment programs are using them for various WBA tools.[23]

It is important to note that, while WBA tools with entrustment anchors have seemed to improve on previous WBA tool anchors, they are not perfect. A dearth of evidence exists for how, when, and why entrustment anchors work or do not work. The objective of this qualitative study was to better understand the experience of residents and faculty with respect to traditional WBA anchors and O-SCORE entrustment anchors. This understanding is key to assisting medical educators in optimizing the use of entrustment anchors in WBA.

What was known and gap
Research suggests that workplace-based assessment tools using entrustment anchors provide more reliable assessments than traditional anchors, but there is a need for evidence that describes how and why entrustment anchors work.

What is new
Semistructured interviews with residents and faculty as to why they think entrustment anchors are better than traditional anchors.

Limitations
Interviews were conducted at a single site and focused on specific anchors used at that site, limiting generalizability.

Bottom line
The pragmatic language used in entrustment anchors makes them more concrete and transparent, and aligns them with training outcomes.

Methods

This study was conducted at a single Canadian institution where some residency training programs have adopted WBA tools using the O-SCORE entrustment anchors (O-SCORE, OCAT, OBAT, and a daily assessment tool in anesthesia that has not been published). These programs continue to use some WBA tools with traditional anchors as well. Other programs only use traditional WBA tools. We invited physicians involved in residency education, as well as residents (via the postgraduate office), to participate in 1-hour semistructured interviews exploring their experiences with WBA tool anchors (March–December 2017). We used constructivist grounded theory because it is well-suited to generate a social theory operating as a relevant explanation for a phenomena based on iterative analysis and generation of themes.[24]

All participants completed a short questionnaire to help establish whether they were from specialties using traditional anchors only (phase 1) or specialties using traditional and entrustment anchors (phase 2). Twelve participants agreed to an interview in phase 1 and 10 agreed in phase 2. The 22 participants were from 12 different specialties (see TABLE 1 for participant demographics).

A research assistant trained in qualitative data collection techniques[25] led participants through the phase 1 and phase 2 semistructured interview scripts (each with minor variations for resident versus faculty participants; provided as online supplemental material). Both phase 1 and phase 2 interviews began with a discussion of traditional anchors used in WBA tools (BOX 1, SCALES A&B), followed by a discussion of entrustment anchors (BOX 1, SCALE C). The traditional and entrustment anchors used for this study (BOX 1) have been used previously at this institution, so that all participants would be familiar with the 2 traditional anchor examples and phase 2 participants would have used WBA tools with these specific entrustment anchors (BOX 1).

The research team analyzed the data using constructivist grounded theory methodology.[26] The study's primary investigator (N.D.) and a qualitative methodologist (A.M.) analyzed data with the support of a trained qualitative research assistant (R.A.). Our analytic process featured 6 steps: (1) complete independent reading of each transcript (N.D., A.M., R.A.); (2) open/line-by-line coding (N.D., A.M., R.A.); (3) thematic coding (N.D., A.M., R.A.); (4) iterative thematic review (N.D., A.M., R.A.); (5) peer debriefing (co-investigators W.G. and J.R.); and (6) producing findings (N.D., A.M., W.G., J.R.).[27,28] Data were analyzed iteratively over the course of 1

TABLE 1
Participant Demographics

Demographic	Phase 1 Residents (n = 6)[a]	Phase 1 Faculty (n = 6)[a]	Phase 2 Residents (n = 5)[b]	Phase 2 Faculty (n = 5)[b]
Postgraduate year level	1–6	N/A	3–4	N/A
Years supervising residents	N/A	2–26	N/A	5–21
Procedure-based specialty	1	0	3	2

[a] No experience with entrustment anchors.
[b] Experience with both traditional and entrustment anchors.
Abbreviation: N/A, not applicable.

year using NVivo 11 qualitative data analysis software.[29] The team met repeatedly to discuss progressive analytic insights throughout the year.

Rigorous grounded theory requires a documented reflective process for data analysts.[28] The 3 data analysts also captured reflective memos throughout the analytic process. One analyst (N.D.) is actively involved with research on assessment anchors. We frequently discussed her interpretive lens as an expert on entrustment-based assessment in medical education versus the other 2 analysts' experiences making entrustment-based decisions outside of a medical education context. We ceased recruitment when participants' reflections sufficiently informed our initial inquiry.

The Ottawa Health Science Network Research Ethics Board approved the study.

BOX 1 Ratings Scales

Scale A (Traditional)

1 – Rarely meets expectations

2 – Inconsistently meets expectations

3 – Meets expectations

4 – Sometimes exceeds expectations

5 – Consistently exceeds expectations

Scale B (Traditional)

1 – Poor

2 – Fair

3 – Good

4 – Very good

5 – Excellent

Scale C (Entrustment)

1 – Requires complete guidance: "I had to do."

2 – Able to perform but requires repeated direction: "I had to talk them through."

3 – Some independence but intermittent prompting required: "I had to direct them from time to time."

4 – Independent for most things but requires assistance for nuances: "I had to be there just in case."

5 – Complete independence: "I did not need to be there."

Results

Pragmatic Language

Participants expressed that entrustment anchors work because they use pragmatic language, which expresses clinically contextual judgments. Because entrustment anchors are grounded in clinical judgments, they are (1) *concrete*, making them more justifiable to trainees; (2) *transparent*, making explicit the link between clinical assessment and learning progress; and (3) *align with training outcomes*, enabling better feedback conversations (TABLE 2).

All participants agreed that the vague language of traditional anchors challenged raters to interpret the true meaning of the scale, which in turn made it difficult to express their assessments accurately. On the other hand, entrustment anchors were described as more *concrete*, because they are anchored in clinical judgments (TABLE 2). Phase 2 participants felt that the *transparency* of entrustment anchors led to WBA tools that were easier to understand as evaluations of actual clinical competence, which makes low scores more acceptable to trainees. These sentiments were echoed by faculty who expressed that they did not fear residents' reactions to receiving a lower score (TABLE 2).

Phase 1 participants who had not been exposed to entrustment anchors felt that these anchors had the potential to provide more useful feedback to learners because they describe their current performance relative to the professional goal of independent practice. Phase 2 residents agreed and expressed that the feedback was in fact more useful because it is *aligned with training outcomes*. Phase 2 faculty also suggested that the conversations they were having with trainees based on these anchors were more productive, given that trainees accepted where they had been rated on the scale (TABLE 2).

Contextual Concerns

Phase 1 and phase 2 participants expressed different opinions on the extent to which entrustment anchors would productively operate across 3 distinct assessment contexts: (1) procedural versus non-procedural;

TABLE 2
Perspectives on Why Entrustment Anchors Work

Feature of Entrustment Anchors	Participant Quotes
Concrete: making them more justifiable to trainees	[Does it mean] expectations according to their level of training, according to a competent psychiatrist? There's just so much room for adjusting the bar that it's hard to know what it really means... (phase 1 faculty) ... The strengths are that it has clear anchors. It's very transparent in terms of what you're seeing. (phase 1 faculty)
Transparent: making explicit the link between clinical assessment and learning progress	... You can get a 1 or a 2 and that doesn't mean that you're not where you're supposed to be ... (phase 2 resident) ... It's okay for me to rank them at the left-hand end of the scale. And that's not bad, right. So that automatically kind of takes the badness out of the lower end of the scale. (phase 2 faculty)
Align with training outcomes: enabling better feedback conversations	... "Okay, well what do I need to do to be independent?" It sort of allows you to, not just like have a discussion about it, but come up with a plan of action and be more objective in the feedback that's being given. (phase 2 resident) ... Any time I've sat down with residents and given them feedback and based it on this, they seem much more receptive to feedback. (phase 2 faculty)

(2) junior versus senior trainees; and (3) direct versus indirect observation. TABLE 3 illustrates these varying perspectives.

A common sentiment among phase 1 participants was that entrustment anchors would work well in procedural contexts but not in other clinical settings. However, participants who had been using entrustment anchors in outpatient clinics reported that these anchors worked well. Faculty described how they operationalized the anchors in a clinic setting (TABLE 3).

Another area of concern raised by the phase 1 participants was that junior residents would feel

TABLE 3
Perspectives on Use of Entrustment Anchors in Different Contexts

Context	Phase 1	Phase 2
1. Procedural versus non-procedural	I think it would be easier to kind of complete for procedural tasks versus some of the more like conversational and patient interaction things. (resident)	... We have not used it at all for procedures, I think it definitely works for procedures, but I think that the same thing applies for a clinical environment. So, you know, I think that a lot of people feel that these can only be applied to technical skills, but you know, if you think about time management in a clinic, which is a non-technical skill, if the clinic is running late and I have to step in and see a patient that's assigned to the resident to keep the clinic on track, that's 'I had to direct them from time to time,' right. I had to step in to do something. So, I think that we can still apply this. (faculty)
2. Junior versus senior trainees	... As a senior I could see how you would want to nearing independence and you would start higher up the scale. But as a junior, when you're at the bottom of the scale, that's probably not as encouraging. (resident)	... If you're PGY-1 you're supposed to get ones and twos. There's nothing wrong with that. But of course we've grown up in a system where ones and twos means that you failed. (resident)
3. Direct versus indirect observation	... You go see a patient. Then they come back and report to us and tell us what they did and we say oh, I would have done this. Let's go back and clarify some of these things. So but to use this it's almost like you have to actually observe the whole interaction because there's a lot more observation that's involved in order to be able to complete this... (faculty)	In clinic, I'm not going to directly observe them take a history but they're going to report back to me. And I know they didn't ask the question if I ask them again, "Did you ask about this, this and this?" And they haven't, then they know and you can see them than go, "Oh, yeah I forgot to ask about that," and "No, I didn't think about that diagnosis so I didn't ask about that line of questioning." So... I don't feel like I need to watch them take a history to know they missed important details. (faculty)

frustrated at consistently receiving lower scores, as they would not be expected to achieve "top scores" until later on in their training. They acknowledged that, while there would be some clinical activities that they would quickly become independent on, many activities would take more time to achieve scores on the higher end of the scale. They thought that residents would feel bad when they received a score on the low end of the scale. However, phase 2 residents expressed that because the lower scores made sense, they were not disheartening: "... It's preposterous that a first-year resident would be able to do all this independently" (phase 2 resident). Given that, they did not feel bad when receiving a score of 1 or 2. Notably, phase 2 participants commented that entrustment anchors represent a substantial change from their prior assessment experience where "low scores" had a negative connotation, but given the clarity afforded by the entrustment anchors, they readily adapted to receiving lower scores (TABLE 3).

The question of whether entrustment anchors require more direct observation was raised by phase 1 participants with the majority indicating that "... to use [entrustment anchors] it's almost like you have to actually observe the whole interaction" (phase 1 faculty). However, phase 2 faculty did not seem to identify any difference with regard to the amount of direct observation required to use either type of anchor (TABLE 3).

Unforeseen Gaps

When comparing entrustment anchors to traditional anchors, participants noted that traditional anchors provide information about how a trainee is progressing relative to their peers while entrustment anchors do not. Many residents in this study said they desire this type of information. Faculty in this study placed far less value on peer comparison and seemed to suggest that we should not be promoting this type of comparison (BOX 2).

In addition to lacking peer comparison anchors, participants also expressed that entrustment anchors do not provide residents with information about their expected rate of progress. Residents noted that this lack of information is a challenge. Residents and faculty with experience using these entrustment anchors identified that, while we likely should try to divert the culture away from peer comparison, there is value in knowing if they are progressing appropriately through their training. Residents described alternate approaches to obtaining this information (BOX 2). Some faculty suggested that they would provide context to their ratings by providing adjacent normative information (BOX 2).

BOX 2 Perspectives on the Role of Normative Data

- ... You don't want to be below your peers, I think that for me is more important than being above my peers... I want to be in line with everyone else moving forward, because if there's clear gaps... that's what I want to work on. So, yeah definitely that part is missing from entrustability scales. (phase 2 resident)
- People say what's the typical PGY-2 supposed to be, well I don't think that there's anyone who's typical truthfully, so I think people have to stop worrying about that. (phase 2 faculty)
- ... You don't have a real sense of where you should be and whether or not you're doing well or not doing well. (phase 1 resident)
- ... If I feel like my evaluation was less than I expected, that would be an opportunity that I would probably speak to them and say, "I thought I was doing better, but is my performance appropriate for my level?" (phase 2 resident)
- ... In the last 6 months... you should be 4 and 5. If you're below that, then you're below where you should be... (phase 2 faculty)

Although the philosophy of entrustment is shifting trainees away from relativistic, individual assessment and toward competency-based assessment, resident participants expressed a persistent desire to know that they were progressing as expected. They felt that this information was provided when they were compared against a "typical" trainee with similar experience as them. This information is not provided in a WBA tool using entrustment anchors.

Discussion

This study provides valuable insight into why residents and faculty theorize that entrustment anchors are better than traditional anchors. The more pragmatic language of entrustment anchors makes them *concrete,* and therefore more easily justifiable, and more *transparent*, enabling use of the entire scale as both faculty and residents have a better understanding of how rating levels relate to clinical practice. The *alignment* of the language of entrustment anchors *with training outcomes* promotes better feedback conversations as it emphasizes progress toward independence over personal evaluation. These 3 key advantages may help explain the improved reliability of WBA tools that use entrustment anchors, as noted in the literature.[11–18,20–22]

Recent research supports our study participants' perceptions that entrustment anchors promote use of the entire scale and better feedback. In a study examining the impact of adding entrustment

language to a rating scale, Dolan and colleagues noted that raters used the lower end of the scale, whereas they previously had noted a restriction to using only the "higher end" of the scale.[30] This research also noted an increase in the written feedback once entrustment language was incorporated.[30] There is a concern that resistance to giving and receiving low scores in medical education can lead to biased assessments (ie, use of only the higher end of the scale),[31] potentially contributing to a "failure to fail."[32] It is important to find approaches to mitigate this impact. The combined results of the study by Dolan et al[30] and our current study suggest that entrustment anchors promote use of the entire scale, and therefore, may offer a piece of the solution.

Given the improved reliability of WBA tools using entrustment anchors, many residency training programs are moving to adopt these tools. Our study provides insight into issues that faculty and residents who are unfamiliar with these new anchors may raise. In this study, there were concerns raised by participants without experience using entrustment anchors about the context in which these scales can be effectively used (procedural versus non-procedural, junior versus senior trainees, and direct versus indirect observation). However, those with experience provided contrasting perspectives for these contextual concerns. Our study results suggest that these concerns need to be acknowledged and addressed. Training may be helpful for both faculty and residents prior to the introduction of WBA tools using entrustment anchors to maximize their effectiveness.

While the majority of data from this study provides evidence to support the use of entrustment anchors, participants did express that a lack of normative data is a significant unforeseen gap. Residents with and without experience being assessed using WBA tools with entrustment anchors expressed a desire for peer comparison, which they felt was lost when using entrustment anchors. It is interesting to note that residents seem to believe that when they receive an assessment of "above expectations," it is an accurate depiction of their performance, indicating they are progressing beyond their peers. Most faculty in the study questioned the reliability of assessments using these types of anchors, as their personal experience was consistent with the literature, which demonstrates a significant range restriction using only the high end of the scale.[9,10,32]

The majority of residents also expressed a desire to know if they were progressing appropriately. Residents exposed to entrustment anchors discussed alternate methods of receiving this information, such as asking faculty to comment as to whether or not their performance was appropriate for their level of experience. Faculty provided examples explaining why a learner received a certain score, but framed it for the learner with a verbal statement that reassured them that their performance was appropriate for their trainee level. Although competency-based medical education curricula decrease the emphasis on the time spent acquiring abilities, there is clearly still a time-based element to residency training. Residents in this study expressed a strong desire to receive the information that either their progress is appropriate, or that they should spend more time and effort on this milestone. Thus, it would seem imperative for programs using WBA tools with entrustment anchors to explicitly identify where and how residents will receive information about their overall rate of progression.

This study has limitations. The use of a single site and the assessment culture at this institution does not necessarily represent other institutions, limiting the generalizability of these findings. This study also focused on the specific anchors that were used at our institution; therefore, the results may not generalize to WBA tools that use another specific type of entrustment anchor such as milestone-based rating scales. Finally, the sample size was small.

This research presents an opportunity for further exploration into the experience of different types of entrustment anchors, such as milestone-based rating scales describing the specific performance required at each level on the way toward competent independent practice. Although the levels for these scales are often described in greater detail, the rater must still make a judgment that the trainee has performed at the particular level described. Our core themes may offer a framework for survey-based approaches aimed at larger sample populations to assist in deepening the understanding of how entrustment anchors are used by faculty and residents and for comparison studies at institutions using a different type of entrustment anchor.

Conclusion

Insights gained from this analysis contribute to a theoretical framework of benefits and challenges related to the adoption of entrustment anchors. A richer understanding of faculty and resident perspectives on entrustment anchors can assist WBA developers in creating more acceptable tools, including optimizing the language for entrustment anchor descriptors. Faculty and residents who have not used WBA tools expressed concerns that those exposed to the WBA tools do not have. As residents in this study expressed concerns regarding the lack of normative

data for the expected rate of progression, there may be a need for training programs to ensure that this information continues to be provided.

References

1. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA*. 2002;287(2):226–235.
2. Bullock G, Allen T, Chapin J, Cole G, Harris KA, Hodder R, et al. Assessment. In*: Competence by Design: Reshaping Canadian Medical Education*. Ottawa, Canada: Royal College of Physicians and Surgeons of Canada; 2014:106–117. http://www.royalcollege.ca/rcsite/educational-initiatives/cbd-medical-education-ebook-e. Accessed April 22, 2019.
3. Embo M, Driessen E, Valcke M, van der Vleuten CP. Integrating learning assessment and supervision in a competency framework for clinical workplace education. *Nurse Educ Today*. 2015;35(2):341–346. doi:10.1016/j.nedt.2014.11.022.
4. Bok HG, Teunissen PW, Favier RP, Rietbroek NJ, Theyse LF, Brommer H, et al. Programmatic assessment of competency-based workplace learning: when theory meets practice. *BMC Med Educ*. 2013;13:123. doi:10.1186/1472-6920-13-123.
5. Govaerts M, van der Vleuten CP. Validity in work-based assessment: expanding our horizons. *Med Educ*. 2013;47(12):1164–1174. doi:10.1111/medu.12289.
6. Carline JD, Wenrich M, Ramsey PG. Characteristics of ratings of physician competence by professional associates. *Eval Health Prof*. 1989;12(4):409–423. doi:10.1177/016327878901200403.
7. Kreiter CD, Ferguson K, Lee WC, Brennan RL, Densen P. A generalizability study of a new standardized rating form used to evaluate students' clinical clerkship performances. *Acad Med*. 1998;73(12):1294–1298.
8. van der Vleuten C, Verhoeven B. In-training assessment developments in postgraduate education in Europe. *ANZ J Surg*. 2013;83(6):454–459. doi:10.1111/ans.12190.
9. Turnbull J, Van Barneveld C. Assessment of clinical performance: in-training evaluation. In: Norman GR, Van der Vleuten CPM, Newble DI, eds. *International Handbook of Research in Medical Education*. London, UK: Kluwer Academic Publishers; 2002:793–810.
10. Gray JD. Global rating scales in residency education. *Acad Med*. 1996;71(suppl 1):55–61.
11. Crossley J, Johnson G, Booth J, Wade W. Good questions, good answers: construct alignment improves the performance of workplace-based assessment scales. *Med Educ*. 2011;45(6):560–569. doi:10.1111/j.1365-2923.2010.03913.x.
12. Durkin GJ. Development and implementation of an independence rating scale and evaluation process for nursing orientation of new graduates. *J Nurses Staff Dev*. 2010;26(2):64–72. doi:10.1097/NND.0b013e3181d47991.
13. Weller JM, Jones A, Merry AF, Jolly B, Saunders D. Investigation of trainee and specialist reactions to the mini-Clinical Evaluation Exercise in anaesthesia: implications for implementation. *Br J Anaesth*. 2014;103(4):524–530. doi:10.1093/bja/aep211.
14. Gofton WT, Dudek NL, Wood TJ, Balaa F, Hamstra SJ. The Ottawa Surgical Competency Operating Room Evaluation (O-SCORE): a tool to assess surgical competence. *Acad Med*. 2012;87(10):1401–1407. doi:10.1097/ACM.0b013e3182677805.
15. MacEwan MJ, Dudek NL, Wood TJ, Gofton WT. Continued validation of the O-SCORE (Ottawa Surgical Competency Operating Room Evaluation): use in the simulated environment. *Teach Learn Med*. 2016;28(1):72–79. doi:10.1080/10401334.2015.1107483.
16. George BC, Teitelbaum EN, Meyerson SL, Schuller MC, DaRosa DA, Petrusa ER, et al. Reliability, validity, and feasibility of the Zwisch scale for the assessment of intraoperative performance. *J Surg Educ*. 2014;71(6):e90–e96. doi:10.1016/j.jsurg.2014.06.018.
17. Warm EJ, Mathis BR, Held JD, Pai S, Tolentino J, Ashbrook L, et al. Entrustment and mapping of observable practice activities for resident assessment. *J Gen Intern Med*. 2014;29(8):1177–1182. doi:10.1007/s11606-014-2801-5.
18. Beard JD, Marriott J, Purdie H, Crossley J. Assessing the surgical skills of trainees in the operating theatre: a prospective observational study. *Health Technol Assess*. 2011;15(1):i–xxi, 1–162. doi:10.3310/hta15010.
19. Rekman J, Gofton W, Dudek N, Gofton T, Hamstra SJ. Entrustability scales: outlining their usefulness for competency-based clinical assessment. *Acad Med*. 2016;91(2):186–190. doi:10.1097/ACM.0000000000001045.
20. Rekman J, Hamstra S, Dudek N, Wood T, Seabrook C, Gofton W. A new instrument for assessing resident competence in surgical clinic: the Ottawa Clinic Assessment Tool (OCAT). *J Surg Educ*. 2016;73(4):575–582. doi:10.1016/j.jsurg.2016.02.003.
21. Voduc N, Dudek N, Parker CM, Sharma KB, Wood TJ. Development and validation of a bronchoscopy competence assessment tool in a clinical setting. *Ann Am Thorac Soc*. 2016;13(4):495–501. doi:10.1513/AnnalsATS.201508-548OC.
22. Halman S, Rekman J, Wood T, Baird A, Gofton W, Dudek N. Avoid reinventing the wheel: implementation of the Ottawa Clinic Assessment Tool (OCAT) in internal medicine. *BMC Med Educ*. 2018;18(1):218. doi:10.1186/s12909-018-1327-7.
23. Gofton W, Dudek N, Barton G, Bhanji F. Workplace-Based Assessment Implementation Guide: Formative tips for medical teaching practice. 1st ed. Ottawa, 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 April 22, 2019.
24. Charmaz K. Grounded theory methods in social justice research. In: Denzin NK, Lincoln YS, eds. *The SAGE Handbook of Qualitative Research*. Thousand Oaks, CA: SAGE Publications; 2011:359–380.
25. Creswell JW. *Qualitative Inquiry & Research Design: Choosing Among Five approaches*. 3rd ed. Thousand Oaks, CA: SAGE Publications; 2013.
26. Charmaz K. *Constructing Grounded Theory*. 2nd ed. Thousand Oaks, CA: SAGE Publications; 2014.
27. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77–101.
28. Lincoln YS, Guba EG. *Naturalistic Inquiry*. Beverly Hills, CA: SAGE Publications; 1985.
29. Edhlund B, McDougall AG. *Nvivo 11 Essentials: Your Guide to the Leading Qualitative Data Analysis Software*. Stallarholmen, Sweden: Form & Kunskapab; 2016.
30. Dolan BM, O'Brien CL, Green MM. Including entrustment language in an assessment form may improve constructive feedback for student clinical skills. *Med Sci Educ*. 2017;27(3):461–464.
31. Gomez-Garibello, C, Young, M. Emotions and assessment: considerations for rater-based judgements of entrustment. *Med Educ*. 2018;52(3):254–262. doi:10.1111/medu.13476.
32. Dudek NL, Marks M, Regehr G. Failure to fail: the perspectives of clinical supervisors. *Acad Med*. 2005;80(suppl 10):84–87.

Nancy Dudek, MD, MEd, is a Professor, Department of Medicine, University of Ottawa, Ottawa, Canada; **Wade Gofton, MD, MEd,** is Associate Professor, Department of Surgery, University of Ottawa, Ottawa, Canada; **Janelle Rekman, MD, MAEd,** is Clinical Hepatobiliary Surgery Fellow, Department of Surgery, Virginia Mason Medical Center; and **Allan McDougall, MA, PhD,** is Adjunct Professor, Department of Education, University of Ottawa, Ottawa, Canada.

Funding: This study was funded by a Physicians' Services Foundation Inc Health Research Grant.

Conflict of interest: The authors declare they have no competing interests.

Research was presented in part at the International Conference on Residency Education, Halifax, Nova Scotia, Canada, October 18–20, 2018.

The authors would like to thank Katherine Scowcroft for her excellent work on this project as a research assistant, Tanya Horsley and the Research Unit of the Royal College of Physicians and Surgeons of Canada for their early guidance, and Tyson Gofton for his thoughtful review and suggestions regarding the manuscript.

Corresponding author: Nancy Dudek, MD, MEd, The Ottawa Hospital, Room 1105D, 505 Smyth Road, Ottawa, ON K1H 8M2 Canada, 613.737.7350 ext 75596, fax 613.739.6974, ndudek@toh.ca

Received November 29, 2018; revision received March 14, 2019; accepted April 9, 2019.