



Design of a new competency-based entrustment scale for the evaluation of resident performance

Janissardhar Skulsampaop^{1,2} · Jessica Rabski¹ · Ashirbani Saha^{1,3} · Michael D. Cusimano¹

Received: 27 May 2022 / Revised: 10 February 2023 / Accepted: 8 April 2023 / Published online: 10 July 2023
© The Author(s) 2023

Abstract

Purpose Recent changes in the design and evaluation of residents to a more competency or mastery-based framework requires frequent observation, evaluation and documentation of residents by busy clinician teachers. Evaluating and determining competent performance is essential for formative evaluation and must be defensible and sound for summative purposes. We sought out experienced Faculty perspectives regarding: (1) important resident performance markers for demonstrating competent attainment of an EPA; (2) the standard of performance expected of graduating residents; (3) evidence for the validity of our purposed entrustment scale; and (4) necessary components required to provide feedback to residents in guiding the development of competent performance of an EPA.

Methods We surveyed Canadian 172 neurosurgical Faculty who had publicly available email addresses and received 67 questionnaire responses, 52 of which were complete responses regarding resident performance markers and our proposed entrustment scale (ES) which consisted of five levels of graded achievement focused on resident performance.

Results Being able to “perform safely” was consistently rated as the most important element of competence that Faculty stated was the critical marker of competence that should be rated, and was found in the D and E Levels of our scale. Our scale does not include any commentary on “performing without supervision” which was rated as the least important marker of performance. For the graduating neurosurgical resident, 90.4% of Faculty indicated that residents should be capable of adapting performance or decisions in response to contextual complexities of the activity independently and safely (Level E) (67.3%) or being able to perform a procedure safely without complexities independently (level D) (21.3%). Eighty percent indicated that the descriptions of competence levels described in our ES (Level A through Level E) represent the appropriate progression of entrustment required demonstrating competent attainment of an EPA. Forty-four percent of Faculty had considerable concern about liability issues with certification of competence based on an ES that is based on descriptions of decreased or no supervision of residents. “Documenting a few weaknesses,” “providing contextual comments of the case,” “providing suggestions for future learning,” and “providing a global assessment for an EPA with one-rating” were rated as the most necessary components in providing effective feedback.

Conclusion Our proposed entrustment global rating scale is easily understood by Faculty who indicate that its graded levels of competence reflect appropriate surgical resident progression in a feasible way. Faculty clearly indicated that the standard of a graduating resident should reflect the ability to perform safely beyond simply performing a case and be able to apply clinical judgments to be able to respond and alter behaviour in response to the clinical and contextual complexities of a case. Our scale focuses on evaluation of resident performance, rather than assessing the supervisor’s degree of involvement. This study has implications for the certification of competence of surgeons and physicians.

Keywords Competency-based medical education · Entrustment scale · Resident education · Entrustable professional activities · Assessment · Evaluation · Standards

Introduction

Postgraduate medical education has undergone considerable transformation in an attempt to address ongoing concerns regarding the readiness of medical graduates for independent

This study was presented in the 42nd Annual Meeting of the Association for Surgical Education (ASE) at Surgical Education Week, San Antonio, Texas on May 4, 2022.

Extended author information available on the last page of the article

practice [1–7]. Many countries have sought to define necessary competencies required of physicians for effectively meeting the healthcare needs of the public they serve in an attempt to increase transparency and accountability of medical training [8–10]. Eventually, competency-based educational frameworks were gradually integrated into many postgraduate medical education programs internationally [11]. Following the adoption, medical education accreditation and regulatory bodies were faced with how to operationalize such overarching competencies to create effective and reliable outcomes-based assessments of residents by Faculty raters.

One solution introduced the concept of an entrustable professional activity (EPA) which was defined as a specific task or responsibility representative of clinical activities required of professional practice, which can be used to monitor and assess a resident's progression of entrustment until a required level of competence is achieved to allow for independent practice [12]. By performing an EPA, residents are required to use a combination of competencies and as such, assessment of competencies can be inferred through the observation of a resident's performance of an EPA [13]. Ten Cate has further clarified that, "competencies are person descriptors, as they signify what individuals are able to do, whereas EPAs are work descriptors and only reflect the work, tasks and activities that are to be carried out in healthcare, irrespective of who does that work" [14]. EPAs therefore coincide with the normal operations of clinical practice, where Faculty supervisors are responsible for assigning an appropriate level of responsibility or "entrustment" to a resident when performing an activity. Importantly, EPAs formalize this process of entrustment and enable the tracking of resident competence attainment across defined sets of professional tasks and activities, promoting accountability and transparency of graduating residents.

Furthermore, resident education is moving from time-based to competency-based medical education (CBME). As valid and reliable assessments for designated EPAs that provide trainees with valuable feedback are the cornerstone of CBME, the entrustment scale plays important role in determining the success of CBME. In Canada, the Ottawa Surgical Competency Operating Room Evaluation (O-SCORE), based on 5 levels, is recommended to use in resident EPA assessment by the RCPSC [15, 16]. The ACGME in the United States developed "Milestones," which are performance levels used in resident/fellow assessment of each subcompetency. It has five levels, proceeding from lower to higher competency [9, 17]. The performance in each ACGME milestone refers to the expected performance in each stage of training, and each specialty indicates details on the procedures that trainees can perform and how well they perform for each ACGME milestones level. Level 4 is considered the graduation target [9].

The Intercollegiate Surgical Curriculum Programme (ISCP) of the United Kingdom (UK) developed "Supervision Level" which assesses resident performance on how much supervision is required in each task [10].

Most of the currently used entrustment scales depend on the supervisor's degree of involvement in the case or level of supervision. These levels do not take into account the complexities that different clinical scenarios may present to the graduating resident surgeon. For example, The Royal College of Emergency Medicine of UK (RCEM) entrustment scale has four main levels (1–4), that focus on the extent of supervisor involvement in residents performing the activity. In level 2, supervisor is "off-site" but still within the hospital in case the trainee needs assistance, whereas in level 3, the supervisor is on standby at home [18]. This could raise concern regarding patient safety and may increase liability [19, 20].

In practice, entrustment decisions given by Faculty raters can be influenced by many different factors, including a supervisor's clinical and teaching experience, the Faculty member's case mix (are they operating on routine cases or more complex cases related to an EPA), personality, propensity to trust, their own expectations, activities they are engaged with outside the operating room, and, issues like existing rapport with residents [21–27]. As such, entrustment scales (ES) that highlight a supervisor's level of involvement in a case rather than focusing on explicit markers of resident performance, minimizes transparency behind Faculty decisions regarding resident entrustment and what a resident needs to improve to gain the desired level of entrustment. Should a graduating resident later face legal challenges based on competency, Faculty raters who certify competence achievement based on their judgment "when they left the resident alone" may raise questions such as how they actually decided performance was competent.

We posit, that assessment of resident competence should be focused on observation of performances that take into account the resident's ability to think and act in a way that responds to new challenges they face as a case evolves. It moves away from a supervision-based scale to one that values an ability to respond to the contextual demands of a case and to provide trainees with effective feedback to help guide the development of competency of each EPA. The objectives of our study were to: (1) identify important resident performance markers for demonstrating competent attainment of an EPA; (2) identify the standard of performance expected of graduating residents; (3) collect evidence for the validity of our purposed entrustment scale; and (4) identify necessary components required to provide feedback to residents in guiding the development of competent performance of an EPA.

Methods

We developed a questionnaire with questions regarding years of practice, number of staff physicians and residents working in the respondents' specialty at their hospital, number of residents the respondents supervise per day, as well as the important resident performance markers for demonstrating competence attainment of an EPA, evidence of validity of our novel entrustment scale, standard expected of EPA respective to our proposed entrustment scale, important components of EPA assessment forms, importance of accessing all of a resident's previous EPA evaluations and respondent's perspectives on certifying competence based on level of resident supervision on additional risk of liability.

For questions regarding the expected standard competence performance, we asked the respondents to choose the level they thought should be a graduation target of residents from one of the proposed five entrustment scales ranging from level A to E, representing lower to higher competency. Levels A-D included aspects of the levels in the O-SCORE, but we excluded the aspects in the scale relating to supervision and purposefully added a fifth level, which explored the ability of the resident to competently respond to novel complexities that might present to the practicing surgeon and their ability to do the procedure competently, even when complexities arise.

Contextual complexities that are referred to in the scale include not only common complications, but also rare complications, and other events/environments that deviate from the normal situation. For example, not having the instrument that the surgeon needs, meaning they must adapt other available instruments to successfully accomplish the procedure, working with newly identified unexpected events or findings during a case such as dealing with an inadvertently injured blood vessel, or, working in difficult team dynamics, where the trainee needs to use a variety of competencies, e.g., leadership and communication, to be able to handle the situation and provide excellent patient care.

The details of our proposed entrustment scale and how it compares to other ES is outlined in Table 1.

Survey

An online questionnaire was sent via email to all Canadian neurosurgery Faculty with publicly available email addresses using the SurveyMonkey platform. Emails inviting neurosurgical Faculty to participate were sent out three times over a four-month period. Participation was completely voluntary. To ensure participant confidentiality, no identifiable information was collected. Each respondent provided informed consent for participation.

This study reports on the questions related to our proposed entrustment scale. Approval for the study was obtained from the Research Ethics Board of Unity Health-St. Michael's Hospital.

Statistical analysis

Only complete responses of questions regarding important performance markers, necessary components of an EPA assessment form, face validity of our proposed entrustment scale and standard of performance expected of graduating residents were analyzed.

Descriptive data of respondent's demographics (proportion for categorical responses, median and ranges for numerical responses) regarding working experience, along with descriptive data about programs, including: number of physicians, number of residents at respondent's hospital and number of residents under the respondent's supervision per day were calculated.

The responses of 5-point Likert scale questions were summarized using proportion, means (strongly agree/extremely important = 5, agree/very important = 4, neither agree nor disagree/not so important = 3, disagree/not so important = 2, strongly disagree/not at all important = 1), and summation proportion of "strongly agree" and "agree" or "extremely important" and "very important."

For question regarding Faculty's Perspectives on degree of importance of resident performance markers for demonstrating competence attainment of an EPA, we ranked the degree of importance of each performance marker by using the calculated means of perceived importance.

We noted the proportion of responses for all questions that required categorical responses and used a probability level of 0.8 for tests of proportionality. Medians, minimums, maximums and/or IQRs were calculated for numerical responses. Categorical responses were analyzed using Fisher exact test and numerical responses were analyzed using Wilcoxon rank-sum test for sub-group analyses involving the respondent's years of Faculty experience (up to 10 years and more than 10 years). A p value of <0.05 was used for statistical significance. All analyses were carried out in R [33].

Latent class analysis (using package *poLCA*) [34] was conducted using the agreeability (coded as agreed, neutral, and disagreed) of a respondent in considering each of the eight performance markers as important for evaluating competency. Latent refers to an unobserved variable consisting of different categories/subgroups/classes. A set observed categorical variables (in our case, the eight performance markers) is analyzed through latent class analysis. Such analysis allows to characterize a latent (unobserved) variable such that the parameters of some of the observed variables vary across the classes of the latent variable [35, 36].

Table 1 Comparison of our proposed entrustment scale with other currently used entrustment scales

Our proposed entrustment scale	O-SCORE [15, 16]	ACGME milestones [9, 17]	ISCP supervision level [10]	ISCP competence level [10, 28–32]	RCEM entrustment scale [18]
0				0 No experience expected	
1	A Allowed to observe only, as requires complete guidance or is not currently permitted to perform the activity yet	Level 1 The resident/fellow demonstrates milestones expected of an incoming resident/fellow	I Able to observe only: no execution	1 Has observed	1 Direct supervisor observation/involvement, able to provide immediate direction/assistance
2	B Requires constant guidance and instruction when performing the activity	Level 2 The resident/fellow is able to perform some tasks with limited autonomy but is not yet performing at a mid-training level	II Able and trusted to act with direct supervision	2 Can do with assistance	
3	C Requires occasional guidance or instructional prompting when performing the activity	Level 3 It is considered as mid-training competency. Trainee can perform common tasks with autonomy	III Able and trusted to act with indirect supervision The supervisor does not need to guide all aspects of the activity. For those aspects which do need guidance, this may be given from another setting	3 Can do whole but may need assistance	2a Supervisor on the 'shop-floor' (e.g. ED, theatres, AMU, ICU), monitoring at regular intervals 2b Supervisor within hospital for queries, able to provide prompt direction or assistance and trainee knows reliably when to ask for help
4	D Requires assistance or instruction only when responding to contextual demands of the activity to ensure patient safety	Level 4* Resident/fellow has advanced and substantially demonstrates the milestones targeted for residency/fellowship. This level is designed as the graduation target	IV* Able and trusted to act at the level of a day-one consultant	4* Competent to do without assistance including complications	4* Would be able to manage with no supervisor involvement (all trainees practice with a consultant taking overall clinical responsibility)

Table 1 (continued)

Our proposed entrustment scale	O-SCORE [15, 16]	ACGME milestones [9, 17]	ISCP supervision level [10]	ISCP competence level [10, 28–32]	RCEM entrustment scale [18]
5 E	5	Level 5 Resident/fellow has advanced beyond performance targets set for residency/fellowship and is demonstrating “aspirational” goals which might describe the performance of someone who has been in practice for several years. It is expected that only a few exceptional residents will reach this level	V Able and trusted to act at a level beyond that expected of a day-one consultant		

ACGME Accreditation Council for Graduate Medical Education, *ISCP* The Intercollegiate Surgical Curriculum Programme, *O-SCORE* The Ottawa Surgical Competency Operating Room Evaluation, *RCEM* The Royal College of Emergency Medicine

*Level represents standard of performance expected of graduating resident

Finally, open-ended comments were thematically analyzed to help understand the reasoning behind Faculty’s responses.

Results

We received a total of 67 questionnaire responses, of which 52 responses (usable response rate of 77.6%) were considered complete responses with fully usable data regarding our objectives and included into the analysis for this study. Of the respondents, 40% had been in a neurosurgical Faculty position for 10 or less than 10 years and 60% for 11 years or longer. Demographics of Faculty are shown in Table 2. By applying Latent Class analysis to the responses related to the importance of each performance marker, we found two classes of Faculty that described their perspectives on necessary performance markers for demonstrating competence attainment of an EPA (Fig. 1). Table 3 demonstrates characteristics of each LCA class. The first group included Faculty who had a mixture of agreeability on every performance marker. In contrast, Faculty in the class 2 were found to have 100% agreeability for being able to “perform safely,” “perform effectively,” “adapt performance or decisions in response to unexpected events,” “adapt performance or decisions in response to contextual complexities of the case,” and “perform independently,” which are similar to the five performance markers with the highest average degree of importance for demonstrating competence attainment of an EPA.

No significant differences in these two classes were found when considering Faculty’s working experience or selected level E as a graduation target.

Aim 1: Resident performance markers of EPA competence attainment

When Faculty were asked to rate the degree of importance of each performance marker for demonstrating competence attainment of an EPA, the most important resident performance marker was to perform the procedures safely (Fig. 2). Moreover, being able to “perform safely” is the marker that had highest proportion of strongly agree or agree (96.1%). The other four performance markers with the highest average degree of importance were: being able to “perform effectively,” “adapt their performance or decisions in response to unexpected events,” “adapt their performance or decisions in response to contextual complexities of the case,” and “perform independently.” On the other hand, the least important resident performance marker, on average, was to “perform without supervision.” In addition, being able to “perform without supervision” is the marker that had greatest proportion of strongly disagree or disagree of (19.3%).

Table 2 Demographics of respondents ($N=52$)

		<i>N</i> (%)
Years of Practice Number (%)	<5 years	11 (21.2%)
	5–10 years	10 (19.2%)
	11–20 years	19 (36.5%)
	>20 years	12 (23.1%)
# Faculty Practicing at Respondent’s Hospital (median [Range])		9.5 [3–32]
# Residents on Participant’s Specialty Service at Respondent’s Hospital (median [Range])		5 [1–32]
# Residents Supervised by Respondent per Day (median [Range])		1 [1–5]

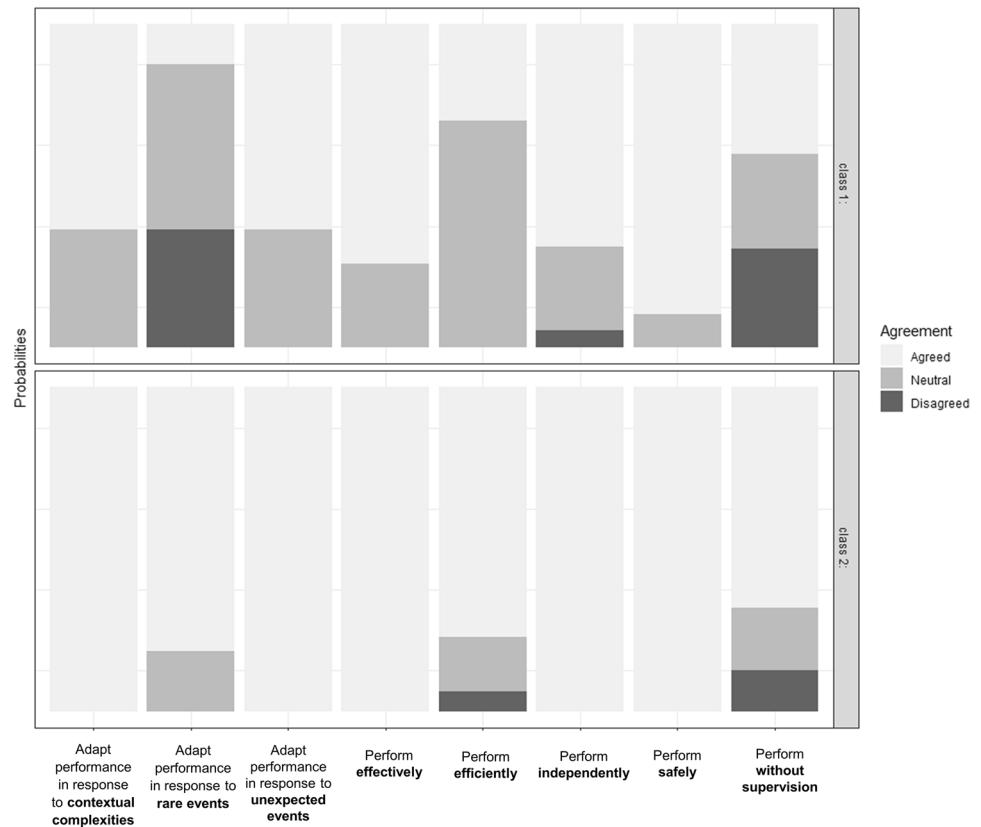
Fig. 1 Findings from latent class analysis

Figure 3 shows the results of the performance markers Faculty indicated to be important for demonstrating resident competence attainment of an EPA from their perspective as a Faculty member (in blue-colored bar) and the performance markers that would be important to them if they were a patient (green-colored bar). Being able to “perform safely” and “adapt performance or decision in response to contextual complexities of the case” were the two most frequently chosen performance markers by Faculty from their perspective as both Faculty and a patient point of view. Performing without supervision was considered far less important when respondents considered the level from the perspective of Faculty or the perspective as patients when answering the question.

The finding that being able to “perform without supervision” was seen as less important than other resident performance markers in demonstrating resident competence attainment of an EPA were supported by the following Faculty comments:

A neurosurgeon who strongly disagreed with being able to “perform without supervision” said:

“No student pilot transports passengers without supervision....an important analogy.”

Another neurosurgeon who disagreed with this performance marker stated:

Table 3 Characteristics of faculty in class 1 and 2 from latent class analysis

	Class1 (N=19) N (%)	Class 2 (N=33) N (%)	P value
<i>Working experience</i>			
<5 years	3 (15.8%)	8 (24.2%)	<i>p</i> = 0.75
5–10 years	3 (15.8%)	7 (21.2%)	
11–20 years	7 (36.8%)	12 (36.4%)	
20 years	6 (31.6%)	6 (18.2%)	
# Faculty Practicing at Participant’s Hospital (median)	10	9	<i>p</i> = 0.68
# Residents on Participant’s Specialty Service at Participant’s Hospital (median)	4	5	<i>p</i> = 0.43
# Residents Supervised by Participant per Day (median)	1	2	<i>p</i> = 0.08
# Faculty considered entrustment Level E as standard of performance expected of graduating residents	14 (73.7%)	22 (66.7%)	<i>p</i> = 0.76
# Faculty who strongly agree/agree with the following performance markers for demonstrating competence attainment of an EPA			
Perform safely	17 (89.5%)	33 (100%)	<i>p</i> = 0.002*
Adapt their performance or decisions in response to contextual complexities of the case	12 (63.2%)	33 (100%)	
Perform independently	13 (68.4%)	33 (100%)	
Perform effectively	14 (73.7%)	33 (100%)	
Adapt their performance or decisions in response to unexpected events	12 (63.2%)	33 (100%)	
Perform without supervision	7 (36.8%)	23 (69.7%)	
Adapt their performance or decisions in response to rare events	2 (10.5%)	27 (81.8%)	
Perform efficiently	5 (26.3%)	26 (78.8%)	

EPA entrustable professional activity; N number

*Statistically significant

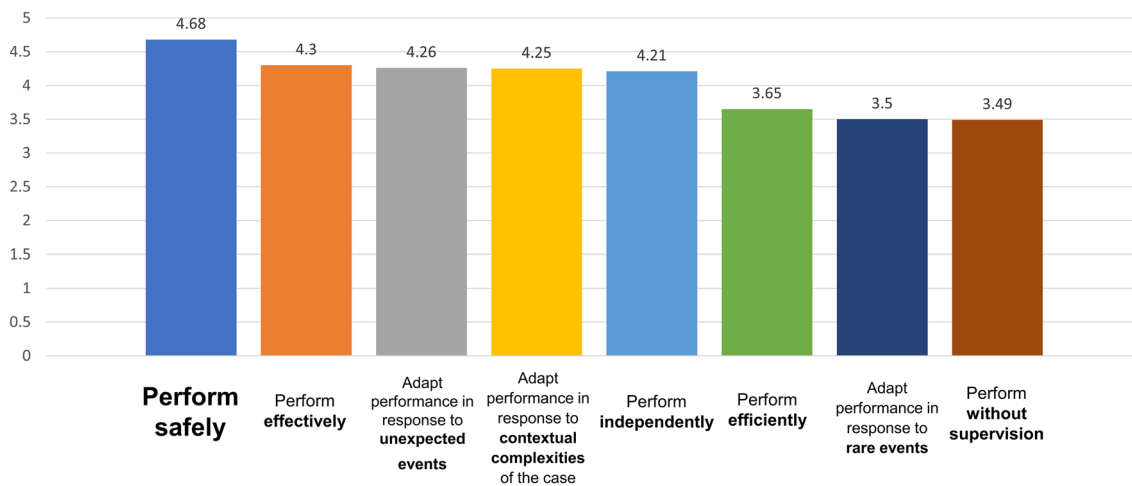


Fig. 2 Faculty’s perspectives on degree of importance of resident performance markers for demonstrating competence attainment of an EPA

“It cannot be a marker, if it is not observed. Who marks? The resident will say: me. However, that has a clear observer-bias built in.... The only control tool then is the postsurgical outcome. One may argue: if the patient is fine, surgery was probably fine.

However - that does not assess the quality of care delivered.... That makes such an evaluation system unreliable.”

Another neurosurgeon who neither agreed nor disagreed with this marker remarked:

“Should be ‘could perform’ instead of ‘can perform’ because I will always be present during surgeries involving patients under my care.”

Liability with documentation of faculty level of supervision

The issue regarding liability was raised as many currently used entrustment scales assess resident competency with the degree of supervision. Nearly half of the Faculty surveyed (44.2%, 23/52) strongly agreed or agreed that, “certifying competence based on decreased levels of resident supervision (e.g., documenting level of supervision) places Faculty at additional risk of liability should a patient later complain about their quality of treatment.”

In contrast, 28.8% of Faculty strongly disagreed/disagreed, and 19.2% neither agreed nor disagreed with this sentiment.

No significant difference of opinion was appreciated based on Faculty’s working experience.

A neurosurgeon who agreed that, “certifying competence attainment using entrustment scales based on decreasing levels of resident supervision placed additional risk of liability on Faculty” said:

“Agree based on experience with a College complaint (that was dismissed).”

Aim 2: The standard of performance expected of graduating residents

Over two-thirds of Faculty (67.3%, 35/52) believed that level E, which is defined as residents being capable of adapting performance or decisions in response to contextual complexities of the activity independently and safely, as opposed to being able to solely perform a task without complexities, represents the standard expected of competent performance of graduating residents. Twelve respondents (23.1%) considered being able to simply perform a procedure without complexities independently to be representative of an appropriate level of competence for graduating residents. Level C which is being able to perform the activity with occasional guidance required was selected by only 3.8% (2/52).

No respondent indicated that the lowest two levels of competence (A and B) were sufficient for graduating residents. Three Faculty selected others and provided comments. No significant differences in responses were found as a result of Faculty’s working experience.

Aim 3: Evidence of validity of proposed entrustment scale

The results supported that our proposed entrustment scale has good evidence of validity as 80.8% of Faculty agreed or strongly agreed to the appropriateness of the proposed progression of entrustment for ensuring resident competence in performing a given EPA. 78.9% believed that five different levels of entrustment would be ideal for such a scale, and 76.9% agreed the provided descriptions for each progression of entrustment were easy to understand (Table 4).

Our scale can discriminate the different levels of performance expected at graduation as over 90% Faculty believed that level E or D should be a graduation target, and no Faculty selected the two lowest levels of the scale as sufficient.

Aim 4: Necessary components required to provide effective feedback to residents in guiding the development of competent performance of an EPA

Table 5 shows that 71.1% of Faculty believed that “documenting a few weaknesses” was an extremely important or very important component, followed by “providing contextual comments of the case” (67.3%), “providing suggestions for future learning” (57.7%), and “providing a global assessment for an EPA with one-rating” (50%). “Documenting 2–3 strengths” was only perceived as very or extremely important by 40.4%. Finally, “providing an evaluation for each checklist item of an EPA” was deemed important by only 27% of Faculty, and when compared to the other components, a significantly higher percentage of Faculty indicated that it was not so or not at all important (9.6%) ($p < 0.0002$).

A neurosurgery Faculty with less than 5 years’ experience provided the following comment:

“The most important feedback and evaluation is the personal, immediate and oral format. At the end of each rotation, one should give a global written assessment. That would be my preferred format.”

Another neurosurgeon with 11–20 years of working experience indicated an important element of providing resident feedback:

“Knowledge of management principles for the case.”

Discussion

To support the CBME’s goals of increasing transparency and accountability of its residency graduates, this study has identified performance markers for standardizing Faculty expectations for guiding entrustment decisions and documenting

Table 4 Faculty perspectives on evidence for validity of our proposed entrustment scale ($N=52$)

	Degree of agreement					Average degree of agreement	Strongly agree/agree
	Strongly agree N (%)	Agree N (%)	Neither agree nor disagree N (%)	Disagree N (%)	Strongly disagree N (%)		
Appropriate progression	8 (15.4%)	34 (65.4%)	8 (15.4%)	1 (1.9%)	1 (1.9%)	3.90	80.8%
5 different level (A-E) is ideal number of levels	8 (15.4%)	33 (63.5%)	8 (15.4%)	1 (1.9%)	2 (3.8%)	3.84	78.9%
Descriptions of each level is easy to understand	6 (11.5%)	34 (65.4%)	6 (11.5%)	4 (7.7%)	2 (3.8%)	3.73	76.9%

and tracking residents' successful attainment of a required level of competence for an EPA. Our scale is compared to others in use in Table 1. Most of the other scales assess resident performance based on some indication of their supervisor's degree of involvement in the case. The results in our study revealed that being able to "perform without supervision" was rated as the least important performance marker for demonstrating resident competence attainment of

an EPA. Additionally, 44.2% of Faculty believed that certifying competence based on decreased levels of resident supervision (e.g., documenting that resident is competent enough and the supervisor does not need to come on-site to supervise) would place Faculty at additional risk of liability. Our entrustment scale addresses the need to avoid the level of supervision as a performance marker.

Table 5 Faculty perspectives on necessary components required to provide feedback to residents in guiding the development of competence performance of an EPA ($N=52$)

Components required to provide resident feedback	Degree of Importance						Average Degree of Importance (Rank)	Extremely Important/Very Important (Rank)
	Extremely Important N (%)	Very Important N (%)	Somewhat Important N (%)	Not So Important N (%)	Not At All Important N (%)	I don't know N (%)		
Documenting 2–3 weaknesses	10 (19.2%)	27 (51.9%)	12 (23.1%)	1 (1.9%)	1 (1.9%)	1 (1.9%)	3.86 (1)	71.1% (1)
Providing contextual comments (unique or challenging case details)	9 (17.3%)	26 (50%)	13 (25%)	2 (3.8%)	1 (1.9%)	1 (1.9%)	3.78 (2)	67.3% (2)
Providing suggestions for future learning	8 (15.4%)	22 (42.3%)	15 (28.8%)	5 (9.6%)	0 (0%)	2 (3.8%)	3.66 (3)	57.7% (3)
Providing a global assessment for an EPA (one-rating)	4 (7.7%)	22 (42.3%)	18 (34.6%)	4 (7.7%)	1 (1.9%)	3 (5.8%)	3.48 (4)	50% (4)
Documenting 2–3 strengths	4 (7.7%)	17 (32.7%)	25 (48.1%)	4 (7.7%)	1 (1.9%)	1 (1.9%)	3.37 (5)	40.4% (5)
Providing an evaluation for each checklist item of an EPA	3 (5.8%)	11 (21.2%)	22 (42.3%)	11 (21.2%)	3 (5.8%)	2 (3.8%)	3.0 (6)	27% (6)

EPA entrustable professional activity

We found that being able to “adapt performance in response to contextual complexities of the case” was highly important from both Faculty perspectives as a patient and perspectives as a Faculty member in rating graduating residents. This supports the inclusion of Level E in any ES. While the ISCP’s competence scale also assesses the ability of the resident to be able to manage without assistance, including potential common complications, our scale is not limited to solely considering complications [10, 28–32]. We focus on contextual complexities and the demands that these complexities place on the resident to perform safely. It also includes environments or events that deviate from a “normal” situation. Our Level E indicates that the graduating resident needs to be able to handle new events that occur during a case, not only able to do the sort of case that presents any challenging by going beyond the ordinary case or situation. Being able to handle new events that arise and to be able to “think of one’s feet” and adapt a solution that meets the needs of the patient to get them through surgery in a safe way is considered an important feature of a competent graduating resident. Over 90% of Faculty selected entrustment level E (67.3%) and D (23.1%) as a graduation target and no Faculty chose the lowest two levels, indicating that our scale discriminates between resident performance and so provides evidence for the validity of our proposed entrustment scale.

Being able to “perform efficiently” was ranked the eighth necessary performance marker selected by Faculty from both Faculty perspectives and Faculty as a patient perspective (Fig. 3). These findings challenge entrustment scales based on residents performing without supervision or other anecdotal definitions that highlight the need to perform efficiently [37], and instead agree with some studies suggesting

that efficiency should be reserved for experts [38]. It also reveals that though it is important to operate independently, this does not necessitate a lack of supervision.

Moreover, 80.8% agreed that our scale’s descriptions were easy to understand and 78.9% corresponded that the five different levels of entrustment (A-E) represented an ideal number, which contrasts the use of some existing 9-point or 4-point entrustment scales [39, 40]. The highest level of entrustment, “E”, was designed to include important features of performance as identified in the literature (performing independently, safely, while also being able to respond to contextual complexities that could arise) [41–44]. Though a majority selected “E” as the level of entrustment meant to represent a resident’s competence attainment, about one quarter selected lower entrustment levels of “D” or “C”. This could possibly be due to the variations in the personal comfort levels or attributes of supervisors with entrusting residents with professional activities [25, 45–47].

It is important to note that the standard of performance expected of graduating residents using our proposed entrustment scale is EPA-dependent. For example, 79% of neurosurgery Faculty indicated that level E should be the graduation target for performing burr hole drainage of a chronic subdural hematoma while only 57% of them believed that level E should be graduation target for performing peripheral nerve decompression procedures [48].

We found that “documenting a few weaknesses,” “providing contextual comments of the case,” “providing suggestions for future learning,” and “providing a global assessment for an EPA with one-rating” were the most necessary components in providing effective feedback. Our scale helps guide residents to develop competent performance of an EPA, as our scale helps identify weaknesses, includes

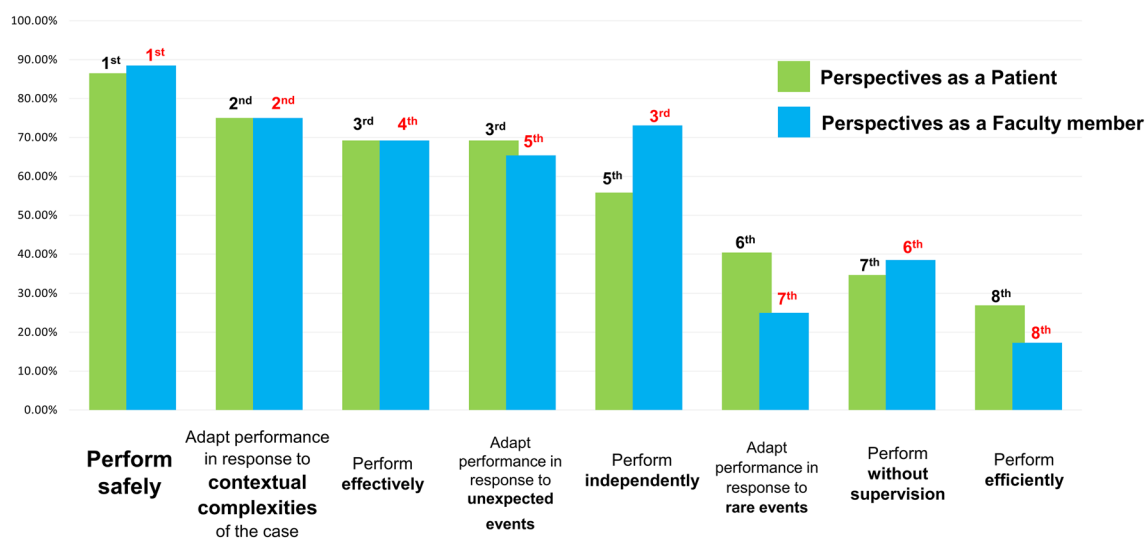


Fig. 3 Perspectives of faculty as a patient versus faculty as a faculty member

“adapting their performance or decisions in response to contextual complexities of the case,” as a performance marker, and is a global rating scale. These features of the scale would support Faculty in providing residents with feedback involving contextual comments which are considered necessary components required for effective feedback. Moreover, our scale is learner-centred, as the scale focuses solely on trainee’s performance rather than primarily highlighting the level of supervision required. It could be used on index cases, defined by the committee tasked with defining the competency levels expected of residents at the time of graduation. These cases would differ based on the specialty, and the expectations of patients and regulatory bodies at the time that the competency level is defined.

Limitations and future studies

Our study shares the limitations of other survey-based methods. We recognize that our results are representative of academic neurosurgeons in Canada, and as such, may not represent the perspectives of neurosurgeons elsewhere or of Faculty from other specialties. Given the electronic nature of the survey, and the difficulty of ascertaining if all emails sent were actually received, our response rate is a lower-bound estimate of the response rate. In addition to having a vast scope of practice, the specialty of neurosurgery has many overlapping procedures with other surgical specialties, regularly requires both medical and surgical management of unstable and critically ill patients and necessitates significant emotional and interpersonal demands of its trainees [49]. Thus, this study’s results likely represent reasonable insights and targets that could be generalized to other specialties looking to integrate CBME frameworks into their curriculum.

We acknowledge that we are in an early phase of the collection of evidence for the validity of our measure. Future studies will need to be conducted over a number of institutions, and, different levels of training, to determine the validity and reliability of this study’s proposed entrustment scale. In addition, given that the importance of providing contextual comments or responding to contextual complexities of a case were emphasized, future studies should be performed to see how the different contexts of the same EPA affect Faculty expectations for standard of performance expected of graduating residents.

Conclusion

To facilitate transparency and accountability of resident training within CBME curriculum, this study has defined Faculty expectations of resident performance markers deemed important for guiding entrustment decisions and

acknowledging competence attainment of an EPA. We have provided an alternative entrustment scale that is learner and feedback-centred and focuses on the development of competency in residents that are being supervised and evaluated. The entrustment scale is a global rating scale that is supervision independent. It has good evidence of validity, as it has appropriate progression levels, is easy to understand, and can discriminate between the levels of training.

Acknowledgements Thanks to Rachel Lamont who assisted in the editing of the manuscript.

Author contributions Conception and design: MDC. Acquisition of data: JR. Analysis and interpretation of data: all authors. Drafting the article: all authors. Critically revising the article: all authors. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: JS. Statistical analysis: AS. Administrative/technical/material support: MDC. Study supervision: MDC.

Funding No funding was received for conducting this study.

Data availability The raw datasets generated during the current study are not publicly available, however, written requests for aggregated data from researchers will be considered.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Smith F, Goldacre MJ, Lambert TW. Adequacy of postgraduate medical training: views of different generations of UK-trained doctors. *Postgrad Med J.* 2017;93(1105):665–70. <https://doi.org/10.1136/postgradmedj-2016-13456>.
2. Napolitano LM, Savarise M, Paramo JC, Soot LC, Todd SR, Gregory J, et al. Are general surgery residents ready to practice? A survey of the American College of Surgeons Board of Governors and Young Fellows Association. *J Am Coll Surg.* 2014;218(5):1063–72e31. <https://doi.org/10.1016/j.jamcollsurg.2014.02.001>.
3. Kempenich JW, Willis RE, Rakosi R, Wiersch J, Schenarts PJ. How do perceptions of autonomy differ in general surgery training between faculty, senior residents, hospital administrators, and the general public? A Multi-Institutional Study. *J Surg Educ.* 2015;72(6):e193–201. <https://doi.org/10.1016/j.jsurg.2015.06.002>.
4. Mattar SG, Alseidi AA, Jones DB, Jeyarajah DR, Swanstrom LL, Aye RW, et al. General surgery residency inadequately prepares

- trainees for fellowship: results of a survey of fellowship program directors. *Ann Surg.* 2013;258(3):440–9. <https://doi.org/10.1097/SLA.0b013e3182a191ca>.
5. Guntupalli SR, Doo DW, Guy M, Sheeder J, Omurtag K, Kondapalli L, et al. Preparedness of obstetrics and gynecology residents for fellowship training. *Obstet Gynecol.* 2015;126(3):559–68. <https://doi.org/10.1097/AOG.0000000000000999>.
 6. Dijkstra IS, Pols J, Remmelts P, Brand PL. Preparedness for practice: a systematic cross-specialty evaluation of the alignment between postgraduate medical education and independent practice. *Med Teach.* 2015;37(2):153–61. <https://doi.org/10.3109/0142159X.2014.929646>.
 7. Fonseca AL, Reddy V, Longo WE, Udelsman R, Gusberg RJ. Operative confidence of graduating surgery residents: a training challenge in a changing environment. *Am J Surg.* 2014;207(5):797–805. <https://doi.org/10.1016/j.amjsurg.2013.09.033>.
 8. The Royal College of Physicians and Surgeons of Canada. What is Competence by Design? The Royal College of Physicians and Surgeons of Canada. <https://www.royalcollege.ca/rcsite/cbd/what-is-cbd-e>. Accessed 18 May 2022.
 9. Edgar L, McLean S, Hogan S, Hamstra S, Holmboe E. The Milestones Guidebook. 2020. The Accreditation Council for Graduate Medical Education. <https://www.acgme.org/globalassets/milestonesguidebook.pdf>. Accessed 20 Apr 2022.
 10. Intercollegiate Surgical Curriculum Programme. The New Surgical Curriculum for August 2021. <https://www.iscp.ac.uk/iscp/curriculum-2021/>. Accessed 18 May 2022.
 11. Miller P. Competency based education and training: a world perspective. In: Arguelles A, Goncz A (eds) *Competency based education and training: a world perspective*. Balderas, Mexico: Editorial Limusa S.A. de C.V. Grupo Noriego Editores 2000. ISBN: ISBN 968-18-6113-2. *Journal of Workplace Learning.* 2001;13:260–1. <https://doi.org/10.1108/jwl.2001.13.6.260.2>.
 12. Ten Cate O. Entrustability of professional activities and competency-based training. *Med Educ.* 2005;39(12):1176–7. <https://doi.org/10.1111/j.1365-2929.2005.02341.x>.
 13. ten Cate O, Scheele F. Competency-based postgraduate training: can we bridge the gap between theory and clinical practice? *Acad Med.* 2007;82(6):542–7. <https://doi.org/10.1097/ACM.0b013e31805559c7>.
 14. Ten Cate O. When I say ... entrustability. *Med Educ.* 2020;54(2):103–4. <https://doi.org/10.1111/medu.14005>.
 15. The Royal College of Physicians and Surgeons of Canada. The Ottawa Surgical Competency Operating Room Evaluation (O-SCORE): A tool to assess surgical competence. <https://www.royalcollege.ca/rcsite/documents/canmeds/keylime-ep32-js-goffton-et-al-e.pdf>. Accessed 18 May 2022.
 16. Goffton 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):1401–7. <https://doi.org/10.1097/ACM.0b013e3182677805>.
 17. The Accreditation Council for Graduate Medical Education and the American Board of Plastic Surgery. The Plastic Surgery Milestone Project. 2015. <https://www.acgme.org/globalassets/pdfs/milestones/plasticsurgerymilestones.pdf>. Accessed 20 Apr 2022.
 18. The Royal College of Emergency Medicine. RCEM Curriculum. 4 August 2021. https://res.cloudinary.com/studio-republic/images/v1633447432/RCEM_Curriculum_2021_Master/RCEM_Curriculum_2021_Master.pdf?_i=AA. Accessed 18 May 2022.
 19. Carraccio C, Englander R, Holmboe E, Kogan J. Driving care quality: aligning trainee assessment and supervision through practical application of entrustable professional activities, competencies, and milestones. *Acad Med.* 2015. <https://doi.org/10.1097/ACM.0000000000000985>.
 20. ten Cate O. Entrustment decisions: bringing the patient into the assessment equation. *Acad Med.* 2017;92(6):736–737.
 21. ten Cate O, Snell L, Carraccio C. Medical competence: the interplay between individual ability and the health care environment. *Med Teach.* 2010;32(8):669–75. <https://doi.org/10.3109/0142159X.2010.500897>.
 22. Hauer KE, Boscardin C, Fulton TB, Lucey C, Oza S, Teherani A. Using a curricular vision to define entrustable professional activities for medical student assessment. *J Gen Intern Med.* 2015;30(9):1344–8. <https://doi.org/10.1007/s11606-015-3264-z>.
 23. Sterkenburg A, Barach P, Kalkman C, Gielen M, ten Cate O. When do supervising physicians decide to entrust residents with unsupervised tasks? *Acad Med.* 2010;85(9):1408–17. <https://doi.org/10.1097/ACM.0b013e3181eab0ec>.
 24. Hoops HE, Haley C, Cook MR, Lopez O, Dewey E, Brasel KJ, et al. Factors influencing amount of guidance in the operating room during laparoscopic cases. *Am J Surg.* 2019;217(5):979–85. <https://doi.org/10.1016/j.amjsurg.2019.03.019>.
 25. Kogan JR, Hess BJ, Conforti LN, Holmboe ES. What drives faculty ratings of residents' clinical skills? The impact of faculty's own clinical skills. *Acad Med.* 2010;85(10 Suppl):S25–8. <https://doi.org/10.1097/ACM.0b013e3181ed1aa3>.
 26. Govaerts MJ, van der Vleuten CP, Schuwirth LW, Muijtens AM. Broadening perspectives on clinical performance assessment: rethinking the nature of in-training assessment. *Adv Health Sci Educ Theory Pract.* 2007;12(2):239–60. <https://doi.org/10.1007/s10459-006-9043-1>.
 27. Kennedy TJ, Regehr G, Baker GR, Lingard L. Point-of-care assessment of medical trainee competence for independent clinical work. *Acad Med.* 2008;83(10 Suppl):S89–92. <https://doi.org/10.1097/ACM.0b013e318183c8b7>.
 28. Bowditch M, Ryan W, Gregory R. Trauma and Orthopaedic Surgery Curriculum. The Intercollegiate Surgical Curriculum Programme. 4 August 2021. <https://www.iscp.ac.uk/media/1331/trauma-orthopaedic-surgery-curriculum-2021-minor-changes-for-august-2022.pdf>. Accessed 18 May 2022.
 29. Cook T, Lund J. General Surgery Curriculum. The Intercollegiate Surgical Curriculum Programme. 4 August 2021. <https://www.iscp.ac.uk/media/1327/general-surgery-curriculum-2021-minor-changes-for-august-2022.pdf>. Accessed 18 May 2022.
 30. Whitfield P, Thomson S, Brown J, Kitchen N, Edlmann E. Neurosurgery Curriculum. The Intercollegiate Surgical Curriculum Programme. 4 August 2021. <https://www.iscp.ac.uk/media/1328/neurosurgery-curriculum-2021-minor-changes-for-august-2022.pdf>. Accessed 18 May 2022.
 31. Cornford P, Forster L, Hodgson D, Morley R, Lane T, Parkinson R. Urology Curriculum. The Intercollegiate Surgical Curriculum Programme. 4 August 2021. <https://www.iscp.ac.uk/media/1325/urology-curriculum-2021-minor-changes-for-august-2022.pdf>. Accessed 18 May 2022.
 32. McCarthy M, Jones K, Travers H, Flett M, Suttie S, Davies R. Vascular Surgery Curriculum. The Intercollegiate Surgical Curriculum Programme. 4 August 2021. <https://www.iscp.ac.uk/media/1113/vascular-surgery-curriculum-aug-2021-approved-oct-20.pdf>. Accessed 18 May 2022.
 33. The R Project for Statistical Computing. The R Foundation. 2017. <https://www.r-project.org/>. Accessed 20 May 2022
 34. Package 'poLCA'. The R Foundation. <https://cran.r-project.org/web/packages/poLCA/poLCA.pdf>. Accessed 20 May 2022.
 35. Vermunt JK, Magidson J. Latent class analysis. In: *The sage encyclopedia of social sciences research methods 2*; 2004. pp. 549–553.
 36. McCutcheon AL. Latent class analysis. No. 64. Sage; 1987.

37. Mellinger JD, Williams RG, Sanfey H, Fryer JP, DaRosa D, George BC, et al. Teaching and assessing operative skills: from theory to practice. *Curr Probl Surg*. 2017;54(2):44–81. <https://doi.org/10.1067/j.cpsurg.2016.11.007>.
38. Guo y, Findley-Bhatta K, Huang CC, Lu S, Wang Y. Competency-based education: evidence from social work postgraduates from five universities in China. *Asian Soc Work Policy Rev*. 2016. <https://doi.org/10.1111/aswp.12096>.
39. DaRosa DA, Zwischenberger JB, Meyerson SL, George BC, Teitelbaum EN, Soper NJ, et al. A theory-based model for teaching and assessing residents in the operating room. *J Surg Educ*. 2013;70(1):24–30. <https://doi.org/10.1016/j.jsurg.2012.07.007>.
40. Weller JM, Castanelli DJ, Chen Y, Jolly B. Making robust assessments of specialist trainees' workplace performance. *Br J Anaesth*. 2017;118(2):207–14. <https://doi.org/10.1093/bja/aew412>.
41. 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(5):740–3. <https://doi.org/10.1007/s11606-019-04878-y>.
42. Chen XP, Harzman A, Cochran A, Ellison EC. Evaluation of an instrument to assess resident surgical entrustable professional activities (SEPA). *Am J Surg*. 2020;220(1):4–7. <https://doi.org/10.1016/j.amjsurg.2019.08.026>.
43. Happel JP, Ritter JB, Neubauer BE. Optimizing the balance between supervision and autonomy in training. *JAMA Intern Med*. 2018;178(7):959–60. <https://doi.org/10.1001/jamainternmed.2018.1250>.
44. Stucke RS, Sorensen M, Rosser A, Sullivan S. The surgical consult entrustable professional activity (EPA): defining competence as a basis for evaluation. *Am J Surg*. 2020;219(2):253–7. <https://doi.org/10.1016/j.amjsurg.2018.12.056>.
45. Santen SA, Wolff MS, Saxon K, Juneja N, Bassin B. Factors affecting entrustment and autonomy in emergency medicine: “How much rope do I give them?” *West J Emerg Med*. 2019;20(1):58–63. <https://doi.org/10.5811/westjem.2018.10.39843>.
46. Govaerts MJ, Schuwirth LW, Van der Vleuten CP, Muijtjens AM. Workplace-based assessment: effects of rater expertise. *Adv Health Sci Educ Theory Pract*. 2011;16(2):151–65. <https://doi.org/10.1007/s10459-010-9250-7>.
47. Hauer KE, Ten Cate O, Boscardin C, Irby DM, Iobst W, O'Sullivan PS. Understanding trust as an essential element of trainee supervision and learning in the workplace. *Adv Health Sci Educ Theory Pract*. 2014;19(3):435–56. <https://doi.org/10.1007/s10459-013-9474-4>.
48. Rabski JE, Saha A, Cusimano MD. Setting standards of performance expected in neurosurgery residency: A study on entrustable professional activities in competency-based medical education. *Am J Surg*. 2021;221(2):388–93. <https://doi.org/10.1016/j.amjsurg.2020.12.014>.
49. Lipsman N, Khan O, Kulkarni AV. “The actualized neurosurgeon”: a proposed model of surgical resident development. *World Neurosurg*. 2017;99:381–6. <https://doi.org/10.1016/j.wneu.2016.12.039>.

Authors and Affiliations

Janissardhar Skulsampaopol^{1,2} · Jessica Rabski¹ · Ashirbani Saha^{1,3} · Michael D. Cusimano¹

✉ Michael D. Cusimano
injuryprevention@smh.ca

¹ Division of Neurosurgery, Department of Surgery, University of Toronto, Toronto, ON, Canada

² Department of Surgery, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand

³ Department of Oncology, McMaster University, Hamilton, ON, Canada