Canadian Medical Education Journal

Scientific Reports

Assessment burden by design: exploring the variability in competence by design assessment forms

Romaisa Ismaeel,[1] Jamie Riggs,[2] Mary Ott,[3] Andrew K Hall,[4,5] Catherine Patocka,[6] Andrew Petrosoniak,[7] Adam Szulewski,[5,8] Teresa Chan,[9] Brent Thoma[5,9,10]

[1]College of Medicine, University of Saskatchewan, Saskatchewan, Canada; [2]Department of Emergency Medicine, University of British Columbia, British Columbia, Canada; [3]Faculty of Education, York University, Ontario, Canada; [4]Department of Emergency Medicine, University of Ottawa, Ontario, Canada; [5]Royal College of Physicians and Surgeons of Canada, Ontario, Canada; [6]Department of Emergency Medicine, University of Calgary Cumming School of Medicine, Alberta, Canada; [7]Division of Emergency Medicine, University of Toronto, Ontario, Canada; [8]Departments of Emergency Medicine and Psychology, Queen's University, Ontario, Canada; [9]School of Medicine, Toronto Metropolitan University School of Medicine, Ontario, Canada; [10]Department of Emergency Medicine, University of Saskatchewan, Saskatchewan, Canada

Correspondence to: Brent Thoma, Department of Emergency Medicine, College of Medicine, University of Saskatchewan. Saskatoon, Saskatchewan, Royal University Hospital, Room 2684, Saskatoon SK S7N0W8; telephone: 306-881-0112; email: brent.thoma@usask.ca; X: @Brent_Thoma.

Published ahead of issue: Dec 8, 2025; CMEJ 2025 Available at https://doi.org/10.36834/cmej.81428

Abstract

Background: The Royal College of Physicians and Surgeons of Canada's (RCPSC) Competence by Design (CBD) framework has been criticized for increasing assessment burden due to the high number of required Entrustable Professional Activity (EPA) assessments. Another contributing factor may be the inefficient design of assessment forms. We explored variability in form design to identify differences that could impact learners' and assessors' experience with CBD.

Methods: Annotated assessment forms for the 'Core 1' (C1) EPA were collected in March 2023 from all (*n* = 14, 100%) RCPSC emergency medicine (EM) residency programs in Canada that had implemented CBD. Forms were divided into six sections to compare their design. The variability between form sections was described relative to RCPSC recommendations on form design.

Results: EPA assessments were completed within six learning management systems. Variability was found throughout the form including the number of context variables, included milestones, milestone rating criteria, and text boxes for narrative feedback; phrasing of narrative feedback prompts, milestone descriptions, and entrustment score criterion; visual presentation of the entrustment score; arrangement of form components; and the components' selection format. The mandatory completion of form components was inconsistent. Some forms could be partially completed by residents. One form added a global performance rating scale. The number of clicks required to complete a form ranged from 12 to 47.

Conclusion: We found considerable variability in the design of the EM C1 EPA assessment form. Variations that make completion more challenging could increase assessment burden. CBD programs should be aware of this and seek to optimize the design of their forms.

Résumé

Résumé français à venir.

Introduction

Competence By Design (CBD), the Royal College of Physicians and Surgeons of Canada's (RCPSC) approach to Competency Based Medical Education (CBME),[1] has been criticized for increasing assessment burden in medical training.[2–5] Assessment burden refers to the increased time and cognitive load required to complete assessments and the resulting negative impact that this has on residents' and attending physicians' wellbeing and on assessment quality.[4,6] Cognitive load is the mental effort required to process information and perform tasks within the limited capacity of working memory. Suboptimal presentation of information increases extraneous cognitive load, which surpasses the capacity of the working memory and imposes constraints on the amount of information that can be handled at one time.[7] Increased assessment burden has potential to compromise resident self-regulatory learning, assessor-learner engagement, and overall resident competence.[3,4] It can be increased by ambiguity in assessment content, the use of dysfunctional electronic learning platforms, and sheer volume of EPA assessments requiring completion.[3,4] As CBD is centered around the frequent assessment of Entrustable Professional Activities (EPAs), poorly designed EPA assessment forms could also increase assessment burden and decrease assessment quality.

The RCPSC standardized the content of CBD EPA assessments within a specialty nationally and provided instruction to inform EPA form design through the Observation Template 1[8] (introduced in 2017), EM EPA Guide[9] (introduced in 2018), and Technical Guide 1 for EPA Observation Forms[10] (introduced in 2020). However, we identified frequent deviations from the suggested electronic form design within our own programs and institutions. Poor form design could make form completion more time-consuming and frustrating,[11] and variability in technology likely produce differences in how CBD is experienced across program sites. While multi-site evaluations of CBD have occurred, they have focused on high-level outcomes such as the experience of and impact on residents,[3] fidelity of implementation at the program level,[12] and large-scale strategy for reforms in education.[13] More detailed aspects, such as institution-specific variations in form design and educational outcomes, have received comparatively less attention. In this study, we sought to quantify the variability in local form design of CBD EPA assessment forms across 14 Canadian academic institutions.

Methods

This study utilized a multi-centre, cross-sectional comparative methodology to identify differences in EPA form design between academic institutions. We selected the RCPSC emergency medicine (EM) 'Core of Discipline' (Core) EPA 1 as the model EPA for the comparative analysis (Figure 1).

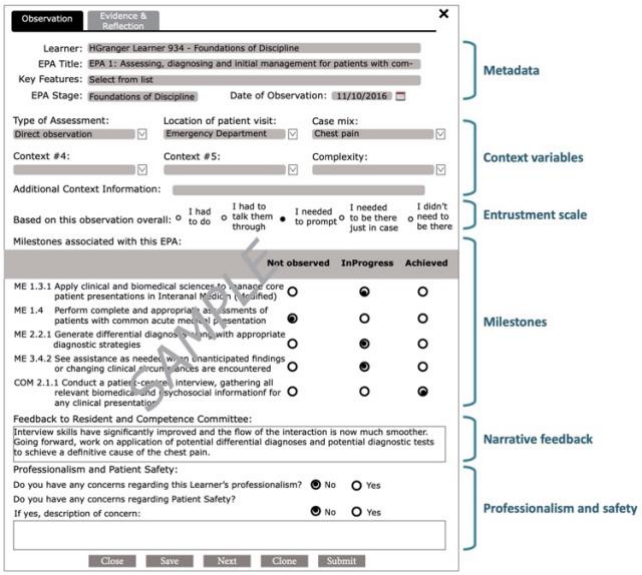

Figure 1. RCPSC EPA 1 observation template

This EPA was selected to enhance the generalizability of our findings to other specialties, as its focus on the resuscitation of a patient is similar to that of EPAs used by many medical and surgical disciplines.

In March 2023, we contacted the program directors or faculty members on Competence Committees of each of the 14 institutions with RCPSC emergency medicine residency programs (Dalhousie University, Université Laval, McGill University, Université de Montréal, Queen's University, University of Ottawa, University of Toronto, McMaster University, Western University, University of Manitoba, University of Saskatchewan, University of Alberta, University of Calgary, University of British Columbia) via email and collected C1 EPA assessment forms with annotations explaining their functionality. Specifically, we requested a screen capture of the form as a whole and additional screen captures of any particularly distinct parts of the form, with any identifying information of learners or assessors redacted. In the annotations of their forms, we asked them to comment on:

1. The learning management system used by their institution
2. The items that were prepopulated, the presence of any 'form logic' (e.g., a component of the form can

only be filled out if a particular response is selected for a preceding question), and any forcing functions and form validation (e.g., *x* must be filled out to be able to submit the form).

The collected forms were divided into six sections for analysis as per the RCPSC Technical Guide 1[10] and Observation Template 1:[8] metadata, context variables, milestones, entrustment score, narrative feedback, and professionalism and safety. These are referred to as "Technical Guide" and "Observation Template" throughout the remainder of the text. Table 1 defines each of these sections and provides examples of the information the sections were meant to gather.

Two authors (RI, BT) reviewed each form and compared the forms' designs against each other and to the guidance provided by the RCPSC. The authors quantified differences in form content such as the number and wording of components and recorded any other differences in form design. Aspects for comparison were identified and defined through the review process. If any annotations were unclear, the individual who provided the form from that site was contacted by email for clarification.

Table 1. Formatting and content recommendations for each section of EPA assessment forms as described by the RCPSC CBD Observation Template[8], EM EPA Guide[9], and CBD Technical Guide[10].

Form Section	Information Collected	RCPSC Guidance for Formatting
Metadata	Helps users understand, manage, and organize data, making it easier to search, retrieve, and analyze information. Examples of metadata include the date the assessment occurred and/or was submitted, the assessor filling out the form, and the learner being assessed.	*Guidance on the collection of metadata not provided.*
Context variables	Describe the clinical scenario of the observation. Tracked by competence committee to ensure variety in clinical exposure that mirror independent practice. Intended to orient assessor to provide feedback targeted to the environment in which the case took place. Enables a more manageable number of EPAs for each discipline by allowing a single EPA to be completed for multiple contexts. Context variables differ for each EPA.	Four context variables included in the Observation Template 1:[8] • Type of assessment: direct; other criterion not specified • Location of patient visit: emergency department; other criterion not specified • Case mix: chest pain; other criterion not specified • Complexity: criterion not specified Four context variables included in EPA EM Guide:[9] • Patient demographic: infant; child; youth; adult; senior • Setting: emergency department; simulation • Complexity: simple; complex patient characteristic; complex clinical characteristic; complex environmental characteristic • Clinical presentation: shock; cardiorespiratory arrest; respiratory failure; severe sepsis; other Textbox for additional context information.[8,9]
Milestones	Drawn from the CanMEDS Physician Competency Framework and attached to each EPA with the goal of outlining the skill set required to perform the EPA, inform curriculum development, and provide a scaffold for coaching and assessment. Specialty committees select a small number (typically < 8) of these milestones to focus the observation on the key aspects of that task and to guide feedback.	13 milestones specified;[9] however, only "prioritized" milestones need to be listed.[10] "Prioritized" milestones must be bolded.[10] Each milestone does not have to be scored.[10] Wording of milestones must be exactly what was decided on by the specialty committee.[10] Acceptable to have single scale on form for entire EPA.[10]
Entrustment Scale	Used to rate the trainee's performance of the EPA activities, providing feedback for both the trainee and competence committee. Anchors are intended to align with the judgement of the observer/assessor regarding the degree of support that the trainee required to perform the EPA.	Recommendation is O-SCORE.[9] Any rating scale with entrustment anchors may be utilized; '5-point' scale does not need to be used.[10] Use of alternative entrustment scale will not be perceived as a weakness by accreditation surveyors (so long as provided the alternate scale facilitates the appropriate assessment conversations as intended by CBD).[10]
Narrative Feedback	Comments about residents' performance that provide details Ideally specific, actionable, and timely to inform trainee coaching.	One textbox with prompt "Feedback to Resident and Competence Committee."[8] Must include a space for narrative comments.[10]
Professionalism and Safety	Concerns or 'red flags' related to trainee professionalism and patient safety.	Two questions included with selection options 'yes' or 'no' for both:[8] • Do you have any concerns regarding this Learner's professionalism? • Do you have any concerns regarding patient safety? Textbox with prompt "If yes, description of concern"[8]

Results

We collected 16 forms from the 14 active RCPSC EM residency programs (100%) in March 2023. Two programs were transitioning between to new form versions and provided both their previous and new form. The four programs using the RCPSC ePortfolio software had identical forms, so annotations were not requested from three programs. Annotated forms were provided by program directors (*n* = 6), assistant/associate program directors (*n* = 2), or frontline faculty members familiar with the assessment forms (*n* = 4). The differences in content and format between each of the forms is outlined in Supplemental File 1.

The comparative analysis was organized into eight categories based on the six sections outlined in Table 1 and visualized in Figure 1, with two extra aspects: navigation of content (comparing the approaches used to navigate through the form and select responses) and additional elements (describing unique additions to the EPA assessment forms)**.**

The ePortfolio (*n* = 4, 25.0%), CBME.med (*n* = 1, 6.3%), Elentra (*n* = 7, 43.8%), MedSIS (*n* = 1, 6.3%), McMAP (*n* = 1, 6.3%), and One45 (*n* = 2, 12.5%) learning management systems were used across the 14 programs. Significant differences were found in the design of assessment forms, both across and within the same learning management system.

Metadata

Metadata provides information related to the circumstances around the assessment such as the title, learner, date of assessment, and a short description of the EPA. It was generally consistent across forms, with most forms automatically attaching details regarding the assessor, learner, date of assessment, EPA title and description. One form (6.3%) included an external link to the EM EPA Guide[9] under the EPA title. Four forms (25.0%) also contained images of the resident being assessed and/or the assessor with details such as level of training and their email (Supplemental File 2, Image 1).

Context variables

Context variables request information about the details specific to the clinical case. The Observation Template[8] and EM EPA Guide[9] collectively described five context variables to be included on the C1 form related to the setting of the observation, demographics of the patient, complexity of the case, and the clinical presentation, and the type of assessment. The "Type of Assessment" context variable was one of four included in the Observation Template.[8] It was meant to specify whether the learner was directly or indirectly observed. The EM EPA Guide[9] later clarified that all assessments of the C1 EPA needed to be directly observed and did not include it in the list of context variables to be on the form. It was replaced with "Patient Demographic." The EM EPA Guide[9] detailed that either a supervising physician or a physician at TTP (i.e., at the Transition to Practice stage of CBD stage) could complete the EPA assessment, but did not create a separate context variable on the form to track this.

The number of context variables included in the other forms ranged from four to seven, with most forms including four (62.5%). The context variables in four forms (25.0%) totaled five, with either the inclusion of a context variable related to the type of assessment (i.e., direct or indirect) or the type of assessor (i.e., supervising physician or resident at TTP). One form (6.3%) included both, resulting in a total of six context variables. A seventh context variable was included in one form (6.3%) to indicate the rotation the resident was on when it was completed. This form replaced a question about the setting of the observations with the site/hospital that the observation took place. Similar to the Observation Template[8] and EM EPA Guide,[9] the majority of the forms (*n* = 9, 56.3%) contained an additional narrative textbox along with the context variables selection that allowed users to elaborate on the clinical case. Completion of this textbox was optional across all of the forms.

The selection options for the "Age" demographic varied, with some forms containing a drop-down numerical list of ages and others having categories such as "infant, child, adult, elderly." The wording of the selection options for the other context variables was in keeping with RCPSC guidance.

Milestones

The RCPSC EM EPA Guide[9] described 13 milestones for the EM C1 EPA, each corresponding to the CanMEDS physician competency framework.[14] In one form (6.3%), the 13 milestones were summarized and re-worded into 11 items referred to as ''Dimensions." Conversely, there were three additional milestones included on one form (6.3%), resulting in 16 total milestones that could be rated. Another (6.3%) kept all 13 milestones but shortened their descriptions. These three forms deviated from RCPSC guidance provided in the Technical Guide[10] which instructed that the milestones be written with the exact wording specified in the EM EPA Guide.[9] Some forms (*n* = 3, 18.8%) mapped each milestone to their associated

CanMEDS role,[14] for example, "COM 3.1" to refer to "Communicator" in the framework.

There was variability in the criterion used to rate milestones in comparison to the Observation Template[8] which detailed a 3-point scale (Not Observed, In Progress, or Achieved); other forms utilized scales with two, three, four, or six criteria (Table 2). One form (6.3%) did not include a rating scale and instead listed the milestones checkboxes with the option of selecting the ones relevant to the assessment or none at all (Supplemental File 2, Image 2). These variations were acceptable based on the guidance provided in Technical Guide[10] which states that each milestone does not have to be rated (Table 1).

Table 2. Variations in milestone scale criterion across 16 EM C1 EPA assessment form designs

2-point scale (*n* = 1)	In Progress	Achieved				
3-point scale (*n* = 11)	Not Observed	In Progress	Achieved			
4-point scale (*n* = 1)	N/A	Not Done	Done, but needs attention	Done		
6-point scale (*n* = 2)	Not Observed	I had to do	I had to talk them through	I had to prompt them from time to time	I needed to be there in the room just in case	I did not need to be there
	No rating	Intervention: Requires others' action for completion	Direction: Requires supervision and others' guidance for completion	Support: Requires minimal supervision or guidance for completion	Autonomy: Does not require guidance or supervision for completion	Excellence: Demonstrates excellence; is a good role model
No scale (*n* = 1)	Select all that apply					

One form (6.3%) emphasized the milestones by requesting narrative comments about each, and another (6.3%) by requiring that they all be rated prior to form submission. Milestones were de-emphasized in some forms (*n = 5*, 31.3%) by allowing them to be collapsed/expanded through clicking a button that toggled their visibility on the form (i.e., visible or not visible) and/or by making these textboxes optional. As per the Technical Guide,[10] milestones must be visible but do not have to be rated or individually commented on and only eight milestones need to be "emphasized" (i.e., focused on) in the EPA (Table 1).

Entrustment Scale

As outlined in Table 3, all programs used a five-point entrustment scale and all but one (6.3%) used some variation of the 5-point Ottawa Surgical Competency Operating Room Evaluation (O-SCORE).[15,16] Two forms explicitly numbered the criterion from 1-5. The scale on the form that did not use the O-SCORE had a vertical orientation, was colorized with an ombre gradient (from darkest at the top to lightest at the bottom), and had a horizontal line that divided the third and fourth criterion. RCPSC guidance recommends the O-SCORE but allows for the use of an alternative entrustment score,[15] however there are no suggestions provided on other modifications. One (6.3%) other scale also used a red, orange, blue, yellow, and green color sequence corresponding to each of the five criteria, with red as the color for the lowest entrustment score and green for the highest. Images of these variations can be found in Supplemental File 2, Images 3 and 4.

Narrative feedback

The number of textboxes for narrative feedback ranged from one to three with variable prompts (Table 4). Every form had uniquely worded prompts with varying degrees of which/how many responses were required to be completed to submit the form (mandatory prompts are denoted with an asterisk in Table 3). The Observation Template[8] presents a single narrative feedback textbox with the prompt "Feedback to Resident and Competence Committee." The nature of the feedback prompts was guided by three types of framing to focus faculty attention on the assessment encounter. Six forms (37.5%) contained one feedback textbox with a prompt focused generically on feedback to the resident and competence committee, which was congruent with RCPSC guidance. Seven forms (43.8%) contained two feedback textboxes with the prompts for six of these forms focusing on strengths and areas for improvement. Three forms (18.8%) had three feedback textboxes, while different in focus, aimed to elicit more detailed feedback. An outlier to this framing was a form (6.3%) that contained two feedback textboxes with prompts to gather feedback specific to the EPA and on global performance.

Table 3. Entrustment scales used to rate EPA performance across 16 EM C1 EPA assessment forms

Ottawa O-SCORE (*n* = 4)	I had to do	I had to talk them through	I needed to prompt them from time to time	I needed to be there in the room just in case	I didn't need to be there
Variation 1 (*n* = 6)	I had to do	I had to talk them through	I needed to prompt them	I needed to be there just in case	I did not need to be there
Variation 2 (*n* = 1)	I had to do	I had to talk them through	I needed to prompt them occasionally	I needed to be there just in case	I didn't need to be there
Variation 3 (*n* = 1)	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there (in theory)
Variation 4 (*n* = 1)	I had to do	I had to talk them through	I needed to prompt	I had to provide minor direction	I did not need to provide direction for safe and independent care
Variation 5 (*n* = 1)	1 - I had to do	2- I had to talk them through	3 - I needed to prompt them occasionally	4 - I needed to be there in the room just in case	5 - I did not need to be there
Variation 6 (*n* = 1)	1 - I had to do	2 - I had to talk the resident through it	3 - I had to direct the resident from time to time	4 - I had to be there just in case	5 - I didn't need to be there (or was in the room but didn't need to be)
Variation 7 (*n* = 1)	Intervention I had to complete some/all of the EPA task/activity	Direction I needed to provide major redirection	Support I provided minor redirection to ensure safety/minimal competence	Competent I didn't need to act (coaching aside) for safety/minimal consequences	Proficient I viewed this performance as exemplary

**Shaded boxes indicate prompts from the original O-SCORE. All programs used variations of the third ("I had to prompt them from time to time") and fourth ("I needed to be in the room just in case") O-SCORE prompts.*

Table 4. Number of narrative textboxes and associated prompts used across 16 EM C1 EPA assessment forms

1 Textbox (2 variations across 16 forms)	Feedback to Resident and Competence Committee.*~		
	Briefly describe what you observed and suggest 1 thing the resident could work on to progress to safe and independent practice for this EPA.		
2 Textboxes (7 variations across 16 forms)	Comments (mandatory)*		Next Steps
	Strengths		Actions or areas for improvement
	Comments specific to this EPA*		Comments on residents' performance on shift today?*
	COMMENT: on something the resident did well.*		COMMENT: on something the resident needs to improve.*
	Please describe what the resident is doing well and why they should keep doing it.*		Please describe something that the resident can do to improve for next time.*
	Please provide comments about how the resident performed on this specific EPA.*		Please provide evidence to support your global rating of the resident. Also, consider writing one suggestion for improvement.*
	Please provide comments about how the resident performed on THIS specific EPA.*		Please provide evidence to support you global rating of the resident. Also consider writing one suggestion for improvement.*
3 Textboxes (3 variations across 16 forms)	Continue…	Consider…	Feedback to Resident and Competence Committee
	Comments (mandatory)*	Global Feedback	Next Steps
	Context Briefly describe the case/ content.*	Performance Please describe your observation of what the resident did.*	Recommendations/Reinforcement Please describe what the resident should do in future cases +/- what the resident did well and should continue doing.*

Asterisks (*) indicates a mandatory textbox.

Professionalism and safety

All forms except three (18.8%) included questions related to professional and safety concerns. In keeping with the Observation Template,[8] the majority of (the) forms included two questions for this section, with one addressing professionalism and the other addressing patient safety. These questions were often pre-selected to 'No' (*n* = 10, 62.5%). In some forms (*n* = 6), selecting 'Yes' in response to either question caused a mandatory textbox to appear below the question which required completion before the form could be submitted. However, there were also forms where this box did not appear and/or was not mandatory.

Navigation of content

There was variability in the function and selection format of responses for context variables, milestones, narrative feedback textboxes, and questions for professionalism and safety. Some forms required selection of certain items before the form could be submitted, while other forms pre-populated responses. Variability in the requirement for narrative feedback is described in Table 4. The degree of completion also varied for milestones, with one form not

requiring ratings for milestones at all, a second having milestone ratings pre-selected, and a third populating milestone ratings based on the entrustment score that is selected (for example, selecting the highest level of entrustment on the entrustment scale would populate "Achieved" on the milestone scale). All forms required the context variables to be selected prior to form submission.

Three types of selection controls were utilized across the forms: drop-down menus, checkboxes to "select all that apply," and radio buttons (Supplemental File 2). The criterion in the entrustment scale, milestones, and professionalism and safety concerns were selected exclusively through radio buttons. Context variable questions generally used drop-down menus in keeping with the Observation Template.[10] Apart from this, two variations were the use of checkboxes in the 'complexity' context variable which allowed more than one criterion to be selected and another in which context variables criterion were selected through buttons (as opposed to drop-down menus) (Supplemental File 2, Image 5). The latter approach required only a single click to select a response, halving the number of clicks required to complete the context variable section.

The date on one form was selected by clicking on a calendar icon which then expanded to a date picker. These variations led to significant differences in the number of clicks required to complete an assessment between forms with a minimum of 12 clicks, and up to a maximum of 47 clicks (Figure 2). None of the RCPSC guidance documents provided directions regarding the navigation of content within EPA assessment forms.

Additional elements

Some forms included additional elements at the end of the form that were not included in the Observation Template.[8] These included:

- A textbox requesting generalized feedback on the form: "*Feedback about this form?*"
- A textbox to describe any 'red flags' that may have occurred during the observation: "*Do you have reasons to flag this assessment?*"
- A global performance rating scale and corresponding feedback textbox to elicit information unrelated to the specific objectives of the EPA
- A question inquiring about in-person feedback: *"Did you have an opportunity to meet with this trainee to discuss their performance?"*

Specific learning management systems had unique capabilities. For example, the Elentra/Entrada system allowed assessment forms to be partially completed by the resident before they were forwarded to their assessor while the ePortfolio system gave assessors the option to upload additional documents to support the assessment.

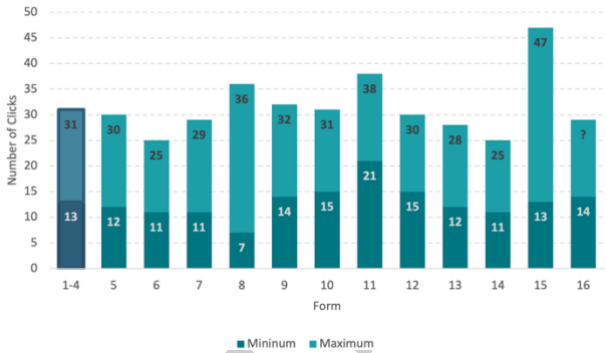

*Figure 2. Minimum and maximum number of clicks require to complete 16 EM C1 EPA assessment forms *Dark blue bar represents RC ePortfolio*

Discussion

This study sought to find and quantify variability in the design of EPA assessment forms across Canadian emergency medicine residency programs. Our comparative analysis identified substantial variability. Some of the most notable discrepancies included:

- Context variables: selection format and the number of variables included
- Milestones: choice of wording, and degree of interaction, with one form not requiring any milestone ratings while another required response to a six-point scale for each milestone
- Entrustment scales: wording of criteria, presentation with colour grading, and the use of a line to divide criteria denoting low and high entrustment
- Narrative feedback textboxes: both their wording of prompts and the number of required narrative responses
- Inclusion of additional components: global rating scale and textbox for global feedback
- The number of clicks required to complete the form: minimum of 12 and maximum of 47

It is reasonable to assume that some of the characteristics of these form components increase (or decrease) assessment burden. For example, ambiguously worded prompts, forcing engagement with uninformative components, and a high number of required clicks can increase assessors' extraneous cognitive load and increase

time for form completion. This could reduce the number of assessments completed, potentially impacting resident progression and development, the validity of assessment decisions, and the experience of CBD. In addition, variations in design can cause inconsistencies in how assessments are being completed at different residency programs, resulting in residents at different sites experiencing different assessment systems despite being at the same level of training. Developing and adopting best practices in form design across sites could limit the negative impact of poor design and increase consistency in the experience of the national assessment program. However, too much standardization could reduce flexibility in a way that impedes programs' adoption of changes that are necessary for particular learning environments.

For many programs, one of the main changes in the implementation of CBD was the addition of electronic EPA assessment forms that had to be completed frequently.[17] As a result, many frontline faculty and residents may conflate negative experiences navigating through EPA assessment forms with CBD, making efficient assessment form design particularly important to the rollout of CBD.

CBD has been criticized for its rigidity,[2–5] but institutions have taken advantage of the flexibility afforded to EPA assessment form design to modify their own forms. Unfortunately, as most residents and faculty only use the version of EPA forms developed for their own institution, it is difficult for variations to be identified or best/worst practices in form design to be established and transferred between institutions. This variability could allow future studies to be conducted that compare the impact of various approaches to form design. For example, adding additional narrative textboxes may increase the time and number of clicks required to complete a form, but if the resulting feedback is of greater quality this may be a reasonable tradeoff.[18]

We believe that one of the key messages that medical education leaders should take from this research is that the design of their EPA forms can be flexible even within the same learning management system, and usability testing can improve these forms. Every effort should be made to optimize form design for the local context with heavy consultation from key frontline users. Recognizing budgetary constraints, ideally forms would be subject to regular iteration with emphasis on end-user feedback through defined and explicit pilot testing rather than implementing with a sense of finality.[3,17] Design-based research processes that have been used to develop tools to both present CBD assessment data and identify generalizable user needs could provide a framework for the scholarly development and iteration of assessment forms.[19–22]

Limitations and strengths

Our study has several limitations. First, our analysis was completed on EPA forms in use during March 2023. Some programs modified their EPA forms when the second version of the EM EPA Guide was released for July 1, 2024. However, given the lack of research or guidance on EPA form design, we do not anticipate that the forms would have been modified in a way that reduced the variability identified in the study. We also did not analyze the form used at the recently established medical school at Toronto Metropolitan University (TMU), however, this form would have contributed little to this analysis as its design was informed by the work outlined in this manuscript.

Secondly, we only compared versions of the C1 EPA assessment form for emergency medicine in our study. As we identified substantial variability within a single form, we anticipate that the inclusion of additional emergency medicine forms and those from other specialties would identify more, rather than less, variability in form design. As the design of forms at an institution is generally determined based upon the limitations of its Learning Management System and/or decisions regarding form design that are made centrally, we believe that our results are generalizable to other disciplines.

Thirdly, our study design was purely comparative and did not allow for the exploration of the implications of the variability between forms on assessment burden or identify the most efficient and effective EPA form design. We are in the process of investigating this through a follow up study that interviews residents, faculty, and program directors to better understand the implications of different design choices.

One of the key contributions of this study is that it highlights the degree of variability between CBD programs, which has not been explored until now. Previous studies have compared the perspectives of participants at multiple sites, but minimally investigated differences in implementation as specific as form design. We anticipate that our findings may lead to further exploration of the impact of this variability and its impact on the impact and implementation of CBD.

Conclusion

Despite the existence of a standard EPA form outlined by RCPSC through the Observation Template, Technical

Guide,[10] and EM EPA Guide,[9] there was considerable variability in the design of all components of the EM C1 EPA assessment form found by our study. Most residents and educators complete EPAs within single institutions and are likely unaware that this variability could be impacting their experience of CBD. We recommend that programs review the design variations described in this study and consider how their assessment forms could be modified with the goal of reducing assessment burden and improving assessment quality. Further study investigating the positive and negative impacts of assessment form design on assessment burden and quality is warranted to characterize and inform an ideal assessment for design.

Conflicts of Interest: Brent Thoma, Andrew Hall, and Adam Szulewski are contracted to conduct administrative and educational work for the Royal College of Physicians and Surgeons of Canada. Brent Thoma is an editor for the CMEJ. He adhered to the CMEJ policy on editors as authors.
Funding/Support: Romaisa Ismaeel received a research grant through the Office of the Vice-Dean Research from the University of Saskatchewan, Saskatoon, Canada.
Acknowledgements: The authors wish to acknowledge Mann Parikh, a student at McMaster University, for their assistance in reviewing and summarizing the collected forms. The authors also wish to acknowledge Maria Azam, a Master of Design in Interaction Design graduate at Emily Carr University of Art + Design, Vancouver, Canada, for reviewing the design terminology used in this manuscript.
Previous presentations: The preliminary results of this research were presented at the Canadian Association of Emergency Physicians (CAEP) conference, June 11, 2024, Saskatoon, Saskatchewan, Canada; and the International Conference on Residency Education (ICRE), September 20, 2024; Ottawa, Ontario, Canada.
Authorship: All authors have each signed their own attestation statement that they meet the requirements of authors. The lead author, in addition to their own attestation as an author, has signed an attestation that all authors listed on this paper meet the requirements for authors.
Edited by: Anita Acai (senior section editor); Marcel D'Eon (editor-in-chief)

References

1. Frank JR, Karpinski J, Sherbino J, et al. Competence by design: a transformational national model of time-variable competency-based postgraduate medical education. *Perspect Med Educ.* 2024;13(1):201–23. https://doi.org/10.5334/pme.1096
2. van Mil S. *Intended and unintended consequences: the impact of competence by design implementation on the experiences of trainees in a pediatric residency program* [Thesis]. 2024 Available from: https://macsphere.mcmaster.ca/items/478cf026-ea43-4387-bd86-f51670cc0768 [Accessed on Apr 8, 2025].
3. Ott MC, Pack R, Cristancho S, et al. "The most crushing thing": understanding resident assessment burden in a competency-based curriculum. *J Grad Med Educ*. 2022 Oct 14;14(5):583–92. https://doi.org/10.4300/JGME-D-22-00050.1
4. Szulewski A, Braund H, Dagnone DJ, et al. The assessment burden in competency-based medical education: how programs are adapting. *Acad Med*. 2023 Nov;98(11):1261. https://doi.org/10.1097/ACM.0000000000005305
5. Royal College of Physicians and Surgeons of Canada. *Competence by design: resident physician pulse check*. Royal College of Physicians and Surgeons of Canada; Sept 2022. Available from: https://www.royalcollege.ca/content/dam/document/standards-and-accreditation/2022-resident-pulse-check-final-report-e.pdf [Accessed on Nov 17, 2025].
6. Tierney AA, Rosner BI. Clinical assessment of residents: a survey of clinician educators regarding resident assessment burden and modifiable factors. *J Grad Med Educ*. 2023 Feb 1;15(1):92–7. https://doi.org/10.4300/JGME-D-22-00188.1
7. Sweller J, van Merrienboer JJG, Paas FGWC. Cognitive architecture and instructional design. *Educ Psychol Rev.* Sept 1998;10(3):251–96. https://doi.org/10.1023/A:1022193728205
8. Royal College of Physicians and Surgeons of Canada. *Competence by design observation template 1.* Royal College of Physicians and Surgeons of Canada; 2017. Available from: https://www.royalcollege.ca/content/dam/documents/ibd/shared/cbd-observation-templates-e.pdf [Accessed on Nov 17, 2025].
9. Royal College of Physicians and Surgeons of Canada. *Entrustable professional activity guide: emergency medicine.* Royal College of Physicians and Surgeons of Canada; 2018. 2018. Available from: https://www.royalcollege.ca/content/dam/documents/ibd/shared/epa-guide-emergency-med-e.pdf [Accessed on Nov 17, 2025].
10. Royal College of Physicians and Surgeons of Canada. *Competence by design technical guide 1: EPA observation forms.* Royal College of Physicians and Surgeons of Canada; 2020. Available from: https://cumming.ucalgary.ca/sites/default/files/teams/6/CBD/cbd-technical-guide-1-epa-observation-forms-e%20(1).pdf. [Accessed on Nov 17, 2025].
11. Ismaeel R, Riggs J, Hall A, et al. Assessment burden by design: Exploring variability in assessment forms within competence by design. *Int Conf Resid Educ ICRE.* 2024;
12. Hall AK, Rich J, Dagnone JD, et al. It's a marathon, not a sprint: rapid evaluation of competency-based medical education program implementation. *Acad Med*. 2020 May;95(5):786–93. https://doi.org/10.1097/ACM.0000000000003040
13. Hall AK, Oswald A, Frank JR, et al. Evaluating competence by design as a large system change initiative: readiness, fidelity, and outcomes. *Perspect Med Educ*. 2024 Feb 6;13(1):95–107. https://doi.org/10.5334/pme.962
14. Frank JR, Snell L, Sherbino J, editors. *CanMEDS 2015 physician competency framework.* Ottawa: Royal College of Physicians and Surgeons of Canada; 2015.
15. 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 Oct;87(10):1401–7. https://doi.org/10.1097/ACM.0b013e3182677805
16. MacEwan MJ, Dudek NL, Wood TJ, Gofton WT. Continued validation of the O-SCORE (Ottawa Surgical Competency Operating Room Evaluation): use on the simulated environment. *Teach Learn Med*. 2024;16. http://doi.org/10.1080/10401334.2015.1107483

17. Ott M, Apramian T, Cristancho S, Roth K. Unintended consequences of technology in competency-based education: a qualitative study of lessons learned in an OtoHNS program. *J Otolaryngol - Head Neck Surg*. 2016 Jan;52(1):s40463-023-00649–2. http://doi.org/10.1186/s40463-023-00649-2
18. Woods R, Bandi V, Thoma B, et al. E-Mail nudges, faculty development workshops and template changes…what impacts the quality of supervisor narrative comments in competency based medical education? *Int Conf Resid Educ ICRE Program Eval Summit* 2024.
19. Thoma B, Bandi V, Carey R, et al. Developing a dashboard to meet competence committee needs: a design-based research project. *Can Med Educ J*. 2020 Jan 7 https://doi.org/10.36834/cmej.68903
20. Carey R, Wilson G, Bandi V, et al. Developing a dashboard to meet the needs of residents in a competency-based training program: A design-based research project. *Can Med Educ J.* 2020 Sept 14 https://doi.org/10.36834/cmej.69682
21. Yilmaz Y, Carey R, Chan T, et al. Developing a dashboard for faculty development in competency-based training programs: a design-based research project. *Can Med Educ J.* 2021 June 30 https://doi.org/10.36834/cmej.72067
22. Yilmaz Y, Carey R, Chan T, et al. Developing a dashboard for program evaluation in competency-based training programs: a design-based research project. *Can Med Educ J.* 2022 Apr 27. https://doi.org/10.36834/cmej.73554