# How are Entrustable Professional Activities Assessed within Simulation and Clinical Settings?

Hajar Al-Mughairi, MD
Department of Experimental Surgery
McGill University, Montreal, Quebec, Canada
April 2021

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science in Experimental Surgery- Surgical Education.

# Table of Contents

**Section 4: Thesis conclusion**

## Abstract 1: Assessment of Entrustable Professional Activities within Simulation and Clinical Settings: A Scoping Review

**Introduction:**

Entrustable professional activities (EPAs) are a fundamental unit in the competency-based education model (CBME). Assessment of EPAs is a crucial step in the implementation process of CBME, which remains challenging due to a lack of consistency in meaningful assessment methods. There is great variability in the assessment approaches used to evaluate EPAs among residency training programs, which might influence successful CBME implementation. This scoping review seeks to explore how EPAs are being assessed and implemented within residency training programs.

**Methods:**

We conducted a review of MEDLINE, PubMed, Embase, PsychINFO databases, and Google Scholar by searching the full-text articles and abstracts published between January 2000 and January 2020. Additional studies were identified through reference lists. We included the publications addressing the EPA assessments as defined by the Royal College which were written or translated to English only.

**Results:**

A total of 635 articles were screened for eligibility, and ultimately, twenty-eight full-text articles and six abstracts were enrolled in the final review. 122 EPAs were identified, with 86 (70.5%) EPAs corresponding to non-technical skill alone, 26 (21.3%) technical skill alone, and 10 (10.2%) combined technical and non-technical skills.

Out of 122 EPAs, 86 (69%) were assessed clinically and 38 (31%) were assessed in a simulated environment. In simulation assessment, a standardized patient was the commonly

utilized simulation modality, and checklists were the most used assessment tool, either alone or with other tools. However, the Entrustability scale was the dominant assessment tool of EPA in the clinical setting (77%).

Seven rater types were identified: attending physician, trained rater, senior resident, peer, self, nurse, and patient. Moreover, around half of EPAs (52%) were assessed by the attending physicians.

**Conclusion:**

This review identified that EPAs are being assessed diversely, in multiple settings, by raters at various experience levels, and using different assessment tools. As the methods of assessment are not consistent across institutions, future research should seek to establish standards for the most efficient and effective way to assess the EPAs, which are mandatory by the medical accreditation bodies.

## Résumé 1 : Évaluation d'activités professionnelles dignes de confiance dans des contextes de simulation et cliniques : un examen de la portée

**Introduction:**

Les activités professionnelles confiables (APE) sont une unité fondamentale dans le modèle d'éducation par compétences (CBME). L'évaluation des APE est une étape cruciale dans le processus de mise en œuvre de la CBME, qui reste difficile en raison d'un manque de cohérence dans les méthodes d'évaluation significatives. Il existe une grande variabilité dans les approches d'évaluation utilisées pour évaluer les APE parmi les programmes de formation en résidence, ce qui pourrait influencer la réussite de la mise en œuvre de la CBME. Cet examen de la portée vise à explorer comment les APE sont évalués et mis en œuvre dans les programmes de formation en résidence.

**Méthodes :**

Nous avons effectué une revue des bases de données MEDLINE, PubMed, Embase, PsychINFO et Google Scholar en recherchant les articles en texte intégral et les résumés publiés entre janvier 2000 et janvier 2020. Des études supplémentaires ont été identifiées grâce à des listes de références. Nous avons inclus les publications traitant des évaluations EPA telles que définies par le Collège royal qui ont été écrites ou traduites en anglais uniquement.

**Résultats:**

Au total, 635 articles ont été sélectionnés pour leur éligibilité et, finalement, vingt-huit articles en texte intégral et six résumés ont été inscrits dans l'examen final. 122 EPA ont été identifiés, avec 86 (70,5%) EPA correspondant à des compétences non techniques seules, 26 (21,3%) à des compétences techniques seules et 10 (10,2%) à des compétences techniques et non techniques combinées.

Sur 122 EPA, 86 (69 %) ont été évalués cliniquement et 38 (31 %) ont été évalués dans un environnement simulé. Dans l'évaluation par simulation, un patient standardisé était la modalité de simulation couramment utilisée, et les listes de contrôle étaient l'outil d'évaluation le plus utilisé, seul ou avec d'autres outils. Cependant, l'échelle de confiance était l'outil d'évaluation dominant de l'EPA en milieu clinique (77 %).

Sept types d'évaluateurs ont été identifiés : médecin traitant, évaluateur qualifié, résident senior, pair, soi-même, infirmière et patient. Par ailleurs, environ la moitié des EPA (52 %) ont été évaluées par les médecins traitants.

**Conclusion:**

Cet examen a identifié que les EPA sont évalués de manière diverse, dans de multiples contextes, par des évaluateurs à divers niveaux d'expérience et en utilisant différents outils d'évaluation. Comme les méthodes d'évaluation ne sont pas cohérentes entre les établissements, les recherches futures devraient chercher à établir des normes pour la manière la plus efficace et la plus efficace d'évaluer les EPA, qui sont obligatoires par les organismes d'accréditation médicale.

## Abstract 2: Assessment of EPAs within Surgical Foundations (SF) Program at McGill University, Montreal, Canada


**Introduction:**

Competency by design (CBD) has been introduced by the Royal College of Physicians and Surgeons of Canada (RCPSC) as a novel educational model for residency training programs. Within the CBD competence continuum, Surgical Foundations (SF) program is a special training program where the residents from nine surgical specialties learn the basic surgical knowledge and skills. CBD focuses on mastery of competencies and uses entrustable professional activities (EPAs) as the organizing assessment framework. The purpose of this study is to explore how EPAs are assessed within the SF program, during both the Foundation of discipline and the Transition to discipline stage of training.

**Methods:**

Thirteen residents enrolled in SF program from six surgical subspecialties participated in the study after obtaining informed consent to access their EPA assessment data. All the assessment data (resident's demographic, EPAs, assessment and rater- related characteristics) was provided by a data analyst in the medical education system at McGill University.

**Results:**

A total of 595 EPA assessment forms were completed between July 2019 and August 2020. The vast majority of the received assessment data (95%) was conducted in a clinical setting. Sixteen EPAs assessed, with 7 corresponding to Transition to discipline and 9 to Foundation of discipline stage. There was great variability in the number of completed assessments per resident, where the lowest was 14 and the highest was 65. Three assessment forms were utilized to assess the EPAs in SF (EPA observation, procedural competencies and

multi-source feedback). About 70% (411) of the assessment forms were completed by residents and 27% (165) by the attending physicians. The overall assessment using the entrustment scale ranged from level 2 (I had to talk them through) to level 5 (I did not need to be involved) where the level 4 (I needed to be there just in case) was the most commonly assigned entrustment level. Around 85% of EPA assessments were achieved as their assigned entrustment levels were level 4 and 5. The EPA 2.1b: Critically Ill Surgical and EPA 2.5a: Fundamental Surgical Procedures were the most challenging EPAs and were frequently assigned low ratings.

**Conclusion:**

Within thirteen months, 13 residents in the SF program completed 595 EPA assessment forms related to 16 EPAs identified by RCPSC. There was an apparent variability in completed EPA assessments between residents and sub-specialties where the residents marked as the most prevalent raters. Moreover, two assessment tools were utilized: the Entrustability scale and the Global Rating Scale. Further study could be conducted to investigate the causes behind the variability and the impact of residents being the most common assessors. This study conducted among a small group and can provide minor trends which would be interesting to follow in a larger study. Moreover, it is worth exploring how EPAs are assessed within simulation and correlate with the results of clinical-based assessment.

## Résumé 2 :Évaluation des EPA dans le cadre du programme Fondements chirurgicaux (SF) à l'Université McGill, Montréal, Canada


**<u>Introduction:</u>**

La compétence dès la conception (CDB) a été introduite par le Collège royal des médecins et chirurgiens du Canada (CRMCC) en tant que nouveau modèle éducatif pour les programmes de résidence. Dans le continuum de compétences CBD, le programme Surgical Foundations (SF) est un programme de formation spécial où les résidents de neuf spécialités chirurgicales apprennent les connaissances et compétences chirurgicales de base. CBD se concentre sur la maîtrise des compétences et utilise les activités professionnelles confiables (APE) comme cadre d'évaluation organisateur. Le but de cette étude est d'explorer comment les EPA sont évalués au sein du programme SF, à la fois au cours de l'étape Fondation de la discipline et de la transition vers la discipline de la formation.

**<u>Méthodes :</u>**

Treize résidents inscrits au programme SF de six surspécialités chirurgicales ont participé à l'étude après avoir obtenu un consentement éclairé pour accéder à leurs données d'évaluation EPA. Toutes les données d'évaluation (démographie du résident, EPA, caractéristiques liées à l'évaluation et à l'évaluateur) ont été fournies par un analyste de données du système d'éducation médicale de l'Université McGill.

**<u>Résultats:</u>**

Au total, 595 formulaires d'évaluation de l'EPA ont été remplis entre juillet 2019 et août 2020. La grande majorité des données d'évaluation reçues (95 %) a été réalisée dans un cadre clinique. Seize EPA évalués, dont 7 correspondant au stade Transition vers la discipline et 9 au stade Fondement de la discipline. Il y avait une grande variabilité dans le nombre d'évaluations complétées par résident, où la plus faible était de 14 et la plus élevée était de 65.

Trois formulaires d'évaluation ont été utilisés pour évaluer les EPA en SF (observation de l'EPA, compétences procédurales et rétroaction multi-sources). Environ 70 % (411) des formulaires d'évaluation ont été remplis par les résidents et 27 % (165) par les médecins traitants. L'évaluation globale à l'aide de l'échelle d'attribution allait du niveau 2 (je devais les expliquer) au niveau 5 (je n'avais pas besoin d'être impliqué) où le niveau 4 (j'avais besoin d'être là juste au cas où) était le plus souvent attribué niveau de confiance. Environ 85 % des évaluations de l'EPA ont été réalisées car les niveaux d'habilitation qui leur étaient assignés étaient les niveaux 4 et 5. L'EPA 2.1b : Critically Ill Surgical et EPA 2.5a : Fundamental Surgical Procedures étaient les EPA les plus difficiles et ont souvent reçu des notes faibles.

**Conclusion:**

En treize mois, 13 résidents du programme SF ont rempli 595 formulaires d'évaluation EPA liés à 16 EPA identifiés par le CRMCC. Il y avait une variabilité apparente dans les évaluations EPA complétées entre les résidents et les sous-spécialités, où les résidents étaient les évaluateurs les plus courants. De plus, deux outils d'évaluation ont été utilisés : l'échelle de confiance et l'échelle d'évaluation globale. Une étude plus approfondie pourrait être menée pour étudier les causes de la variabilité et l'impact des résidents étant les évaluateurs les plus courants. Cette étude menée auprès d'un petit groupe peut fournir des tendances mineures qu'il serait intéressant de suivre dans une étude plus large. De plus, il vaut la peine d'explorer comment les EPA sont évalués dans le cadre de la simulation et sont en corrélation avec les résultats de l'évaluation clinique.

# Acknowledgment

I would like to express my great gratitude and appreciation to Dr.Kevin Lachapelle for his continuous guidance and support throughout my Master's study. His outstanding mentorship and valuable suggestions made this study possible and successful.

I would like to offer my special thanks to Elif Bilgic for her remarkable contribution in developing and planning my study. I worked with her side by side, and I learned a lot from her extensive experience as a researcher. She gave her time generously and provided me continuous support and valuable advice during all the steps of my research journey.

I would also like to extend my thanks to Ekaterina Lebedeva, with whom I worked to search the literature. Her willingness to search all databases and provide me with the latest publications in my field of study has been very much appreciated.

My special thanks are extended to Hanna Taleb, who provided me with EPA assessment data. Without her great help and collaboration, this study would not have been possible.

I would like to show my profound gratitude to my husband, Ahmed, and my two lovely children (Loay and Leen) for being in my life. They provide me with love, energy, and encouragement to finish this project.

Finally, special thanks are extended to my parents and family for their endless love and support to achieve my goals. Without their encouragement, this work would not have been beautiful!

**Thank you everyone.**

# Conflicts of Interest:

I have no conflicts of interest to disclose.

# Preface and Contribution of Authors

**1- Hajar Al-Mughairi, MD:**

Methodology, Design and planning the Study

Conducted the Study and Collected the Data

Interpreted the Data

Performed the Manuscript and Editing

**2- Ekaterina Lebedeva:**

Academic Associate at Montreal General Hospital, McGill University

Performed the literature search

Provided the full-text articles and abstracts for the scoping review

**3- Hanna Taleb:**

Data analyst- Medical Education Systems team, McGill University

Provided the residents assessment data

**4- Elif Bilgic, BSc:**

Methodology, Design and planning the Study

Supervised the Study

Contributed to the Data analysis

Contributed to the Manuscript and Editing

**5- Kevin Lachapelle, MD:**

Methodology, Design and planning the Study

Supervised the Study

Contributed to the Data analysis

Contributed to the Manuscript and Editing

## Abbreviations:

**ACGME:** Accreditation Council of Graduate Medical Education

**Can MEDS:** Canadian Medical Education Directions for Specialties

**CBD:** Competency by design

**CBME:** Competency-based medical education

**CCC**: Clinical competence committee

**EDM:** Entrustment decision-making

**EPA:** Entrustable Professional Activities

**GRS:** Global Rating Scale

**RCPS**: Royal College of Physicians and Surgeons of Canada

**SF:** Surgical Foundations

**WBA:** Workplace assessment

# Chapter 1: Introduction

## 1.1Background:

In the last two decades, medical education has evolved to meet public expectations regarding professional accountability. One way of ensuring accountability is to ensure that the public is cared for by competent, well-trained physicians through a robust medical, educational program. Competency-Based Medical Education (CBME) is a new paradigm to train healthcare professionals, which has been adopted by many medical educational systems worldwide and more specifically in North America. CBME refers to a learner-centered instructional model of medical education programs that focuses on the mastery of competencies and predefined abilities as learning outcomes (1-2). Moreover, this model de-emphasizes the time-based advancement during residency training and focuses on progression based on the achievement of competencies (1-2) while ensuring good alignment between the curriculum objectives, instructional strategies, and the assessment tools. Therefore, it overcomes the limitations of the traditional approach which is based on generating knowledge and skills over time which is less prescriptive regarding the overall level of competency required of a healthcare professional (table1) (3).

According to the International CBME Collaborators (ICBME), there are five core components of CBME: “an outcomes competency framework, progressive sequencing of competencies, learning experiences tailored to competencies, teaching tailored to competencies, and the programmatic assessment (4)”. Identification of these components is fundamental, leading to a better understanding of the organizing framework of the CBME program and the related implementation process (4).

## 1.2 Competency, Entrustable professional activities (EPAs) and milestones:

There are common terminologies related to the CBME that will deepen the understanding of CBME model. There is consensus amongst the international experts in medical education regarding the definition of these terms, which ensures a shared language to describe CBME, enhances effective communication, and overcomes the limitations to adopt the new model (5).

### 1.2.1 Competence and competency:

The pioneers in medical education addressed competence and its related terms. Englander et al. referred to competence as “the array of abilities (knowledge, skills, and attitudes) across multiple domains or aspects of performance in a certain context. Statements about competence require descriptive qualifiers to define the relevant abilities, context, and stage of training. It is multi-dimensional and dynamic, and it changes with time, experience, and setting” (5).

Frank et al. defined competency as “an observable ability of a health professional, integrating multiple components such as knowledge, skills, values, and attitudes. Since competencies are observable, they can be measured and assessed to ensure their acquisition” (1). Competencies have been incorporated in various medical training programs and are considered the foundations of graduate medical education. There are two common competency-based frameworks adopted by many educational institutions, the Accreditation Council of Graduate Medical Education (ACGME) in the United States and the Canadian Medical Education Directions for Specialties (Can MEDS) in Canada. Each competency framework encompasses a set of core competencies that are required to be fulfilled by the medical training programs. For example, there are seven domains endorsed by Can MEDS framework: Medical Expert, Communicator, Collaborator, Leader, Health Advocate, Scholar, Professional (6). Each competency domain encompasses a group of tasks through which the trainer competence can be assessed and judged. These tasks are known as Entrustable Professional Activities (EPAs).

**1.2.2 Entrustable professional activities (EPAs) and milestones:**

Decisions about competence may be predicted on multiple, longitudinal workplace-based assessment of EPAs (7). EPAs refer to the essential tasks of a discipline (profession, specialty, or subspecialty) that an individual can be trusted to perform without direct supervision in a given healthcare context, once sufficient competence has been demonstrated.

EPA concept should address one or more competencies and be observable and measurable in the process and outcomes (5). Ten Cate described the EPAs as a bridge between a competency framework and daily clinical practice (8).

EPAs and competencies are not interchangeable terms and there are few distinctions between them as addressed by Ten Cate. Competencies are characteristic of clinical health care professionals while the EPAs describe the clinical tasks. Moreover, each EPA reflects a set of competencies where the physicians must achieve proficiency across all the EPAs and the associated competencies (9-10).

Milestone is another emerging term in CBME and is closely related to EPAs. Each EPA bundles a group of milestones which are defined as observable markers of an individual's ability along a developmental continuum (5) (Figure1).

## 1.3 Overview of Competency by Design

Many postgraduate and undergraduate medical training programs in North America, Europe, and Australia adopt the CBME as a training approach (11). The Royal College of Physicians and Surgeons of Canada (RCPS) has implemented a customized version of CBME known as the Competency by Design model (CBD). Although CBD incorporates the same principles and core elements of CBME (12), the overarching goal of CBD is to provide the best patient care through a well-designed, learner-centered learning model and residency training program (13). Consequently, CBD divides the residency training program into four stages: Transition to discipline, Foundations of discipline, Core of discipline, and Transition to practice. Each stage has predefined outcomes, including sets of milestones and (EPAs), whereby the resident has to demonstrate competency and proficiency while performing these activities throughout the residency journey (Figure2) (12). The promotion from one stage to another is based on completeness and achieving competency in all EPAs identified at each level.

## 1.4 Surgical Foundations program:

The Surgical Foundations (SF) program is an initial period of the post-graduate training program at McGill University, Montreal. All the junior surgical residents at McGill University belong to the disciplines of cardiac surgery, general surgery, neurosurgery, orthopaedic surgery, otolaryngology, plastic surgery, urology, vascular surgery, and obstetrics and gynecology are enrolled in 2-years surgical foundations training program. During the residency training, residents learn the fundamental skills of surgery and acquire basic knowledge and attitudes essential to practice general surgery and its related subspecialties. Figure 3 demonstrates the position of SF in relation to CBD stages as per the Royal College rubric. The residents complete the SF and surgical specialty training simultaneously prior to entering the Core of discipline stage.

Consequently, upon completing training, the residents will provide the initial assessment and management of critically ill patients. Furthermore, they will be competent to participate as effective members of their disciplinary surgical team and assume greater accountability for perioperative management. Moreover, the SF program incorporates all the Can MEDS roles in the residents' training requirements; thus, the residents will be equipped with all Can MEDS related competencies (13).

The SF program encompasses the first two stages of CBD: Transition to discipline and Foundations of discipline. Transition to discipline contains seven EPAs, each with identifiable core milestones rated on a 5- points competency scale whereby the Foundation of discipline contains nine EPAs with different core milestones. Diagram 1 illustrates the EPAs list in each stage. For instance, EPA#6 in Transition to discipline is '*Repairing simple skin incisions/lacerations'* skill (form1). This EPA aims to measure the resident's ability to perform simple wound repair with different sizes and the required degree of supervision while performing this task. To assess this task, there is a list of associated precautions, steps and measures that must be achieved called milestones. This EPA should be observed and assessed directly by the supervisor either in simulation or clinical settings. The resident's proficiency in performing the task unsupervised is identified and rated on a 5-point entrustment scale using the assessment form designed by the Royal College.

Another example is EPA#4 in the Foundation of discipline, which is "*Providing patient education and informed consent in preparation for surgical care*" (form2). This non-technical skill aims to evaluate the resident's communication skills with patients and families while providing them a right and clear surgical management plan. There are predefined milestones related to this skill that must be achieved while taking consent. The resident will be observed and assessed directly by the supervisor in different clinical and simulated case scenarios.

Accordingly, the overall performance and entrustment decision will be documented on the EPA's assessment form designed by the Royal College.

The resident has to complete and achieve competency in all the associated EPAs in each stage to be promoted to the next complex stage of training, the core of discipline. All the EPAs assessment focuses on workplace assessment (WBA), an informal assessment where the resident is observed and assessed in real clinical scenarios with no unified standardisation and structure. In WBA, the supervising clinicians will direct, assess and document the resident's performance in a given task according to a supervision scale and provide formative feedback to allow the resident to correct their performance in the next clinical encounter.

In the SF program, the final decision regarding the resident's clinical performance, readiness to practice unsupervised and the promotion to the next level of training is reviewed by the clinical competence committee (CCC). The residents in Surgical Foundations will be reviewed by two committees: Surgical Foundations Competence Committee and their surgical specialty's Competence Committee. The two committees respectively will monitor progress in Surgical Foundations EPAs and surgical specialty's EPAs.

The CCC is a specialized clinical training committee that aims to review the resident's assessment data and provide the final decision about the resident's clinical performance. It is chaired by a member of the residency training faculty but not the program director, while the other committee members are selected based on regulations determined by the university (14). Moreover, there is a variable number of CCCs in each program depending on the program size (e.g. a large program has multiple committees). Each CCC must meet at least twice per year to review each resident's files, though some programs may have frequent meetings depending on the program size, resident's number, and training stages. Based on the performance evaluation of multiple observations over time, the committee can assess the resident's progression and identify the residents lagging behind in a milestone's achievement. Therefore, the CCC allows a group decision-making process to determine when the trainee can advance to the next stage. The final decisions about resident progression are reported to the residency training committee.

## 1.5 Entrustment decision-making in competency-based education:

The decision to trust a medical trainee with the critical responsibility to care for patients is fundamental to clinical training**.** In the clinical training environment, entrustment decision-making (EDM) is not an easy process as it can affect the trainees' learning journey and impact patient safety. Moreover, it is considered the key element for assessing entrustment professional activities (EPAs) within the context of competency-based education. To understand the EDM process, it is worth mentioning the definition of trust and entrustment in healthcare training. Trust refers to the reliance of the medical team on a trainee successfully performing the professional task with the ability to ask for help promptly as needed (15). However, entrustment is a conscious decision to trust a trainee to correctly execute the task unsupervised based on prior observations and what is known about the trainee (16).

In the clinical setting, there are three recognizable entrustment modes: presumptive trust, initial trust, and grounded trust (16). Presumptive trust is based solely on credentials, without prior interaction with the trainee, whereas initial trust is built based on the supervisor's first impression of the trainee. Additionally, with time and after frequent observations and multiple contacts with the trainee, grounded trust is developed, which is considered a fundamental step in the EDM process (16). To get a better understanding of the EDM process, Ten Cate and his colleagues propose two categories of EDM, either ad hoc or summative EDM (16). An ad hoc entrustment decision happens continuously in healthcare situations with no impact on the final judgment for independence. In contrast, the summative decision is based on grounded trust. It is developed upon multiple clinical observations and evaluations in various medical care scenarios, thus leading to awarding the trainee certification and the privilege to practice unsupervised (16).

Although this decision about trustworthiness of the trainee appears straightforward, several factors might influence the decision. For instance, the trainee's and supervisor's characteristics, the trainee-supervisor relationship, and the context, nature, and complexity of the assessed task can play significant roles in this judgment (17). All of these interplaying factors might influence the supervisor's decision and lead to imperfect entrustment. For example, novice supervisors may be hesitant to rely on the trainees to perform a specific task and unable to decide how much autonomy should be given to trainees. Therefore, a supervisor's experience can influence the ability to judge the trainee's performance. Moreover, the relationship between the supervisor and trainee may bias the entrustment decision. Strong interpersonal relationships, shared expectations and experiences, and frequent contact between the supervisor and trainee facilitates trust formation and influences the judgment process (17). On the other hand, the unavailability of opportunities to perform clinical works, excessive workload hindering direct observation, and high-risk procedures requiring a high degree of trust may impede the EDM process (17).

In summary, the EDM process remains subjective with many contributing factors that should be taken into account, especially if the decision is established for high-stake assessment.

## 1.6 Assessment of competency and EPA:

The meaningful assessment of competence is critical for implementing effective competency-based medical education (18); thus, the assessment of EPAs has received special attention in the medical arena. It is worth mentioning, EPA and competency are not interchangeable terms, hence their assessment modes are not the same. Assessment of different elements within EPA can provide information on competency. Since competency is observable, it can be measured and assessed. However, the question arises as to how we measure and assess competency.

In general, inconsistency exists regarding the appropriate assessment methods of EPAs (19) within CBME, with lots of arguments regarding the EDM process. Establishing a valid, reliable, and authentic assessment tool plays a central role in the successful implementation of CBME and helps achieve the desired outcomes. Moreover, it provides strong insight into the trainees' progression and competence in each stage before advancing to the next level of training.

The measurement of competencies has been widely debated among medical educators. Carraccio et al. noted the complexity of competencies assessment which required multiple assessment tools to measure all the aspects of competency (20). This complexity led the educators to break down the competencies into multiple smaller tasks and behaviors that can be directly observed. However, Malone et al. and Vleuten et al.  have criticized this approach as reductionist (21- 22). They highlighted that this breakdown strategy might threaten the validity of the utilized assessment tools, and mastery of the individual tasks does not necessarily reflect the overall competence. Moreover, judging the capability of the learners to integrate all the tasks to patient care is more important than judging their performance in one simple task (20).

Overall, competence assessment is achieved through a holistic assessment of three constituents: knowledge, clinical judgment and reasoning, and technical skills. All the available evaluation data formats that have been employed to assess the trainee's competency should be analyzed collectively and carefully before giving the green light for indirect clinical supervision. Assurance about trainees' performance in each component can be obtained through multimodal assessment methods, including real-time direct observations, simulations, oral structured examination (OSCE), and written exams (23). Therefore, simulation could complement the workplace assessment of competence, where different aspects of competence may be assessed in both settings: simulation and clinical environment. Murto et al. emphasize the importance of using quantitative and qualitative assessment tools as they complement each other (24). Also, Carraccio et al. suggest an assessment framework integrating competencies, EPAs, and milestones in order to judge the trainees' performance and level of supervision (10).

EPA is a fundamental element within the CBME model. The introduction of this concept enabled the holistic assessment of learners and provided a potential solution to assess competencies (20). The EPA reflects multiple competencies and integrates multiple milestones contained in the clinical activity and can be observed and assessed as a total activity. As per Cate, the EPAs "should be observable and measurable" and "EPAs have a holistic nature" (25). However, the detailed EPA assessment framework in terms of assessment tools, settings, and raters should be standardized with clear guidance on implementing the direct observation to ensure accurate assessment (20).

Workplace assessment (WBA) is the central mode to assess competence and EPAs within CBME. It relies on direct observation of trainee performance by the supervisor in authentic clinical activities. It needs to be collected continuously on multiple different clinical occasions by multiple assessors. Therefore, this method provides a collection of assessment data, which can be used to measure the competency and make the entrustment judgment. There are different forms of WBA, including direct clinical observation, multisource feedback (MSF), end -of-rotation evaluation, mini-clinical Evaluation exercise (mini-CEX), and Case-based Discussion (CbD) (26-27).

One important advantage of WBA is its ability to provide timely and direct feedback, which guides the trainee in the next clinical encounter and improves performance (27). Furthermore, it enables assessment of non-technical competencies, such as professionalism, in an authentic clinical environment, which might be difficult to assess otherwise (28). In contrast, WBA limitations have been addressed frequently by medical educators and highlighted by several studies (16,27-29). One main limitation is that WBA is a formative assessment that is ill-structured and lacks standardization as it happens in different contexts with different patients and variable supervisors. Therefore, WBA tends to have low inter, intra-rater and across occasion consistency (16,28). Another concern is busy clinical practice leading to poor faculty participation in the assessment process and feedback. In addition, a rater's subjectivity may lead to assessment bias and the inability to discriminate between the trainees (16). Green et al. addressed the role of faculty in competency assessment and how they can impact the assessment. Most significantly, they conclude that "the biggest problem in evaluating competencies is, in our opinion, not the lack of adequate assessment instruments but, rather, the inconsistent use and interpretation of those available by unskilled faculty" (30).

Since there are many challenges with direct observations, can simulation facilitate the EPA assessment and offer insight into the competency of trainees? Nowadays, simulation-based training (SBT) routinely complements actual clinical learning and patient care experiences in surgical education (31). It plays an essential role in the training and assessment of healthcare professionals as it provides a controlled, safe clinical environment and overcomes the limitations of using humans as training resources. It enables the trainees to practice the task multiple times without endangering patients. Moreover, it is an efficient tool to train novices, enhances skill retention, and prevents skill decay (32). Furthermore, simulation assessment overcomes the reliability issue encountered in WBA as the exam delivery can be standardized to evaluate a specific problem using the same simulator to evaluate every examinee by the same examiner. (33).

Despite all the strengths mentioned above of simulation training and assessment, there are some important drawbacks. The costs of simulation activity should be considered especially if we add up all the contributing factors to the successful implementation of simulation, such as technology, simulators modality and faculty (34). In addition, although simulation is a novel aid to mirror real-life circumstances, it is still not a real experience (34). The learners still need to see a real patient and practice in a real clinical environment. Finally, the transfer of simulation training to clinical practice is not guaranteed. (34-35). Although the transfer of learning to real patient setting can be facilitated by using high fidelity or VR simulation, it is still not guaranteed. (34-35).

Nevertheless, many studies proposed that simulation-based training can be an effective tool to assess the EPAs and trainee's proficiency in both the technical (26,27,36) and non-technical skills (36). Wu et. el. conducted a study to determine simulation's utility to assess the non-technical EPAs (situational awareness, decision making, communication, teamwork, and leadership) in otolaryngology emergencies (36). They reported that simulation is a successful and useful assessment modality to assess non-technical EPAs. Johnston et al. documented a successful development and implementation of assessment tools to assess neonatal intubation EPA using a neonatal simulator (37). The assessment tools included a checklist, global rating scale, and EPAs entrustment scale. Moreover, simulation allows evaluation of more than one EPA in one complex simulated case scenario rather than one EPA in an isolated clinical task. For instance, one simulated case about the resuscitation of a critically ill patient can be used to assess non-technical EPAs like communication skills, leadership, and teamwork as well as technical EPAs like recognizing a critically ill patient and the ability to initiate the necessary management.

Of note, various assessment tools have been utilized to assess EPAs including a checklist, global rating scale (GRS), and Entrustability scale. Following is a brief introduction to give an overall understanding of each tool.

The Entrustability scale is “behaviorally anchored ordinal scales based on progression to competence” (38). It guides the assessor to judge the trainees' readiness for independent practice. Moreover, it provides the assessors with a structured assessment tool that can be used across multiple tasks and track the learner's performance over time. Multiple observations are assessed collectively before providing the final trust judgment. This final decision impacts the trainee’s future training and progression; Ten Cate explained, “each 5 level has direct consequences for the trainee and patient care” (39). Therefore, promoting the trainee to the next complex stage of training is highly dependent on the final entrustment level.

A checklist is a task-specific, objective measurement of the trainee’s actions. That is used to identify whether the key steps of the task have been completed or not. Although a checklist provides timely feedback to an observable task based on a dichotomous response format, it is not an ideal tool to assess the actions and tasks that cannot be converted to a dichotomous format (40).

GRS is a scale characterized with clearly defined and detailed statements about the task with an anchored rating scale, enabling the rater to judge the trainee's performance. It indicates the degree and range of accomplishment rather than yes or no.

After the previous overview of various assessment tools utilized to assess EPAs, it is worth investigating which EPAs are assessed with each assessment tool? Which assessment tools utilized to assess the different components of EPA (technical, non-technical)? What are the assessment settings of each assessment tool?

In summary, CBME is a novel medical educational model in which the EPA concept and assessment is a new field. EPA is an essential unit in the CBME framework and plays a fundamental role in successfully implementing the competency model in any program. Yet, the lack of a meaningful assessment tool of EPA can directly impact the implementation process. Although the holistic assessment is a helpful approach to assess EPA, there is still no consensus around EPA measurements. Hence, EPA assessment merits thorough research to understand how EPAs are being assessed. What are the available assessment strategies in the literature that have been implemented to assess EPAs? What are the utilized assessment settings, clinical or simulation? What are the utilized assessment tools in each setting? We decided to conduct this study to address the questions around this new concept and explore how EPAs have been assessed.

## 1.7 Study hypothesis

In this paper, we propose that EPA may be assessed in clinical settings as well as in simulation settings. Also, EPA assessments in a simulated environment correlate with those obtained in the clinical environment. Initially, we did a scoping review to explore the available literature in this field. Then, we focused on understanding the EPA implementation and assessment within a well-defined training program at McGill University (Surgical Foundations program).

## 1.8 Thesis division and objectives:

The specific objectives of this thesis are:

1.To identify the types of assessed EPAs.

2.To explore how EPAs are being assessed (assessment settings, assessment tools, and rater's role).

3.To explore how EPAs are being assessed within a University Surgical Foundations Program during both the Foundation of discipline and the Transition to discipline stage of training.

To fulfill the above objectives, this paper has been divided in two parts: a scoping review of EPA assessments within both simulation and clinical settings, and EPA assessment within the Surgical Foundations program at McGill University.

First, we conducted a scoping review to identify the types of assessed EPAs and understand how the EPAs have been assessed in terms of assessment tools, settings, and assessor's roles. We decided to choose a scoping review over a systematic review because EPA assessment is a relatively new concept, and hence, the assessment strategies might vary across institutions. Therefore, we sought to overview the types of assessed EPA and identify the various assessment tools utilized to assess EPA within simulation and clinical settings. Rather than it aims to produce or synthesize evidence about the effective assessment method to assess EPA.

After performing the review, we got an overall understanding of EPA assessment. Subsequently, we decided to correlate what was published in the literature regarding EPA assessment with the EPA assessment plan implemented within SF training program. Therefore, in the second manuscript, we described how EPAs were being assessed at the Surgical Foundation Program at McGill University, Montreal, Canada.

# SECTION 2:

# Manuscript 1

# Assessment of Entrustable Professional Activities within Simulation and Clinical Settings: A Scoping Review

Hajar Al-Mughairi, MD[1]
Elif Bilgic, BSc[2]
Etkatrina Lebedeva, MILS[3]
Kevin Lachapelle, MD, FRCPS(C), FACS[1,4]

[1] Department of Experimental Surgery, McGill University, Montreal, QC

[2] Steinberg-Bernstein Centre for Minimally Invasive Surgery and Innovation, McGill University Health Centre, Montreal, QC, Canada

[3] Montreal General Hospital Medical Library, McGill University Health Centre, Montréal, QC Canada

[4] Division of Cardiac Surgery, McGill University, Montreal, QC

**Conflicts of Interest and Source of Funding:** No conflicts of Interest to declare. No funding sources to declare.

## Abstract:

**Introduction:**

Entrustable professional activities (EPAs) are a fundamental unit in the competency-based education model (CBME). Assessment of EPAs is a crucial step in the implementation process of CBME, which remains challenging due to a lack of consistency in meaningful assessment methods. There is great variability in the assessment approaches used to evaluate EPAs among residency training programs, which might influence successful CBME implementation. This scoping review seeks to explore how EPAs are being assessed and implemented within residency training programs.

**Methods:**

We conducted a review of MEDLINE, PubMed, Embase, PsychINFO databases, and Google Scholar by searching the full-text articles and abstracts published between January 2000 and January 2020. Additional studies were identified through reference lists. We included the publications addressing the EPA assessments as defined by the Royal College which were written or translated to English only.

**Results:**

A total of 635 articles were screened for eligibility, and ultimately, twenty-eight full-text articles and six abstracts were enrolled in the final review. 122 EPAs were identified, with 86 (70.5%) EPAs corresponding to non-technical skill alone, 26 (21.3%) technical skill alone, and 10 (10.2%) combined technical and non-technical skills.

Out of 122 EPAs, 86 (69%) were assessed clinically and 38 (31%) were assessed in a simulated environment. In simulation assessment, a standardized patient was the commonly utilized simulation modality, and checklists were the most used assessment tool, either alone or

with other tools. However, the Entrustability scale was the dominant assessment tool of EPA in the clinical setting (77%).

Seven rater types were identified: attending physician, trained rater, senior resident, peer, self, nurse, and patient. Moreover, around half of EPAs (52%) were assessed by the attending physicians.

**Conclusion:**

This review identified that EPAs are being assessed diversely, in multiple settings, by raters at various experience levels, and using different assessment tools. As the methods of assessment are not consistent across institutions, future research should seek to establish standards for the most efficient and effective way to assess the EPAs, which are mandatory by the medical accreditation bodies.

## 2.1 Background:

The widespread adoption of CBME by many medical graduate educations mandated a standardized implementation process. Assessment is a critical step in any educational model that needs to be evaluated carefully to ensure successful implementation. Carraccio et al. highlighted four major steps for successful CBME execution: "(1) competency identification, (2) determination of competency components and performance levels, (3) competency assessment, and (4) overall evaluation" (20). According to Carraccio et al., assessment of competencies was the most challenging step and was responsible for the unsuccessful implementation's attempts. Therefore, establishing diverse, accurate and reliable assessment strategies is necessary to meaningfully evaluate the trainees' competency across its three domains: knowledge, skills, and attitudes.

EPAs is a core element of CBME, described by Ten Cate as the unit of professional activities (9). The outcome of EPAs measurement leads to entrustment decisions about the capability of the trainees to practice unsupervised. However, there is still a debate about the appropriate setting to implement and assess the EPAs, and decisions that can be made based on the results. Moreover, EPAs are being used to make decisions about trainees' skill level, so it is important to understand how EPAs are being assessed. Furthermore, few available studies in the literature addressing the EPA assessments and infrequent publications about EPA implementation and assessment were also noted by O'Dwod et al. in the systematic review of EPAs conducted between 2011 and 2018 (41). Therefore, we conducted a scoping review to get an overall understanding how the EPAs have been assessed within simulation and clinical settings.

Of note, there are few main differences between a literature review, a systematic and a scoping review. A systematic review is a type of research synthesis aiming to identify the literature on a particular condition. It then evaluates and appraises the available evidence about this condition according to a structured process with a pre-defined rigid protocol. It is conducted to answer specific questions about the effectiveness, appropriateness, meaningfulness, and feasibility of a certain practice (42-43). On the other hand, a scoping review is a broader approach to map the existing literature about a particular concept according to a predefined protocol with clear objectives and searching methods. It aims to address broader questions about the key concepts and definitions of a particular topic and clarify the type of available evidence to identify the gap in the research knowledge (42-43). In contrast, the traditional literature review seeks to summarize the studies on a particular topic and examine all the relevant aspects of the research field focusing on history, concept, importance, problems, and the reported evidence (43).

We decided to choose a scoping review over a systematic review because EPA assessment is a relatively new concept and this scoping review does not intend to produce or synthesize evidence about the effective methods to assess EPA. Rather, we aim to provide an overview of different EPA assessment approaches. In this scoping review we sought to investigate how the EPAs have been implemented across institutions within the competency framework. In addition, we aim to understand how EPAs are being assessed and identify the various assessment strategies and platforms utilized for EPA assessments.

## 2.2 Materials and Methods:

### 2.2.1 Search strategy:

With the assistance of a medical librarian (E.L), we searched the published articles between January 2000 and January 2020 in MEDLINE, PubMed, Embase, PsychINFO, and Google Scholar. Medical subject headings terms (MeSH) and keywords related to CBME and EPA were used to develop the search strategies (table 2). Additional searches were conducted using the citations of the relevant publications. The search terms included "Competency-based education,"" Entrustable professional activities (EPAs)" and "competency assessment".

**2.2.2 Study selection:**

The titles and abstracts of EPA-related articles were screened. Full-text articles and the abstracts related to EPA assessment in graduate medical education (GME) were included in the review. We included all the studies that recognized the skills as EPAs as defined by RCPS "authentic tasks of a discipline" (44) and found to be EPAs within the RCPS collection. On the other hand, we excluded the articles that (1) mentioned CBME only (concepts, process, challenges), (2) were not about the residents and interns (UGM, fellows, nurses, other healthcare professionals), (3) were the assessed skills are not EPAs as per RCPS recommendations, (4) were missing details about the assessed EPA or skill, participants, assessment methods or tools, (5) were not translated to English (figure 4).

Studies assessing milestones and competencies were included in the review. Moreover, the validation studies about a specific assessment tool to assess the EPA or skill were also included.

**2.2.3 Data extraction:**

Data was extracted from included full-text articles and abstracts. Two separate data extraction sheets were created for both the full-text articles and abstracts. Extracted information was almost similar in both forms, with fewer heading in the abstract chart. The extracted data include the article's characteristics (e.g. title, country and date of publication), participant-related characteristics (e.g. specialty, level of training), rater characteristics (e.g. rater type and number), EPAs- related data (e.g. type of EPA, number of EPA), assessment-related data (e.g. setting of assessment, assessment tools), and the study result (table 3).

#### 2.2.4 Abstract reviews:

Six abstracts were included in the review (full texts are not published yet). These studies assessed residents from variable levels ranging from PGY-1 to PGY-4 and were belonging to different residency training programs (1 General surgery, 1 Obstetrics and gynecology,1 primary care,1 Emergency medicine, and 1 Psychiatry).

Four abstracts assessed EPAs through simulation using OSCE scenarios, and 2 assessed clinically via direct observation and self-assessment.

## 2.3 Results:

The initial search recorded 635 publications. After the primary screening and application of the inclusion and exclusion criteria, 28 full-text articles and 6 abstracts were included in the review. A total of 122 EPAs were assessed in full-text articles. The study characteristics are illustrated in Table 4. Most of the full-text studies were conducted in USA (n=12, 42%) followed by Canada (n=11, 39%). Out of 28, there were five pilot studies, 3 validation studies and one study described the process of development and implementation of EPAs in their institution. Moreover, analysis of the publications reveals two synonyms for EPAs: Capabilities in Practice (CiP) (45) and Observable Practice Activities (OPA) (46). The following results will be for 28 full-text articles.

### 2.3.1 Participant characteristics:

We found 24 studies conducted among residents, 3 among interns, and 1 among both residents and interns. Out of 122 EPAs, residents assessed in 111 (90%) EPAs and interns assessed in 11 EPAs. While most of the populations in the reviewed studies were interns or residents from different levels of training and various medical and surgical backgrounds, two studies investigated other groups *along with* the residents (fellows, medical students, experts). Several studies were conducted among residents from multiple residency training programs belonging to different training centers (36,47-48). Table 4 shows the distribution of the studies according to the participants' specialties. The training level of the residents enrolled in the studies ranged from PGY1 to PGY5. Furthermore, 20 publications examined trainees with different levels of experience in order to explore the change in the proficiency over the training' years and measure the ability of assessment methods to discriminate between senior and junior trainees.

### 2.3.2 Rater characteristics:

The type of the rater assessing the EPAs was varied across the studies. Seven rater types were identified: attending physician, trained rater, senior resident, peer, self, nurse, and patient. Around half of the identified EPAs were assessed by the attending physician (52%). In addition, multiple assessors evaluated the same EPA in 44% of total EPAs, while trained raters assessed 4% of total EPAs.

Notably, the type of the assigned rater was primarily influenced by the assessment setting. For instance, all the simulation-based assessment of EPAs were evaluated by attending physicians or trained raters.

### 2.3.3 EPA characteristics and assessment setting:

122 EPAs were assessed by 28 studies, although some studies assessed more than one EPA and some EPAs were assessed by more than one study. According to the elements of EPA they assessed, we identified three categories of EPAs: non-technical skill alone 86 (70.5%), technical skill alone 26 (21.3) and combined technical and non-technical skills 10 (8.2%).

Most of the assessed EPAs were communication-type non-technical skills (e.g. communication with the patient, family and other healthcare workers, presentation skills, counseling and consultation skills, informed consent, handover). Moreover, the EPA "recognizing the emergency and initiating its management" was the most frequently assessed EPA. We found it assessed 13 times: 6 times clinically and 7 times using simulation.

Two assessment settings were utilized: clinical or simulation. There were 86 (69%) EPA assessments conducted in a clinical environment compared with 38 (31%) EPAs assessed in a simulated environment (table 5).

### 2.3.4 Clinical-based assessment:

Three assessment platforms were recognized as a means to evaluate and document the resident's performance a: paper copies, electronic portfolio (49), and mobile application (50). The distribution of clinical assessments based on EPA type was as follows: 81.4% non-technical skills, 14% technical skills, 4.6% technical and no-technical together. Out of 86 clinical-based assessment of EPAs, 5 assessment tools were identified. The Entrustability scale was the dominant assessment tool (67,78%), followed by Ottawa Clinic Assessment Tool (OCAT) (10,11.6%) (51) and Likert Global Rating Scale (6, 7%). Other utilized assessment tools were Operative Performance Rating System (OPR) (2, 2.3%) (52) and Psychotherapy Process Q-sort (PQS) scale (1,1.1%) (53). However, some EPAs were assessed clinically by several studies using different tools. For instance, documentation skill was assessed by two separate studies, with one study using the Entrustability scale as an assessment tool (49) and the other using OCAT (51). All the identifiable metrics were Global Rating Scales (GRS).

### 2.3.5 Simulation-based assessment:

All of 38 simulation assessments followed an objective structured clinical examination (OSCE) style. However, depending on the utilized simulators, the exam delivery was variable. Five simulation modalities were identified: standardized patients, task trainers, high-fidelity simulation, screen-based simulation and hybrid simulation. A standardized patient was the most commonly utilized simulation activity, which is used to assess 15 EPAs. The high-fidelity simulation and task trainers were the second most utilized activities, with each used to assess 10 EPAs.

Like the scenario within clinical assessment, several studies assessed the same EPA using different simulation modalities. For example, the EPA about "recognizing the emergency and initiating its management" was assessed seven times in a simulated environment by separate studies; five times by high-fidelity simulation, and twice by standardized patients.

Moreover, eight assessment tools were identified in simulation assessment: checklist, Entrustability scale, Non-Technical Skills in Surgery (NOTSS), Objective Structured Assessment of Technical Skills (OSATS), Direct Observation of Procedural Skills assessment tool (DOPS), Ottawa Global Rating Scale (Ottawa GRS), Queen's Simulation Assessment Tool (QSAT) and Objective measurement. Checklists were the most prevalent assessment tool, either alone or in combination with other tools.

Notably, there was a tendency to use multiple assessment tools for simulation-based assessment of EPA in the same setting. For instance, three assessment tools were used in one setting to assess the neonatal intubation EPA: checklist, Entrustability scale, and Global skills assessment (GSA) (37). In addition, except for one EPA, the Entrustability scale was not used solely as an assessment tool, although was used alone in the majority of EPA clinical assessments.

## 2.4 Discussion:

This review identified the studies that have conducted EPA assessments in various residency and internship training programs. Although few available studies in the literature addressed the EPA assessments in the residency and internship programs, in this review we could identify 122 EPA assessments across 28 full-text publications. We recognized different EPA components, and diverse EPA assessment methods and tools have been utilized.

The first point revealed by this review is that non-technical EPAs are readily assessed (70.5%) compared with technical skills, particularly communication-type skills such as presentation skills, counseling, handover and informed consent. This observation can be explained by the great emphasis by the CBME model on providing the best learning experience to the trainee while ensuring good patient care and safety. Also, communication skills are commonly encountered, easy to observe and evaluate. It is generally agreed that communication skills play a vital role in patient management, which ultimately impacts patient safety. In addition, CanMEDs and ACGME recognized communication and interpersonal skills as core competencies that are required for medical education programs accreditation.

Secondly, there is a tendency to assess both aspects of EPA (technical and non-technical) in one assessment setting, like resuscitation skills (54). It is well-known that technical and non-technical skills are strongly associated with each other, and both types have to be trained and implemented simultaneously (55). Alken et al. describe various approaches to effectively integrate technical and non-technical surgical skills training, which enable the trainees to learn and apply both sets concurrently (55). Consequently, this training strategy can also be adopted as an assessment strategy depending on the curriculum objectives that have to be aligned with instructional strategies and assessment methods.

Our review reveals that the utilized EPA assessment methods vary considerably. Three-quarters of the EPA clinical assessments were assessed by the Entrustability scale alone, whereas the rest used other specific assessment tools. All of these assessment tools were GRS. However, while various assessment tools in the simulation assessment were used alone, others were combined with other tools. Checklists were the predominant simulation assessment tool either alone or combined with GRS. Moreover, the Entrustability scale was used in simulation assessment to assess three EPAs only (out of 38 EPAs), once solely and the rest with other measurements.

The Entrustability scale is a global rating scale with 5 behavioral anchored levels of supervision where the rater can judge the trainee's readiness to practice unsupervised. This scale enables the assessors to evaluate trainee performance and make the entrustment decisions according to real-world daily clinical observations. Compared with checklists and other GRS tools, the Entrustability cale aims to judge the learners' readiness for safe independent practice rather than focusing on learners' deficiencies or comparing their performance with their peers (38). According to the above result about the use of Entrustability scale in EPA assessment, it seems that the Entrustability scale was the preferable EPA assessment tool in the clinical setting. This observation may be attributed to the subjectivity of the Entrustability Scale, which makes this tool more feasible and appropriate in the work-place assessment. This scale is characterized by construct alignment with competency progression, thus making the assessment meaningful and enabling the raters to easily translate their judgment guided by the narrative description (38). However, the subjectivity of the Entrustability Scale can lead to poor reliability.

Furthermore, the GRS represents the majority of the identified assessment tools, especially those utilized for clinical assessment. One main advantage of this scale is that the rater can use it across multiple tasks. In contrast, its limitations include subjectivity, low inter-rater reliability, and high rater training requirements (40). GRS would be the best choice for high-stake summative assessments, with the presence of qualified raters, and could be the preferable assessment for surgical competency (56).

On the other hand, checklists were predominantly utilized to assess EPAs in simulation settings. Since checklist is a task- specific assessment tool and its items provide clear task descriptions, it has higher inter-rater reliability than GRS and does not require extensive rater training. On the other hand, the checklist is not suitable for non-observable actions, and a separate checklist is required for each task. Nevertheless, since the checklist can provide direct feedback with clear areas of weakness and improvement, it is preferable for formative assessment of technical skills rather than surgical competency (56).

The tendency to use multiple assessment tools in a simulation setting can be attributed to the granularity of simulation, meaning that different aspects of EPAs (technical and non-technical components) can be assessed in one simulation session using various assessment tools. For instance, assessment of EPA about lumbar puncture can be conducted through a checklist to assess the technique-related element and the Entrustability scale to assess the communication skills and overall performance.

In sum, our findings support the observation about a lack of standardized and unifying EPA's assessment plan, which is also reported by Mubuuke et al. as a challenge to implement CBME in the undergraduate training programs (57).

Furthermore, our review revealed that clinical assessment was the most adopted assessment mode for all the three identified EPA types, mainly the non-technical ones. Despite the challenges of work-place assessment addressed by several studies (16, 27- 29**)**, our observation revealed the feasibility of clinical assessment. Since most assessed EPAs are frequently encountered during daily clinical encounters making clinical assessment more prevalent despite the busy clinical practice, clinical practice creates more learning opportunities.

Our review also supports the growing importance of simulation and its vital role in complementing the clinical training and assessment. Notably, there was an effective integration of simulation in the EPA assessments, which contributed significantly to a successful EPA implementation process. This is supported by the results of the studies conducting EPA assessment using simulation (32,37,47-48,51,58). Moreover, Dwyer et al. addressed the role of simulation in the implementation and assessment of EPAs. He advocated using the simulation as a supplement to the work-place assessment due to the ability of the simulation to identify the resident who needed more focused training to achieve the required level of proficiency (58). Additionally, simulation can alleviate the challenges faced during direct clinical observation, especially the administrative burden to the program and the remarkable time and effort spent by attending and trainees to carry out the assessment process. In fact, one assessment tool cannot evaluate all the aspects of clinical competency, and as suggested by Murto et al., a combination of work-placed and simulation assessment strategies would be the best assessment option (24).

In a simulation setting, five different simulation modalities were utilized to assess the EPA. Standardized patients (SPs), a high-fidelity simulation, and task trainers were the predominant activities. Virtual patient simulation and hybrid simulation were also used efficiently to assess some EPAs.

Hybrid simulation is defined as using two or more simulation modalities within the same simulation session. Typically, integration of standardized patients and part-task trainers conjugates technical and non-technical skills training allowing holistic competence assessment (59). Additionally, in more complex cases and procedures like delicate surgeries and endoscopies, virtual reality simulations can be considered as means of EPA training and assessment. Augmented reality superimposes computer-generated images and sounds onto the real world, while virtual reality is an entirely computer-generated simulated environment. These novel simulation techniques can simulate endovascular, laparoscopic procedures, and crisis resource management training (59-60). Moreover, they allow the assessment of more than one EPA in one simulated task.

Furthermore, in any assessment process the raters play a central role, particularly in CBME, where the supervisors are required to make the entrustment decision about the trainees' performance. Although trustworthiness is a subjective decision, and it appears straightforward, several factors might be influenced such as the trainee's and supervisor's characteristics, supervisor–trainee relationship, the context, nature, and complexity of the assessed task, as discussed by Hauer et al. (17).

Moreover, our review identifies multiple rater types, with some studies using more than one rater types depending on the assessed EPAs and assessment setting. All the contributing factors of entrustment decision should be considered, especially if the decision is established for high-stake assessment. For example, rater experience should be taken into account according to the importance and complexity of the task and the assessment's purpose. Experienced supervisors can interpret the performance and make inferences on the trainee's actions, whether novice supervisors or senior residents focus on reporting rather than analyzing the actions and behaviors (17). In addition, senior residents tend to rate trainees higher than attending physicians.

Another notable observation in this review is inconsistency in the use of the term EPAs. Two studies referred to EPAs as Capabilities in Practice (CiP) (45) and Observable Practice Activities (OPA) (46). In their review, Meyer et al., addressed the lack of consistency around EPA related terminology and raised the importance of using a shared language to have a shared and focused understanding of the new concepts. Standardized nomenclatures facilitate results generalization and interpretation without confusion. He further explained "adoption of consistent language to clarify topics of study surrounding the use of EPAs will enable the field to move forward" (61).

## 2.5 Conclusion:

In conclusion, EPA assessment within the context of CBME is an evolving topic. Given the busy service in most medical and surgical training programs, the implementation process of EPA assessments remains challenging. Although many of the reviewed publications follow the RCPS standards regarding the application of the Entrustability scale, there is still a diverse and wide range of assessment tools being used in which the passing standard is likely quite different, especially those conducted in simulation.

This review supports the need to establish a standardized and unified process of EPA assessments to facilitate the implementation of EPA within the training settings and enhance the advancement of CBME. A future systematic review may be conducted to examine the effectiveness and feasibility of the assessment tools identified in this scoping review. Such a follow-up systematic review can synthesize our evidence regarding the meaningful and appropriate approach to measure and assess EPAs.

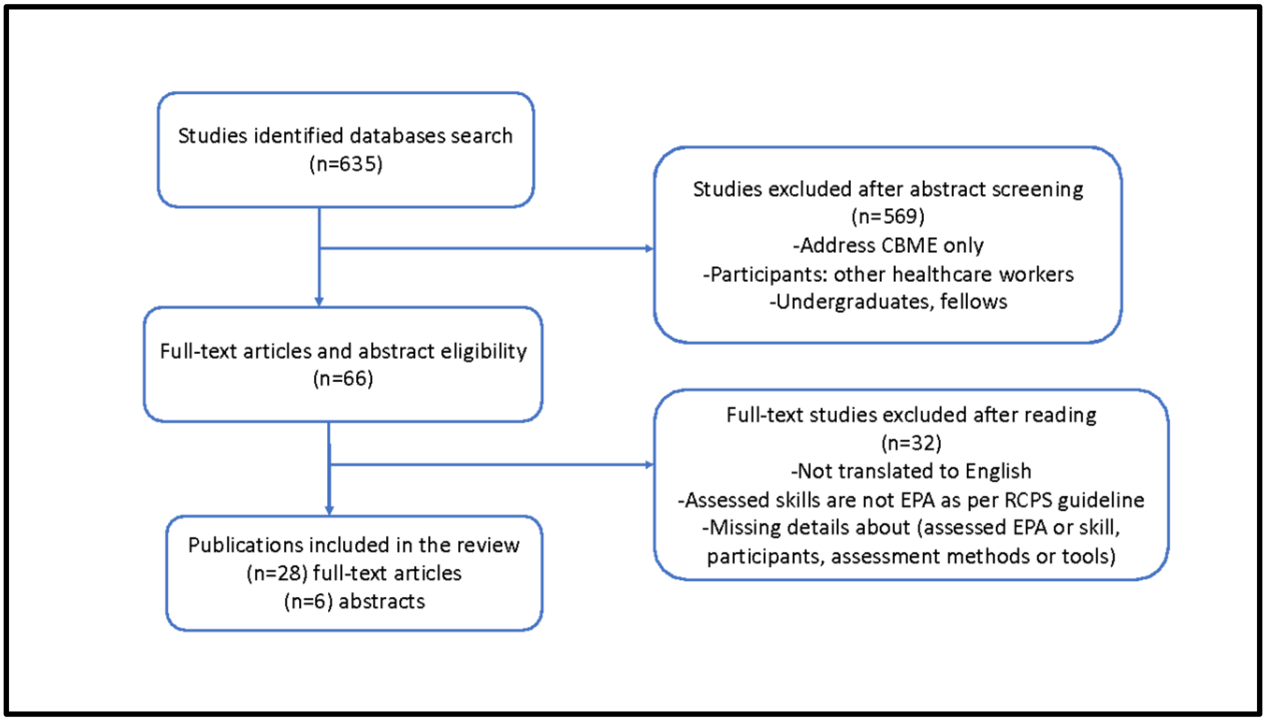


**Figure 4:** Flow chart illustrating process of study identification and selection.

**<u>Key Words / MeSH terms:</u>**

| **Keywords** | **MeSH terms** |
|---|---|
| simulation-based training | • Clinical competence/standards |
| simulation-based assessment | Competency-based education/methods |
| competency-based medical education | Competency-based education/standards |
| simulation | Competency-based education/trends |
| assessment | Education, medical, graduate/methods |
| Medical education | curriculum |
| Residency education | Internship and residency |
| OSCE | |
| Validity | |
| WBA | |
| simulation technology | |
| competency assessment | |
| Milestones | |
| Entrustable professional activities (EPAs) | |
| Surgical assessment | |
| Surgical education | |
| CanMEDS roles | |
| Transition to discipline | |

**Table 2**: MeSH terms and keywords used for the search strategy to identify the relevant studies.

# SECTION 3:

# Manuscript 2

## EPAs within Surgical Foundations (SF) Program

## at McGill University, Montreal, Canada

## 3.1: Pre-amble:

In section 2, the scoping review provided us with an overall understanding of how EPAs have been assessed. Different types of EPA elements were assessed in both workplace and simulation settings using diverse assessment tools. Moreover, various simulation modalities were used to facilitate the assessment. Assessors play a fundamental role in any assessment process, and the scoping review gave an overview of the roles of different raters who participated in EPA assessment. EPAs are still a new concept within the CBED framework, and their assessment has been evaluated by many studies as highlighted by our recent scoping review.

EPAs have been implemented within the CBD model by the RCPS of Canada. It is a new field, and many programs are starting to adapt and implement them in their curriculum. One of these programs is the Surgical Foundations (SF) program. The next section will focus on understanding the initial implementation and results of using EPA in the Surgical Foundations (SF) program at McGill University. We then aimed to correlate the results about EPA assessment between the published literature and what has been done in the SF program.

# EPAs within Surgical Foundations (SF) Program

# at McGill University, Montreal, Canada

Hajar Al-Mughairi, MD[1]
Elif Bilgic, BSc[2]
Hanna Taleb[3]

[1] Department of Experimental Surgery, McGill University, Montreal, QC

[2] Steinberg-Bernstein Centre for Minimally Invasive Surgery and Innovation, McGill University Health Centre, Montreal, QC, Canada

[3] Medical Education Systems, Faculty of Medicine, McGill University, Montreal, QC

**Conflicts of Interest and Source of Funding:** No conflicts of Interest to declare. No funding sources to declare.

## Abstract:

**Introduction:**

Competency by design (CBD) has been introduced by the Royal College of Physicians and Surgeons of Canada (RCPSC) as a novel educational model for residency training programs. Within the CBD competence continuum, Surgical Foundations (SF) program is a special training program where the residents from nine surgical specialties learn the basic surgical knowledge and skills. CBD focuses on mastery of competencies and uses entrustable professional activities (EPAs) as the organizing assessment framework. The purpose of this study is to explore how EPAs are assessed within the SF program, during both the Foundation of discipline and the Transition to discipline stage of training.

**Methods:**

Thirteen residents enrolled in SF program from six surgical subspecialties participated in the study after obtaining informed consent to access their EPA assessment data. All the assessment data (resident's demographic, EPAs, assessment and rater- related characteristics) was provided by a data analyst in the medical education system at McGill university.

**Results:**

A total of 595 EPA assessment forms were completed between July 2019 and August 2020. The vast majority of the received assessment data (95%) was conducted in a clinical setting. Sixteen EPAs assessed, with 7 corresponding to Transition to discipline and 9 to Foundation of discipline stage. There was great variability in the number of completed assessments per resident, where the lowest was 14 and the highest was 65. Three assessment forms were utilized to assess the EPAs in SF (EPA observation, procedural competencies and multi-source feedback). About 70% (411) of the assessment forms were completed by residents

and 27% (165) by the attending physicians. The overall assessment using the entrustment scale ranged from level 2 (I had to talk them through) to level 5 (I did not need to be involved) where the level 4 (I needed to be there just in case) was the most commonly assigned entrustment level. Around 85% of EPA assessments were achieved as their assigned entrustment levels were level 4 and 5. The EPA 2.1b: Critically Ill Surgical and EPA 2.5a: Fundamental Surgical Procedures were the most challenging EPAs and were frequently assigned low ratings.

**Conclusion:**

Within thirteen months, 13 residents in the SF program completed 595 EPA assessment forms related to 16 EPAs identified by RCPSC. There was an apparent variability in completed EPA assessments between residents and sub-specialties where the residents marked as the most prevalent raters. Moreover, two assessment tools were utilized: the Entrustability scale and the Global Rating Scale. Further study could be conducted to investigate the causes behind the variability and the impact of residents being the most common assessors. This study conducted among a small group and can provide minor trends which would be interesting to follow in a larger study. Moreover, it is worth exploring how EPAs are assessed within simulation and correlate with the results of clinical-based assessment.

## 3.2 Background:

Medical education in Canada witnessed rapid changes after the introduction of Competency by Design (CBD) by the Royal College of Physicians and Surgeons of Canada (RCPSC) in 2015. CBD is a customized version of competency-based medical education (CBME), that adapts the same principles and core components as CBME. This initiative is a learner-centered and an outcome-based approach to learning, aiming to ensure the physician's accountability to meet the health care system needs. Furthermore, the CBD framework encompasses a set of competencies and milestones which follow the Canadian standard for medical training, CanMEDS roles (12). Figure 2 illustrates the competency continuum within the context of CBD. It describes the different stages of residency training programs starting from the transition from medical school to residency, then progressing through the different stages of residency and its related predefined outcomes, and finally, the transitioning from residency to practice. The advancement from one stage to another is based on the achievement of predefined educational outcomes rather than time.

The SF program is a special training program encompassing the initial two stages of residency training (Transition to discipline and Foundations of discipline). It was introduced in 2018 and any resident matching one of the nine surgical subspecialties (Cardiac Surgery, General Surgery, Neurosurgery, Obstetrics and Gynecology, Orthopedic Surgery, Otolaryngology Head and Neck Surgery, Plastic Surgery, Urology and Vascular Surgery) is automatically enrolled in this program (62). This program is generally no longer than four blocks for Transition to discipline and 12 blocks for Foundations of discipline (each block constitutes 4 weeks) (62). Over the first two stages of residency training, the resident completes the SF training and the surgical specialty training simultaneously (figure3).

In this program, the resident learns the fundamental surgical skills and achieves the competence to provide pre-operative and post-operative management (62). The progression toward the Core of discipline of surgical specialty is based on achieving proficiency in all the EPAs. Each EPA is an essential task in the discipline (e.g. consent), which is anchored to an Entrustability scale ranging from 1 (I had to do it) to 5 (I did not need to be involved) according to RCPSC recommendations (figure 5). Moreover, EPA is the organizing assessment framework in CBD that guides and reflects each resident's performance across the various clinical activities. There are 16 EPAs that have been developed by RCPSC for SF program and implemented to guide and plan the teaching and assessment activities during this stage of training (diagram1).

EPA assessment focuses on frequent direct and indirect observations of a resident's performance in authentic clinical activities by the clinical supervisors (63). This workplace-based assessment (WBA) facilitates timely and constructive feedback, thus guiding and facilitating the resident's learning process toward an independent practice (63). The Royal College established four observation templates for EPA assessments to assist the documentation of WBA: EPA observation (form1), procedural competencies (form 2), multi-source feedback (form 3), narrative feedback (form 4) (64) (templates 1-4). Although, each template has been designed to assess a specific set of EPAs with different milestone anchors and assessment tools, all of these templates encourage narrative feedback, support WBA and provide the competence committee with sufficient information to reach the final decision.

In section 2, the scoping review explored how EPAs have been assessed, types of assessed EPAs and the various assessment methods including the assessment settings, tools and assessor roles. Since the scoping review provided an overview of EPA assessments, this study aimed to explore how EPAs have been assessed within a university SF program. It sought to identify the different types of skills assessed, the type and number of submitted EPA assessment forms, the utilized assessment tools, and the roles of raters who conducted the assessment.

## 3.3 Materials and methods:

### 3.3.1 Ethics approval:

This Study was approved by our institutional ethics research board.

IRB study number: A10-E66-19A (19-10-031).

### 3.3.2 Participants:

Thirteen surgical residents enrolled in the SF program participated in the study. Informed consent was obtained from all participants to access their EPA assessment data (form3). In order to obtain the consent, we attended the SF classes twice to explain the study objectives and methods and to clarify any queries about the study. Furthermore, consent forms included two detailed pages explaining about the study objectives and procedures. All the EPA assessments were conducted at McGill University Health Center (MUHC) and the other McGill affiliated hospitals.

## 3.3.3 Methods:

All the EPA assessment data was requested by filling out a data request form on the medical education system website at McGill University https://www.mcgill.ca/meded-systems/data-request-form. The requested information was about the resident characteristics (name, gender, PGY level, surgical subspecialty, assigned rotation at time of assessment), EPA related characteristics (EPA form name and code, type of EPA: technical or non-technical), assessment-related characteristics (assessment setting, date and place of assessment, type of assessment form, number of completed assessment per EPA, milestones and overall assessment's scores), type of rater who assessed the resident (attending physician, resident, fellow, nurse, etc.).

All the data was provided by the data analyst (HT) in a password protected excel sheet format. The study included EPA assessment data completed between July 1st, 2019, and August 5th, 2020, and belongs to residents from two CBD stages: Transition to discipline and Foundation of discipline. After receiving the data, all resident names were anonymized by the researcher, and each resident and evaluator was assigned a specific identifier. No instructors or administrators have access to any names or have any method of identifying a participant or an evaluator.

## 3.4 Results:

From July 2019 to August 2020, a total of 595 EPA assessment forms were submitted by 13 residents in the SF program. More than half of the collected EPA assessment (377, 63%) was completed during the Foundation of discipline stage, compared with 37% of assessment completed during Transition of discipline stage. The total assessed EPAs were 16, with 7 belonging to Transition to discipline and 9 to Foundation of discipline stage. Out of 16, twelve EPAs (75%) had two components of technical and non-technical skills, while four EPAs were about non-technical skills only. This categorization of EPA components was based on the RCPSC's description of individual EPA and the associated milestones.

All the participants were in the first year of residency training, with 9 (69%) males, and 4 (31%) females. According to surgical subspecialties, the distribution of residents was as follows: 4 General Surgeons (GS), 3 from Obstetrics and Gynecology (OB/GYN), 2 from Orthopedic Surgery, 2 from Urology, 1 from Plastic Surgery, and 1from Cardiac Surgery (table 6).

The vast majority of the received assessment data was conducted in a clinical setting 566 (95%). Apart from 11 data points, we could not have access to the assessment data carried out in a simulated environment. Moreover, the assessment setting was not specified in eighteen assessment forms. Although the clinical assessment venues were not specified in 326 out of 566 assessment forms, six clinical settings were identified: day clinic, emergency room, intensive care unit (ICU), inpatient, outpatient clinic and operating room (OR). Inpatient and OR were the most prevalent clinical assessment venues with 94 and 93 assessment forms, respectively.

There was a significant variation in the total completed assessment forms per resident as well as the number of filled assessment forms per EPA. The mean filled assessment forms per resident was 45, the median was 48, and the maximum number of submitted forms was 66, while the lowest was 14 (diagram 2). The EPA 2.1(providing initial management for critically ill surgical patients) was the most evaluated EPA, followed by EPA 2.7 (managing uncomplicated postoperative surgical patients), and then EPA 2.5 (demonstrating the fundamental aspects of surgical procedures). All these EPAs belonged to the foundation of discipline stage. Moreover, the variation was also observed between the surgical subspecialties. For instance, 4 residents from GS completed 135 assessment forms compared with 161 forms completed by 3 residents from OB/GYNs (table 6).

There were three EPA assessment templates identified: EPA observation (form 1), procedural competencies (form 2), and multi-source feedback (form 3). All the EPAs within Transition of discipline were assessed using form 1. In contrast, multiple forms have been used to assess EPAs in Foundation of discipline as follows: 5 EPAs using form 1, 1 EPA using form 2, 1 EPA using forms 1 and 2, and 2 EPAs using forms 1 and 3.

Form 1 was the most prevalent assessment form, used in 80% (475) of total EPA observations compared with 16% (95) and 3.5% (21) of observations using forms 2 and 3, respectively. Both forms 1 and 2 have a 5- level entrustment scale with descriptive entrustment anchors. However, form 3 has a global rating scale where the EPA is assessed on a 5-point scale with descriptive anchors ranged from “not observed” to “always”. The majority of observations (94%) contained narrative feedback (two or more words written in the comment section). Assessment form 4 was not used in any of the received assessment data since it was not recommended in the assessment plan provided by RCPSC.

The lowest assigned overall entrustment level was level 2 (I had to talk them through), and the highest was level 5 (I did not need to be involved). According to the RCPSC recommendations, the EPA is considered achieved if the overall score is level 4 or 5. Based on this recommendation, about 85% of EPA assessments were achieved as the assigned entrustment levels were 4 and 5. In contrast, 14.8% of submitted EPA assessments were not achieved (level 2 and 3) (table 7). EPA 2.1b: Critically Ill Surgical and EPA 2.5a: Fundamental Surgical Procedures, were frequently assigned low entrustment levels.

Three rater types were identified: attending physician, resident, and fellow. About 69% (411) of assessment forms were filled by residents, 28% (165) by the attending physicians, and 3% (19) by fellows. The residents assessed 77% of clinical observations in Transition of discipline and 64% of Foundation of discipline compared with 22% and 31% of assessment completed by the attending physician in Transition of discipline and of Foundation of discipline, respectively.

## 3.5 Discussion:

This study explored how EPAs were assessed among the SF residents at McGill University during the period between July 1$^{st}$, 2020, and August 3$^{rd}$, 2021. It identified the assessed components of EPA and described the assessment plan of each EPA designed by

RCPS.As per RCPS recommendations, each resident has to fill out a specific number of assessment forms for each EPA within the 16 block period of the SF training program (62). The assessed EPA is considered achieved if the overall score in the submitted assessment form was 4 or higher.

This study revealed a considerable variation in the number of completed EPA assessments per EPA as well as between residents and surgical sub-specialties. Although RCPS has advised a specific number of EPA assessment forms to be completed before the end of 16 months of training, many residents are still lagging with many overdue EPA assessments. This variability in the number of submitted assessment forms can be attributed to many reasons. For example, the residents might not submit the non-achieved EPA assessments or document their performance due to the busy practice or staff burden. In addition, to complement our analysis of clinical assessments, more details about the residents' performance could potentially emerge from a future analysis of simulation-based assessment data.

Furthermore, some residents, like those in OB/GYN, needed to fill out more assessment forms to achieve the passing entrustment level. We observed that of the 161 assessment forms collected by OB/GYN residents, 35 were not achieved compared with 11 non-achieved assessments out of 135 forms completed by GS residents.

It is worth mentioning that the COVID-19 pandemic within the period from March to August 2020 might impact the SF training, as clinical opportunities decreased to perform the procedures and fulfill the RCPS requirements.

Regarding the assessment tools, this study identified two assessment metrics that were implemented into the assessment forms recommended by the RCPS: the Entrustability Scale and the Global Rating Scale. According to our scoping review, both scales were utilized in the clinical assessment of the EPA in the published results. Usage of GRS as an assessment tool enabled the assessment of all EPA types using the same scale without customization. Nevertheless, the assessors need to be trained on how to use the scale.

Analysis of the overall performance revealed that 85.2% of EPA assessments were achieved (assigned entrustment level were 4 or 5). This data represents the EPA assessed using the Entrustability scale. Notably, three EPAs (EPA 2.1b, EPA 2.5a, EPA 2.6) were frequently assigned low ratings. These three EPAs belonged to the Foundation of discipline stage and their associated milestones involve big chunks of fundamental procedural skills required to perform the task. The involvement of technical skill components might explain the struggle by some residents to achieve these three EPAs, because acquisition of technical skills is a prolonged process where the trainee goes through three distinct stages, as described in the 3-stage Motor skill learning theory described by Fitts and Posner (1967) (65). According to Motor skill learning theory, there are three phases to acquire motor skills: the cognitive phase, the associative phase, and the autonomous phase. The cognitive stage is the initial stage where the learner gathers knowledge about the task and tries to have an overall understanding of the individual steps required to perform the procedure. In the subsequent associative phase, the learner moves from the ‘what to do’ to ‘how to do" stage and demonstrate an ability to practice the task facilitated by feedback. Finally, in the autonomous phase, executing the skill becomes automatic with less cognitive involvement. Therefore, the achievement of EPA with technical skills needs more practice when workplace training may not be enough. Simulation training could help the residents by providing more opportunities for practice and facilitating the achievement of this EPA.

Moreover, our study of assessment in the McGill SF program has shown that residents were the main EPA assessors in both Transition to discipline and Foundation of discipline stages. In contrast, the published results in our scoping review revealed that attending physicians were the prevalent raters of EPAs. This discrepancy can be explained by RCPS recommendations regarding the assessor's roles described in the EPA assessment plans for both Transition to discipline and Foundation of discipline stages. RCPS described the assessment plan of each EPA and the rater's role who should perform the assessment. All the Transition to discipline EPAs were advised to be assessed by the supervisor without specifying the role. However, the assessor roles (faculty, junior or senior resident, nurse. etc.) were specified in some EPAs belonging to the Foundation of the discipline stage. Therefore, since the supervisor roles were not specified in many EPA assessment plans for both Transition to discipline and Foundation of discipline, the majority of residents preferred to choose their colleagues to assess them. This preference could possibly be attributed to the interpersonal relationships between residents and busy or unavailable staff. Hauer et al. have described how supervisor characteristics and supervisor-trainee relationship can influence the entrustment decision process (17). Furthermore, Green et al. highlighted that successful workplace assessment requires a trained and an experienced assessor to evaluate the trainee in an authentic clinical environment (30). Moreover, supervisor experience impacts the degree of autonomy provided to the trainee, especially while performing complex or critical procedures, ultimately influencing the entrustment level (17).

The above observation about the EPA assessors raises the question as to whether the residents were trained to use the assessment tools, as the utilized tools were GRS requiring rater training (40). Rater training has been proven to increase the standardized interpretation of assessment tools (38). Moreover, it would be a good future direction to study the impacts of rater roles on the assigned overall performance.

## 3.6 Conclusion and future directions:

To conclude, the SF study demonstrated an apparent variability in the number of completed assessments per EPA as well as per resident. Our study also reveals the different assessment forms and tools that were utilized as per RCPS recommendations, as well as where the residents performed the majority of EPA assessments.

Exploring the reasons behind the variability in the completed EPA assessments would be a promising future direction, enabling medical education programs to ensure consistent training and assessment. Moreover, enrollment of more residents in the study or including two cohorts can explain more robustly the above observations. Furthermore, analyzing more assessment data will allow observing the trend of progression of each resident over time. Finally, simulation assessment data would enrich our understanding and implementation of EPA assessment within the SF program.

# Section 4: Thesis Summary and Conclusion

## 4.1 General Findings:

Assessment is a complex and crucial step in any learning process. In the medical field, it should reflect what the trainee learns and determines the readiness to be a safe health care professional capable of serving the community's needs. Although various assessment methods have been implemented to assess the trainee, in order to achieve the benefit from the assessment methods, the appropriate assessment activities should align with curriculum objectives and instructional strategies. CBME is a novel educational paradigm with comprehensive strategies to teach, train and assess the medical trainee. Even so, since it is a new model, the assessment strategies within this training paradigm are not yet standardized, with lots of debate around the appropriate assessment methods to assess competence and EPAs.

At the first step, we did a scoping review to enrich our understanding of EPA assessment within the CBME framework. This review identified the three aspects of assessed EPAs. Moreover, it has shown that diverse assessment tools have been utilized to assess the EPA within simulation and clinical settings. It highlighted that simulation complements clinical assessment since each setting can assess a different EPA element. Further research could be conducted to evaluate the utilized EPA assessment tools and assess their validity and reliability.

In section 3, we explored how EPAs are assessed within the SF training program at McGill University. We aimed to correlate what we found in our assessment data and what was published in the literature. As in our review, three aspects of EPA have been assessed as per RCPS guidelines. The Entrustability Scale and GRS were utilized to assess the residents within the SF program, which were also identified as assessment metrics according to the scoping review. The analysis of the rater roles in SF assessment data has shown that residents were the most prevalent assessors which was not the case according to our scoping review in which the attending physicians and the trained raters were the common assessors. It is well-known that raters play an essential role in EDM, which is a fundamental process in CBME assessment. Furthermore, Hauer et al. highlighted that rater role and interpersonal relationships might impact the final judgment about trainee's performance (17). Since the busy clinical practice might have hindered the attending physician from completing the assessment, simulation could ensure that both the attending physician and trainees have a dedicated time of learning and teaching. In addition, training the residents and fellows on how to use and interpret the Entrustabilty Scale could overcome this limitation. This discrepancy between the published result about the assessor's role and what we found in our SF study should be taken into account and studied thoroughly to ensure trainee competence. The RCPS of Canada could also consider reviewing the assessor role in the assessment plan belonging to each EPA.

## 4.2 Conclusion:

In conclusion, I have explored how EPAs have been assessed within CBME. This thesis provided an overall understanding of EPA assessments and highlighted various assessment tools which were utilized to assess trainees' competence. This thesis demonstrates the need for a unifying assessment plan to assess EPA, which can measure the trainee's readiness to be a safe health care professional capable of serving the community's needs.

# Appendices and Figures

**Surgical Foundations: Transition to Discipline EPA #6**

**Repairing simple skin incisions/lacerations**

Key Features:
- This EPA may be observed in simulation

Assessment plan:
Direct observation by supervisor

Use Form 1. Form collects information on:
- Type of scenario: clinical; simulation
- Wound size: < 2 cm, 2-5 cm, >5 cm

Collect 1 observation of achievement
- Wound must be at least 5 cm long

Relevant milestones

1 **TD ME 2.4.2 Use appropriate prophylaxis**
2 **TD ME 3.2.2 Obtain and document informed consent for simple wound closure**
3 **TD ME 3.4.4 Perform pre-procedural tasks for a simple wound closure**
- **Apply aseptic technique**
- **Gather and manage the availability of appropriate instruments and materials for minor procedures**
- **Obtain appropriate assistance**
- **Position the patient appropriately**
- **Prepare the operative site**
- **Hand-cleanse, gown and glove**
- **Demonstrate appropriate draping of the patient**
- **Deliver pre-procedural local anesthesia if appropriate**

4 **TD ME 3.4.5 Perform procedural tasks in a timely, skillful and safe manner**
- **Use common surgical instruments, including but not limited to needle drivers, retractors, forceps, clamps, and scissors**
- **Select and use suture materials**
- **Assess the quality of the closure**

5 **TD COM 3.1.4 Plan and discuss appropriate postoperative care and issues with patients and families**

**Form 1:** EPA#6: Repairing simple skin incisions/lacerations. Source: Royal College of Physician and Surgeon website.

**Surgical Foundations:** Foundations EPA #4

**Providing patient education and informed consent in preparation for surgical care**

Key Features:

- This EPA focuses on the communication that occurs with patients and families to inform and discuss plans for surgical care

Assessment plan:
Direct observation by supervisor

Use Form 1. Form collects information on:

- Type of procedure: emergency; elective
- Setting: clinical; simulation

Collect 3 observations of achievement

- At least 2 different assessors
- At least one emergency procedure
- At least one elective procedure
- At least two in clinical setting

Relevant milestones

1 **F ME 2.3.1** Work with patients and their families to understand relevant options for care

2 **F ME 2.4.2 Ensure that the patient and family are informed about the risks and benefits of each treatment option in the context of best evidence and guidelines, addressing fears and concerns**

3 **F ME 2.4.3** Discuss clinical uncertainty with the patient and family

4 **F ME 3.2.1 Obtain informed consent for commonly performed procedures and therapies, under supervision**

5 **F ME 3.2.2** Assess patients' decision-making capacity

6 **F ME 4.1.3** Apply standardized care paths, including patient education components

7 **F COM 1.2.1** Optimize the physical environment for patient comfort, privacy, engagement, and safety

8 **F COM 1.5.2** Manage challenging conversations

9 **F COM 1.6.1** Encourage discussion, questions, and interaction to validate understanding during the encounter

10 **F COM 3.1.2** Plan and discuss appropriate post-operative, immediate and/or long-term care and issues with patients and families as appropriate

11 **F COM 4.1.2 Communicate with cultural awareness and sensitivity**

12 **F COM 5.1.1 Document information about patients and their medical conditions in a manner that enhances intra- and interprofessional care**

13 **F HA 1.2.1** Select patient education resources related to surgical practice

14 **F P 1.3.2** Recognize and respond appropriately in situations where consent is obtained under constraints of emergency circumstances

15 **F P 1.4.3 Manage conflicts of interest related to surgical care, including consent issues related to the duality of the learner as surgeon**

**Form 2:** EPA#4 Providing patient education and informed consent in preparation for surgical care. Source: Royal College of Physician and Surgeon website.

## Consent form for participation
## Statement of Consent

I agree to take part in the described McGill University research. I have had the project explained to me, and I have read the consent form, which I keep for my records. Being part of the research is voluntary and I can leave at any stage without any consequences or penalties.

**Further use of data**

I agree that the information provided can be used in the context of research theses, conference presentations and/or publication in academic journals associated with the McGill University. Upon completion of this project, the researcher may want to use words and data collected from this project for other educational purposes including but not limited to presentations to peers at conferences, for further research or to students in lectures. By participating in this research, I consent to the following statements:

- I agree to participate in this research project **YES NO**
- I consent that my final results belonging to the final clinical and simulation assessment will be retrieved and used ONLY for research analysis purpose **YES NO**
- I consent to have the materials (assessment forms, videotape) related to my participation to be used for research analysis purpose **YES NO**
- I consent that the data and findings can be used in further research projects that have ethics approval. **YES NO**

**If you have any ethical concerns or complaints about your participation in this study, and want to speak with someone not on the research team, please contact the McGill REB Ethics Officer, Ms.Ilde Lepore, at 514-398-8302 or at <u>ilde.lepore@mcgill.ca</u>**

Student researcher: Hajar Mohammed Al-Mughairi ,hajar.al-mughairi@mail.mcgill.ca

<u>Supervisor:</u> Dr.Kevin Lachapelle , kevin.lachapelle@mcgill.ca

*Please sign below if you have read the above information and consent to participate in this study. Agreeing to participate in this study does not waive any of your rights or release the researchers from their responsibilities. A copy of this consent form will be given to you and the researcher will keep a copy.*

*Participant's Name:* ______________________________________________

*Participant's Signature:* ___________________________________

*Date:* ___________

**Thank you for your time**

**Form 3:** Consent form

## EPA OBSERVATION - TEMPLATE 1 (EXAMPLE)

Observation | Evidence & Reflection

Learner: HGranger Learner 934 - Foundations of Discipline

EPA Title: EPA 1: Assessing, diagnosing and initial management for patients with com-

Key Features: Select from list

EPA Stage: Foundations of Discipline    Date of Observation: 11/10/2016

| Type of Assessment: | Location of patient visit: | Case mix: |
|---|---|---|
| Direct observation | Emergency Department | Chest pain |
| Context #4: | Context #5: | Complexity: |
| | | |

Additional Context Information:

Based on this observation overall: ○ 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

Milestones associated with this EPA:

| | Not observed | InProgress | Achieved |
|---|---|---|---|
| ME 1.3.1 Apply clinical and biomedical sciences to manage core patient presentations in Interanal Medicin (Modified) | ○ | ● | ○ |
| ME 1.4 Perform complete and appropriate assessments of patients with common acute medical presentation | ● | ○ | ○ |
| ME 2.2.1 Generate differential diagnosis along with appropriate diagnostic strategies | ○ | ● | ○ |
| ME 3.4.2 See assistance as needed when unanticipated findings or changing clinical circumstances are encountered | ○ | ● | ○ |
| COM 2.1.1 Conduct a patient-centred interview, gathering all relevant biomedical and psychosocial informationf for any clinical presentation | ○ | ○ | ● |

SAMPLE

Feedback to Resident and Competence Committee:

Interview skills have significantly improved and the flow of the interaction is now much smoother. Going forward, work on application of potential differential diagnoses and potential diagnostic tests to schieve a definitive cause of the chest pain.

Professionalism and Patient Safety:

Do you have any concerns regarding this Learner's professionalism? ● No ○ Yes

Do you have any concerns regarding Patient Safety? ● No ○ Yes

If yes, description of concern:

Close | Save | Next | Clone | Submit

Note: This document is a representation of a CBD observation form available in the Royal College's ePortfolio system.



**Template 1:** EPA observation

## PROCEDURAL COMPETENCIES - TEMPLATE 2 (EXAMPLE)

Observation | Evidence & Reflection

Learner: HGranger Learner 934 - Foundations of Discipline

EPA/IM Title: EPA 4: Perform a lumbar puncture

Key Features: Select from list

EPA Stage: Foundations of Discipline    Date of Observation: 11/10/2016

Complexity:

Based on this observation overall: ○ 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

Milestones associated with this EPA:

| | 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 |
|---|---|---|---|---|---|
| 1.1 Informed consent obtained and placed on chart | ○ | ○ | ○ | ● | ○ |
| 1.2 Appropriate lumbar puncture kit obtained, and set up properly using sterile technique | ○ | ○ | ● | ○ | ○ |
| 1.3 Time out performed prior to procedure | ○ | ○ | ● | ○ | ○ |
| 1.4 Patient placed in appropriate position and landmarks identified | ○ | ○ | ● | ○ | ○ |
| 1.5 Site cleaned using proper technique | ○ | ○ | ● | ○ | ○ |

(Additional milestones)



Feedback to Resident and Competence Committee:

Second lumbar puncture completed by resident. Required prompting through the procedure.
Some difficulty locating L3 - L4 - L5 interspaces following the palpation of the posterial iliac crests.
For the next LP, I want you to work on techniques to access the CSF.

Professionalism and Patient Safety:

Do you have any concerns regarding this Learner's professionalism? ● No ○ Yes

Do you have any concerns regarding Patient Safety?

If yes, description of concern: ● ○ Yes

Close | Save | Next | Clone | Submit

Note: This document is a representation of a CBD observation form available in the Royal College's ePortfolio system.



**Template 2:** Procedural competencies

**MULTIPLE SOURCE FEEDBACK (MSF) - TEMPLATE 3 (EXAMPLE)**

Observation | Evidence & Reflection

Learner: HGranger Learner 934 ☑ Date of Observation: 11/10/2016

Framing

This template is intended to capture your Observation Rating of a learner, based on your multiple encounters with that learner over time. Observation Ratings are provided anonymously and collated prior to presentation to the learner. Please complete this observation within 2 weeks of receiving this request. If you require further assistance, please contact the Program Administrator.

The following Milestones were demonstrated:



| | Not observed | Never | Sometimes | Usually | Always |
|---|---|---|---|---|---|
| Recognize when to seek help in providing clear explanations to the patient and family | ○ | ○ | ○ | ◉ | ○ |
| Synthesize patient information including symptoms, differential diagnosis, and treatment plan clearly and concisely | ◉ | ○ | ○ | ○ | ○ |
| Identify patients requiring handover to other physicians or health care professionals | ○ | ○ | ○ | ◉ | ○ |
| Elicit a basic history | ○ | ○ | ○ | ○ | ◉ |

Feedback to Resident and Competence Committee:

Resident is able to provide information to patients and their families in a clear and accurate manner. Further development may be required in order to 'check-in', verify, and validate as to whether there is complete understanding with regard to diagnosis, prognosis, and management plan.

Professionalism and Patient Safety:

| | | |
|---|---|---|
| Do you have any concerns regarding this Learner's professionalism? | ◉ No | ○ Yes |
| Do you have any concerns regarding Patient Safety? If yes, description of concern: | ◉ No | ○ Yes |

Close | Save | Next | Clone | Submit

Note: This document is a representation of a CBD observation form available in the Royal College's ePortfolio system.



**Template 3:** Multi-source feedback

**NARRATIVE OBSERVATION - TEMPLATE 4 (EXAMPLE A)**

Narrative | Evidence

LearnerStage: Foundations of Discipline

Learner: HGranger Learner 934

Date of Observation: 11/10/2016

Feedback to Resident and Competence Committee:

Tendency to use too much medical jargon when explaining issues to patients. Ex. In the patient with an abnormal lesion on the ches x-ray you said, "It could be an infiltrate, a granuloma, a malignancy..." In the future you should consider the patient's background, unless they work in healthcare they are probably unfamiliar with those types of words. You need to find ways to explain things without using medical jargon, for example, a malignancy would be better understood as cancer for many people.



Professionalism and Patient Safety:

Do you have any concerns regarding this Learner's professionalism? ◉ No ○ Yes

Do you have any concerns regarding Patient Safety? ◉ No ○ Yes

If yes, description of concern:

Close | Save | Next | Clone | Submit

Note: This document is a representation of a CBD observation form available in the Royal College's ePortfolio system.



**Template 4:** Narrative observation

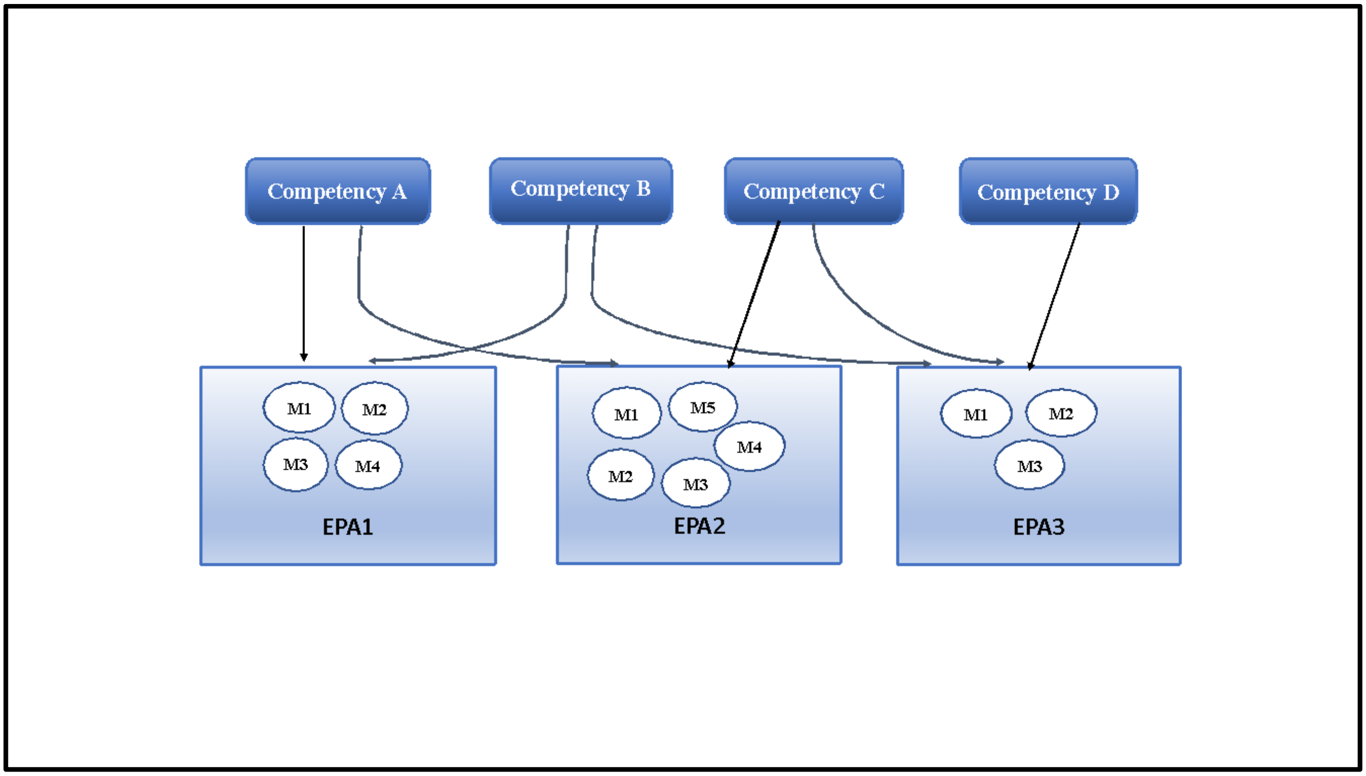


**Figure1:** Illustrates the relationship between the Competency, Entrustable Professional Activity (EPA) and milestone (M).

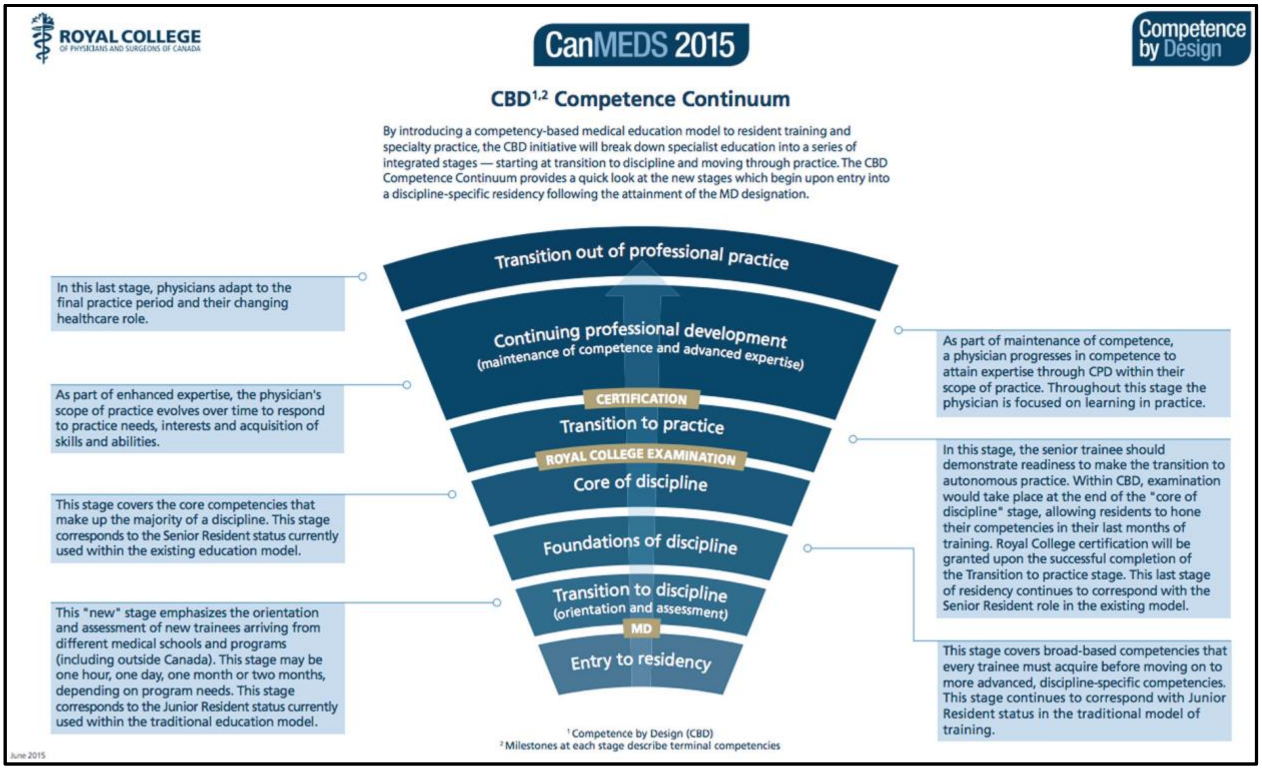


**Figure 2:** Competency by design model.

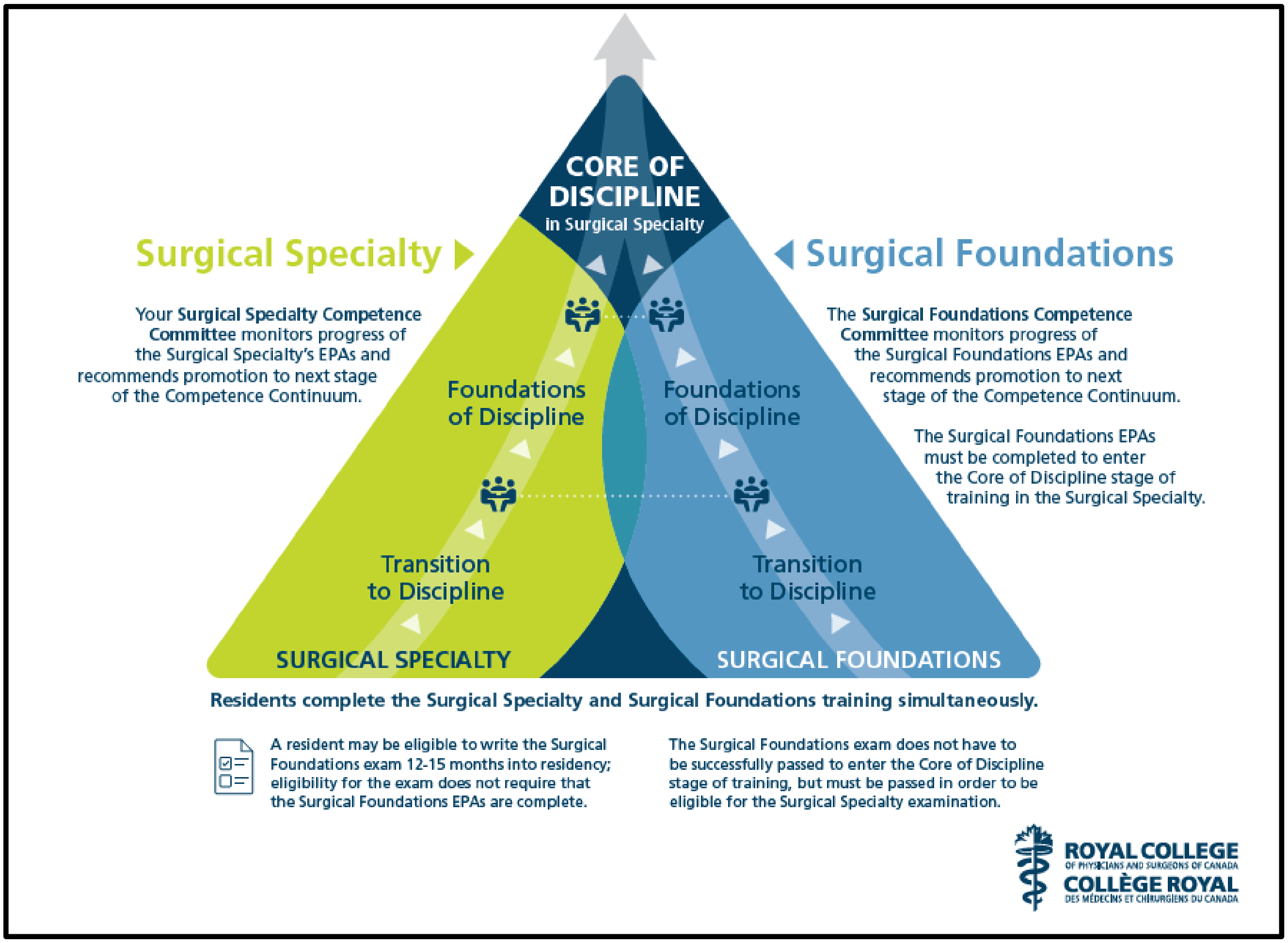


**Figure 3:** Surgical Foundations within CBD.

## O-SCORE Entrustability Scale

| Level | Descriptor |
|---|---|
| 1 | **"I had to do"**<br>i.e., requires complete hands on guidance, did not do, or was not given the opportunity to do |
| 2 | **"I had to talk them through"**<br>i.e., able to perform tasks but requires constant direction |
| 3 | **"I had to prompt them from time to time"**<br>i.e., demonstrates some independence, but requires intermittent direction |
| 4 | **"I needed to be in the room just in case"**<br>i.e., independence but unaware of risks and still requires supervision for safe practice |
| 5 | **"I did not need to be there"**<br>i.e., complete independence, understands risks and performs safely, practice ready |

Gofton WT, Dudek NL, Wood TJ, Balaa F, Hamstra SJ. The Ottawa surgical competency operating room evaluation (O-SCORE): a tool to assess surgical competence. Acad Med. 2012;87(10);1401-7. http://journals.lww.com/academicmedicine/Fulltext/2012/10000/The_Ottawa_Surgical_Competency_Operating_Room.24.aspx Reproduced with permission of the authors.

**Figure 5**: Demonstrates a 5-level Entrustability Scale.

**Transition to Discipline**

- **EPA1**: Performing the preoperative preparation of patients for basic surgical procedures
- **EPA2:** Recognizing and initiating early management for critically ill surgical patients
- **EPA3:** Documenting clinical encounters
- **EPA4:** Demonstrating handover technique
- **EPA5:** Demonstrating ability to function in the operating room
- **EPA6:** Repairing simple skin incisions/lacerations
- **EPA7**: Managing tubes, drains and central lines

**Foundations of Discipline**

- **EPA1:** Providing initial management for critically ill surgical patients
- **EPA2:** Providing initial management for trauma patients
- **EPA3:** Assessing and performing risk optimization for preoperative patients in preparation for surgery
- **EPA4:** Providing patient education and informed consent in preparation for surgical care
- **EPA5:** Demonstrating the fundamental aspects of surgical procedures
- **EPA6:** Participating in surgical procedures
- **EPA7:** Managing uncomplicated postoperative surgical patients
- **EPA8:** Managing postoperative patients with complications
- **EPA9:** Supervising junior learners in the clinical setting

**Diagram1:** List of EPAs associated with Transition of discipline and Foundation of discipline. (Source: Royal College website).

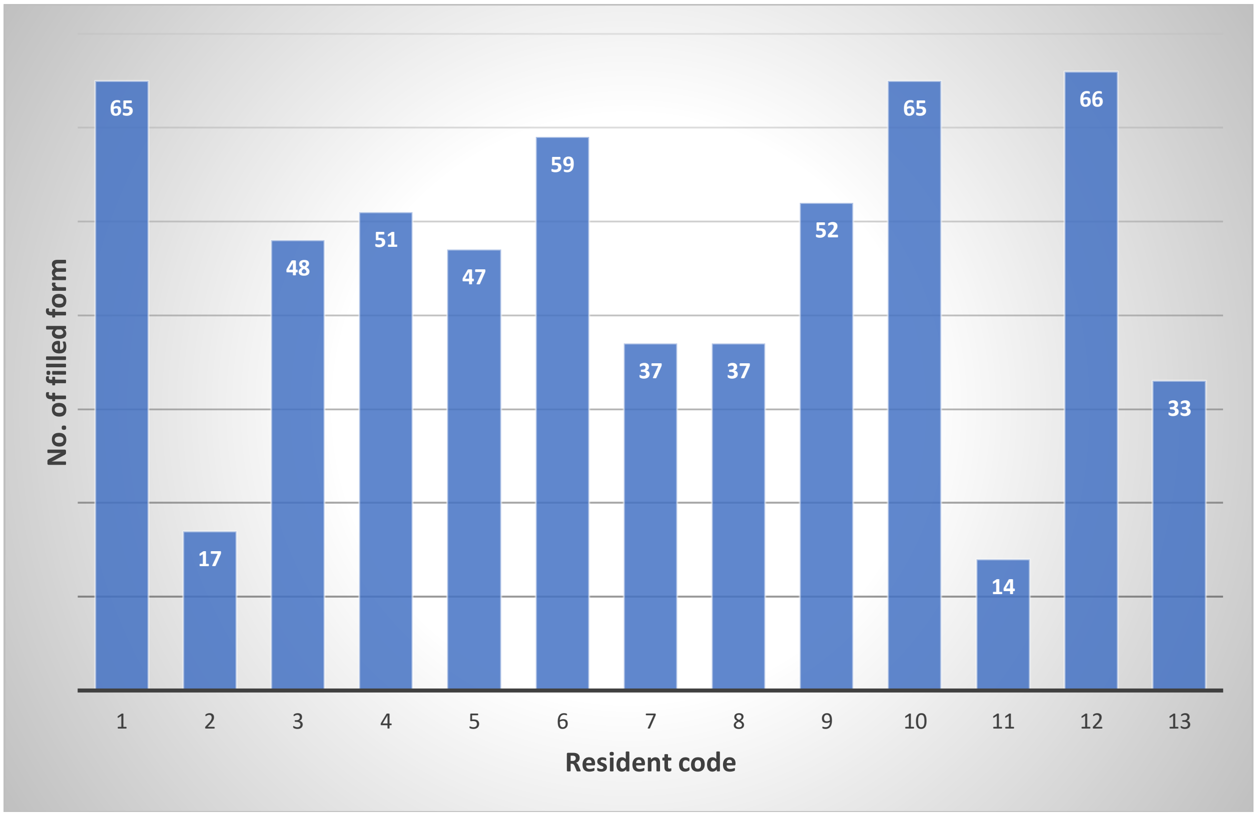


**Diagram 2:** Illustrates the number of filled EPA forms per resident.

| A Comparison of the Elements of Structure- and Process-based Versus Competency-based Educational Programs | | |
|---|---|---|
| | Educational Program | |
| Variable | Structure- and Process-based | Competency-based |
| Driving force for curriculum | Content—knowledge acquisition | Outcome—knowledge application |
| Driving force for process | Teacher | Learner |
| Path of learning | Hierarchical (teacher ⇒ student) | Non-hierarchical (teacher ⇔ student) |
| Responsibility for content | Teacher | Student and teacher |
| Goal of educational encounter | Knowledge acquisition | Knowledge application |
| Typical assessment tool | Single subjective measure | Multiple objective measures (“evaluation portfolio”) |
| Assessment tool | Proxy | Authentic (mimics real tasks of profession) |
| Setting for evaluation | Removed (gestalt) | “In the trenches” (direct observation) |
| Evaluation | Norm-referenced | Criterion-referenced |
| Timing of assessment | Emphasis on summative | Emphasis on formative |
| Program completion | Fixed time | Variable time |

**Table1:** Illustrates comparison between the traditional educational model and CBME model.

| Study characteristics | | |
|---|---|---|
| **Country** | | **Number** |
| Canada (11) | Hamilton | 1 |
| | Kingston | 4 |
| | Ottawa | 1 |
| | Montreal | 1 |
| | Toronto | 2 |
| | Multicentre | 2 |
| USA | | 12 |
| USA and Canada | | 1 |
| Australia | | 1 |
| United Kingdom | | 2 |
| India | | 1 |
| Total | | **28** |
| **Year of publication** | | |
| 2000-2010 | | 0 |
| 2010-2015 | | 4 |
| 2015-2020 | | 24 |

**Table 3:** Demonstrates the study characteristics.

| **Specialty /program** | **No. of studies** |
|---|---|
| Foundations of discipline | 1 |
| ENT | 2 |
| General surgery | 3 |
| Internal medicine | 4 |
| Emergency medicine | 5 |
| Obstetrics and Gynecology | 2 |
| Psychiatry | 1 |
| General pediatric | 2 |
| Orthopedic | 1 |
| Primary care | 2 |
| Dermatology | 1 |
| Multiple specialties | 4 |
| **Total** | **28** |

**Table 4**: Shows the distribution of the study according to the participants' background.

| Assessment setting | No. of assessed EPAs | Technical skills | Non-technical skills | Technical and non-technical together |
|---|---|---|---|---|
| **Clinical-based assessment** | 86 (69%) | 12 (14%) | 70 (81.4%) | 4 (4.6%) |
| **Simulation-based assessment** | 38 (31%) | 14 (37%) | 16 (42%) | 8 (21%) |
| **Total** | 124 | 26 | 86 | 12 |

**Table 5:** Distribution of assessment setting according to EPA type.

| Specialty | No. of residents | Total EPA assessments | No. of assessments rated < level 4 | Average completed assessment per resident |
|---|---|---|---|---|
| General Surgery | 4 | 135 | 11 | 33 |
| Obstetrics and Gynecology | 3 | 161 | 35 | 54 |
| Orthopedic | 2 | 102 | 9 | 51 |
| Urology | 2 | 118 | 22 | 59 |
| Cardiac Surgery | 1 | 33 | 3 | 33 |
| Plastic Surgery | 1 | 48 | 5 | 48 |
| Total | 13 | 595 | 85 | |

**Table 6:** Distribution of residents according to surgical subspecialty and the total completed assessment forms.

| **Entrustment level** | **Frequency** |
|---|---|
| Level 1 | 0 |
| Level 2 | 16 (2.8%) |
| Level 3 | 69 (12%) |
| Level 4 | 273 (47.6%) |
| Level 5 | 216 (37.6%) |
| **Total** | **574** |

**Table 7**: Analysis of overall performance according to the assigned entrustment level.

# References

1. Jason R. Frank, Linda S. Snell, Olle Ten Cate, Eric S. Holmboe, Carol Carraccio, Susan R. Swing, Peter Harris, Nicholas J. Glasgow, Craig Campbell, Deepak Dath, Ronald M. Harden, William Iobst, Donlin M. Long, Rani Mungroo, Denyse L. Richardson, Jonathan Sherbino, Ivan Silver, Sarah Taber, Martin Talbot & Kenneth A. Harris (2010). Competency-based medical education: theory to practice, Medical Teacher, 32:8, 638-645.

2. Frank, J. R., Snell, L., Englander, R., Holmboe, E. S., & ICBME Collaborators (2017). Implementing competency-based medical education: Moving forward. Medical teacher, 39(6), 568–573.

3. Carraccio, C., Wolfsthal, S. D., Englander, R., Ferentz, K., & Martin, C. (2002). Shifting paradigms: from Flexner to competencies. Academic medicine: journal of the Association of American Medical Colleges, 77(5), 361–367.

4. Van Melle, E., Frank, J. R., Holmboe, E. S., Dagnone, D., Stockley, D., Sherbino, J., & International Competency-based Medical Education Collaborators (2019). A Core Components Framework for Evaluating Implementation of Competency-Based Medical Education Programs. Academic medicine: journal of the Association of American Medical Colleges, 94(7), 1002–1009.

5. Englander, R., Frank, J. R., Carraccio, C., Sherbino, J., Ross, S., Snell, L., & ICBME Collaborators (2017). Toward a shared language for competency-based medical education. Medical teacher, 39(6), 582–587.

6. Royal College of Physicians and Surgeons of Canada. http://www.royalcollege.ca/rcsite/canmeds/canmeds-framework-e. Accessed on 24 August 2020.

7. Veale, P., Busche, K., Touchie, C., Coderre, S., & McLaughlin, K. (2019). Choosing Our Own Pathway to Competency-Based Undergraduate Medical Education. Academic medicine: journal of the Association of American Medical Colleges, 94(1), 25–30.

8. Cate O. T. (2018). A primer on entrustable professional activities. Korean journal of medical education, 30(1), 1–10.

9. Ten Cate O. (2013). Nuts and bolts of entrustable professional activities. Journal of graduate medical education, 5(1), 157–158.

10. Carraccio, C., Englander, R., Holmboe, E. S., & Kogan, J. R. (2016). Driving Care Quality: Aligning Trainee Assessment and Supervision Through Practical Application of Entrustable Professional Activities, Competencies, and Milestones. Academic medicine: journal of the Association of American Medical Colleges, 91(2), 199–203.

11. Nguyen, V. T., & Losee, J. E. (2016). Time- versus Competency-Based Residency Training. Plastic and reconstructive surgery, 138(2), 527–531.

12. Pinsk, M., Karpinski, J., & Carlisle, E. (2018). Introduction of Competence by Design to Canadian Nephrology Postgraduate Training. Canadian journal of kidney health and disease, 5, 2054358118786972.

13. Competence by Design. The Royal College of Physicians and Surgeons of Canada, 2017. Available at: http://www.royalcollege.ca/rcsite/cbd/competence-by-design-cbd-e. Accessed on July 1, 2020.

14. Royal College of Physicians and Surgeons of Canada. http://www.royalcollege.ca/rcsite/cbd/assessment/committees/managing-competence-committees-e
Accessed on August 31, 2020.

15. Ahle, S. L., Gielissen, K., Keene, D. E., & Blasberg, J. D. (2020). Understanding Entrustment Decision-Making by Surgical Program Directors. The Journal of surgical research, 249, 74–81.

16. Ten Cate, O., Hart, D., Ankel, F., Busari, J., Englander, R., Glasgow, N., Holmboe, E., Iobst, W., Lovell, E., Snell, L. S., Touchie, C., Van Melle, E., Wycliffe-Jones, K., & International Competency-Based Medical Education Collaborators (2016). Entrustment Decision Making in Clinical Training. Academic medicine: journal of the Association of American Medical Colleges, 91(2), 191–198.

17. Hauer, K. E., Ten Cate, O., Boscardin, C., Irby, D. M., Iobst, W., & O'Sullivan, P. S. (2014). Understanding trust as an essential element of trainee supervision and learning in the workplace. Advances in health sciences education: theory and practice, 19(3), 435–456.

18. Carraccio, C., Englander, R., Van Melle, E., Ten Cate, O., Lockyer, J., Chan, M. K., Frank, J. R., Snell, L. S., & International Competency-Based Medical Education Collaborators (2016). Advancing Competency-Based Medical Education: A Charter for Clinician-Educators. Academic medicine: journal of the Association of American Medical Colleges, 91(5), 645–649.

19. Wang, P., Chan, E., Forster, A., De Dios, J. V., Sener, A., Dave, S., & Chahine, S. (2018). Perceptions on Competence by Design in urology. Canadian Urological Association journal = Journal de l'Association des urologues du Canada, 13(7), E183–E189.

20. Carraccio, C. L., & Englander, R. (2013). From Flexner to competencies: reflections on a decade and the journey ahead. Academic medicine: journal of the Association of American Medical Colleges, 88(8), 1067–1073.

21. van der Vleuten, C. P., & Schuwirth, L. W. (2005). Assessing professional competence: from methods to programmes. Medical education, 39(3), 309–317.

22. Malone, K., & Supri, S. (2012). A critical time for medical education: the perils of competence-based reform of the curriculum. Advances in health sciences education: theory and practice, 17(2), 241–246.

23. Holmstrom, A. L., & Meyerson, S. L. (2019). Obtaining Meaningful Assessment in Thoracic Surgery Education. Thoracic surgery clinics, 29(3), 239–247.

24. Humphrey-Murto, S., Wood, T. J., Ross, S., Tavares, W., Kvern, B., Sidhu, R., Sargeant, J., & Touchie, C. (2017). Assessment Pearls for Competency-Based Medical Education. Journal of graduate medical education, 9(6), 688–691.

25. Ten Cate O. (2005). Entrustability of professional activities and competency-based training. Medical education, 39(12), 1176–1177.

26. Holmboe E. S. (2017). Work-based Assessment and Co-production in Postgraduate Medical Training. GMS journal for medical education, 34(5), Doc58.

27. Norcini, J., & Burch, V. (2007). Workplace-based assessment as an educational tool: AMEE Guide No. 31. Medical teacher, 29(9), 855–871.

28. Govaerts, M. J., van der Vleuten, C. P., Schuwirth, L. W., & Muijtjens, A. M. (2007). Broadening perspectives on clinical performance assessment: rethinking the nature of in-training assessment. Advances in health sciences education: theory and practice, 12(2), 239–260.

29. Fitzgerald, J. T., Burkhardt, J. C., Kasten, S. J., Mullan, P. B., Santen, S. A., Sheets, K. J., Tsai, A., Vasquez, J. A., & Gruppen, L. D. (2016). Assessment challenges in competency-based education: A case study in health professions education. Medical teacher, 38(5), 482–490.

30. Green, M. L., & Holmboe, E. (2010). Perspective: the ACGME toolbox: half empty or half full? Academic medicine: journal of the Association of American Medical Colleges, 85(5), 787–790.

31. Chauvin S. W. (2015). Applying Educational Theory to Simulation-Based Training and Assessment in Surgery. The Surgical clinics of North America, 95(4), 695–715.

32. Sullivan, A., Summer, E., Ades, A., & Sawyer, T. (2019). Acquiring and maintaining technical skills using simulation: Initial, maintenance, booster, and refresher training. Cureus, 11(9),e5729.

33. Scalese, R. J., Obeso, V. T., & Issenberg, S. B. (2008). Simulation technology for skills training and competency assessment in medical education. Journal of general internal medicine, 23 Suppl 1(Suppl 1), 46–49.

34. Walsh K. (2013). Simulation: the need for a balanced view. Journal of biomedical research, 27(3), 243–244.

35. McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation-based medical education research: 2003–2009. Medical Education. 2010;44:50–63.

36. Wu KY, Kim S, Fung K, Roth K. Assessing nontechnical skills in otolaryngology emergencies through simulation-based training. Laryngoscope. 2018;128(10):2301–2306.

37. Bruno, C., Levit, O., Gangadharan, S., Scherzer, D., Moussa, A., Auerbach, M., & INSPIRE Research Network (2019). Neonatal Intubation Competency Assessment Tool: Development and Validation. Academic pediatrics, 19(2), 157–164.

38. Rekman, J., Gofton, W., Dudek, N., Gofton, T., & Hamstra, S. J. (2016). Entrustability Scales: Outlining Their Usefulness for Competency-Based Clinical Assessment. Academic medicine: journal of the Association of American Medical Colleges, 91(2), 186–190.

39. Ten Cate O. (2016). Entrustment as Assessment: Recognizing the Ability, the Right, and the Duty to Act. Journal of graduate medical education, 8(2), 261–262.

40. Adler, M. D., Vozenilek, J. A., Trainor, J. L., Eppich, W. J., Wang, E. E., Beaumont, J. L., Aitchison, P. R., Pribaz, P. J., Erickson, T., Edison, M., & McGaghie, W. C. (2011). Comparison of checklist and anchored global rating instruments for performance rating of simulated pediatric emergencies. Simulation in healthcare: journal of the Society for Simulation in Healthcare, 6(1), 18–24.

41. O'Dowd, E., Lydon, S., O'Connor, P., Madden, C., & Byrne, D. (2019). A systematic review of 7 years of research on entrustable professional activities in graduate medical education, 2011-2018. Medical education, 53(3), 234–249.

42. Peters, M. D., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. International journal of evidence-based healthcare, 13(3), 141–146.

43. Munn, Z., Peters, M., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC medical research methodology, 18(1), 143.

44. EPAs and CanMEDS milestones. Royal College of Physicians and Surgeons of Canada. https://www.royalcollege.ca/rcsite/cbd/implementation/cbd-milestones-epas-e. Accessed on April 13, 2021.

45. Quraishi, S., Wade, W., & Black, D. (2019). Development of a GMC aligned curriculum for internal medicine including a qualitative study of the acceptability of 'capabilities in practice' as a curriculum model. Future healthcare journal, 6(3), 196–203.

46. Warm, E. J., Mathis, B. R., Held, J. D., Pai, S., Tolentino, J., Ashbrook, L., Lee, C. K., Lee, D., Wood, S., Fichtenbaum, C. J., Schauer, D., Munyon, R., & Mueller, C. (2014). Entrustment and mapping of observable practice activities for resident assessment. Journal of general internal medicine, 29(8), 1177–1182.

47. Chaplin, T., Egan, R., Cofie, N., Gu, J. J., McColl, T., & Thoma, B. (2018). The Implementation of a Multi-institutional Multidisciplinary Simulation-based Resuscitation Skills Training Curriculum. *Cureus*, *10*(11), e3593.

48. Dagnone, J. D., Hall, A. K., Sebok-Syer, S., Klinger, D., Woolfrey, K., Davison, C., Ross, J., McNeil, G., & Moore, S. (2016). Competency-based simulation assessment of resuscitation skills in emergency medicine postgraduate trainees - a Canadian multi-centred study. Canadian medical education journal, 7(1), e57–e67.

49. Karthikeyan, P., & Pulimoottil, D. T. (2019). Design and Implementation of Competency Based Postgraduate Medical Education in Otorhinolaryngology: The Pilot Experience in India. Indian journal of otolaryngology and head and neck surgery : official publication of the Association of Otolaryngologists of India, 71(Suppl 1), 671–678.

50. Stahl, C. C., Collins, E., Jung, S. A., Rosser, A. A., Kraut, A. S., Schnapp, B. H., Westergaard, M., Hamedani, A. G., Minter, R. M., & Greenberg, J. A. (2020). Implementation of Entrustable Professional Activities into a General Surgery Residency. Journal of surgical education, 77(4), 739–748.

51. Rekman, J., Hamstra, S. J., Dudek, N., Wood, T., Seabrook, C., & Gofton, W. (2016). A New Instrument for Assessing Resident Competence in Surgical Clinic: The Ottawa Clinic Assessment Tool. Journal of surgical education, 73(4), 575–582.

52. Steiman, J., Sullivan, S. A., Scarborough, J., Wilke, L., Pugh, C., Bell, R. H., & Kent, K. C. (2018). Measuring Competence in Surgical Training through Assessment of Surgical Entrustable Professional Activities. Journal of surgical education, 75(6), 1452–1462.

53. Ravitz, P., Lawson, A., Fefergrad, M., Rawkins, S., Lancee, W., Maunder, R., Leszcz, M., & Kivlighan, D. M., Jr (2019). Psychotherapy Competency Milestones: an Exploratory Pilot of CBT and Psychodynamic Psychotherapy Skills Acquisition in Junior Psychiatry

Residents. Academic psychiatry : the journal of the American Association of Directors of Psychiatric Residency Training and the Association for Academic Psychiatry, 43(1), 61–66.

54. Hagel, C. M., Hall, A. K., & Dagnone, J. D. (2016). Queen's University Emergency Medicine Simulation OSCE: an Advance in Competency-Based Assessment. CJEM, 18(3), 230–233.

55. Alken, A., Luursema, J. M., Weenk, M., Yauw, S., Fluit, C., & van Goor, H. (2018). Integrating technical and non-technical skills coaching in an acute trauma surgery team training: Is it too much?. American journal of surgery, 216(2), 369–374.

56. Labbé, M., Young, M., & Nguyen, L. (2018). Toolbox of assessment tools of technical skills in otolaryngology-head and neck surgery: A systematic review. The Laryngoscope, 128(7), 1571–1575.

57. Mubuuke, A. G., Mwesigwa, C., Maling, S., Rukundo, G., Kagawa, M., Kitara, D. L., & Kiguli, S. (2014). Standardizing assessment practices of undergraduate medical competencies across medical schools: challenges, opportunities and lessons learned from a consortium of medical schools in Uganda. The Pan African medical journal, 19, 382.

58. Dwyer, T., Wadey, V., Archibald, D., Kraemer, W., Shantz, J. S., Townley, J., Ogilvie-Harris, D., Petrera, M., Ferguson, P., & Nousiainen, M. (2016). Cognitive and Psychomotor Entrustable Professional Activities: Can Simulators Help Assess Competency in Trainees?. Clinical orthopaedics and related research, 474(4), 926–934.

59. Yaw Tun, J., Granados, A., Mavroveli, S., Nuttall, S., Kadiyala, A. N., Brown, R., Bello, F., & Kneebone, R. L. (2012). Simulating various levels of clinical challenge in the assessment of clinical procedure competence. Annals of emergency medicine, 60(1), 112–20.e5.

60. Bradley P. (2006). The history of simulation in medical education and possible future directions. Medical education, 40(3), 254–262.

61. Meyer, E. G., Chen, H. C., Uijtdehaage, S., Durning, S. J., & Maggio, L. A. (2019). Scoping Review of Entrustable Professional Activities in Undergraduate Medical Education. Academic medicine: journal of the Association of American Medical Colleges, 94(7), 1040–1049.

62. Surgical foundations. Royal College of Physicians and Surgeons of Canada. http://www.royalcollege.ca/rcsite/ibd-search-e?N=10000033+10000034+4294967069. Accessed on October 10, 2020.

63. CBD and assessment, Royal College of Physicians and Surgeons of Canada. http://www.royalcollege.ca/rcsite/cbd/cbd-assessment-e. Accessed on October 10, 2020.

64. CBD observation templates. Royal College of Physicians and Surgeons of Canada. http://www.royalcollege.ca/rcsite/cbd/assessment/wbas/cbd-assessment-templates-e .Accessed on October 10, 2020.

65. Reznick, R. K., & MacRae, H. (2006). Teaching surgical skills--changes in the wind. The New England journal of medicine, 355(25), 2664–2669.