Evaluation of the Characteristics of a Workplace Assessment Form to Assess Entrustable Professional Activities (EPAs) in an Undergraduate Surgery Core Clerkship

Vernon R. Curran, MEd, PhD, Diana Deacon, MSc, MEd,† Henry Schulz, MEd, PhD,‡ Katherine Stringer, MBChB,§ Craig N. Stone, MD,¶ Norah Duggan, MD,§ and Heidi Coombs-Thorne, MA, PhD†*

*Faculty of Medicine, Memorial University, St. John's, Newfoundland, Canada; †Medical Education Scholarship Centre (MESC), Faculty of Medicine, Memorial University, St. John's, Newfoundland, Canada; ‡Faculty of Education, Memorial University, St. John's, Newfoundland, Canada; §Discipline of Family Medicine, Faculty of Medicine, Memorial University, St. John's, Newfoundland, Canada; and ¶Discipline of Surgery, Faculty of Medicine, Memorial University, St. John's, Newfoundland, Canada

OBJECTIVE: Entrustable Professional Activities (EPAs) are explicit, directly observable tasks requiring the demonstration of specific knowledge, skills, and behaviors that learners are expected to perform without direct supervision once they have gained sufficient competence. Undergraduate level implementation of EPAs is relatively new. We examined the characteristics of a workplace assessment form (clinic card) as part of a formative programmatic assessment process of EPAs for a core undergraduate surgery rotation.

DESIGN: A clinic card was introduced to assess progression towards EPA achievement in the clerkship curriculum phase. Students completing their core eight (8) week clerkship surgery rotation submitted at least 1 clinic card per week. We compiled assessment scores for the 2015 to 2016 academic year, in which EPAs were introduced, and analyzed relationships between scores and time, EPA, training site, and assessor role. We surveyed preceptors and students, and conducted a focus group with clinical discipline coordinators of all core rotations.

SETTING: This study took place at the Faculty of Medicine, Memorial University in St. John's, Newfoundland, Canada.

PARTICIPANTS: Third year medical students ($n = 79$) who completed their core eight (8) week surgery clerkship rotation during the 2015 to 2016 academic year, preceptors, and clinical discipline coordinators participated in this study.

RESULTS: EPAs reflecting tasks commonly performed by students were more likely to be assessed. EPAs frequently observed during preceptor-student encounters had higher entrustment ratings. Most EPAs showed increased entrustment scores over time and no significant differences in ratings between teaching sites nor preceptors and residents. Survey and focus group feedback suggest clinic cards fostered direct observation by preceptors and promoted constructive feedback on clinical tasks. A binary rating scale (entrustable/pre-entrustable) was not educationally beneficial.

CONCLUSIONS: The findings support the feasibility, utility, catalytic and educational benefits of clinic cards in assessing EPAs in a core surgery rotation in undergraduate medical education. (J Surg Ed 75:1211-1222. © 2018 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: entrustable professional activities, workplace assessment, programmatic assessment, undergraduate, surgery, education

COMPETENCIES: Practice-Based Learning and Improvement, Patient Care, Systems-Based Practice, Medical Knowledge, Interpersonal and Communication Skills, Professionalism

Correspondence: Inquiries to Vernon R. Curran, MEd, PhD, Faculty of Medicine, Memorial University, St. John's, Newfoundland, Canada A1B 3V6; e-mail: vcurran@mun.ca

  1931-7204/$30.00
https://doi.org/10.1016/j.jsurg.2018.02.013

INTRODUCTION

An important strategy for assessing learners in competency-based medical education (CBME) is workplace assessment. With such an approach, experts are engaged in the direct observation of learners to assess their competence in the context of actual care delivery.[1] Entrustable professional activities (EPAs) have been put forward as a means to improve assessment in CBME by operationalizing competencies and milestones in the clinical setting, and focusing assessment on key activities performed by the learner that demonstrate competence. EPAs have been described as tasks or responsibilities that learners are expected to perform without direct supervision once they have gained sufficient competence.[2] EPAs represent explicit, directly observable tasks that require the demonstration of specific knowledge, skills, and behaviors. It has been suggested that EPAs focus on the qualities of the work to be completed, ground outcomes in the tasks of physicians, and offer an approach to CBME that better addresses concerns around integration of competency domains.[3]

As an offshoot of approaches to CBME, EPAs provide a framework to standardize medical education outcomes and advance competency-based assessment.[4] Although primarily implemented at the postgraduate level, undergraduate programs are examining the use of EPAs in clinical assessment.[5] While competencies describe a student's abilities, EPAs are units of professional practice.[6] ten Cate explains that while "competencies are descriptors of physicians, EPAs are descriptors of work," and are therefore "not an alternative for competencies, but a means to translate competencies into clinical practice."[7] Competencies and EPAs can be viewed as interrelated, as one is necessary to explain the other. Together, they form an intertwined matrix for medical education.[8] The Association of American Medical Colleges (AAMC) describes in the 2014 guide "*Core Entrustable Professional Activities for Entering Residency*" the benefits of EPAs as a conceptual framework for medical education, and defines them as 'activities' that make sense to faculty, trainees, and the public.[9] Undergraduate medical education EPAs, such as those proposed by the AAMC, may help to focus student assessment more directly on workplace activities.[3]

Given the novelty of EPAs, particularly with implementation at the undergraduate level, current interest includes examining best-practice approaches and methods for assessing EPA progression and achievement. It has been suggested that EPAs can provide a holistic perspective on learner assessment and a graded supervision approach for learners is recommended in which they progress through EPAs with decreasing levels of supervision.[7,10] Others have called for a more granular approach to progression through EPAs, especially in the earlier years of training,[3,11,12] while some have emphasized the Dreyfus and Dreyfus model of skill acquisition for assessing learners as they progress from novice to expert.[3,4,10] The challenge of effectively assessing EPAs at all levels of medical education has been highlighted by several researchers as traditional assessment tools can be overly subjective and highly variable.[13] Norcini et al.[14] have described assessment of work and assessment of newer competencies, such as EPAs, as a category of assessment that requires further evidence and research to inform future practices and approaches of assessment in these areas.

Norcini et al.[14] have introduced a set of criteria for good assessment that includes: validity or coherence, reproducibility or consistency, equivalence, feasibility, educational effect, catalytic effect, and acceptability. For formative assessment, "catalytic effect" is highlighted as a key criterion of effective assessment that is concerned with providing results and feedback that enhances and fosters education, or in other words supports further learning. Validity-coherence, educational effect, feasibility, and acceptability are also important criteria to consider in the evaluation of formative assessment systems.[14] Educational effect concerns the way in which the assessment motivates learners to prepare in a way that fosters learning. Feasibility evaluates the extent to which the assessment is practical and realistic, while acceptability is concerned with stakeholders' views on the credibility of the assessment process.

A key element for the implementation of core EPAs is learner assessment based on frequent formative assessments that inform entrustment decisions based on aggregate evidence.[15] At our institution, we introduced a new programmatic assessment process across the clerkship phase of the curriculum that involved the adoption and assessment of EPA achievement. We based this new assessment process on the concept of programmatic assessment developed by Cees van der Vleuten; a longitudinal form of assessment using regular aggregates of different low and high stakes assessment methods to assess progression toward competency. These aggregates then form the basis of customized learning plans.[16,17] It is believed that this assessment approach will foster the acquisition of competencies within specific disciplines and the longitudinal progression of competencies across disciplines throughout clerkship, resulting in a competent undifferentiated graduating physician. Central to assessing whether a learner can be entrusted to perform a particular EPA is direct observation of the learner in clinical settings by faculty supervisors.[4] We developed clinic card forms to enable the assessment of medical student progression towards EPA achievement. The purpose of this study was to evaluate the characteristics of this new workplace assessment form and the adoption of EPAs as a means of formative programmatic assessment in our undergraduate medical education curriculum.

MATERIAL AND METHODS

Memorial University of Newfoundland's Faculty of Medicine offers a 4-year undergraduate medical program leading

to the degree of Doctor of Medicine (MD). The program consists of 4 phases: phase 1 covers normal health and development, phase 2 covers acute reversible or modifiable health issues, phase 3 covers chronic disease, and phase 4 involves integration into practice. The first 3 phases employ a variety of teaching and learning experiences, including self-directed learning, to learn about all aspects of health (physical, mental, social, and psychological), disruptions in health that can lead to disease, and all aspects of science, community health, ethics, and clinical skills related to identifying and describing disease and its diagnosis. During phase 4 (clerkship phase), students take courses that will allow them to experience major disciplines in hospital and community settings throughout affiliated teaching sites in Newfoundland and Labrador and community and hospital settings in New Brunswick and other jurisdictions. Core clinical rotations are included in surgery, anesthesia, emergency medicine, internal medicine, obstetrics/gynecology, pediatrics, psychiatry, and rural family medicine. Students participate as members of the health care team, gaining the knowledge and experience necessary to assume the responsibilities associated with patient care.

We introduced a formative programmatic assessment system in the core eight (8) week surgery rotation in the clerkship phase of the undergraduate medical education curriculum. This rotation involves work-based learning experiences in a variety of clinical environments. Five weeks are spent on a general surgery service and 3 weeks on orthopedic surgery in several hospitals in Newfoundland and Labrador and New Brunswick. There is a mix of tertiary and community hospital experiences with an interactive didactic teaching schedule that is delivered to all students via webcast. The students attend outpatient clinics, assist in the operating room, and see emergency room surgical consultations. The clinic card forms could be completed after any experience in any of these environments.

Assessment specialists developed clinic card forms (Fig. 1) to enable the assessment of medical student progression towards EPA achievement. The cards included itemized EPA statements (Table 1) identified by the AAMC[9] and the choice of specific EPAs to be assessed in the core surgery rotation was decided by the clinical discipline coordinator (CDC) in conjunction with the surgery clerkship curriculum committee. Assessors selected between 1 to 4 EPA items to assess based on the type of activity each student was observed performing (e.g., operating room, clinic, and emergency room) and rated their perceived level of entrustment of student performance using a binary rating scale of either "*pre-entrustable*" or "*entrustable*." The cards also included open-ended items to allow assessors to provide narrative feedback to encourage further student development in the EPA areas. Assessors made 2 types of narrative comments on the clinic cards: (1) feedback on student achievement during the period of time covered by the card; and (2) coaching comments on ways students could improve in the future. Assessors included postgraduate medical residents and preceptors (clinician-educators) in the surgery rotation. Assessors were provided with faculty development through formal webinars on the use of the new clinic cards and one-on-one coaching by the CDC with responsibility for coordination of the clerkship rotation. An online "Faculty Handbook" was produced and disseminated that provided general information on the role of the clinic cards in clerkship assessment and examples of "pre-entrustable" and "entrustable" behaviors for each EPA. The clinic cards also included a QR code that that could be scanned to take assessors to the online handbook.

Medical students undertaking their core surgery rotation ($n = 79$) completed a minimum of 1 clinic card per week of the rotation; thus, there were assessments for 8 times. The class represented approximately 43% ($n = 34$) male and 57% female ($n = 45$) students. Completed clinic card forms assured weekly formative feedback from preceptors to all students and the CDC reviewed all cards. The CDC followed-up with assessors on any negative comments to develop any required remediation plan before the end of the rotation. Orientation sessions on EPAs and use of the clinic cards were presented to students before the start of the core clerkship phase of the curriculum. Students were also able to access an online "Student Handbook" with a section on the "pre-entrustable" and "entrustable" behaviors for each EPA and how to use the clinic card feedback received. The office of undergraduate medical education collated all summative assessments throughout the core clerkship rotations and prepared an assessment summary for each student at 6, 9, and 12 months. These summaries were provided to students and also reviewed by the phase 4 lead (clerkship coordinates) to monitor longitudinal development in each EPA, as well as for the EPAs chosen by each discipline. Based on this review, students not progressing as expected were discussed by the faculty members of the phase 4 management team at comprehensive review meetings at 6, 9, and 12 months. These students received personalized feedback and a learning plan after the reviews.

At the end of each core surgery rotation, we compiled and aggregated the assessment scores for completed clinic card forms for all students during the initial academic year in which EPAs were introduced. We conducted statistical analysis of the scores to determine their relationships with time, EPA, training site, and assessor role (resident or preceptor). Web-based evaluation surveys were also distributed to preceptors across core clerkship rotations, including surgery preceptors, and all medical students enrolled in the core clerkship phase of the curriculum.[18] The preceptor survey included a mix of 10 closed and open-ended items. Closed-ended items were rated using a Likert-type scale and respondents were asked to rate their opinions of the effectiveness of the clinic card form for workplace assessment and utility of a "pre-entrustable" vs "entrustable"

NARRATIVE COMMENTS ON ASSESSED EPAs

Entrustable behaviours demonstrated: _____

Describe opportunities to help learner become entrustable: _____

History: I personally observed this student conduct a focused or general history on a patient. O Yes O No

Physical: I personally observed this student conduct a focused or general physical examination of a patient. O Yes O No

Procedures: I personally observed this student successfully conduct the following procedures _____

Casting (when applicable) O Pre-entrustable O Entrustable

I am a: faculty member / resident (circle one)

Preceptor/Supervisor Signature _____

MEMORIAL UNIVERSITY
Faculty of Medicine

STUDENT CLINICAL ASSESSMENT CARD

SURGERY

Site: _____

Observation period:
O Weekly O Daily

Student name (print):

_____

Print name of assessor (preceptor/supervisor):

_____

Date: (yyyy/mm/dd)
20____/____/____

Assessor must provide feedback on a **minimum of 1** or **maximum of 4** EPAs per clinic card.

ENTRUSTABLE PROFESSIONAL ACTIVITY (EPA) Relevent EPAs	Pre-entrustable	Entrustable
EPA 1 – Gather a history and perform a physical examination.		
EPA 2 – Prioritize a differential diagnosis following a clinical encounter.		
EPA 3 – Recommend and interpret common diagnostic and screening tests.		
EPA 6 - Provide an oral presentation for a clinical encounter.		
EPA 7 – Form clinical questions and retrieve evidence to advance patient care.		
EPA 10 – Recognize a patient requiring urgent or emergent care and initiate evaluation and management.		
EPA 11 – Obtain informed consent for tests and/or procedures.		
EPA 12 – Perform general procedures of a physician.		
EPA 13 – Identify systems failures and contribute to a culture of safety and improvement.		

The student is responsible for obtaining feedback on all EPAs above by rotation's end.

For information on EPA descriptors see (or scan QR code): www.med.mun.ca/epa

FIGURE 1. Clinic card form.

TABLE 1. AAMC EPAs Adopted for Core Surgery Rotation

EPA 1	Gather a history and perform a physical examination
EPA 2	Prioritize a differential diagnosis following a clinical encounter
EPA 3	Recommend and interpret common diagnostic and screening tests
EPA 6	Provide an oral presentation of a clinical encounter
EPA 7	Form clinical questions and retrieve evidence to advance patient care
EPA 10	Recognize a patient requiring urgent or emergent care, initiate evaluation and management
EPA 11	Obtain informed consent for tests and/or procedures
EPA 12	Perform general procedures of a physician
EPA 13	Identify system failures and contribute to a culture of safety and improvement

binary rating scale for assessing EPA achievement. The student survey included a mix of 10 closed and open-ended items. Closed-ended items were rated using a Likert scale and respondents were also asked to rate their opinions of the effectiveness and usefulness of the clinic card forms. A focus group was conducted with a sample of CDCs, including the surgery CDC, to explore perceptions and experiences with the adoption and use of the clinic card form across core clerkship rotations, including the core surgery clerkship rotation.[18] Focus group comments were recorded and a thematic analysis was undertaken to identify key emerging themes. The study received ethics approval from the Health Research Ethics Authority (HREA), Memorial University.

RESULTS

We compared the numbers of ratings submitted for students completing their surgery rotation for the 8 times across the 9 EPAs to determine if some were more commonly rated than others. Figure 2 presents the percentage of EPAs rated across the total forms; for example, EPA 1 was rated 528 times for the 79 students and 8 times, 83.5% of the total of 632 ratings. The 99% CI for the average of 56.9% for all EPAs based on 79 independent observations is 41.9% to 71.8% (Wilson interval). Higher percentages were obtained for EPA 1 "*Gather a history and perform a physical examination*" (83.5%) and EPA 6 "*Provide an oral presentation of a clinical encounter*" (78.5%). There were lower percentages for EPA 11 (33.9%), EPA 13 (38.4%), and EPA 10 (41.8%). These differences indicate that some of the EPAs were rated more frequently than others depending on the student activities that preceptors were most commonly able to observe during the rotation.

The percentage of entrustment ratings for each of the 9 EPAs averaged across the 8 times for the surgery rotation ranged from 61.5% to 89.2% (Fig. 3). We compared the percentage for each EPA against the average of 77.1% for all EPAs, again using the 99% Wilson confidence interval

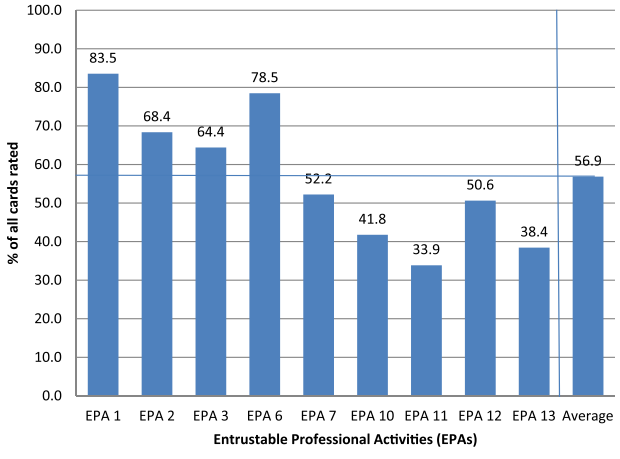

FIGURE 2. Percentage of EPAs rated across EPA cards ($n = 632$).

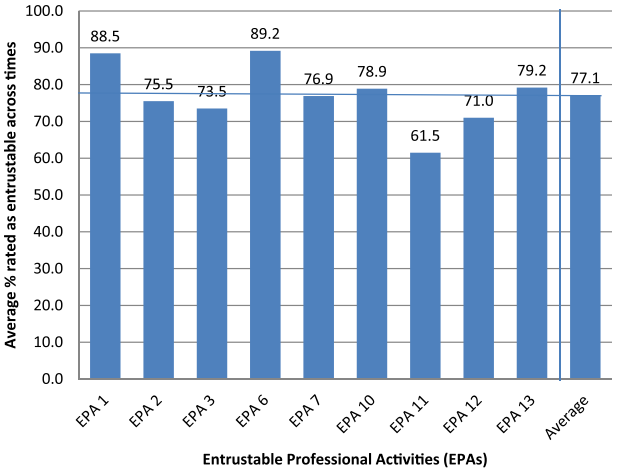

FIGURE 3. Average percentage of EPAs rated as "entrustable".

based on 79 independent observations: 64.3% to 90.0%. We found no significantly higher percentages of "entrustable" ratings for the 99% CI, but did find 2 that were significant for the 95% CI: EPA 1 (88.5%) and EPA 6 (89.2%). However, we did find a significantly lower percentage of "entrustable" ratings for EPA 11 "*Obtain informed consent for tests or procedures*" (61.5%). This suggests that most EPAs were assessed as "entrustable" for the majority of assessments made, with the exception of EPA 11. This activity may not occur as frequently as the other EPAs in the surgery rotation.

We conducted further analyses of proportions to determine if there were any significant changes in entrustment ratings over the course of the rotation; that is, over the eight times that students were rated. The proportions of entrustment ratings for the eight times were compared to the average proportion for the particular EPA (Table 2). This resulted in a large number of statistical tests so we again used the .01 level. Our analyses found statistically significant differences for one of the nine EPAs. The proportion of "entrustable" ratings for EPA 3 "*Recommend and interpret common diagnostic and screening tests*" ranged from 57.1% to 81.5% and indicated a significant difference in "entrustable" ratings from the first to the average of all 8 times ($z = -2.600$, $p = .009$). This suggests an increase in proportion of entrustable ratings from the first to latter times. EPAs 2, 10, 11, and 12 were significantly different at the .05 level, and again time 1 ratings were consistently lower. The proportion of "entrustable" ratings for EPA 2 "*Prioritize a differential diagnosis following a clinical encounter*" ranged from 60.8% to 86.0% and indicated a significant increase in proportions of "entrustable" ratings ($z = -2.450$, $p = .014$) from the first card submitted to subsequent cards. For EPA 10 "*Recognize a patient requiring urgent or emergent care and initiate evaluation and management*" the proportion ranged from 61.3% to 90.0% and indicated a significant increase in proportions of "entrustable" ratings from the first card

TABLE 2. Percentages of "Entrustable" Ratings Across the 8 Times, and Significance as Compared to Average Percentage Across the 8 Times for the Particular EPA (2 Tailed: p)

Time	1	2	3	4	5	6	7	8	Average
EPA 1									
n	71	68	66	65	64	70	62	62	
%	84.5	92.6	86.4	89.2	89.1	85.7	90.3	90.3	88.5
p	.288	.286	.582	.858	.892	.462	.656	.656	
EPA 2									
n	51	49	57	59	60	55	50	51	
%	60.8†	81.6	71.9	79.7	75	72.7	86	76.5	75.5
p	.014*	.320	.528	.460	.924	.630	.080	.876	
EPA 3									
n	49	46	51	54	54	46	52	55	
%	57.1†	80.4	70.6	75.9	81.5	69.6	75.1	80	73.5
p	.009**	.288	.634	.690	.186	.542	.942	.276	
EPA 6									
n	63	59	67	70	58	60	65	54	
%	88.9	89.8	85.1	91.4	89.7	90	87.7	90.7	89.2
p	.944	.870	.282	.542	.904	.834	.702	.710	
EPA 7									
n	40	36	43	46	54	34	40	37	
%	70	80.6	67.4	80.4	85.2	70.6	80	81.1	76.9
p	.300	.604	.140	.570	.150	.382	.642	.548	
EPA 10									
n	31	30	36	41	46	37	31	12	
%	61.3†	90	75	87.8	84.8	78.4	71	83.3	78.9
p	.016*	.138	.562	.164	.332	.932	.276	.710	
EPA 11									
n	22	24	28	30	30	27	25	28	
%	40.9†	62.5	64.3	73.3	63.3	51.9	64	71.4	61.5
p	.048*	.916	.758	.182	.832	.306	.794	.278	
EPA 12									
n	36	36	43	45	44	34	40	42	
%	52.8†	69.4	72.1	68.9	72.7	73.5	77.5	81	71.0
p	.016*	.838	.874	.756	.800	.744	.364	.154	
EPA 13									
n	26	28	33	35	40	28	27	26	
%	65.4	82.1	81.8	85.7	85	78.6	74.1	80.8	79.2
p	.080	.700	.710	.342	.364	.936	.514	.842	

*$p < .05$.
**$p < .01$.
†All significant differences were lower than the average.

submitted to subsequent cards ($z = -2.410$, $p = .016$). The proportions of "entrustable" ratings for EPA 11 "*Obtain informed consent for tests or procedures*" ($z = -1.980$, $p = .048$) and EPA 12 "*Perform general procedures of a physician*" ($z = -2.410$, $p = .016$) also indicated significant increases in proportions of "entrustable" ratings. "Entrustable" ratings on the first card submitted compared to the average were: EPA 11, 40.9% vs 61.5%; and EPA 12, 52.8% vs 71.0%. This suggests that students were more likely to be rated "entrustable" as they gained more experience in the surgery rotation from the first to subsequent times.

We compared the proportions of "entrustable" and "pre-entrustable" ratings between students in the 2 main hospital-based teaching sites hosting students during their rotation in surgery. Table 3 summarizes a Pearson χ^2 analysis comparing ratings between the 2 sites. For the majority of EPAs there were no significant differences. Based on the .01 level, only 1 of the 9 EPAs on which students were rated during their surgery rotation was significant and this only occurred on 1 time: EPA 10 indicated a significant difference for time 6 ($\chi^2 = 6.769$, $p = .009$). EPA 10 also had a significant difference at the .05 level for time 7 ($\chi^2 = 5.143$, $p = .023$). There were several further significances at the .05 level: EPA 3 for time 5 ($\chi^2 = 6.476$, $p = .011$); EPA 6 for time 5 ($\chi^2 = 4.191$, $p = .041$); and EPA 13 for time 5 ($\chi^2 = 6.469$, $p = .011$). There was only 1 difference in entrustment ratings between the 2 hospital sites significant at the .01 level (4 additional differences at the .05 level) of the 72 possible differences, suggesting that differences, if any, were only minimal. In general, this indicates that the rating of EPAs as "entrustable" or "pre-entrustable" were reasonably consistent across these 2 main sites.

We combined the 2 main hospital-based teaching sites (which were compared in the previous paragraph) and

TABLE 3. Pearson χ^2 Analysis Comparing Entrustable vs Pre-Entrustable Ratings Between 2 Main Hospital-Based Teaching Sites (2 Tailed: *p*)

Time	1	2	3	4	5	6	7	8
EPA 1								
n	50	47	46	43	44	48	42	41
χ^2	.163	.345	1.795	.208	.211	.321	.078	.715
p	.686	.557	.180	.648	.646	.571	.780	.398
EPA 2								
n	32	34	36	39	40	36	31	29
χ^2	.163	.186	2.864	.163	2.129	2.363	3.480	.514
p	.686	.666	.091	.686	.145	.124	.062	.473
EPA 3								
n	31	29	33	33	34	28	31	35
χ^2	.682	.013	.122	1.153	6.476	.000	1.697	2.197
p	.409	.909	.727	.283	.011*	1.000	.193	.088
EPA 6								
n	44	40	47	46	40	39	42	34
χ^2	.023	.162	.288	.215	4.191	1.579	.573	.672
p	.879	.687	.592	.643	.041*	.209	.449	.412
EPA 7								
n	24	23	27	28	31	22	24	23
χ^2	.023	.403	.326	1.257	.211	.060	1.186	.403
p	.879	.526	.568	.262	.646	.806	.276	.526
EPA 10								
n	16	18	22	26	28	24	18	22
χ^2	.019	1.800	.430	.580	1.564	6.769	5.143	.002
p	.890	.180	.512	.446	.211	.009**	.023*	.964
EPA 11								
n	12	16	18	17	17	18	13	16
χ^2	.010	.152	.222	.476	.858	.900	.660	3.200
p	.920	.697	.638	.490	.354	.343	.417	.074
EPA 12								
n	23	25	30	29	28	21	25	28
χ^2	.017	.001	2.222	.018	.491	.788	.146	.776
p	.896	.975	.136	.893	.483	.375	.702	.378
EPA 13								
n	12	15	21	19	23	17	13	14
χ^2	3.704	.268	.687	.532	6.469	1.022	.325	1.75
p	.054	.605	.407	.466	.011*	.312	.569	.186

*$p < .05$.
**$p < .01$.

compared their proportions of "entrustable" and "pre-entrustable" ratings to those of periphery sites hosting students during their rotation in surgery. Table 4 presents the results of a Pearson χ^2 analysis comparing ratings. There was a significant difference at the .01 level for 3 EPAs for only one of the times, time 3: EPA 7 ($\chi^2 = 8.032$, $p = .005$); EPA 10 ($\chi^2 = 7.636$, $p = .006$); and EPA 12 ($\chi^2 = 7.213$, $p = .007$). All 3 favored the periphery sites. A further 5 were significant at the .05 level: EPA 2 (time 7, $\chi^2 = 4.989$, $p = .026$); EPA 3 (time 3, $\chi^2 = 4.488$, $p = .034$); EPA 6 (time 3, $\chi^2 = 5.002$, $p = .025$); EPA 11 (time 3, $\chi^2 = 4.480$, $p = .034$); and EPA 13 (time 3, $\chi^2 = 4.190$, $p = .041$). Again, all favored the periphery sites. Interestingly, there was a trend for the significant differences to occur for time three across the majority of EPAs. These results suggest the majority of ratings across EPAs were generally consistent across the main hospital and periphery sites. Where there was a significant difference, ratings were higher for peripheral sites. This may reflect a greater amount of contact between students and 1 or 2 assessors in these smaller sites leading to a greater likelihood of rating a student as entrustable.

Finally, we compared the proportions of "entrustable" and "pre-entrustable" ratings between preceptors vs residents across students' rotation in surgery. Table 5 summarizes a Pearson χ^2 analysis comparing ratings. With the exception of 1 time on 1 EPA, ratings of entrustment did not differ significantly at the .01 level between preceptors and residents. Our analysis found a significant difference for time 5 on EPA 13 ($\chi^2 = 7.559$, $p = .006$). Two EPAs showed significant differences at the .05 level; EPA 11 (time 5, $\chi^2 = 4.979$, $p = .026$) and EPA 13 (Time 8, $\chi^2 = 5.769$, $p = .016$). These results indicate generally that there was consistency across the 2 types of assessors.

The web-based student survey was distributed to all students in the clerkship phase of the curriculum ($n = 79$)

TABLE 4. Pearson χ^2 Analysis Comparing Entrustable Ratings Between Main Hospital-Based and Periphery Teaching Sites

Time	1	2	3	4	5	6	7	8
EPA 1								
n	71	68	66	65	64	70	62	62
χ^2	.033	2.411	1.817	1.341	1.053	.011	3.163	.001
p	.856	.120	.178	.247	.305	.916	.075	.975
EPA 2								
n	51	49	57	59	60	55	50	51
χ^2	.844	1.974	1.341	1.996	.400	.566	4.989	.615
p	.358	.160	.247	.158	.527	.452	.026*	.433
EPA 3								
n	49	46	51	54	54	46	52	55
χ^2	.029	3.208	4.488	1.801	.046	.942	.174	.000
p	.865	.073	.034*	.180	.830	.332	.677	1.000
EPA 6								
n	63	59	67	70	58	60	65	54
χ^2	.009	3.178	5.002	3.424	3.012	.985	.430	.686
p	.924	.075	.025*	.064	.083	.321	.512	.408
EPA 7								
n	40	36	43	46	54	34	40	37
χ^2	2.401	1.794	8.032	.158	.100	1.451	3.151	2.036
p	.121	.180	.005**	.691	.752	.228	.076	.154
EPA 10								
n	31	30	36	41	46	37	31	35
χ^2	.776	.062	7.636	.675	.386	.460	.966	.001
p	.378	.803	.006**	.411	.534	.498	.326	.975
EPA 11								
n	22	24	28	30	30	27	25	28
χ^2	3.316	.000	4.480	.151	.032	.074	1.963	.233
p	.069	1.000	.034*	.698	.858	.786	.161	.629
EPA 12								
n	36	36	43	45	44	34	40	42
χ^2	.358	1.143	7.213	.433	.201	.200	.086	1.930
p	.550	.285	.007**	.511	.654	.655	.769	.165
EPA 13								
n	26	28	33	35	40	28	27	26
χ^2	.910	.101	4.190	.077	.162	.113	.306	.478
p	.340	.751	.041*	.781	.687	.737	.580	.489

*$p < .05$.
**$p < .01$.

and 20 responded (response rate 25%). Table 6 summarizes key results pertaining to student opinions on the effectiveness and usefulness of the clinic card forms.[18] A majority of students (55.0%) did indicate that the forms were not effective, in the current format, for evaluating their performance of an EPA. A majority of students (90.0%) also reported that the forms were "not useful" or only "somewhat useful" as a learning tool. A thematic analysis of the students' open-ended comments revealed that a main issue was students' concern with preceptors' interpretation and understanding of the concept of "entrustment," particularly within the context of assessment of clerks in undergraduate medical education.[18] Students did report that the clinic cards prompted regular direct observation by preceptors and helped to facilitate constructive feedback as part of the assessment process. However, the use of a binary rating scale for assessing "entrustment" was not perceived as helpful.

Eighty four preceptors ($n = 84$) responded to the web-based preceptor survey (Table 6). A majority of preceptors (74.8%) reported that the clinic cards were either "not effective" or only "somewhat effective" in assessing student performance across EPAs. A majority also reported (71.4%) it was either "very difficult" or "somewhat difficult" to interpret the notion of entrustment, and a majority (74.7%) also found the clinic cards as either "not useful" or "somewhat useful". A thematic analysis of open-ended comments indicated that preceptors acknowledged the important role of the clinic card form as a component of programmatic assessment, however, a lack of developmental milestone targets over the clerkship curriculum did make interpretation of the EPAs and the use of an entrustment scale challenging. Preceptors, like students, also felt it was difficult to assess competency progression with a binary rating scale. Some faculty also reported challenges with the adoption of a new system that relied heavily on required observation and the provision of specific, narrative feedback.

Four CDCs participated in a focus group discussion, including the surgery CDC, and the general perception was

TABLE 5. Pearson χ^2 Analysis Comparing Entrustable Ratings Between Attendings and Resident Assessors

Time	1	2	3	4	5	6	7	8
EPA 1								
n	67	64	66	63	60	68	59	61
χ^2	.154	.497	.041	1.454	2.183	2.189	2.485	.317
p	.695	.481	.840	.228	.140	.139	.115	.573
EPA 2								
n	49	47	56	57	56	53	47	50
χ^2	1.180	.481	.738	.475	.468	.212	1.176	.000
p	.277	.488	.390	.491	.494	.645	.278	1.000
EPA 3								
n	47	42	51	52	50	44	49	54
χ^2	.303	1.930	.085	.113	.081	.210	1.054	1.327
p	.582	.165	.771	.737	.776	.647	.305	.249
EPA 6								
n	59	55	66	68	55	59	62	53
χ^2	.413	1.321	1.144	3.147	3.554	.605	.004	2.115
p	.520	.250	.285	.076	.059	.437	.950	.146
EPA 7								
n	39	35	43	44	51	34	37	36
χ^2	.003	.476	.604	1.801	.092	.174	.059	.567
p	.956	.490	.437	.180	.762	.677	.808	.451
EPA 10								
n	30	29	36	39	44	37	29	34
χ^2	1.094	.090	.000	.042	1.874	.257	.286	3.278
p	.296	.765	1.000	.838	.171	.612	.593	.070
EPA 11								
n	22	23	28	28	27	26	22	27
χ^2	2.200	.175	.622	1.765	4.979	.170	.566	.938
p	.138	.676	.430	.184	.026*	.680	.452	.333
EPA 12								
n	34	34	43	43	42	34	37	41
χ^2	.083	.458	.795	1.454	.831	.021	.120	.006
p	.773	.499	.373	.228	.362	.885	.729	.938
EPA 13								
n	24	26	33	33	37	27	25	25
χ^2	.548	1.262	1.331	3.406	7.559	.045	.041	5.769
p	.459	.261	.249	.065	.006**	.832	.840	.016*

*$p < .05$.
**$p < .01$.

the clinic card form and EPAs were reflective of clinical practice tasks and a competency-based assessment approach, however, there were limitations with the form. Again, the use of a binary rating scale was viewed as limiting the potential benefit of a formative, programmatic assessment process in fostering competency development across the clerkship curriculum. Respondents recognized the value of clinic card forms in documenting immediate formative feedback, however a lack of adequate preparation by way of faculty development led to a lower engagement of preceptors resulting in lower quality written coaching comments.

DISCUSSION

Carraccio and Englander[10] report that acceptability of a tool has a critical impact on faculty buy-in of workplace-based assessment, particularly as the shift to CBME has expanded the domains of skills to be assessed. As a result, studying the implementation of assessment tools in real-world settings—what works and what doesn't for the faculty using the tools—is a critical part of establishing their validity and supporting their intended use.[19] EPA-based assessment is gaining momentum, among significant concerns regarding the feasibility of implementation. Englander and Carraccio[20] emphasize the need for both implementation of EPAs as an educational intervention and evaluation of such implementations to understand what works, in what contexts and under what conditions. These are key elements to encourage uptake of EPAs beyond the early adopters and in different settings. They call for further medical education research and publication about the use of EPAs.

It has been suggested that the challenges to assessment in CBME can be addressed by increasing acceptability through building assessment into our daily work, studying the issues in implementation, and providing validity by aligning what we measure with what we do.[10] For the surgery rotation, we

TABLE 6. Summary of Preceptor and Student Evaluation Surveys

	Preceptor				
	Not effective	Somewhat effective	Moderately effective	Very effective	
How effective are the clinic cards for capturing student performance across EPAs?	31 (37.4%)	31 (37.4%)	18 (21.7%)	3 (3.6%)	
	Very difficult	Somewhat difficult	Neither easy or difficult	Somewhat easy	Very easy
How easy is it to interpret what constitutes entrustable vs pre-entrustable?	22 (26.2%)	38 (45.2%)	11 (13.1%)	9 (10.7%)	4 (4.8%)
	Not useful	Somewhat useful	Moderately useful	Very useful	
How useful are the clinic cards as an assessment tool?	24 (28.9%)	38 (45.8%)	17 (20.5%)	4 (4.8%)	
	Student				
	Not effective	Somewhat effective	Moderately effective	Very effective	
How effective were the clinic cards for capturing your performance of an EPA?	11 (55.0%)	8 (40.0%)	1 (5.0%)		
	Not useful	Somewhat useful	Moderately useful	Very useful	
How useful were the clinic cards as a learning tool?	8 (40.0%)	10 (50.0%)	2 (10.0%)		

found particular EPAs were more likely than others to be rated by assessors, specifically EPA 1 "*Gather a history and perform a physical examination*;" EPA 6 "*Provide an oral presentation of a clinical encounter*;" and EPA 2 "*Prioritize a differential diagnosis following a clinical encounter*." These represented the EPAs most commonly performed by students in an outpatient surgery clinic and therefore more likely to be observed and assessed in that clinical environment. Other EPAs may be better assessed in other clinical rotations.

Suggestions made by ten Cate[21] to address concerns of EPA implementation include graded supervision for students in which they progress through EPAs with decreasing levels of supervision. Carraccio and Englander[10] expanded this concept and argued that a supervisory approach should be matched to the developmental level of the learner. Our analysis of EPA implementation in the surgery rotation indicated that a significantly higher proportion of entrustment ratings were given for EPA 1 and EPA 6. A significantly lower proportion of "entrustable" ratings was given for EPA 11 "*Obtain informed consent for tests or procedures*." We suspect that with these EPAs being more frequently observed by preceptors during a fixed encounter (e.g., 1 outpatient clinic), the preceptors may feel more comfortable assigning an entrustment rating. With other less frequently witnessed EPAs, the preceptor may be less confident in assigning an "entrustable" rating.

An analysis of proportions was used to determine if there were any significant changes in the proportion of "entrustable" ratings for each card that was completed over time as compared to the average proportion of "entrustable" ratings across all cards for each EPA. Statistically significant differences were found for the majority of EPAs indicating a trend of increased entrustment scores for these EPAs from the first to latter times. This would be expected as students become more experienced in the rotation. Apart from improving clinical aptitude, there may be other reasons for this finding. Some cards were completed by the same preceptors at different points in time and this may have occurred more frequently in community rotations where there were fewer residents and preceptors. With an increasing volume of interaction between assessor and student over time, the "repeat" assessor may be more confident in assigning "entrustable" to any EPA.

An assessment system that measures competency in an integrated manner, across a variety of clinical situations, and during actual performance is ideal for CBME.[22] For the majority of EPAs there were no significant differences between the "entrustable" and "pre-entrustable" scores between the 2 main hospital-based teaching sites or between the main hospital-based teaching sites and periphery sites. Similarly, for the majority of EPAs, ratings of entrustment did not differ between types of assessors (preceptors vs residents). These findings suggest considerable consistency across teaching sites and assessors, and support the feasibility of EPAs in various settings and with different supervisors.

Overall, students, preceptors, and CDCs believed the introduction of the clinic card form and use of EPAs as a component of a formative programmatic assessment system was important. Generally, it was reported that use of the forms promoted direct observation by preceptors and enabled more helpful and constructive feedback that supported competency development in a number of applicable performance and clinical task areas. However, all groups felt the binary rating scale on the initial iteration of the clinic card form was not useful. A subsequent iteration of the clinic card form introduced in the 2016 to 2017 academic year substituted a 3 level rating scale based on the level of supervision required by the student: "*I guided the student through much of this activity*;" "*I provided corrective coaching tips on critical points during/after the activity*;" "*The activity was successfully and independently completed by the student with coaching points that would be appropriate for a medical graduate*." Chen et al.[3] have discussed the need for a modified scale in the assessment of EPAs at an undergraduate level that includes more gradations of supervision while allowing additional layers of progressive learner autonomy. Proposed scales for EPA assessment at a postgraduate or graduate level have recommended 5 different levels of supervision.[3] Our modified scale focusses on the notion of levels of guidance and coaching provided to the student vs supervision. This scale also corresponds well with the nature of open-ended feedback expected of preceptors through the assessment process and use of the clinic cards. Our ongoing evaluation will provide further insight into the practical interpretation and utility of these concepts in assessment of EPAs at the undergraduate level.

A main limitation of the study may have been the focus on a single disciplinary rotation and the lack of comparison across disciplines to examine the influence of clinical context and experiences on the validity characteristics of the clinic card. Current efforts to increase efficiency in collecting EPA ratings include a pilot of the use of mobile technologies as a means to complete and submit the clinic card assessments in an electronic format.

CONCLUSIONS

Overall, the findings support the catalytic and educational effect, and feasibility of a clinic card form as a means of assessing EPAs in a core surgery rotation in undergraduate medical education. Utility and acceptability of the clinic card form may be improved with the introduction of a new scale that focuses rating toward "level of guidance and coaching" required vs perceptions of trust in the performance of the student. The notion of "trust" as a rating parameter appeared to be very problematic for both students and preceptors. The findings also highlight the importance of validating the implementation of EPAs and evaluating the assessment tools by which these constructs are enacted. It is reasonable to expect entrustment ratings to increase over the eight times that ratings were conducted—students should become more familiar with what is expected. Thus, this suggests a positive effect of the EPA rating procedures. Relatively few differences were found between training sites, which indicates that the cards were used similarly across different hospital settings. There were also few differences between different assessor roles, indicating that both groups gave similar ratings.

These results are encouraging and suggest this assessment method may have broader application across other clinical education settings in undergraduate medical education. For formative assessment, catalytic effect, educational effect, feasibility, and acceptability are key criterion for evaluating effective assessment systems.[14] Overall, the findings suggest that the use of a clinic card form in a core undergraduate surgery rotation promoted greater feedback from preceptors that may have furthered learning and fostered progressive development of learners' competencies. Preceptors and students valued the process as credible and meaningful, and with revisions to the rating scale, should enhance the practicality and acceptability of the clinic card form as a tool in a formative programmatic assessment process.

ACKNOWLEDGMENT

This work was supported by the Dr. Wallace Ingram Award, Faculty of Medicine, Memorial University. The funder had no involvement in study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

REFERENCES

1. Carraccio C, Englander R, Gilhooly J, et al. Building a framework of entrustable professional activities, supported by competencies and milestones, to bridge the educational continuum. *Acad Med*. 2017;92(3):324-330.
2. Touchie C, ten Cate O. The promise, perils, problems and progress of competency-based medical education. *Med Educ*. 2016;50(1):93-100.
3. Chen HC, van den Broek WE, ten Cate O. The case for use of entrustable professional activities in undergraduate medical education. *Acad Med*. 2015;90(4):431-436.
4. Calaman S, Hepps JH, Bismilla Z, et al. The creation of standard-setting videos to support faculty observations of learner performance and entrustment decisions. *Acad Med*. 2016;91(2):204-209.

5. Hauer KE, Boscardin C, Fulton TB, Lucey C, Oza S, Teherani A. Using a curricular vision to define entrustable professional activities for medical student assessment. *J Gen Intern Med*. 2015;30(9):1344-1348.

6. ten Cate O, Chen HC, Hoff RG, Peters H, Bok H, van der Schaaf M. Curriculum development for the workplace using entrustable professional activities (EPAs): AMEE guide no. 99. *Med Teach*. 2015; 37(11):983-1002.

7. ten Cate O. Nuts and bolts of entrustable professional activities. *J Grad Med Educ*. 2013;5(1):157-158.

8. ten Cate O. Entrustability of professional activities and competency-based training. *Med Educ*. 2005;39(12): 1176-1177.

9. Association of American Medical Colleges. Core Entrustable Professional Activities for Entering Residency: Curriculum Developer's Guide. Available at: https://www.aamc.org/initiatives/coreepas/publicationsandpresentations/; 2014 Accessed March 2017.

10. Carraccio CL, Englander R. From Flexner to competencies: reflections on a decade and the journey ahead. *Acad Med*. 2013;88(8):1067-1073.

11. Rashid P. Entrustable professional activities: time to be trusted? *ANZ J Surg*. 2015;85(5):298-299.

12. Kwan J, Crampton R, Mogensen LL, Weaver R, van der Vleuten CP, Hu WC. Bridging the gap: a five stage approach for developing specialty-specific entrustable professional activities. *BMC Med Educ*. 2016;16:117. http://dx.doi.org/10.1186/s12909-016-0637-x.

13. Choo KJ, Arora VM, Barach P, Johnson JK, Farnan JM. How do supervising physicians decide to entrust residents with unsupervised tasks? A qualitative analysis *J Hosp Med*. 2014;9(3):169-175.

14. Norcini J, Anderson B, Bollela V, et al. Criteria for good assessment: consensus statement and recommendations from the Ottawa 2010 Conference. *Med Teach*. 2011;33(3):206-214.

15. Englander R, Flynn T, Call S, et al. Toward defining the foundation of the MD degree: core entrustable professional activities for entering residency. *Acad Med*. 2016;91(10):1352-1358.

16. van der Vleuten CPM, Schuwirth LWT. Assessing professional competence: from methods to programmes. *Med Educ*. 2005;39(3):309-317.

17. van der Vleuten CPM, Schuwirth LWT, Driessen EW, et al. A model for programmatic assessment fit for purpose. *Med Teach*. 2012;34(3):205-214.

18. Stringer K, Deacon D, Coombs-Thorne H, McGrath G, Duggan N, Curran V. Introducing entrustable professional activities into an undergraduate medical education clerkship curriculum. *Cureus*. 2017;9(6): e.

19. American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. Standards for educational and psychological testing. Washington, D.C.: AERA; 2014.

20. Englander R, Carraccio C. From theory to practice: making entrustable professional activities come to life in the context of milestones. *Acad Med*. 2014;89(10): 1321-1323.

21. ten Cate O. AM last page: what entrustable professional activities add to a competency-based curriculum. *Acad Med*. 2014;89(4):691.

22. Dhaliwal U, Gupta P, Singh T. Entrustable professional activities: teaching and assessing clinical competence. *Indian Pediatr*. 2015;52(July):591-597.