Scoping Review of Entrustable Professional Activities in Undergraduate Medical Education

Eric G. Meyer, MD, H. Carrie Chen, MD, PhD, Sebastian Uijtdehaage, PhD, Steven J. Durning, MD, PhD, and Lauren A. Maggio, PhD

Abstract

**Purpose**
Entrustable professional activities (EPAs) are a hot topic in undergraduate medical education (UME); however, the usefulness of EPAs as an assessment approach remains unclear. The authors sought to better understand the literature on EPAs in UME through the lens of the 2010 Ottawa Conference Criteria for Good Assessment.

**Method**
The authors conducted a scoping review of the health professions literature (search updated February 2018), mapping publications to the Ottawa Criteria using a collaboratively designed charting tool.

**Results**
Of the 1,089 publications found, 71 (6.5%) met inclusion criteria. All were published after 2013. Forty-five (63.4%) referenced the 13 Core Entrustable Professional Activities for Entering Residency developed by the Association of American Medical Colleges (AAMC). Forty (56.3%) were perspectives, 5 (7.0%) were reviews, and 26 (36.6%) were prospective empirical studies. The publications mapped to the Ottawa Criteria 158 times. Perspectives mapped more positively (83.7%) than empirical studies (76.7%). Reproducibility did not appear to be a strength of EPAs in UME; however, reproducibility, equivalence, educational effect, and catalytic effect all require further study. Inconsistent use of the term “EPA” and conflation of concepts (activity vs assessment vs advancement decision vs curricular framework) limited interpretation of published results. Overgeneralization of the AAMC’s work on EPAs has influenced the literature.

**Conclusions**
Much has been published on EPAs in UME in a short time. Now is the time to move beyond opinion, clarify terms, and delineate topics so that well-designed empirical studies can demonstrate if and how EPAs should be implemented in UME.

Throughout the last 20 years, competency-based, time-variable education has been hotly debated in medical education.1,2 In response to criticisms that competency-based medical education (CBME) is too reductionist,3,4 ten Cate and Scheele5 have proposed refocusing assessments of individual competencies to, instead, using entrustable professional activities (EPAs) to assess learners. EPAs, each of which requires integration of individual competencies, are “professional activities that together constitute the mass of critical elements that operationally define a profession.”5 Learners are assessed on the basis of whether they can be trusted to carry out an EPA with or without supervision. Through this approach, “patients’ and instructors’ trust in a trainee and their entrustment of responsibility to that trainee are essential concepts.”6 This new approach aligns with entrustment decisions that clinician educators make all the time,7 and it may be less burdensome than other current CBME models, such as the Accreditation Council for Graduate Medical Education (ACGME) Milestones.5 The EPA approach to assessment has been widely described as advantageous for competency-based graduate medical education (GME), and adopting it has the potential to influence every facet of GME—from curricular design to graduation competencies.8–13

In 2014, three perspective or opinion-based articles proposed the use of EPAs in undergraduate medical education (UME).14–16 Since then, there has been intense interest in defining and assessing EPAs in UME. Furthermore, professional associations across the globe have shown interest in EPAs.17,18 For instance, the Association of American Medical Colleges (AAMC) has funded initiatives to define 13 Core Entrustable Professional Activities for Entering Residency (Core EPAs) that medical students should be able to complete without direct supervision upon graduation, and it has sponsored a longitudinal multisite trial to pilot the 13 Core EPAs in UME.19 However, EPAs were designed as a workplace-based approach for assessing learners in GME, and given the differences between GME and UME, the medical education community may need to be cautious applying GME experiences with EPAs to UME.

Given both the possibility that EPAs might not function in UME as they do in GME and the attention and momentum surrounding EPAs, we believe that understanding the state of the science regarding EPAs in UME is vital. To this end, we conducted a scoping review, which is designed to help identify gaps in the research knowledge base, clarify key concepts, and report on the types of evidence that address and inform practice in the field.20,21 Additionally, a scoping review helps clarify working definitions and conceptual boundaries while mapping the key concepts that, in this case, underpin EPAs in UME.20
Because EPAs are fundamentally an approach to assessment, we elected to use the 2010 Ottawa Conference Criteria for Good Assessment (Ottawa Criteria)\(^2\) as a framework to guide our analysis. The Ottawa Criteria allowed us to evaluate the quality of EPAs in seven different domains (see Table 1).

### Method

We completed a scoping review guided by Levac and colleagues’ (2010) six-step approach,\(^3\) as described by Thomas and colleagues\(^4\) (2017), mapping key concepts, delineating types of available evidence, and identifying gaps in research related to EPAs.\(^5\)

First, we collectively agreed on a clear but broad\(^6\) research question: “What is the scope of the literature on EPAs in UME as organized by the Ottawa Criteria?” Next, assisted by a medical librarian, we searched PubMed, Embase, ERIC, SCOPUS, Google Scholar, and Web of Science using relevant medical subject headings (MeSH) and keywords (see Supplemental Digital Appendix 1 at http://links.lww.com/ACADMED/A667 for full list of MeSH terms). We also hand searched personal databases. We conducted our initial search on August 27, 2017, and updated the search on February 5, 2018. We checked the references of EPA-related articles published early (2014) in the literature to confirm that no key publications were missing from the results. We managed all references in EndNote (Clarivate Analytics, Philadelphia, Pennsylvania).

Two of us (E.G.M. and L.A.M.) independently screened titles and abstracts of publications in English (translations accepted) that were specifically about EPAs in UME. We included peer-reviewed publications (including those e-published ahead of print), books, conference proceedings, chapters, reviews, opinion pieces, and blog posts if they mentioned both EPAs and UME. We also included MedEdPORTAL material. We excluded publications that (1) did not mention EPAs in the title or abstract, (2) were not about UME (e.g., publications about residents or EPAs in general, without reference to UME), (3) were not specific to training physicians, or (4) were not translated into English. In September 2017, we sent our literature review to the AAMC Core EPA listserv (1,100 subscribers) to ask for any pertinent articles that we may have missed. Five subscribers responded: Three confirmed the review’s completeness, and two recommended an additional four publications. Abstract review of these four suggested publications did not result in any additions.

Guided by our research question, and informed by the Best Evidence in Medical Education (BEME) synthesis coding sheet,\(^25\) we collectively created a data-charting form using Excel (version 16.16.7; 2011; Microsoft, Redmond, Washington). In addition to identifying typical scoping review results (e.g., the number of authors, author discipline, publication types, populations studied), we reviewed each publication to determine which (if any) of the Ottawa Criteria it discussed—either explicitly or implicitly. We agreed to include implicit references because authors commonly described aspects of the Ottawa Criteria but used different terminology (e.g., “motivational effect” instead of “educational effect”). We chose not to use a specific definition for what constituted an implicit reference, but instead relied on consensus between reviewers (E.G.M. and L.A.M.). For each publication that referenced one or more of these criteria, we documented whether or not the assessment of EPAs met the criteria. If a publication purported that the assessment of EPAs failed to meet a criterion for good assessment, we coded it as “negative.” For example, one paper noted that some of the AAMC’s Core EPAs do not meet the definition of an EPA,\(^26\) which is an example of not meeting the criteria for validity/coherence. However, this same paper also described that the Core EPAs could improve equivalence across programs.\(^26\) Thus, a single publication could be negative for one criterion (e.g., validity/ coherence) and positive for another (e.g., equivalence). We explained our coding as a comment in the data-charting form. Finally, we also used two free-text boxes for a “brief summary of the publication” and “final thoughts.”\(^7\)

Once we all agreed on the data-charting form, each of us independently coded the same publications to improve our consistency and to develop a uniform approach to coding. After coding four randomly selected publications, we reached consensus and confirmed that our approach to data coding was consistent with the research question and purpose. We equally divided the remaining articles for coding, making sure that we did not code papers that we had authored ourselves. One of us (E.G.M.) double-checked all other authors’ coding. We discussed differences until we achieved complete consensus. We summarized the quantitative findings for coding using descriptive statistics, including tabulating the number of times the publications mapped, positively or

### Table 1

| **Table 1** 2010 Ottawa Criteria* |
|-----------------|------------------|
| **Criterion**   | **Explanation**  |
| Validity/coherence | There is a body of evidence that is coherent (“hangs together”) and that supports the use of the results of an assessment for a particular purpose |
| Reproducibility/consistency | The results of the assessment would be the same if repeated under similar circumstances |
| Equivalence | The same assessment yields equivalent scores or decisions when administered across different institutions or cycles of testing |
| Feasibility | The assessment is practical, realistic, and sensible, given the circumstances and context |
| Educational effect | The assessment motivates those who take it to prepare in a fashion that has educational benefit |
| Catalytic effect | The assessment provides results and feedback in a fashion that creates, enhances, and supports education; it drives future learning forward |
| Acceptability | Stakeholders find the assessment process and results to be credible |

negatively, to each of the Ottawa Criteria. We then discussed the implications of study findings to practice and research. Lastly, we shared a preliminary version of this report and the charting results with two leaders in the EPA field, Dr. Olle ten Cate and Dr. Robert Englander, for comment.

Results

We identified 1,089 publications, and after applying the inclusion and exclusion criteria, we included 68 articles, one MedEdPORTAL presentation,27 one 114-page guide,28 and one conference proceeding from an oral presentation29 (see Figure 1). The main reason for excluding other articles was that they were not about UME, but about EPAs in general. A complete list of the 71 included publications14–19,26–29,31–91 is available in Supplemental Digital Appendix 2 at http://links.lww.com/ACADMED/A668.

The chart with coding results is available in Supplemental Digital Appendix 3, available at http://links.lww.com/ACADMED/A669.

Publication characteristics

Publications were authored by 362 unique authors; 297 authors (82.0%) contributed to 1 publication and 61 authors (16.9%) contributed to 2 to 5 of the included publications. Holmboe authored or coauthored 6 (8.5%) of the included publications, Englander and Carraccio each contributed to 11 (15.5%), and ten Cate authored or coauthored 22 (31.0%) of them. The average number of authors per publication was 5 (range: 1–19). The first author of 47 of the included publications (66.2%) was a clinician educator, as determined by degree and job title. Forty of the publications (56.3%) represented multi-institutional collaborations (see Table 2). Although, across all the publications, authors represented 11 countries, a U.S. or Canadian author contributed to 56 of the publications (78.9%). Forty-five (63.4%) of the publications referenced the AAMC’s Core EPAs. Fifty-five (77.5%) publications did not identify a specific specialty focus but, instead, described EPAs in UME broadly. Specialty areas within UME that were specifically discussed included emergency medicine,35 surgery,7,14 psychiatry,7,91 internal medicine,39,48,52 anesthesia,58 obstetrics,66 and pediatrics.61

Publication types

Four perspective publications were written in 2014 on using EPAs in UME.14–16,26 In the following years, the number of publications grew rapidly: 11 in 2015; 23 in 2016; and 33 in 2017. Forty publications were perspective pieces (56.3%), 5 were unstructured literature reviews (7.0%), and 26 were empirical studies (36.6%). Of these 26, all were

---

**Figure 1** Flowchart illustrating the process of identifying, through the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, 71 publications for a review of the literature on entrustable professional activities in undergraduate medical education. The authors initially searched the literature (see Method and Supplemental Digital Appendix 1, available at http://links.lww.com/ACADMED/A667) on August 27, 2017, and they updated their search on February 5, 2018.
prospective studies; 16 (61.5%) used quantitative methods, and the remaining used qualitative (n = 4 [15.4%]) or mixed (n = 6 [23.1%]) methods. While the populations studied in these 26 prospective studies were typically medical students (21 studies [80.1%]), several studies examined other groups. For example, two sets of investigators queried GME faculty and residents to understand their expectations of graduating medical students (21 studies [80.1%]), several narratives included the following: 

**Supporting frameworks**

Eleven studies (15.5%) explicitly referenced a theoretical or educational framework. Miller’s Pyramid of Assessment was also mentioned twice. Other frameworks cited in the publications included the following:

**Table 2**

Demographics of Articles About Entrustable Professional Activities in UME

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No. (% of 71)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Country of origin</strong></td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td>50 (70.4)</td>
</tr>
<tr>
<td>Netherlands</td>
<td>15 (21.1)</td>
</tr>
<tr>
<td>Canada</td>
<td>8 (11.3)</td>
</tr>
<tr>
<td>Germany</td>
<td>4 (5.6)</td>
</tr>
<tr>
<td>Australia</td>
<td>3 (4.2)</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>3 (4.2)</td>
</tr>
<tr>
<td>Israel</td>
<td>1 (1.4)</td>
</tr>
<tr>
<td>Mexico</td>
<td>1 (1.4)</td>
</tr>
<tr>
<td>Hungary</td>
<td>1 (1.4)</td>
</tr>
<tr>
<td>Switzerland</td>
<td>1 (1.4)</td>
</tr>
<tr>
<td>India</td>
<td>1 (1.4)</td>
</tr>
<tr>
<td><strong>Funded</strong></td>
<td></td>
</tr>
<tr>
<td>Not specified</td>
<td>40 (56.3)</td>
</tr>
<tr>
<td>Medical students</td>
<td>21 (29.6)</td>
</tr>
<tr>
<td>UME faculty</td>
<td>6 (8.5)</td>
</tr>
<tr>
<td>GME faculty</td>
<td>5 (7.0)</td>
</tr>
<tr>
<td>Residents/interns</td>
<td>5 (7.0)</td>
</tr>
<tr>
<td>Patients</td>
<td>1 (1.4)</td>
</tr>
<tr>
<td><strong>Study population multisite</strong></td>
<td>8 (11.3)</td>
</tr>
<tr>
<td><strong>Phase of UME studied (if specified)</strong></td>
<td></td>
</tr>
<tr>
<td>Preclerkship</td>
<td>4 (5.6)</td>
</tr>
<tr>
<td>Clerkship</td>
<td>15 (21.1)</td>
</tr>
<tr>
<td>Postclerkship</td>
<td>8 (11.3)</td>
</tr>
<tr>
<td>Recent UME graduates</td>
<td>7 (9.9)</td>
</tr>
</tbody>
</table>

Abbreviations: UME indicates undergraduate medical education; GME, graduate medical education.

The percentages in some sections do not add up to 100% because some articles were not specific, and others fit in multiple categories (multiauthor, multiple populations studied, etc.).

Some articles bylines included both U.S. and Canadian authors.

**Ottawa Criteria mapping**

All seven of the Ottawa Criteria were addressed in multiple publications. Overall, the 71 publications we analyzed explicitly or implicitly referenced the Ottawa Criteria 158 times (see Chart 1).

Of the 158 mentions of the Ottawa Criteria, the one most frequently addressed in publications was “validity and coherence” (n = 42; 26.6%), while “reproducibility/consistency” and “equivalence” were only mentioned, respectively, 5 (3.2%) and 7 (4.4%) times. The number of Ottawa Criteria discussed in publications varied: Some publications mapped EPAs to all seven criteria, and others mapped EPAs to none. Most mentions of the Ottawa Criteria (n = 132; 83.5%) characterized EPAs in UME as “positively” fulfilling one or more of the criteria. Perspective publications were more positive (83.7%; 82/98) than empirical work (76.7%; 33/43) about how well EPAs fulfilled the Ottawa Criteria; however, a dramatic increase in negative perspectives (n = 13) appeared in 2017 compared with only two critical opinions published earlier (see Figure 2).

Although only five studies mapped EPAs to the reproducibility criterion, four of these (80%) reported that EPA assessments in UME were not easily reproducible.

**Validity/coherence.** The validity and coherence criterion was explored extensively—namely, in the methods for developing EPAs and for teaching and assessing them in UME. Several collections of EPAs were developed specifically for UME. One previously mentioned is the AAMC’s set of Core EPAs, which was later followed by an extensive curricular guide. After considering these two publications, ten Cate questioned whether some of the Core EPAs described therein were in fact EPAs and critiqued how those EPAs were defined. Additional collections, often providing evidence for content validity, were developed for acute care, genomics, oncology, internal medicine clerkship, and the UK internal medicine UME curriculum. One article went beyond just proposing a list of EPAs and described an extensive system for developing EPAs for entry into clerkship. Another developed standard-setting videos of
EPAs.44 Several publications discussed methods, in addition to direct patient care, for assessing EPAs through a variety of learning methods, such as physician-mentored patient rounds,46 case-based learning,66 and simulation experiences.35,72,80 Two publications described the assessment of EPAs in, respectively, the preclerkship and postclerkship periods. One set of authors expanded the levels of entrustment/supervision to include finer gradations for UME,31 and another set broadened the information used for entrustment decisions to include student self-assessment and patient input.28 It was noted, however, that students and faculty may intrinsically conceptualize trust differently.20,71 A comprehensive Association for Medical Education in Europe (AMEE) guide explored methods for identifying and validating EPAs along with means for mapping them to competencies and milestones.26

**Reproducibility/consistency.** Discussion of reproducibility and consistency was predominantly negative; many authors indicated that reproducibility could vary depending on the evaluator’s level of training63,68 and available technological support.89 Finally, an opinion piece noted that the layers of inference required to make an entrustment decision will negatively impact reliability.90

**Equivalence.** Authors of the seven papers that mentioned equivalence were predominantly positive, defending the use of EPAs to standardize medical education.26,36,40 Others questioned this suggestion because of variations in the concept of trust and how different evaluators determine trust across institutions and specialties.89,91

**Feasibility.** Authors positively described implementing EPAs in UME in 84.8% of publications addressing feasibility (28/33); however, most of these (19; 67.8%) were perspectives. Publications that explicitly discussed implementation described a process requiring intense faculty development.44,53 For example, a group of Swiss physician leaders leading a national effort to restructure the Swiss medical school curriculum around EPAs noted that doing so would require a great deal of work and prudent application of education theory to be successful.27 Others noted that some of the AAMC’s Core EPAs may not be current components of medical school curricula,57 potentially requiring additional experiences.29 One author highlighted that the potential confusion over the language surrounding the Core EPAs might require careful faculty development.26 Mapping basic science content to EPAs was also described as difficult.29

**Educational effect.** Twenty-three publications (32.4%) explored the educational effect of EPAs on students. The effects were linked to receiving feedback (1) from multiple sources, (2) from a trustworthy supervisor, (3) immediately after the activity, and (4) with a focus on the trainee’s ability to act unsupervised.98 These effects may be enhanced by maintaining transparency in faculty members’ reasons for entrusting students.97 Another publication explored how EPAs might enable medical schools to meet individual medical students’ needs,43 which might involve extra “transitional” training for those students not yet entrustable with EPAs needed for entry into residency.83 The comprehensive roadmap for how the AAMC was planning to study the Core EPAs in a longitudinal multisite trial underscored these educational effects as well.49

**Catalytic effect.** Several authors posited that UME EPAs will have a catalytic effect on UME curriculum design and that, because incorporation of EPAs will inherently refocus curricula, the opportunity to intentionally develop a focused vision should not be missed.33 One such reform included combining the assessment of EPAs and supervision with efforts to improve patient safety.89 One set of authors postulated that if there is general agreement on which EPAs are important for medical school graduates, then all medical schools would, in effect, share the same core curriculum, which could mean, in turn, that the capabilities of medical school graduates across the United States would also become more consistent.78 Authors also considered the effect that such standardization of curricula could have on interdisciplinary training and osteopathic graduates/programs,51 which may prioritize EPAs differently.

**Acceptability.** While 17 out of 22 publications (77.3%) described EPAs as acceptable to learners and educators, several authors cautioned that EPAs need to be studied longitudinally and

---

### Chart 1

**Number of Times the Ottawa Criteria Are Mentioned in Articles on EPAs in UME**

<table>
<thead>
<tr>
<th>Perspective type</th>
<th>Validity/ coherence</th>
<th>Reproducibility/ consistency</th>
<th>Equivalence</th>
<th>Feasibility</th>
<th>Educational effect</th>
<th>Catalytic effect</th>
<th>Acceptability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pos</td>
<td>Neg</td>
<td>Pos</td>
<td>Neg</td>
<td>Pos</td>
<td>Neg</td>
<td>Pos</td>
</tr>
<tr>
<td>Perspective</td>
<td>22</td>
<td>3</td>
<td>—</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td>Review</td>
<td>3</td>
<td>3</td>
<td>—</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td>Empirical study</td>
<td>10</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td><strong>Total per criterion by valence</strong></td>
<td><strong>35</strong></td>
<td>7</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td><strong>Total per criterion</strong></td>
<td><strong>42</strong></td>
<td><strong>5</strong></td>
<td><strong>7</strong></td>
<td><strong>33</strong></td>
<td><strong>23</strong></td>
<td><strong>26</strong></td>
<td><strong>22</strong></td>
</tr>
</tbody>
</table>
bridged to other assessment approaches via sound educational theory.\(^{37,58,65}\) Krupat\(^{90}\) warned that accepting EPAs into UME as a “foregone conclusion” is not well founded. Ten Cate and colleagues\(^{38}\) noted early on that most EPA literature is theoretical, offering no clear links to patient outcomes, and others\(^{64}\) have reiterated similar concerns:

> Although we will not have proof that our proposed road leads to desired patient care outcomes until implementation and study, we have little evidence supporting the current model despite its entrenchment, and we have clear evidence that this discontinuous path leads us away from, rather than closer to, desired outcomes.

Authors noted that a method for interpreting entrustment decisions as graduation requirements would be necessary if promotion committees were to discontinue time-based graduation systems, which would also require a nationally acceptable plan for handling learners who do not meet graduation requirements “on time.”\(^{39}\) Others have observed that answering these and other questions will require multi-institutional, prospective research.\(^{34}\) The AAMC Core EPA project has promised to provide some of these answers,\(^{19}\) including through recent efforts to engage the broader education community to define, collectively, how it will advance the feasibility and generalizability of the pilot beyond the work being done by the pilot schools.\(^{78}\)

### Inconsistency and Conflation

In addition to results related to authorship, article type, theoretical frameworks, and the Ottawa Criteria, we discovered that not all of the publications explored EPAs as an assessment approach. Furthermore, we found inconsistent uses of the term “EPA” and conflation of EPA concepts. Specifically, we observed the term “EPA” to signify all of the following:

- **An EPA**: An activity or unit of professional practice that can be entrusted to an individual.

- **EPA Assessment**: Assessment of an individual’s competence in a single EPA or a set of EPAs using an entrustment or supervisory scale.

- **EPA Advancement Decision**: The use of EPA assessment results (typically combined with other assessment information) to inform student outcomes.

- **EPA Curricular Framework**: The use of EPAs to guide an entire program of learning.

The most basic use of the term “EPA” is, as inferred by the name, simply an activity that may be entrusted to a trainee and that, in conjunction with other EPAs, operationally defines a profession: *An EPA*. This basic term was occasionally confused with skills trainees have (e.g., interpreting an electrocardiogram [EKG] reading). The skill of interpreting an EKG is a competency required for a larger EPA, but it does not define a profession. A second use of “EPA” referred to the application of an entrustment or supervisory scale to directly assess trainee performance of an EPA (e.g., assessing a student’s performance of collecting a history by using a nine-level supervision scale\(^{31}\)): *An EPA Assessment*. The third use of the term “EPA” referred to using entrustment for an EPA based on various sources of assessment information (e.g., knowledge tests, multisource feedback, EPA assessments) to make a summative decision about students (e.g., to determine grades, advancement, or even graduation): *An EPA Advancement Decision*. A final, fourth use of the term “EPA” referred to an entire program of learning and/or assessment designed around EPAs: *An EPA Curricular Framework*. 
Discussion

We found that most publications on EPAs in UME were enthusiastic. However, a few authors (Carraccio, Englander, ten Cate) commanded the field. While the number of empirical articles is increasing, the 26 available studies do not yet sufficiently support the use of EPAs as an assessment approach in UME. This finding is concerning because many have proposed implementing the EPA approach as a cornerstone of medical curricula.17 While the EPA approach may eventually fill this role, more rigorous empirical research is needed. For example, while more than half of the publications referenced how assessments with EPAs align with other educational assessment frameworks, many did not reference a theoretical framework, which can undermine the quality of the results.

Each of the seven Ottawa Criteria for a good assessment have been explored empirically in the publications we examined. More work is needed to map EPA assessments in UME to all seven criteria, but especially reproducibility/consistency, equivalence, educational effect, and catalytic effect. Specifically, research to determine how to improve the reproducibility of assessments with EPAs in UME is needed because the limited empirical work in this area indicated that reproducibility may be a weakness of using EPAs for assessment in UME. The purported shared mental model of entrustment with regard to medical students completing EPAs8 does not appear to be as intuitive as hoped. While many of the publications we reviewed provided potential solutions (e.g., video training, increased faculty development), the lack of reproducibility and shared mental models might be solved with improved clarity of EPA terminology and accurately defining EPAs.

The use of the term “EPA” is inconsistent in the current literature on EPAs in UME. One of the most important tasks in any field of study is to develop a shared nomenclature. It is only through a shared understanding of words that shared concepts can evolve into more focused ideas. When words are used imprecisely, generalizing results is more difficult. Many of the publications we reviewed confounded four uses of the term “EPA”: EPAs as activities (the original definition offered by ten Cate and Scheele6), EPA assessments, EPA advancement decisions, and EPAs as curricular frameworks. This confusion, as predicted,26 may have contributed to results or findings being falsely attributed to a particular construct and misapplying or inappropriately generalizing others’ work. We recommend returning to the original definition of an EPA as only the activity and then adding (as we have done above) additional terms/descriptors for other uses of the term “EPA.” Adoption of consistent language to clarify topics of study surrounding the use of EPAs will enable the field to move forward.

The AAMC’s 13 Core EPAs have received substantial attention (63.4% of all publications we reviewed referenced the project). While some publications propose aligning UME curricula according to the 13 Core EPAs,17,39,79 our scoping review of the current state of science regarding EPAs suggests that doing this may be premature. The AAMC’s 13 Core EPAs and their longitudinal study were developed for U.S. MD-granting medical schools. The findings of the AAMC’s pilot, which have yet to be released, may not generalize to medical education systems outside of the United States or to other programs (e.g., osteopathic programs). Additionally, several tools for designing EPAs47 and assessing the quality of EPAs45,76 were unavailable when the Core EPAs were developed. We share previous concerns26,90 that several of the 13 Core EPAs may need to be revised and may not be as generalizable as some are purporting. As such, educators working on projects that build on the Core EPAs, as if these were the definitive set of EPAs for UME, should proceed with caution.

Several efforts were taken to ensure that the results of this scoping review would benefit the field. The methodology we used in this scoping review was rigorous, following evidence-based guidelines on how to effectively scope a field of study.20,21,23,24 We solicited input from a wide range of stakeholders (AAMC EPA listserv) and sought input on early drafts of this report from leaders in the field. We also used a widely accepted framework of good assessment (the Ottawa Criteria)22 to organize our coding, which helped us present our results in a defensible manner. Lastly, we have diverse expertise from the fields of assessment, EPAs, and knowledge synthesis.

Along with its strengths, this review has some limitations—namely, that we specifically chose to review the use of EPAs in UME as an assessment approach. This decision led to our selection of the Ottawa Criteria,22 which apply to single assessments, as a framework for our coding. We did not realize until we were synthesizing the results of our coding that the literature had not been consistent in referring to EPAs as a tool for single assessments. This lack of consistency may have influenced the degree to which the Ottawa Criteria applied to the publications we reviewed.

Another potential limitation is how we created the search for our review. We chose not to search for each term within the phrase “entrustable professional activities” separately, so we may have missed relevant work on each of these individual concepts. However, searching for component terms would have dramatically changed our research approach because these individual terms would not have fit within the construct of an assessment. We also designed our review to focus on EPA literature specifically applied in UME. This decision was in response to the notable efforts focused on implementing EPAs in UME, where learning and assessments occur outside the workplace. EPAs were designed for workplace-based assessment, so their application during the UME phase of medical training is unique. Thus, while this review captures the scope of literature on EPAs in UME, it likely does not capture some empirical work undertaken in a more general sense, not specific to UME. Lastly, the use of EPAs is a rapidly expanding area of the literature, and additional articles have likely been published since we undertook our review. Despite these limitations, the review demonstrates a clear progression in the field of EPAs in UME and helps illuminate areas of weakness and areas needing further study.

Our review supports the need for increased empirical work founded in theoretical frameworks to further clarify terminology and understand if assessments and summative decisions making with EPAs meet the criteria for a good assessment in a UME environment.22 The new Ottawa Criteria,96 designed for broader systems of assessment, might
be applied to better understand EPAs as a framework for assessment, while the EQual rubric44 will help appraise the quality of EPAs themselves. The development and use of EPAs, EPA assessments, EPA advancement decisions, and EPA curricular frameworks represent novel approaches to several aspects of UME. There is a great deal of excitement and enthusiasm in the field related to EPAs. Much has been published in a short period of time, and the concept of EPAs has permeated the medical education community. This energy is needed to motivate the numerous empirical studies necessary to move the field forward; however, if the terminology and conceptual frameworks surrounding EPAs are not more explicit, established, and universally accepted, results will likely be equivocal, generalizations will be difficult to substantiate, and EPAs could simply be remembered as a fad. The first results of the AAMC’s Core EPA project are soon to be released. The timing could not be better for the medical education community to develop a shared language of EPAs and take the next steps toward a truly competency-based medical curriculum.

Acknowledgments: The authors wish to thank Dr. Olle ten Cate and Dr. Robert Englander for their thoughtful review of a preliminary draft of this review. They also wish to thank Ms. Rhonda Allard, research librarian at Uniformed Services University of the Health Sciences, for her invaluable contributions to their literature search.

Funding/Support: None reported.

Other disclosures: None reported.

Ethical approval: Reported as not applicable.

Disclaimer: The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Uniformed Services University of the Health Sciences, the Department of Defense, or the U.S. Government.


E.G. Meyer is assistant professor, Department of Psychiatry, Uniformed Services University of the Health Sciences, Bethesda, Maryland; ORCID: https://orcid.org/0000-0002-0538-4344.

H.C. Chen is professor, Department of Pediatrics, Georgetown University School of Medicine, Washington, D.C.; ORCID: https://orcid.org/0000-0003-1663-1598.

S. Uijtdehaage is professor, Department of Medicine, Division of Health Professions Education, Uniformed Services University of the Health Sciences, Bethesda, Maryland; ORCID: https://orcid.org/0000-0001-8598-4683.

S.J. Durning is professor, Department of Medicine, Division of Health Professions Education, Uniformed Services University of the Health Sciences, Bethesda, Maryland.

L.A. Maggio is associate professor, Department of Medicine, Division of Health Professions Education, Uniformed Services University of the Health Sciences, Bethesda, Maryland; ORCID: https://orcid.org/0000-0002-2997-6133.

References

6 ten Cate O. Trust, competence, and the supervisor’s role in postgraduate training. BMJ. 2006;333:748–751.
8 Englander R, Carraccio C. A lack of continuity in education, training, and practice violates the “do no harm” principle. Acad Med. 2018;93(35/35 Competency-Based, Time-Variable Education in the Health Professions):S12–S16.
15 Hirsh DA, Holmboe ES, ten Cate O. Time to trust: Longitudinal integrated clerkships and entrustable professional activities. Acad Med. 2014;89:201–204.
Review


31 Chen HC, van den Broek WE, ten Cate O. The case for use of entrustable professional activities in undergraduate medical education. Acad Med. 2015;90:431–436.


be used to assess competency in medical students? J Gen Intern Med. 2017;32: s675–s676.
77 Klapheke M, Johnson T, Cubero M. Assessing entrustable professional activities during the psychiatry clerkship. Acad Psychiatry. 2017;41:345–349.
95 Association of American Medical Colleges Learn Serve Lead: Proceedings of the 56th Annual Research in Medical Education Sessions);S110–S117.