Competencies, Milestones, and Entrustable Professional Activities: What They Are, What They Could Be

As I sat with my faculty colleagues in a conference room reviewing the files of residents for our residency program's clinical competency committee, I couldn't help reflecting upon other ways to rate performance, such as those used in competitive sports. In some cases, like track and field competitions, the rating process is easy; if there is a race, the fastest runner wins and is judged to be the best. Also, a stopwatch provides comparison with other runners across the country. In baseball, there are numerous statistics about batting averages and home runs that can be used to rate and sort the quality of hitting. Even one extra hit every 10 games can make the difference between a good player and an all-star. Sometimes, assessments are more nuanced and require the judgments of experts as well as agreed-upon standards for the experts, to use. For example, ice skaters and gymnasts at the Olympics are scored by experts, based on the performance of various difficult turns, jumps, and landings. As an observer I can recognize the falls but need the observations and judgments of experts to determine the scores. There are even dog shows that rate dogs according to how close they came to an ideal set of characteristics for the breed.

As we discussed each resident, I wondered which methods of judgment we were using. We were trying to identify which residents had demonstrated that they had met various standards, called competencies, based on the data we had collected. In essence, we were trying to use our data to judge who was making progress toward becoming a good doctor and who might represent a substantial risk for causing harm to patients. This was a serious and difficult task because of ambiguities in the standards we used and the limitations in the completeness and accuracy of our data. Although we had several kinds of data, including performance on annual in-service exams, we could have used more information about the residents' professionalism, communication skills, or even their patient care activities. We knew our decisions could have serious consequences to individuals judged to be performing below the standards as well as to patients who would depend on our assessments when they sought care from our graduates.

Our conversations usually began with a review of data—the scores on standardized tests, the completion of required projects, and the observations of faculty in various settings, reflected in evaluation forms. However, the comments on the evaluation forms tended to cluster, with little differentiation between residents. Discussions often veered off into anecdotes. For example, for one resident, one faculty member said, "Joe (name and details changed) is a well-meaning, capable resident, but at times he can be too aggressive, too pushy, especially with the consultants. I got an earful from one of the cardiology fellows about him the other night."

"Really? I did too, but from psychiatry," responded another faculty. "He was pushing for a patient's transfer to a psych facility even though there were no available beds."

"But he usually is doing it for the patients' best interests after they have been waiting a long time. He's just trying to make a decision," said a third.

“Well, I agree that he is usually doing it for the right reasons, but he needs to lower the volume,” responded a fourth faculty member. “We are all under stress.”

“Does anyone have anything to say about his patient care, his systems-based practice, or his practice-based learning, so that we can fill out his milestones?” asked the program director.

“I’ve only worked with him a few times,” answered the first faculty member, “and we were so busy, I hardly had a chance to watch him. But he seemed fine. He seemed to know his medicine. He even filled out one of those patient safety forms when a nurse gave the wrong medication.”

“Yeah, that’s great. We need all the residents to do that. I agree. I’ve worked with him two or three times. He did a good job on documentation of his charts. He knew about how the documentation affected billing. I think he will be okay,” said another faculty member.

We rated him a little above average for his year in some of the milestones, a little below average for some others, checked the various required boxes, made a note of his need to work on interpersonal skills, and moved on to the next resident.

This was how our meeting went for each resident: an initial review of the data and then an unstructured discussion about certain characteristics of the resident that could range from adequacy of knowledge to communication style to personal behavior. In spite of limited data, we reached consensus on almost everyone in this way, deciding that most of the residents were progressing nicely, identifying a few areas for needed improvement for some, and agreeing that two residents had potential problems in either knowledge or communications skills. It seemed that our gut feelings about the residents were relatively consistent even if we could not point to very many specific data to support them.

I left the meeting feeling satisfied that we had fulfilled our responsibilities, but I was troubled by the current assessment process. Had we paid enough careful consistent attention to competencies, milestones, and entrustable professional activities, terms that we used freely but did not seem to fully understand? In fact, many of the faculty had expressed...
frustration with the proliferation of terms and the tasks associated with them that added to their work without any clear improvement in assessment. Because much of our confusion was about the meaning and implementation of certain important terms, I have offered below a description of what the key terms mean to me, which may be helpful as I attempt to develop some recommendations for improving the assessment process. The Alliance for Academic Internal Medicine also provides useful definitions of these and other terms.1

*Competency* is the term often used to refer to a specific area of performance that can be described and measured, such as the competencies (summarized by Swing2) that were identified by the Outcomes Project of the Accreditation Council for Graduate Medical Education (ACGME): patient care, medical knowledge, professionalism, practice-based learning and improvement, systems-based practice, and interpersonal and communication skills. Carraccio et al3 described the four steps of competency development as competency identification, determination of competency components and performance levels, competency evaluation, and overall assessment of the process.

*Competent and competence* are terms often used when describing a global, general impression of the adequacy of knowledge, clinical skills, and attitudes of a health care provider to practice independently and autonomously, usually at the end of residency training. For example, “She demonstrates above-average competence in patient care; her performance shows she is a competent doctor.”

*Milestone* refers to a point along a continuum of a competency or subcompetency; milestones are clearly described and are usually specialty-specific. For example, a milestone for internal medicine might be the ability to gather information to define and diagnose a patient’s problem, and for emergency medicine, a milestone might be the ability to perform airway management skills for a patient with respiratory failure. A milestone will characterize expectations for residents at various stages of the development of expertise in a particular competency.

An *entrustable professional activity* (EPA) refers to a discrete clinical activity that requires the utilization and integration of various competencies and represents an activity associated with a specific clinical event. It often entails some risk to patients and the need for a faculty supervisor to assess the skills of the learner and allocate a level of responsibility that demonstrates a corresponding level of trust (i.e., entrustment) in the resident’s performance. For example, delivering a baby, or resuscitating a patient after a cardiac arrest, are examples of EPAs.

The importance of these terms is that they currently form the foundational language of competency-based assessment and help in the creation of a shared conceptual model of resident or student assessment and feedback. However, because of confusion about the meanings of these terms, they also are at the crux of many of the problems faculty have with the current assessment structure.

Assessment methods in medical education have become tied to our conceptions of the development of expertise in medicine. Unless we know what the continuum of performance looks like, it is difficult to know how to judge and assess a student along that continuum. Dreyfus4 proposed five levels of performance, which he labeled *novice, advanced beginner, competence, proficiency, and expertise*. This model suggests that a medical trainee progresses through the levels of increasing expertise during training and, at some point, would reach a level in which performance was deemed competent. This would presumably correspond with safe, independent, autonomous, high-quality care. However, it should be remembered that according to the Dreyfus model, the level called *competence* occurs in the middle of the continuum and is followed by two additional and higher levels that may be achieved. This placement of competence in the middle of the expertise spectrum can create a problem with the meaning of *competence*, since there are clearly higher levels of performance. It raises the question, Why not expect the provider to reach the next level—proficiency—or even expertise—before independent practice?

Most research on the development of expertise involves the study of the individuals at the two ends of the spectrum, the novice and expert, and attempts to understand how their thought processes differ.5 Although this research is informative of the changes that occur during the development of expertise, it does not answer questions about the relative capabilities of individuals along the continuum to provide high-quality care. Ericsson6 offers a perspective on expertise by examining experts in other fields than medicine, such as music and chess, and attempts to apply his findings to health care. Ericsson suggests that experts require deliberate practice and coaching to reach expert skill levels and that their expertise can be distinguished from the skill levels of others who are experienced but not expert. However, it is not clear whether expert-level performance is necessary for the independent practice of medicine, or if so, under what circumstances. It is also not clear how many physicians would qualify as experts and whether such a standard would result in a deficit of needed providers. While it is likely that not all patients have medical problems that would require the skills of an expert, knowing which patients do require such expertise can usually be done with certainty only in retrospect. A meaningful description of competence and expertise and their relationship to quality of care would have important implications for workforce estimates, licensing decisions, and team-based care as well as the assessment of residents.

Assessments of clinical competence typically involve a combination of direct workplace observation by faculty, observations with standardized patients or simulated patients, and standardized testing of medical knowledge. Although standardized tests such as in-service exams and board certification exams have clear criteria for passing, other assessments are more subjective, and in either case there is not an expectation that the performance level for passing should be that of an expert. Added to this uncertainty about the evaluation of competence is the general lack of consistent data from observations by faculty.

In this month’s *Academic Medicine*, we have several articles that deepen our understanding of competencies, milestones, and EPAs. Chen et al7 describe the rationale for EPAs in undergraduate medical education. EPAs were originally introduced by ten Cate8 as a way to describe the work performed by various resident trainees and the relationship of trust between trainees and faculty supervisors that could be useful in assessing the progressive expertise of the resident. Chen et al suggest that an
undergraduate EPA can be viewed either as a lower level of entrustment of an EPA that has already been defined for graduate medical education—such as management of chronic disease or performance of a normal delivery—or as an entirely different EPA skill set appropriate for a medical student, such as performance of a history and physical exam, that would be accomplished correctly and independently prior to beginning internship.

Caverzagie et al9 describe the development of EPAs for internal medicine and how they will be integrated into milestones and competencies. Because EPAs are clinical activities related to a discrete aspect of patient care, they have competencies associated with them, and each of the competencies can be assessed through milestones.

Both Chen et al and Caverzagie et al envision EPAs as helping to overcome the difficulties in assessment of competencies and milestones that Williams et al10 raise regarding the inadequacy of current global assessment strategies to measure milestones for the competencies.

In a Commentary, Holmboe11 responds to Williams et al with both an acknowledgment of the challenges of workplace assessment and also a plea for greater faculty commitment to workplace assessment so that milestones and competency-based assessments can reach their full potential in providing formative competency-based assessments can reach their full potential.

Although most of the emphasis on competencies and EPAs is focused on learners and teachers and is independent of the care delivery system, it has also been shown that the environment of care provides strong incentives for trainee behaviors and is related to the quality of care provided later.3 9 Competent trainees need to be trained in health care environments that allow them to learn and model high-quality care. Assessment of trainees must include assessment of their training environment.

In one of his early publications on EPAs, ten Cate4 noted the important role of subjectivity and gut feelings in the assessment of learners: “Educators do not fully exploit these gut feelings about trustworthiness for assessment purposes.” I believe this may be the key concept that links all the others. Competency-based assessment requires the investment in time and energy to create trusting relationships with trainees so that we can have well-founded gut feelings about our trainees. Without the investment in time to provide capable faculty assessors, all of the checklists and evaluation forms will be of little value. With trained and motivated faculty, the other sources of data can be placed into context and enrich the conversations between faculty that need to continue to create the three-dimensional pictures that our residents require for their growth and that our patients deserve for their care.

David P. Sklar, MD

Editor’s Note: The opinions expressed in this editorial do not necessarily reflect the opinions of the AAMC or its members.

References

8. ten Cate O. Trust, competence, and the supervisor’s role in postgraduate training. BMJ. 2006;333:749–751.