

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/323495147>

Enhanced Requirements for Assessment in a Competency-Based, Time-Variable Medical Education System

Article *in* Academic Medicine · March 2018

DOI: 10.1097/ACM.0000000000002066

CITATIONS

0

READS

40

5 authors, including:



Olle ten Cate

University Medical Center Utrecht

484 PUBLICATIONS 7,322 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Attaining demonstrable competence for residency [View project](#)



A longitudinal study to determine the effects of flipped classroom [View project](#)

All content following this page was uploaded by [Olle ten Cate](#) on 18 March 2018.

The user has requested enhancement of the downloaded file.

Enhanced Requirements for Assessment in a Competency-Based, Time-Variable Medical Education System

Larry D. Gruppen, PhD, Olle ten Cate, PhD, Lorelei A. Lingard, PhD, Pim W. Teunissen, MD, PhD, and Jennifer R. Kogan, MD

Abstract

Competency-based, time-variable medical education has reshaped the perceptions and practices of teachers, curriculum designers, faculty developers, clinician educators, and program administrators. This increasingly popular approach highlights the fact that learning among different individuals varies in duration, foundation, and goal. Time variability places particular demands on the assessment data that are so necessary for

making decisions about learner progress. These decisions may be formative (e.g., feedback for improvement) or summative (e.g., decisions about advancing a student). This article identifies challenges to collecting assessment data and to making assessment decisions in a time-variable system. These challenges include managing assessment data, defining and making valid assessment decisions, innovating in assessment, and

modeling the considerable complexity of assessment in real-world settings and richly interconnected social systems. There are hopeful signs of creativity in assessment both from researchers and practitioners, but the transition from a traditional to a competency-based medical education system will likely continue to create much controversy and offer opportunities for originality and innovation in assessment.

Competency-based medical education (CBME) has captured considerable attention from teachers, learners, and regulatory bodies. It has changed the perceptions and practices of curriculum designers, faculty developers, clinician educators, and program administrators. CBME has particularly significant implications for assessment.

In 2010, Frank et al^{1,2} defined CBME as “an outcomes-based approach to the design, implementation, assessment, and evaluation of medical education programs, using an organizing framework of competencies” that “de-emphasizes time.” This definition highlights the contrast between traditional educational models, which can be characterized as “fixed time, variable outcomes,” and CBME, which can be described as “fixed outcomes, variable time.” Because fixed outcomes are necessary to ensure standards of

quality care and patient safety, time variability is a logical implication of CBME when one recognizes that individual learners come to medical education with variable preparation and skills, pursue and attain competencies at different rates, and demonstrate their competence while possibly needing remediation or acceleration.^{3–5} Fundamentally, CBME can be seen as an approach to honor the commitment to meet public expectations of the profession.

In this article, we will explore the implications of time variability for assessment. Assessment is even more important in CBME programs than in most traditional programs, particularly postgraduate ones. Traditional programs assume that a specified period in training will lead to competence; thus, assessment is largely a “safety check.” However, this assumption is crumbling under the weight of individual learner cases in which competence was not acquired in the set period of time and not established from the assessment data itself, which identify specific areas of deficiency and needed remediation. In CBME programs, which do not make this assumption, the learners’ progression trajectories are more individualized, and assessment must bear the significant burden of justifying consequential differences in length of training.

Principles of Assessment in CBME

Because CBME treats time as a variable rather than a constant, it requires an alternative method for deciding when a learner is prepared to move on to the next phase of education. This method is predicated on assessing and judging competence. *Competence* is defined as:

the array of abilities across multiple domains or aspects of physician performance in a certain context. Statements about competence require descriptive qualifiers to define the relevant abilities, context, and stage of training. Competence is multi-dimensional and dynamic. It changes with time, experience, and setting.¹

Time variability places considerable demands on the assessment process, demands that are both larger and more varied than is typical for most traditional programs. We will use a modern assessment validity framework^{5–7} to guide our examination of these demands.

A key principle is that assessment always has a target, a construct about which we are trying to make a judgment. Constructs are intangible and based on our theories of education, performance, judgment, etc.⁸ Common constructs for assessment in medicine are knowledge, professionalism, communication, teamwork, and numerous others. CBME posits that *competencies* are the key

Please see the end of this article for information about the authors.

Correspondence should be addressed to Larry Gruppen, Department of Learning Health Sciences, University of Michigan Medical School, 219 Victor Vaughan House, 1111 East Catherine St., Ann Arbor, MI 48109; telephone: (734) 936-1662; e-mail: lgruppen@umich.edu.

Acad Med. 2018;93:S17–S21.

doi: 10.1097/ACM.0000000000002066

Copyright © 2018 by the Association of American Medical Colleges

constructs about which we need to make judgments and for which we need to gather data.

A second key principle is the distinction between *assessment data* and the *assessment judgments* that are made from those data. In a multiple-choice test of knowledge, assessment data (the scores) are often distinct from the assessment judgments and decisions (pass/fail, remediation, commendation). In contrast, assessment judgments based on assessment data in the clinical arena can become less obvious. The faculty observer is often the source of the assessment data when evaluating performance; the coach or advocate of the learner and the judge making use of those data decide whether the learner is competent—that is, has achieved a milestone, can progress, etc.⁹ The measurement *tool* (form, checklist) merely provides data for these judgments and is only a means to this greater end, not the end itself.

A third key principle regards the decision that is the consequence of the judgment. In usual educational testing settings, this decision pertains to student progress or to receiving a diploma. In clinical education, the decision resulting from an assessment may also result in feedback or the permission to act with less supervision in patient care (i.e., to a decision that a learner can be entrusted to act at a specified level of supervision).¹⁰ Assessment decisions may be summative or high-stakes decisions (e.g., graduation) or formative, low-stakes decisions (e.g., performance feedback that a learner may use to guide her or his own learning).

Implications for Gathering Assessment Data

Assessment tasks

Because we cannot assess competencies directly, we must define construct-relevant tasks or activities that produce behaviors that we can observe and measure in some way. These tasks can be as simple as answering a question (in a survey of knowledge test) or as complex as discussing palliative care with a patient and her family. The selection of these tasks has profound consequences for the validity of our assessment judgments. Good assessment judgments require multiple sources of data from a range of relevant tasks that use a variety of

assessment methods.⁹ Thoughtful blueprinting of assessment procedures will improve the quality of assessment judgments.

Despite the fact that expert opinions do not always agree, they are increasingly considered valuable and even necessary to arrive at valid assessment decisions regarding clinical trainees.^{11,12} Such opinions require the alignment of assessment data gathering with how individual clinicians (as assessors) interact with and gain trust in trainees in the authentic work environment.^{13,14} This is a new domain of investigation¹⁵ related in part to entrustment decision making.^{16–18}

Innovation in assessment

CBME programs will need assessment expertise to guide innovation and experimentation in assessment methods and procedures for integrating and using the data. It will become necessary to develop or recruit expertise in assessment, data analytics, information management, and decision support if the full promise of CBME is to be realized. CBME will not be viable if we continue to depend on occasional, prescheduled written or performance examinations and fragmentary faculty evaluations of clinical performance. This will, in turn, require resource reallocation, often in the face of overall budgetary constraints.

If we commit to developing and using a variety of established and innovative assessment methods, CBME can take advantage of naturally occurring assessment data that are presently neglected. These include learner products (e.g., entries in an electronic health record), team-based performance (e.g., based on multisource feedback), administrative data (e.g., prior qualifications and tests, activities performed), faculty judgments (e.g., mini-CEX, other observations, case-based discussions), and other assessment opportunities that may not be presently thought of as “good enough” for high-stakes, summative assessment decisions.⁹ While single observation events require high psychometric standards, multiple measures in combination, each of which may not be standardized or reliably reproducible, may yield generalizable results.¹⁹ In principle, the notion of the entrustable professional activity (EPA)²⁰

illustrates this point. As a naturally occurring task (e.g., conducting a risk factor assessment for a health maintenance examination), an EPA can be both a unit or focus of instruction and an assessment task.

Patient outcomes offer other, critically important sources of assessment data that can be incorporated into competency decisions²¹ and enhance the validity of assessments that seek to address the triple aim of enhancing the patient experience, improving population health, and reducing costs.²² However, linking patient outcomes to educational activities is extremely complex.^{23–26} Gathering such data, understanding trainees’ contributions to the outcomes, and establishing how these data should inform competency decisions about individuals remain part of the research agenda for CBME.²⁷

Time flexibility in assessment

If time is not a fixed quantity in the CBME learning process, it also cannot be fixed in assessment. Because competency judgments can be made at any time, assessment data need to be available for those judgments and thus gathered (more or less) continuously. In particular, the formative uses of assessment data in providing feedback to learners need to be linked closely to the setting and time of the performance. Formative feedback requires a dialogue between teacher and learner²⁸ and, thus, assessments that are tied to those dialogues. Such assessments are time consuming for faculty and constitute a significant potential cost for implementing CBME.

Time-flexible assessment also imposes considerable demands on administrative and logistical resources, particularly in trying to schedule formal assessments (e.g., objective structured clinical examinations required for graduation) for large numbers of learners who are pursuing individual learning sequences or plans. Flexible and continuous assessment is particularly difficult to accommodate with the rigid scheduling of high-stakes assessment. The increased assessment flexibility required by CBME will necessitate significant organizational changes in the relevant examination bodies. However, time variability may also be beneficial in that respect. With the fixed length of training modules, there is

significant pressure to evaluate multiple learners in a short period of time, whereas variable length could spread assessment effort more evenly over time.²⁹

Managing assessment data

Because CBME assessment requires data from multiple sources that are gathered more frequently and on variable schedules, it requires a greater level of data sharing, management, and communication than is typical of more traditional assessment systems.³⁰ The logistics and the ethics of communicating a trainee's assessment data within and across programs is a necessary challenge that has not yet received a great deal of attention.^{31–33} Electronic portfolios of performance and assessment data may be part of the solution, but they lack standardization, being quite variable from one institution to the next. Learning analytics may serve to support such data management and analysis,³⁴ and mobile technology may be used for the collection of data in the natural course of clinical activities. Warm et al³⁵ demonstrated the feasibility of tracking almost 200 internal medicine residents over three years with 360,000 data points. CBME will need to address systems for interpreting and sharing competency assessment information across programs and stakeholders (including learners) that may span more than 30 years of professional practice and learning.

Context specificity

Context specificity is a problem for all assessments of complex performance that take place during real-world practice. A pervasive finding is that an assessment of performance in one setting, situation, or case does not perfectly predict performance in even a similar case. Context specificity requires that multiple assessments be done by multiple observers over multiple cases in a variety of contexts to obtain a meaningful and trustworthy estimate of performance. E-portfolios and mobile technology could help capture natural encounters in the workplace to provide feedback, formative assessment, and summative decisions regarding progress in various contexts.³⁶

Along with context specificity, assessment data may be limited by the existence of “implicit” components of competence that may not be easily definable, such as professional identity formation. The

inherent uncertainty and imprecision of assessment in aiding these decisions must not be forgotten.

Implications for Making Assessment Judgments

Formative and summative judgments

Assessment judgments range in the impact they have on the learner and the educational systems. High-stakes (summative) judgments include decisions about passing or failing a course, graduating or retaining a learner, investment in curricular changes, and decisions about competence. Low-stakes (formative) judgments include feedback to guide student self-regulated learning, self-testing such as is included in many e-learning modules, and progress testing.³⁷

In CBME, formative and summative decisions can be viewed as different ends of the same spectrum. A learner early in her training may be assessed for the purpose of providing formative guidance on how to address gaps and accentuate strengths. A more advanced learner may be assessed with the same method for a summative decision related to his progression or remediation. Importantly, the same assessment data can be used for either formative or summative judgments—an assessment activity, *per se*, is neither formative nor summative. However, summative judgments and decisions generally require greater amounts of higher-quality assessment data than do formative judgments.

In the programmatic assessment literature, a series of formative assessments, each of which separately serves to stimulate learning (an approach labeled *assessment for learning*), together may serve to make summative decisions (or what is called *assessment of learning*).³⁸ Similarly, entrustment decisions have been identified as part of ad hoc assessments about learners for training purposes in health care tasks and as part of summative decisions about certification for health care tasks.^{10,20}

Assessment standards and criteria

Summative judgments require not only solid assessment data but also a standard or criterion for performance that defines the decision (competent vs. not competent). These standards may

be quantitative and derived by formal group judgment procedures, such as Angoff or Hofstee methods,³⁹ or they may be qualitative and describe a given performance level based on needed supervision or milestone achievement.⁴⁰ The complex competencies identified by CBME do not readily lend themselves to traditional standardized psychometrics, so considerable effort has been devoted to systematizing more “subjective” judgments (whether qualitative or quantitative) on the part of faculty. These efforts include behavioral anchors for ratings scales, detailed descriptions of what “competent” performance looks like, and faculty development to calibrate faculty to a common set of criteria, such as milestones.⁴¹ Schuwirth et al^{12,30,42} also have emphasized that multiple “subjective” assessments from individual faculty raters may contribute multiple, meaningful perspectives on a learner, in spite of appearing unreliable according to psychometric theory.

EPAs and entrustment decisions

Although not a logical necessity for implementing CBME, EPAs have emerged as a key aspect of many CBME systems. EPAs have been defined as units of professional practice to be entrusted to learners for unsupervised execution once they have demonstrated adequate performance.⁴³ One of the key features of EPAs is the link between authentic, often-everyday tasks of a profession and the opportunities to observe and assess learners' performance completing those tasks.

A key component of the EPA concept is entrustment. Entrustment requires that faculty and assessors make a judgment that integrates learner performance with assessor expectations and the nature of the task/EPA. It has enormous practical value because it captures the assessor's overall judgment in the context of an important professional criterion—Is the learner “trustworthy” to do this task independently?—and the judgment is based on an estimation of the amount of supervision the learner requires.

Entrustment of a learner to perform a task independently may or may not be the goal for each EPA identified for an educational program. Depending on expectations, some EPAs may not be fully mastered by the end of the program.

However, under a CBME framework, individual learners may exceed program expectations and attain sufficient mastery for full entrustment of “unsupervised practice” long before the end of the program because of particular skills, motivation, or learning opportunities.

EPAs provide a framework for granting responsibility as soon as learners are ready for it. Not all learners will attain the targeted competencies at the same rate because of prior learning and differences in sequencing of learning experiences. Therefore, variation in the time required to master EPAs may require individualized training pathways.⁴⁴ Time variability in mastering individual EPAs is not necessarily the only contribution to variability in total program time; the variable number of EPAs that need to be mastered by an individual learner also may affect the total program time.

Entrustment decisions, as an assessment outcome, can be dichotomous (entrusted vs. not entrusted) or incremental to accommodate a development framework for assessment. Different levels of entrustment may reflect different levels of independence or supervision, such as the learner observing only, participating in the EPA with direct supervision (in the room), participating with indirect supervision (not in the room but quickly available), and with distant supervision not quickly available.⁴⁵

If the purpose of assessment is defined as making entrustment decisions for EPAs in clinical practice and the scale of this assessment is framed as the amount of supervision the learner requires, assessment instruments can be created. Early studies have found that these construct-aligned approaches¹³ show favorable validity evidence.^{14,46,47} However, it is important to acknowledge individual differences in faculty observers’ criteria for less intensive supervision. Variability among faculty judges is a threat to the validity of these entrustment judgments.

Transitions and individual plans

It is common for faculty and researchers to identify inadequate preparation in learners who are transitioning from one phase to another in medical education: from undergraduate to graduate education,^{48,49} from graduate to fellowship programs,⁵⁰ or into unsupervised

practice.⁵¹ These gaps in preparation have led to a variety of responses, including boot camps to help prepare medical students for the transition to internship.⁵² Similarly, many residency programs have immersion experiences at their beginning. CBME assessments could aid these efforts by providing clearer expectations of performance and better data on learners’ strengths and weaknesses so that remediation and early learning plans could be adapted to individual needs, rather than to the perceived needs of the group as a whole.

Systems for making assessment judgments

Clinical competency committees (CCCs or entrustment committees) provide an example of innovation in how faculty collaborate around making assessment decisions.^{53,54} These committees collect, review, and synthesize assessment data from various sources and times, organized around defined competencies. The CCC takes these assessment data and uses them to make assessment judgments about the competence of each student, balancing the risks and benefits of the decisions about trust and progression through the program.¹⁸ Although many CCCs tend to focus on identifying at-risk trainees,⁵³ this decision-making structure and process holds promise for making decisions regarding accelerated learning and early graduation as well.⁵⁵ Of particular importance to CBME, the CCC structure would and should support the implementation of a truly time-variable education process.

Conclusion

The assessment of learning outcomes has always been essential to all levels of medical education, but competency-based, time-variable education places particular demands on assessment quality, frequency, purpose, and management that exceed the traditional requirements. As CBME programs multiply and mature, the investment of time, money, and talent into assessment methods and systems will also grow. We envision that the field of assessment will be a dynamic area of innovation over the next several decades as medical education commits to meeting the obligations of being able to document its educational promises to learners, to patients, and to society.

Acknowledgments: The authors wish to thank the guest editorial team and reviewers of this supplement.

Funding/Support: This article is based on a larger paper that was prepared for and funded by the Josiah Macy Jr. Foundation and used in the “Achieving Competency-Based, Time-Variable Health Professions Education” conference held in Atlanta, Georgia, in June 2017.

Other disclosures: One author (O. ten Cate) is a member of the guest editorial team that coordinated this supplement.

Ethical approval: Reported as not applicable.

Previous presentations: Parts of this article were derived from a paper commissioned by the Josiah Macy Jr. Foundation, entitled “Time-Variable Training in Medicine: Insights Derived From the Literature and From Examples in Practice,” by Olle ten Cate, PhD, Eugene J.F.M. Custers, PhD, Larry D. Gruppen, PhD, Pim W. Teunissen, MD, PhD, Lorelei A. Lingard, PhD, and Jennifer Kogan, MD.

L.D. Gruppen is professor, Department of Learning Health Sciences, University of Michigan Medical School, Ann Arbor, Michigan.

O. ten Cate is professor of medical education, Center for Research and Development of Education, University Medical Center Utrecht, Utrecht, the Netherlands.

L.A. Lingard is professor, Department of Medicine, and director, Centre for Education Research & Innovation, Schulich School of Medicine & Dentistry, Western University, London, Ontario, Canada.

P.W. Teunissen is professor, School of Health Professions Education, Maastricht University, Maastricht, the Netherlands, and maternal fetal medicine specialist, VU University Medical Center, Amsterdam, the Netherlands.

J.R. Kogan is professor of medicine, Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania.

References

- 1 Frank JR, Snell LS, Cate OT, et al. Competency-based medical education: Theory to practice. *Med Teach*. 2010;32:638–645.
- 2 Frank JR, Mungroo R, Ahmad Y, Wang M, De Rossi S, Horsley T. Toward a definition of competency-based education in medicine: A systematic review of published definitions. *Med Teach*. 2010;32:631–637.
- 3 Long DM. Competency-based residency training: The next advance in graduate medical education. *Acad Med*. 2000;75:1178–1183.
- 4 Carraccio C, Wolfsthal SD, Englander R, Ferrentz K, Martin C. Shifting paradigms: From Flexner to competencies. *Acad Med*. 2002;77:361–367.
- 5 American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association; 2014.

- 6 Downing SM. Validity: On meaningful interpretation of assessment data. *Med Educ.* 2003;37:830–837.
- 7 Downing SM, Haladyna TM. Validity threats: Overcoming interference with proposed interpretations of assessment data. *Med Educ.* 2004;38:327–333.
- 8 Lurie SJ. History and practice of competency-based assessment. *Med Educ.* 2012;46:49–57.
- 9 Lockyer J, Carraccio C, Chan MK, et al; ICBME Collaborators. Core principles of assessment in competency-based medical education. *Med Teach.* 2017;39:609–616.
- 10 Ten Cate O. Entrustment as assessment: Recognizing the ability, the right, and the duty to act. *J Grad Med Educ.* 2016;8:261–262.
- 11 Gingerich A, Kogan J, Yeates P, Govaerts M, Holmboe E. Seeing the “black box” differently: Assessor cognition from three research perspectives. *Med Educ.* 2014;48:1055–1068.
- 12 Schuwirth LW, van der Vleuten CP. A plea for new psychometric models in educational assessment. *Med Educ.* 2006;40:296–300.
- 13 Crossley J, Johnson G, Booth J, Wade W. Good questions, good answers: Construct alignment improves the performance of workplace-based assessment scales. *Med Educ.* 2011;45:560–569.
- 14 Weller JM, Misur M, Nicolson S, et al. Can I leave the theatre? A key to more reliable workplace-based assessment. *Br J Anaesth.* 2014;112:1083–1091.
- 15 Gingerich A, Regehr G, Eva KW. Rater-based assessments as social judgments: Rethinking the etiology of rater errors. *Acad Med.* 2011;86(10 suppl):S1–S7.
- 16 Damodaran A, Shulruf B, Jones P. “Trust” versus “competency” in the workplace. *Med Educ.* 2017;51:338.
- 17 Holzhausen Y, Maaz A, Cianciolo AT, Ten Cate O, Peters H. Applying occupational and organizational psychology theory to entrustment decision-making about trainees in health care: A conceptual model. *Perspect Med Educ.* 2017;6:119–126.
- 18 Ten Cate O. Managing risks and benefits: Key issues in entrustment decisions. *Med Educ.* 2017;51:879–881.
- 19 Moonen-van Loon JM, Overeem K, Donkers HH, van der Vleuten CP, Driessen EW. Composite reliability of a workplace-based assessment toolbox for postgraduate medical education. *Adv Health Sci Educ Theory Pract.* 2013;18:1087–1102.
- 20 Ten Cate O, Hart D, Ankel F, et al; International Competency-Based Medical Education Collaborators. Entrustment decision making in clinical training. *Acad Med.* 2016;91:191–198.
- 21 Ten Cate O. Entrustment decisions: Bringing the patient into the assessment equation. *Acad Med.* 2017;92:736–738.
- 22 Berwick DM, Nolan TW, Whittington J. The triple aim: Care, health, and cost. *Health Aff (Millwood).* 2008;27:759–769.
- 23 Cook DA, West CP. Reconsidering the focus on “outcomes research” in medical education: A cautionary note. *Acad Med.* 2013;88:162–167.
- 24 Moore DE Jr, Green JS, Gallis HA. Achieving desired results and improved outcomes: Integrating planning and assessment throughout learning activities. *J Contin Educ Health Prof.* 2009;29:1–15.
- 25 Bzowczyk AS, Dow A, Knab MS. Evaluating the impact of educational interventions on patients and communities: A conceptual framework. *Acad Med.* 2017;92:1531–1535.
- 26 Smirnova A, Ravelli ACJ, Stalmeijer RE, et al. The association between learning climate and adverse obstetrical outcomes in 16 nontertiary obstetrics–gynecology departments in the Netherlands. *Acad Med.* 2017;92:1740–1748.
- 27 Gruppen L, Frank JR, Lockyer J, et al; ICBME Collaborators. Toward a research agenda for competency-based medical education. *Med Teach.* 2017;39:623–630.
- 28 Moonen-van Loon JM, Overeem K, Govaerts MJ, Verhoeven BH, van der Vleuten CP, Driessen EW. The reliability of multisource feedback in competency-based assessment programs: The effects of multiple occasions and assessor groups. *Acad Med.* 2015;90:1093–1099.
- 29 Kogan JR, Hatala R, Hauer KE, Holmboe E. Guidelines: The do’s, don’ts and don’t knows of direct observation of clinical skills in medical education. *Perspect Med Educ.* 2017;6:286–305.
- 30 Schuwirth LW, van der Vleuten CP. Programmatic assessment and Kane’s validity perspective. *Med Educ.* 2012;46:38–48.
- 31 Frelles SL, Baker EA, Papp KK, Durning SJ. Medical students during the internal medicine clerkships: Results of a national survey. *Acad Med.* 2008;83:876–881.
- 32 Cox SM. “Forward feeding” about students’ progress: Information on struggling medical students should not be shared among clerkship directors or with students’ current teachers. *Acad Med.* 2008;83:801.
- 33 Cleary L. “Forward feeding” about students’ progress: The case for longitudinal, progressive, and shared assessment of medical students. *Acad Med.* 2008;83:800.
- 34 van der Schaaf M, Donkers J, Slof B, et al. Improving workplace-based assessment and feedback by an e-portfolio enhanced with learning analytics. *Educ Technol Res Dev.* 2016;65:359–380.
- 35 Warm EJ, Held JD, Hellmann M, et al. Entrusting observable practice activities and milestones over the 36 months of an internal medicine residency. *Acad Med.* 2016;91:1398–1405.
- 36 Jonker G, Hoff RG, Ten Cate OT. A case for competency-based anaesthesiology training with entrustable professional activities: An agenda for development and research. *Eur J Anaesthesiol.* 2015;32:71–76.
- 37 Albanese M, Case SM. Progress testing: Critical analysis and suggested practices. *Adv Health Sci Educ Theory Pract.* 2016;21:221–234.
- 38 van der Vleuten CP, Schuwirth LW, Driessen EW, et al. A model for programmatic assessment fit for purpose. *Med Teach.* 2012;34:205–214.
- 39 Downing SM, Yudkowsky R. *Assessment in Health Professions Education.* New York, NY: Routledge; 2007.
- 40 Apramian T, Cristancho S, Sener A, Lingard L. How do thresholds of principle and preference influence surgeon assessments of learner performance? [published online ahead of print May 1, 2017]. *Ann Surg.* doi: 10.1097/SLA.0000000000002284.
- 41 Holmboe ES, Durning SJ, Hawkins RE, eds. *Evaluation of Clinical Competence.* 2nd ed. Philadelphia, PA: Elsevier; 2017.
- 42 Schuwirth L, Ash J. Assessing tomorrow’s learners: In competency-based education only a radically different holistic method of assessment will work. Six things we could forget. *Med Teach.* 2013;35:555–559.
- 43 ten Cate O. Entrustability of professional activities and competency-based training. *Med Educ.* 2005;39:1176–1177.
- 44 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;37:983–1002.
- 45 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.
- 46 Weller JM, Castanelli DJ, Chen Y, Jolly B. Making robust assessments of specialist trainees’ workplace performance. *Br J Anaesth.* 2017;118:207–214.
- 47 Mink RB, Schwartz A, Herman BE, et al. Validity of level of supervision scales for assessing pediatric fellows on the common pediatric subspecialty entrustable professional activities. *Acad Med.* 2018;93:283–291.
- 48 Minter RM, Amos KD, Bentz ML, et al. Transition to surgical residency: A multi-institutional study of perceived intern preparedness and the effect of a formal residency preparatory course in the fourth year of medical school. *Acad Med.* 2015;90:1116–1124.
- 49 Raymond MR, Mee J, King A, Haist SA, Winward ML. What new residents do during their initial months of training. *Acad Med.* 2011;86(10 suppl):S59–S62.
- 50 Mattar SG, Alseidi AA, Jones DB, et al. General surgery residency inadequately prepares trainees for fellowship: Results of a survey of fellowship program directors. *Ann Surg.* 2013;258:440–449.
- 51 Soper NJ, DaRosa DA. Presidential address: Engendering operative autonomy in surgical training. *Surgery.* 2014;156:745–751.
- 52 Cohen ER, Barsuk JH, Moazed F, et al. Making July safer: Simulation-based mastery learning during intern boot camp. *Acad Med.* 2013;88:233–239.
- 53 Hauer KE, Chesluk B, Iobst W, et al. Reviewing residents’ competence: A qualitative study of the role of clinical competency committees in performance assessment. *Acad Med.* 2015;90:1084–1092.
- 54 Brown DR, Warren JB, Hyderi A, et al; AAMC Core Entrustable Professional Activities for Entering Residency Entrustment Concept Group. Finding a path to entrustment in undergraduate medical education: A progress report from the AAMC Core Entrustable Professional Activities for Entering Residency Entrustment Concept Group. *Acad Med.* 2017;92:774–779.
- 55 Hauer KE, Cate OT, Boscardin CK, et al. Ensuring resident competence: A narrative review of the literature on group decision making to inform the work of clinical competency committees. *J Grad Med Educ.* 2016;8:156–164.