Chapter 19
Evaluating the Paradigm Shift from Time-Based Toward Competency-Based Medical Education: Implications for Curriculum and Assessment

Yoon Soo Park, Brian D. Hodges and Ara Tekian

Abstract In the early twentieth century, most curricula were based on a concept of fixed time. Students who successfully completed a program were judged to be competent. However, a paradigm shift toward competency-based education occurred at the end of the twentieth century, allowing only students who are judged "competent" to move forward in a professional school curriculum. There are significant implications to this paradigm shift, particularly for curricular design, performance assessment, faculty development, and resources. Educators may find challenges addressing individual learning differences—some students are able to progress easily in some subject areas, while some may continue to struggle. Learners can also progress at different rates in competency-based education programs. While it is relatively easy to develop competencies in areas of knowledge and skill, it is more difficult to define milestone assessments in areas such as reasoning and judgment, and to assess complex professional behaviors. The promise in competency-based education is to graduate professionals who are better adapted for the needs of complex and rapidly changing systems. Yet, implementing competency-based curricula raise important questions not only in instruction, but also in the assessment of competencies and outcomes. This chapter synthesizes the existing literature and perspectives that support and critique competency-based education, identifies pressing challenges for educators, and speculates on the future of this still emerging paradigm.

Y.S. Park (✉)
Department of Medical Education, University of Illinois at Chicago,
Chicago, IL, USA
e-mail: yspark2@uic.edu

B.D. Hodges
University of Toronto, Toronto, Canada

A. Tekian
University of Illinois at Chicago, Chicago, USA

Takeaways

- Traditional curricular structure in medical education has been defined through a *time-based model*, corresponding to fixed time spent in training.
- Scholars and practitioners have called for a move toward *competency-based model* that ensures the achievement of competencies and entrustment of skills as outcomes, driven by societal expectations and public accountability.
- Implementing a competency-based model accompanies challenges in instruction and assessment, faculty development, and allocation of resources.
- Designing a multifaceted and learner-centered workplace-based assessment in a competency-based setting requires overcoming traditional notions of psychometrics and interpretation of validity. Narrative assessments and subjective judgment models may offer new insights to address these issues.

19.1 Introduction

The movement toward reforming medical education curriculum has rapidly reshaped the instruction and assessment of learners entering the health professions workforce. Traditionally, the signature of medical education curriculum has been defined through a "structure/process" format, as outlined in Abraham Flexner's legacy—*Medical Education in the United States and Canada. A Report to the Carnegie Foundation for the Advancement of Teaching* (Flexner 1910). While Flexner's model introduced a scientific and evidence-based approach to teaching medicine, it is perhaps most noted for the creation of a binary structure to medical education curriculum in which students are taught basic sciences followed by clinical training. By the turn of the century, traditional curriculum in medical education was being challenged, targeting the need to emphasize competence development and entrustment to better meet public demands (Englander et al. 2015a; ten Cate and Scheele 2007). These challenges are in fact, not so foreign to Flexner's philosophy; Flexner also advocated for other issues in medical training, including societal expectations of the medical profession (Whitehead 2010).

The traditional curricular structure in medical education is commonly associated with a *time-based model*, where training corresponds to a fixed time spent in medical school or graduate medical training (Hodges 2010; Tekian et al. 2015; Gruppen et al. 2012). During the past two decades, the time-based model has received increasing attention and criticism from the health professions education community in North American and around the world. Calls for a shift toward *competency-based model* are receiving greater attention. Scholars, practitioners, and policymakers have voiced concerns over the fixed time approach that is currently

dominant in the education system. For example, in the United States, concerns about lack of sufficient training and inadequate readiness to practice following residency have been reported (Crosson et al. 2011). Ensuring the preparedness of newly trained health professionals with the knowledge and skills required for everyday practice is part of an ongoing movement documented in a recent report by the Institute of Medicine (IOM 2014). Moreover, it is within these initiatives that the Accreditation Council for Graduate Medical Education (ACGME) launched the Next Accreditation System (NAS), whereby resident progress is tracked through the achievement of developmental milestones that target each of the six ACGME core competencies (Nasca et al. 2012). Similar movements toward competency-based models are emerging throughout the world. It is at this tipping point that implications for teaching and assessing of specific competencies require further investigation.

While the concept of "competence" has received increasing attention from the field, a consensus definition of the term and a collective rationale for movement toward competency-based medical education (CBME) curriculum has received varying reactions. In this regard, this chapter aims to review the existing literature on the paradigm shift from time-based toward CBME curriculum; strengths and challenges of both models are identified. Moreover, challenges in implementing CBME and relevant assessment frameworks are discussed, including potential barriers faced with current conceptual notions of validity and assessment standards. Implications for the CBME movement in relation to existing assessment guidelines and notions of composite validity (AERA, APA and NCME 2014) and subjective judgment models are also discussed.

19.2 Two Models of Curriculum in Medical Education: Time and Competency

The current literature on competence development identifies two curricular models: (1) time-based models and (2) competency-based models. This section describes the two models and the current trend toward CBME.

19.2.1 Time-Based Model

Tea-steeping model. During the early twentieth century, Flexner's recommendation for scientific and evidence-based training signaled an alignment between clinical and basic sciences in medical education. This subsequently prompted medical schools to join or affirm relationships with universities (Starr 1982; Gidney and Millar 1994). However, universities at the time already conceptualized competence as a time-based tradition, using fixed length programs of study with the

mastery of knowledge and replication of facts and information as the central focus of assessments. This naturally led to a fixed-time model for medical education, which came to be divided into the foundational, basic sciences phase and a clinical phase. Within the time-based model, undergraduate medical education curriculum was characterized by a 2 + 2 structure in most North American programs or a 4 + 2 structure for programs in Europe for the basic sciences and for clinical years, respectively. The concept of fixed time also extended to graduate medical education, in which trainees graduated after an established duration of training time. Moreover, the notion of "rotations" also originated from universities that were affiliated with clinics and hospitals. As such, the organization and structure of time-based model in medical education curriculum can be seen as being rooted in the university's existing tradition (Hodges 2010).

Based on an assumed sufficiency of time as a condition for competence in medicine, Hodges (2010) coined the "tea-steeping" metaphor for the time-based model. Here, students are likened to tea leaves that are soaked for a fixed duration in hot water (i.e., medical school). In this view, changes to admissions policy are analogous to changing the type of tea, while altering or changing the curriculum or the school environment is simply changing the nature of water that is used to brew the tea.

Challenges to the time-based model. Two fundamental issues arise with the fixed, time-based model. First, the assessment structure of competence and proper entrustment of skills can be challenging, as it lacks the flexibility to meet individual progression toward competence. In time-based models, time is so central to completion that learners may graduate without having properly demonstrated mastery in required competencies. To address this limitation, a system to monitor and provide ongoing feedback for learners to achieve competence within the fixed time interval is needed. However, courses in time-based models are structured into predetermined intervals, making ongoing, individualized feedback structurally difficult; that is, monitoring and providing individualized feedback to learners in time-based models—particularly on achieving competencies—can be limited due to the inherent condition of graduating learners within a fixed time period.

While summative assessments are common in time-based curricular models, they do not necessarily provide information about whether learners have acquired the competencies to *perform* at work. Implementing workplace-based assessments (WBAs) that allow continuous measurement of skills in the clinical workplace can be difficult in a time-based environment, because of a lack of alignment between instruction in time-based modules and assessment of overall educational outcomes. Most curricular structure is oriented to completing a time-based block rather than aiming toward overarching educational outcomes. Furthermore, individual rates at which learners acquire necessary competencies may not be measured or tracked within the time-based model. This can also deter promoting self-regulated and flexible learning environments for learners (Gruppen et al. 2012). Because learners progress at different rates, some may achieve the required competencies prior to graduation. A one-size-fits all assessment policy rooted in time-based curricula may not be an adequate model for learning in the twenty-first century.

Second, changes or reforms are difficult to implement in a time-based model; altering curriculum often means only making minor adjustments to accommodate change, rather than a wholesale reform driving the curriculum. Within the traditional time-based curricula, proposals to modify the curriculum often take an additive approach, whereby new topics such as ethics or communications are simply added, rather than streamlining or removing existing content. Therefore, with learners' schedules already fully committed, modifications to curricula may only result in modest changes.

Given these structural issues in the time-based model, ensuring competence development can be problematic. Moreover, in an era when public accountability and societal expectations drive the need for evidence of competence development, the imperative for curricular reform becomes more pressing. Although the time-based model has adequately served the needs of health professionals since Flexner's era, these concerns have led to a movement toward CBME.

19.2.2 Competency-Based Models

Limitations of the time-based model have fueled discussions about adopting new, non-time-based, competency-based models. In this section, the movement and motivation toward competency-based models are introduced, including defining competence from a historic perspective. Factors that constituted competence of a health professional over a century ago during Flexner's era were quite different from the complex competency frameworks used today; a historic understanding of the evolution and conceptualization of competence will provide background for movement toward CBME.

Transition to "entrustment" as potential outcomes in CBME. The defining elements of competence during the early twentieth century, when the time-based model was being formulated, were knowledge or knowledge-based clinical performance. However, toward the end of the twentieth century, notions of competence evolved. During the latter twentieth century, competence frameworks extended beyond knowledge to include quality of patient care, interpersonal and communication skills, professionalism, and teamwork, among others. Moreover, with the introduction of Miller's Pyramid (Miller 1990), assessment of competence was regarded to include not only what a learner *knows*, but also what he or she *does*. The transition from conceptualizing competence as knowledge to conceiving it as a set of "entrustable" skills was an important shift between views of competence during Flexner's era and today.

Competency-based models of medical education curriculum are closely related to *outcome-based models* that were popular in general education in the mid twentieth century. Within outcome-based models, measurable competencies are defined, for which learners are trained; these measurable competencies can, in theory, lead to the creation of individualized, outcome-based curricula. In competency-based models, the outcomes originate from the needs of the

community; such community-based needs inform the desired competencies and outcomes of training, which in turn lead to constructing the curriculum and assessment (Tekian et al. 2015). Many competency-based models require the identification of *entrustable professional activities* (EPAs; ten Cate and Scheele 2007; Englander et al. 2015b), which mark measurable outcomes or competencies in CBME. It is important to emphasize that within such competency-based models, competence is based on *what learners do*—the highest level of Miller's Pyramid. Placing emphasis on actual practice abilities (including "entrustable professional acts") differs significantly from previous notions of assessments used to determine what learners *know how* to do or even *show how* to do—the latter two functioning as proxies for what learners actually *do* in practice.

19.2.2.1   Recommendations from North America that Motivate Medical Education Reform

The movement toward CBME was fueled by reports released from the United States and Canada. In 2010, both the United States and Canada released recommendations for reforming medical education. These reports were created to advance the state of medical education, noting that in many ways, there were modest changed in the 100 years since Flexner's 1910 report. This section summarizes the North American recommendations for medical education, emphasizing principles of CBME.

Recommendations from the United States and Canada. In the United States, four recommendations toward medical education reform were reported in *Educating Physicians: A Call for Reform of Medical School and Residency* (Cooke et al. 2010). Table 19.1 summarizes these recommendations: (1) standardizing outcomes, (2) integrating, (3) fostering habits of inquiry, and (4) forming an identity. Standardizing outcomes refers to both learning and practice outcomes; the report emphasizes identifying competencies and milestones and the use of multiple forms of assessments. Integration refers to linking knowledge and experience, while engaging in different forms of reasoning (analytic, pattern recognition, creative, and adaptive). For habits of inquiry and improvement, the report recommends developing expertise through deliberate practice and feedback, while engaging in communities of inquiry and practice. Finally, for identity formation, the report recommends a commitment to values, participation in community practice, observing role models, and feedback. A similar set of recommendations was released in Canada by the Association of Faculties of Medicine of Canada (AFMC): *The Future of Medical Education in Canada (FMEC): A Collective Vision for MD Education* (AFMC 2010) for medical students and *A Collective Vision for Postgraduate Medical Education in Canada* (AFMC 2012) for postgraduates (see Table 19.1).

Common themes across the United States and Canada. A common theme across both undergraduate and postgraduate medical education from the United States and Canada is a call for outcomes and competency-based curriculum.

Table 19.1 Recommendations for medical education in the United States and Canada

United States[a] (Cooke et al. 2010)	Canada (AFMC 2010, 2012)	
	Undergraduate[b]	Postgraduate[c]
1. Standardize on outcomes that can allow flexibility in learning 2. Integrate knowledge and experience 3. Foster habits of inquiry and improvement that focus on excellence 4. Create professional identity that carry professional values and dispositions	1. Address individual and community needs 2. Enhance admissions processes 3. Build on the scientific basis of medicine 4. Promote prevention and public health 5. Address the hidden curriculum 6. Diversify learning contexts 7. Value generalism 8. Advance inter- and intra-professional practice 9. Adopt a competency-based and flexible approach 10. Foster medical leadership	1. Right mix, distribution and numbers of physicians 2. Diverse learning and work environments 3. A positive and supportive environment 4. Competency based curricula 5. Transitions along the medical educational continuum 6. Effective assessments systems 7. Support clinical teachers 8. Foster leadership development 9. Collaborative governance in PGME 10. Align accreditation standards

Note
[a]Cooke et al. (2010)
[b]AFMC (2010) report for medical students
[c]AFMC (2012) report for postgraduates

Embedded within this recommendation are three important motivations toward CBME (Hodges 2010):

- Increased efficiency
- Decreasing training time
- Reducing the overall cost of medical education

19.2.2.2 Defining and Identifying the Rationale for CBME

In this section, a collective definition of competency and its core elements are presented. The motivation for CBME and recommendations for its implementation are discussed.

Definition of competency. Although authors such as McGaghie et al. (1978) have argued that CBME cannot have a single definition, Frank et al. (2010a) conducted a systematic review of the literature to identify a unitary twenty-first century definition which is now in widespread usage by medical educators. Frank et al. defined CBME as "an approach to preparing physicians for practice that is fundamentally oriented to

Table 19.2 Rationale for implementing CBME and recommendations for implementation

Rationale for CBME	Recommendations for implementing CBME
1. Focus on curricular outcomes	1. Identify the abilities needed of graduates
2. Emphasis on abilities and competencies	2. Explicitly define the required competencies and their components
3. De-emphasis on time-based training	3. Define milestones along a development path for the competencies
4. Promotion of learner centeredness	4. Select educational activities, experiences, and instructional methods
	5. Select assessment tools to measure progress along the milestones
	6. Design an outcomes evaluation of the program

graduate outcome abilities and organized around competencies derived from an analysis of societal and patient needs. It deemphasizes time-based training and promises greater accountability flexibility, and learner-centeredness" (Frank et al. 2010a, p. 636). Supporting this definition, they identified eight components:

1. Defined outcomes and milestones (developmental levels that correspond to a competency framework)
2. Curriculum of competencies
3. Demonstrable/observable abilities
4. Assessment of competencies that indicate process toward defined outcomes
5. Learner-centered
6. Serving societal needs
7. Contrasting with time-based or process-based model
8. Implementation

Rationale for implementing CBME. Frank et al. (2010b) indicated that based on an international discussion and consensus, the rationale for CBME can be organized into four themes. To translate CBME into actual practice, the article also provided a six-step recommendation for planning CBME curricula. Table 19.2 summarizes these main points.

As scholars and practitioners in medical education consider the practical implications of adopting CBME into practice, an important challenge is the development of an assessment system to measure competencies and outcomes that have requisite quality and validity. The ensuing section discusses ideas that are emerging in the assessment of CBME and some of the challenges that lie ahead.

19.3   Assessments in CBME

To create a curriculum that focuses on outcomes and competencies and engages in learner centeredness, a multifaceted assessment system is needed (Holmboe et al. 2010). Furthermore, to create an assessment that adheres to the clinical work setting where medicine is practiced, WBAs need to be included in the assessment (Norcini

and Burch 2007). The nature of CBME requires assessment to be continuous, frequent, criterion-based, and developmental, while using tools that generate validated inferences about and for learners. Within this context, Holmboe et al. (2010) outlined six components of an effective assessment system in CBME:

1. Assessments need to be continuous and frequent
2. Assessments must be criterion-based, using a developmental perspective
3. Competency-based medical education, with its emphasis on preparation for what the trainee will ultimately do, requires robust work-based assessment
4. Training programs must use assessment tools that meet minimum standards of quality
5. We must be willing to incorporate more "qualitative" approaches to assessment
6. Assessment needs to draw upon the wisdom of a group and to involve active engagement by the trainee

The authors also note that future assessments will need to delve into interactions between competence and clinical practice. Concerns raised include reducing CBME into "checkboxes" (Talbot 2004), assessing an overly large number of milestones, and the lack of appropriate assessment forms. Measuring team-based competence and associated outcomes has also been noted as a challenge that will need to be addressed, as the practice of medicine occurs in collaborative environments. At present, most measures of competence relate to individuals and measurable outcomes and appropriate assessments for teams that are not readily available (Hodges 2010). Finally, faculty development needs to be addressed, helping faculty to understand best practices for assessment in the era of CBME.

19.4 Challenges in CBME in the Face of Modern Psychometrics

As noted thus far in this chapter, CBME poses challenges to traditional notions of assessment and in particular the concepts of psychometrics that up to now formed the basis for assessment standards in medical education (AERA, APA, and NCME 2014). In a time-based curriculum, traditional point-in-time summative assessments can be analyzed using standard psychometric validity frameworks. However, in CBME, where time is no longer fixed, psychometric approaches to analyze data about competencies and outcomes are unclear, when they are gathered over time; to date, there is no general framework for measuring competence along a developmental (longitudinal) framework.

Competencies and entrustable professional activities (EPAs). Based on trends in CBME, new assessment methods will need to be developed to measure outcomes: (1) competencies and (2) EPAs. Assessments in CBME need to ensure standardized levels of proficiency that all graduates of the program have sufficiently attained before being deemed "competent" as defined by a competency framework.

The ACGME core competencies (medical knowledge [MK], patient care [PC], professionalism [PROF], interpersonal and communication skills [ICS], practice-based learning and improvement [PBLI], and system-based practice [SBP]) can serve as one such framework. Within this framework, "milestones" are fine-grained developmental levels associated with the competencies. For example, in Internal Medicine residency programs, "Gathers and synthesizes essential and accurate information to define each patient's clinical problem(s)" (PC1; see ACGME and ABIM 2014) is a milestone corresponding to Patient Care with five distinct developmental levels, ranging from "Critical Deficiencies" (Level 1) to "Ready for Unsupervised Practice" (Level 4), and "Aspirational" (Level 5). In CBME, the goal is for learners to achieve a minimum of Level 4 prior to completing the residency program.

EPAs indicate what the learners can actually do (as opposed to qualities represented in competencies) and are units of professional practice. In this sense, EPAs describe specific work that professionals are entrusted to perform as outcomes of their curricular experience. Some EPAs can require multiple competencies for executing the work. For example, "performing an appendectomy" will require a learner's competence in MK and PC; "executing a patient handover" will require competence in MK, PC, ICS, and SBP (ten Cate 2013).

As is evident from these examples, measuring competencies and EPAs in CBME will require using different assessment methods and combining scores from different assessments. Measuring milestones in PC or EPAs that require PC could involve aggregating information from various assessments, such as written tests, rotation evaluations, and even objective structured clinical examinations (OSCEs). The reason for combining scores from different assessment is that each assessment only provides a partial picture of the learner's performance; a full picture of their competence or entrustment can only be gained by aggregating their performance across a "system" of assessments targeted to measure the learner's ability (Holmboe and Hawkins 2008). Each assessment can have different characteristics including but not limited to psychometric validity evidence. Finding methods to optimally combine these scores in a "system" of assessment, one that measures global concepts of competence and includes entrustment is a challenge that must be resolved urgently given that, in the United States currently, Clinical Competency Committees (CCCs) are tasked with aggregating learner performance to determine overall competence.

Composite reliability and validity. One potential approach to aggregate assessment scores is to use a *composite validity* approach, in which information from multiple assessment tools is used to generate inferences about competence (Park et al. 2016). Psychometricians have for decades studied measurement characteristics of combining constructs from multiple assessments (i.e., assessment system) to create validity evidence, including composite score reliability (e.g., see Kane and Case 2004; Borsboom 2012; Markus and Borsboom 2013; de la Torre and Douglas 2004; Rao and Sinharay 2007). For example, if (1) rotation evaluations, (2) written tests, and (3) OSCEs were used to measure a milestone in PC or an EPA associated with PC, prior literature on the psychometric characteristics of the different

Table 19.3 Assessments used to measure Internal Medicine milestone PC1: example of composite score reliability

Assessment used to measure PC1: "Gathers and synthesizes essential and accurate information to define each patient's clinical problem(s)"	Reliability	Weight (%)	Composite reliability
Rotation evaluations	0.45	50	**0.72**
Written examinations	0.65	25	
OSCE	0.55	25	

Note Reliability and weights based on hypothesized psychometric characteristics and importance judged by the institution. Calculation of composite reliability is based on Kane and Case (2004), assuming correlations between assessments range between 0.35 and 0.50

assessment methods indicates that perhaps only written tests would have sufficiently high reliability, while rotation evaluations or OSCEs may have lower reliability when administered in local medical schools or residency programs (e.g., Park et al. 2014; Yudkowsky et al. 2014). By applying a composite score reliability approach, the composite psychometric characteristics of the overall aggregated PC score may be better than the reliability of the individual assessments. Table 19.3 demonstrates an example for the Internal Medicine milestone PC1 ("Gathers and synthesizes essential and accurate information to define each patient's clinical problem(s)").

As demonstrated in Table 19.3, the individual reliability indices of the three assessments vary between 0.45 and 0.65, which may not satisfy sufficient reliability levels for making decisions on learners (Nunnally 1978). However, when the three assessments are viewed as indicators that form a composite, the composite reliability will increase to 0.72. This idea of composite reliability extends to composite validity, where validity evidence for the composite assessment is maximized by using information from different scores.

While tools such as composite reliability and validity can be translated to CBME, it is worth noting that these methods were not originally developed in the health professions setting. Historically, the existing validity frameworks used in medical education (e.g., Kane or Messick; see Brennan 2006; Kane 2013) were motivated from general education contexts where multiple-choice tests are of prevalent use (and performance-based assessments are beginning to gain further use and appreciation; see Gordon 2008). These contexts differ significantly from WBAs that are a required part of CBME. As such, psychometric research and frameworks for CBME need to be developed or refined. In this regard, the importance of developing new psychometric frameworks should be considered a priority, rather than as a long-term project or left as an initiative in which only testing organizations invest.

Subjective, holistic, and narrative forms of assessment. An alternative approach to the recombination of quantitative test scores and psychometric analysis is the use of expert raters to integrate complex information to make decisions about learner competence. Gingerich (2015) has pointed out that that "entrustment" is a social judgment and as such, may be more amenable to integrated, holistic thinking on the part of supervisors than that are areas of knowledge and skills, domains which are more easily broken down into component scores. Similarly, van der

Vleuten et al. (2010), Kuper et al. (2007), Rawlings et al. (2015), and Whitehead et al. (2015) have argued that much more use can be made of qualitative paradigms and data in creating assessment systems appropriate to integrated notions of competence. Hodges (2013) has argued that it may be possible to use multisource, subjective assessments of competence to create composite profiles of competence that are both richer and more informative for further learning than simply combining a set of quantitative scores from diverse instruments.

19.5 Conclusion

This chapter provides a broad overview of CBME starting with the reforms rooted in Flexner's 1910 report. Origins, concepts, and limitations of time-based medical education curriculum, which began from university-based traditions of fixed time courses, are presented. Time-based medical education and its enduring appeal is presented from a historical perspective, noting that the development of competency frameworks such as the ACGME core competencies and of conceptual models such as Miller's Pyramid has motivated the need to replace time-based models with competency-based models.

Among key limitations of time-based models is the lack of flexibility to measure meaningful, clinical practice outcomes due to the fixed time structure that forces assessment to be standardized, point-in-time and generally at the end of rotations. Moreover, curricular reform has been noted to be difficult in traditional time-based programs. Recognizing these limitations, some educators have suggested that a focus on curricular outcomes, emphasis on abilities and competencies, de-emphasis on time-based training, and promotion of learner centeredness are reasons for moving toward CBME.

Practical recommendations have been identified for transitioning from theory to implementation of CBME in medical school curricula. Some of the guidelines for assessments in CBME have been presented here. We have given particular emphasis to challenges and unresolved issues related to assessment and CBME. Psychometric issues arising from assessments in CBME are among important challenges associated with measuring outcomes when curricular time is no longer fixed. Furthermore, some competencies and EPAs require combining scores from different types of assessments. To this end, the use of composite reliability and validity is proposed and discussed as is the possibility of holistic, subjective judgment models. However, we argue, composite reliability and validity and judgment models need to be better conceptualized and refined by the medical education community before being widely implemented. We recommend further research to advance current understanding of assessment systems within a CBME framework.

Overall, CBME, if widely implemented, will be an important and meaningful change to the way health professionals are trained and assessed. It may overcome some of the limitations of the time-based model, and may better serve public

accountability, increase efficiency, and perhaps reduce unnecessary costs. As such, it may be part of a promising future in medical education. However, how well the competencies and outcomes—whether they are milestones or EPAs—can be assessed remains a challenge. Medical educators, psychometricians and those with expertise in judgment models need to develop new and innovative frameworks for addressing these challenges in measurement, as the CBME movement continues to accelerate.

Issues/Questions for Reflection

- Assessing competencies and EPAs in CBME will need continuous refinement to overcome challenges in meeting standards for validity
- Measuring team-based competence and associated outcomes will need to be examined and addressed in future assessments in CBME
- Psychometric concepts will need to be translated to meet CBME contexts, where assessments are longer administered in traditional point-in-time settings. Emerging solutions from narrative assessments and subjective judgment models may offer new insights
- Methods to allocate resources and support faculty to understand best practices for assessment will continue to require investigation

---

References

Accreditation Council for Graduate Medical Education & American Board of Internal Medicine. (2014). *The internal medicine milestone project*. Chicago, IL: ACGME.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards in educational and psychological testing*. Washington, DC: American Educational Research Association.

Association of Faculties of Medicine of Canada. (2010). *The future of medical education in Canada (FMEC): A collective vision for MD education*. Ottawa, ON: The Association of Faculties of Medicine of Canada.

Association of Faculties of Medicine of Canada. (2012). A collective vision for postgraduate medical education in Canada. Ottawa, ON: The Association of Faculties of Medicine of Canada.

Borsboom, D. (2012). Whose consensus is it anyway? Scientific versus legalistic conceptions of validity. *Measurement: Interdisciplinary Research & Perspective, 10*, 38–41.

Brennan, R. L. (2006). *Educational measurement* (4th ed.). Washington, DC: American Council on Education.

Cooke, M., Irby, D. M., & O'Brien, B. C. (2010). *Educating physicians: A call for reform of medical school and residency*. Stanford, CA: Jossey-Bass.

Crosson, F. J., Leu, J., Roemer, B. M., & Ross, M. N. (2011). Gaps in residency training should be addressed to better prepare doctors for a twenty-first-century delivery system. *Health Affairs, 30*, 2142–2148.

de la Torre, J., & Douglas, J. A. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika, 69*, 333–353.

Englander, R., Cameron, T., Addams, A., Bull, J., & Jacobs, J. (2015a). Understanding competency-based medical education. *Academic Medicine Rounds*. Retrieved from http://academicmedicineblog.org/understanding-competency-based-medical-education/

Englander, R., Cameron, T., Addams, A., Bull, J., & Jacobs, J. (2015b). Developing a framework for competency assessment: Entrustable professional activities (EPAs). *Academic Medicine Rounds*. Retrieved from http://academicmedicineblog.org/developing-a-framework-for-competency-assessment-entrustable-professional-activities-epas/

Flexner, A. (1910). *Medical education in the United States and Canada. A report to the Carnegie Foundation for the Advancement of Teaching* (Bulletin No. 4). Boston, MA: Updyke.

Frank, J. R., Mungroo, R., Ahmad, Y., Wang, M., de Rossi, S., & Horsley, T. (2010a). Toward a definition of competency-based education in medicine: A systematic review of published definitions. *Medical Teacher, 32*, 631–637.

Frank, J. R., Snell, L. S., ten Cate, O., Holmboe, E. S., Carraccio, C., Swing, S. R., et al. (2010b). Competency-based medical education: Theory to practice. *Medical Teacher, 32*, 638–645.

Gidney, R. D., & Millar, W. P. J. (1994). *Professional gentlemen: The professions in nineteenth-century Ontario (Ontario Historical Studies Series)*. Toronto: University of Toronto Press.

Gingerich, A. (2015). What if the 'trust' in entrustable were a social judgment? *Medical Education, 49*, 750–752.

Gordon, E. W. (2008). The transformation of key beliefs that have guided a century of assessment. In C. A. Dwyer (Ed.), *The future of assessment* (pp. 3–6). New York, NY: Taylor & Francis Group LLC.

Gruppen, L. D., Mangrulkar, R. S., & Kolars, J. C. (2012). The promise of competency-based education in the health professions for improving global health. *Human Resources for Health, 10*, 1–7.

Hodges, B. D. (2010). A tea-steeping or i-Doc model for medical education? *Academic Medicine, 85*, S34–S44.

Hodges, B. D. (2013). Assessment in the post-psychometric era: Learning to love the subjective and collective. *Medical Teacher, 35*(7), 564–568.

Holmboe, E. S., & Hawkins, R. E. (2008). *Practical guide to the evaluation of clinical competence*. Philadelphia, PA: Mosby Elsevier.

Holmboe, E. S., Sherbino, J., Long, D. M., Swing, S. R., & Frank, J. R. (2010). The role of assessment in competency-based medical education. *Medical Teacher, 32*, 676–682.

Institute of Medicine. (2014). *Graduate medical education that meets the nation's health needs*. Washington, DC: The National Academies Press.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*, 1–73.

Kane, M. T., & Case, S. M. (2004). The reliability and validity of weighted composite scores. *Applied Measurement in Education, 17*, 221–240.

Kuper, A., Reeves, S., Albert, M., & Hodges, B. D. (2007). Assessment: Do we need to broaden our methodological horizons? *Medical Education, 41*(12), 1121–1123.

Markus, K. A., & Borsboom, D. (2013). *Frontiers of test validity theory: Measurement, causation, and meaning*. New York: Routledge.

McGaghie, W. C., Miller, G. E., Sajid, A. W., & Telder, T. V. (1978). *Competency-based curriculum development in medical education*. Switzerland: World Health Organization.

Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine, 65*, S63–S67.

Nasca, T. J., Philibert, I., Brigham, T., & Flynn, T. C. (2012). The next GME accreditation system. *New England Journal of Medicine, 366*, 1051–1056.

Norcini, J., & Burch, V. (2007). Workplace-based assessment as an educational tool: AMEE guide no. 31. *Medical Teacher, 29*, 855–871.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Park, Y. S., Riddle, J., & Tekian, A. (2014). Validity evidence of resident competency ratings and the identification of problem residents. *Medical Education, 48*, 614–622.

Park, Y. S., Zar, F., Norcini, J., & Tekian, A. (2016). Competency evaluations in the Next Accreditation System: Contributing to guidelines and implications. *Teaching and Learning in Medicine*.

Rao, C. R., & Sinharay, S. (2007). *Handbook of Statistics: Psychometrics* (Vol. 26, pp. 979–1030). Amsterdam, Netherlands: Elsevier.

Rawlings, A., Knox, A., Park, Y. S., Reddy, S., Williams, S. R., & Issa, N. (2015). Development and evaluation of standardized narrative cases depicting the general surgery professionalism milestones. *Academic Medicine, 90*, 1109–1115.

Starr, P. (1982). *The social transformation of American medicine: The rise of a sovereign profession and the making of a vast industry*. New York: Basic Books.

Talbot, M. (2004). Monkey see, monkey do: A critique of the competency model in graduate medical education. *Medical Education, 38*, 587–592.

Tekian, A., Hodges, B. D., Roberts, T. E., Schuwirth, L., & Norcini, J. (2015). Assessing competencies using milestones along the way. *Medical Teacher, 37*(4), 399–402.

ten Cate, O. (2013). Nuts and bolts of entrustable professional activities. *Journal of Graduate Medical Education, 5*(1), 157–158.

ten Cate, O., & Scheele, F. (2007). Competency-based postgraduate training: Can we bridge the gap between theory and clinical practice? *Academic Medicine, 82*, 542–547.

van der Vleuten, C. P. M., Schuwirth, L. W. T., Scheele, F., Driessen, E. W., & Hodges, B. D. (2010). The assessment of professional competence: Building blocks for theory development. *Best Practice & Research Clinical Obstetrics & Gynaecology, 24*(6), 703–719.

Whitehead, C. (2010). Recipes for medical education reform: Will different ingredients create better doctors? *Social Science and Medicine, 70*, 1672–1676.

Whitehead, C. R., Kuper, A., Hodges, B., & Ellaway, R. (2015). Conceptual and practical challenges in the assessment of physician competencies. *Medical Teacher, 37*(3), 245–251.

Yudkowsky, R., Park, Y. S., Riddle, J., Palladino, C., & Bordage, G. (2014). Limiting checklist items to clinically-discriminating items: Improved validity of test scores. *Academic Medicine, 89*(7), 1–6.