

What's Next? Developing Systems of Assessment for Educational Settings

John Norcini, PhD

Systems of assessment, in the form of multiple examinations, have existed at least as far back as 650 CE when the Chinese offered a battery of tests for recruiting civil servants.¹ More recently, adequate performance on a variety of different assessments has been required at each step along the educational continuum to both establish competence and support learning. Although these examples constitute systems of assessment, explicit and orderly thought has not always been given to the relationships among the individual tests used in medical schools, residency, and continuing education or to the various purposes they serve. Instead, each individual assessment has been considered on its own merit, and the system has been viewed as the sum of its parts. Over the past few decades, this has begun to change.

In 2001, the National Research Council recommended the development of “new systems of multiple assessments,” which it followed with a series of reports elaborating on the idea.²⁻⁴ In 2005, van der Vleuten and Schuwirth also argued for a focus on programs or systems of assessment.⁵ In 2012, they followed with a specific system (“programmatically assessment”) that proposed ongoing feedback and assessment with periodic decisions based on an aggregation of the data points that were available.⁶ Starting in 2006, this movement was nurtured by the Harvard Macy course titled “A systems approach to assessment in health professions education.” The course was designed to encourage the application of systems thinking to the development of assessment programs that support a variety of purposes including student, faculty, and program accountability and improvement.⁷

In 2018, a consensus framework was proposed as part of the Ottawa Conference.⁸ It was built on the efforts of others, and it included 7 elements of a good system of assessment: coherent, continuous, comprehensive, feasible, purpose driven, acceptable, and transparent and free from bias.^{2-4,9-11} The aim of this work was to offer ways of thinking about important qualities of systems of assessment that might factor into judgments about best practice.

The increased emphasis on systems of assessment has occurred at a time when there have been fundamental changes to medical education. Competency- or outcomes-based education relies on the use of a variety of different methods of assessment, many of which go beyond knowledge to include skills and attitudes. At the same time, there is a need at various points in training to assess progress on the individual competencies as well as performances that require integration across them. Finally, there is growing acknowledgment of the power of formative assessment and the importance of including it in educational systems.

In the context of these developments, a number of questions need to be addressed for systems of assessment in undergraduate, postgraduate, and continuing medical education. These systems are remarkable in their complexity because they need to serve multiple purposes (e.g., formative, summative) for multiple users and stakeholders (e.g., students, faculty, patients, programs, institutions, health care systems), covering multiple competencies and unfolding over time. This Commentary will touch briefly on 4 areas that require work, although there are many others as well.

First, historically, virtually all assessments in education have been summative, and they have not necessarily been deployed in a coordinated fashion. Consequently, a key question for the future relates to how much summative assessment is actually needed to protect patients, establish accountability, and serve the

quality control needs of educational programs. Are there ways these summative assessments can be reduced, coordinated, and/or integrated to achieve better outcomes more efficiently? Can the resources this liberates be used for other purposes within the system?

Second, the recent interest in formative assessment is driven by research indicating that it has significant positive effects on learning and that it is the basis for any feedback provided to students. It faces significant challenges, however, in that it is underused, students do not necessarily value it, resource constraints often encourage the use of the same assessment for both formative and summative purposes, and faculty/institutions are neither expert in creating it nor in integrating it into their curricula, classes, and quality control systems. Underlying these challenges is a lack of research into fundamental issues. What are the effects of using the same assessment for both formative and summative purposes? What are the characteristics of effective formative assessment and the feedback provided on the basis of it? How much of what kind of formative assessment is needed and when? How can learners be motivated to use the feedback they receive? How do we train faculty to produce effective assessments of this type?

Third, the individual learner is often the object of measurement in assessment today. However, an integrated system of assessment would appropriately encompass other potential targets including the curriculum, the faculty, the effectiveness of particular educational interventions, the identification of strengths and weaknesses, the team, and so on. This raises several questions including how data can be shared across these various objectives, when data can be useful for more than one purpose, when and which assessments need to be developed specifically for these “other” objects of measurement, and so on.

Finally, assessment is an essential part of the training and feedback

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

Correspondence should be addressed to John Norcini; email: john.norcini@gmail.com.

Acad Med. 2019;94:S7-S8.

First published online July 30, 2019

doi: 10.1097/ACM.0000000000002908

Copyright © 2019 by the Association of American Medical Colleges

mechanisms in the complex adaptive system of medical education. It will need to support continuous monitoring of the system, rapidly signal problems, provide relevant information in a timely fashion to both users and decision makers, make efficient use of resources, and support needed change. To achieve these ends, it is necessary to apply “systems thinking” to the design of assessments. How do we bring this design approach to the currently decentralized assessment process of many educational programs and institutions? What type of expertise is needed? How do we know that the systems are working well?

There are the beginnings of research, development, and experience for all these questions, and some progress has been made, but further work is needed. Ultimately, it will be useful to create a set of best practices, recognizing that

they will vary by mission, context, and purpose.

Funding/Support: None reported.

Other disclosures: None reported.

Ethical approval: Reported as not applicable.

J. Norcini is research professor, SUNY Upstate Medical School, Syracuse, New York, and president emeritus, Foundation for the International Advancement of Medical Education and Research, Philadelphia, Pennsylvania.

References

- 1 Cheng L. Berkshire Encyclopedia of China. Great Barrington, MA: Berkshire Publishing Group; 2009.
- 2 National Research Council. Knowing What Students Know: The Science and Design of Educational Assessment. Washington, DC: National Academies Press; 2001.
- 3 National Research Council. Systems for State Science Assessment. Washington, DC: National Academies Press; 2006.
- 4 National Research Council. Developing Assessments for the Next Generation Science

Standards. Washington, DC: National Academies Press; 2014.

- 5 van der Vleuten CP, Schuwirth LW. Assessing professional competence: From methods to programmes. *Med Educ.* 2005;39:309–317.
- 6 van der Vleuten CP, Schuwirth LW, Driessen EW, et al. A model for programmatic assessment fit for purpose. *Med Teach.* 2012;34:205–214.
- 7 Harvard Macy Institute. A systems approach to assessment in health professions education. <https://www.harvardmacy.org/index.php/hmi-courses/assessment>. Accessed July 17, 2019.
- 8 Norcini J, Anderson MB, Bollela V, et al. 2018 Consensus framework for good assessment. *Med Teach.* 2018;40:1102–1109.
- 9 Clarke MM. What Matters Most for Student Assessment Systems: A Framework Paper. Washington, DC: World Bank; 2012. Working paper no 1.
- 10 Dijkstra J, Galbraith R, Hodges BD, et al. Expert validation of fit-for-purpose guidelines for designing programmes of assessment. *BMC Med Educ.* 2012;12:20.
- 11 Institutional Effectiveness and Assessment, St. Olaf College. Assessment of student learning. <https://wp.stolaf.edu/ir-e/assessment-of-student-learning-2>. Accessed July 17, 2019.