

Evaluation: Measuring Performance, Ensuring Competence, Achieving Long-term Excellence

THE education of medical students, residents, and practicing physicians is necessary to achieve high-quality health care. For this education to be effective, it must begin with a clear delineation of the goals that learners need to achieve. Evaluation of their subsequent mastery of these goals is an essential element of the educational process. In 1999, the Accreditation Council for Graduate Medical Education changed the focus of residency and fellowship program reviews to emphasize educational outcomes as a key measure of the quality of training programs. An objective of this change in focus is to improve educational outcomes and thus enhance the quality of medical care and increase patient safety. These educational outcomes can only be judged through effective evaluation.

Educational evaluation has two distinct aspects, generally classified as formative and summative evaluation. Formative evaluation, which can also be called feedback, is used to guide learners during the educational process, identifying those areas they have mastered and those areas that need additional work. With formative evaluation, a teacher's relationship with a learner is similar to that of a coach's relationship with an athlete. Learners, like athletes, need to invest the necessary time and effort to achieve their goals. Teachers are responsible for monitoring the learners' progress and providing guidance to help them achieve peak performance.

For formative evaluation to be effective, it must be provided close to the time of the observed evaluation and must contain specific observations that a learner can use to improve performance. This feedback must also be delivered in an appropriate setting and in a constructive manner. Unfortunately, most medical educators have not received formal training in providing evaluation and feedback. The result is often feedback that is deficient in both quality and quantity. To help learners achieve their maximum potential, medical educators must be skilled in providing feedback that helps learners change unsatisfactory performance into satisfactory performance and satisfactory performance into outstanding performance.

Summative evaluation is used to assess the learners' knowledge and/or performance at the completion of an educational activity, and it allows for judgment of the learners' success in achieving their educational goals. Examples of

summative evaluation include final examinations, end-of-rotation faculty evaluations, and patient surveys. Summative evaluation is not intended to change behavior during the educational activity; however, it can be used to develop future activities by identifying areas for improvement. Comparison of the results of summative evaluations with learners' self-evaluations can be used by both teachers and learners to assess the learners' skills in self-evaluation.

In the cognitive domain of learning, summative evaluation has traditionally relied on the use of written and oral examinations. Standardized multiple choice examinations provide an efficient method to assess medical knowledge. Large amounts of material can be tested and large numbers of learners can be assessed and compared using statistically reliable and valid methods. The use of criterion-referenced standards to determine passing performance on an examination allows for objective evaluation. Passing the examination is determined by specific performance criteria and not by the performance of the cohort of learners taking the examination.

Although they are more difficult to standardize and administer than written examinations, oral examinations may allow the assessment of more complex cognitive abilities, including analysis, synthesis, and evaluation. In constructing an oral examination, attention must be given to develop an examination format that provides a uniform and reproducible experience for multiple examinees. This requires careful construction of the questions and establishment of clearly defined criteria for grading. Education and assessment of the examiners are essential to make the oral examination reliable and to minimize examiner bias, thus ensuring an examinee would have the same result if examined by different examiners or given a different set of questions. Measuring the validity of the oral examination requires a metric against which examinee performance can be compared. Identification of such a metric may be difficult and can impair assessment of the examination's validity.

High-fidelity simulation provides a new method to evaluate a learner's performance in all the educational domains—cognitive, affective, and psychomotor—in real time. Anesthesiology has been in the forefront of the development of simulation as an educational tool in medicine. The next frontier in the use of simulation may be the development of summative examinations that can be used in making high-stakes decisions regarding residency education,

Accepted for publication October 13, 2009. The authors are not supported by, nor maintain any financial interest in, any commercial activity that may be associated with the topic of this article.

board certification, and the maintenance of certification. However, before the use of simulation for high-stakes evaluation becomes widely adopted, it will be necessary both to establish standards for the number and content of simulation evaluations required for reliable and valid examinations and to establish the performance criteria necessary to achieve a passing grade. This process will require the anesthesiology educational community to reach a consensus about the proposed standards and to perform pilot studies to validate the proposed methods. Recently, the American Board of Anesthesiology has added the completion of a continuing medical education simulation course as a requirement in the maintenance of certification. Although this does not use simulation for high-stakes evaluation, it is a recognition of the emerging role of simulation. The American Board of Anesthesiology is working with the American Society of Anesthesiologists Committee on Simulation to develop simulation experiences that will provide performance assessment and feedback.

Medical knowledge and practice advance continually. With the majority of a physician's career spent outside of a formal educational environment, development of lifelong learning skills is essential to ensure high-quality medical care. Physicians are expected to identify their own learning needs, develop plans for meeting their educational goals, and assess their success in achieving these goals. However, published data suggest that physicians are not skilled at self-evaluation.

Davis *et al.*¹ reviewed published studies of physicians' ability to self-evaluate. They only included studies that compared physicians' self-assessments using self-ratings with external observations. Seventeen studies met the inclusion criteria. Among the included studies, there were a total of 20 comparisons between self-rating and external observations. In only 35% of the comparisons (7 of 20) was there a positive association between self-ratings and external observations. In some cases, there was actually an inverse relationship between self-assessment and external observations. Indeed, the authors note: "A number of studies found the worst accuracy in self-assessment among physicians who were the least skilled and those who were the most confident." Although

research is limited, the existing data raise serious concerns about the current approach in continuing medical education for physicians.

More effective methods for physicians to identify their learning needs may include the use of self-audits and self-assessments. As an example of self-audit, physicians can perform chart reviews to compare their outcomes with established quality benchmarks. In self-assessments, physicians can take self-administered examinations to objectively identify knowledge deficits. These examinations often provide explanations of right and wrong answers and references for further study. Feedback from simulation experiences can also be used to identify learning needs.

Helping physicians to become skilled in self-evaluation must begin in medical school and continue through residency and beyond. Physicians should be taught how to identify their learning needs through the use of objective methods, such as self-audits and self-assessments. Anesthesiology educators can assist physicians with these tasks by creating valid and reliable tools for self-assessment.

Medical educators have a responsibility to ensure that effective methods of evaluation are used throughout the arc of a physician's career. Educators owe this duty both to the physicians they teach and to the society they serve. However, every physician also has a duty to evaluate rigorously his or her own learning needs and to implement an educational plan that meets these needs. To do anything less is a disservice to the profession and to the patients who entrust their lives to us.

Scott A. Schartel, D.O.,* **David G. Metro, M.D.†** *Department of Anesthesiology, Temple University, Philadelphia, Pennsylvania. schartel@temple.edu. †Department of Anesthesiology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.

Reference

1. Davis DA, Mazmanian PE, Fordis M, Harrison RV, Thorpe KE, Perrier L: Accuracy of physician self-assessment compared with observed measures of competence: A systematic review. *JAMA* 2006; 296:1094-102