

In the Eyes of the Beholder!

WHEN you attended fourth grade, you probably thought the teacher came to school each morning teaching whatever she/he thought might be a good lesson for that day. Now you know better! A great deal of thought and planning go into the development and implementation of the daily lesson plan; even more goes into a curriculum. The essential components of the educational planning process include (1) understanding the characteristics of the learning audience (needs assessment), (2) defining the educational objectives (the desired student outcomes), (3) accounting for the variety of teachers and their teaching styles, (4) implementing the instructional activities in the unique learning environment, and (5) evaluating the learning outcomes.¹ Analogous to medical practice, for example, when you treat hypertension and measure the medicated patient's treated blood pressure, you reaffirm or modify your prescription; with graduate medical education, when you teach

residents and measure their performance, you reaffirm or modify your educational plan. Performance evaluation is key to the success of the teaching-learning process, especially when educating anesthesiologists! Baker has focused on this in his study "Determining Resident Clinical Performance: Getting Beyond the Noise," an investigation not of "competence (what a physician [resident] can do)... [but rather of]...performance (what a physician [resident] actually does in everyday practice)."²

The problem with performance evaluation is that, like beauty, it is in the eyes of the beholder; the biases of the evaluator can't help but shade the assessment! The "yardstick" that one faculty member uses to "measure" an anesthesiology resident's performance is very different from that



“... Baker has provided anesthesiology faculty with an ethnographic tool that has practical utility, recognizing high-performing residents and low performers who will benefit from educational interventions.”

used by another faculty member evaluating the same resident. The yardstick that a faculty member uses to evaluate a CA-1 anesthesiology resident is very different from that used by the same faculty member evaluating that same resident who has graduated to a CA-2 or -3 status. At its core, performance evaluation depends on how “performance” is defined. The dictionary says that performance is “the execution or accomplishment of work, acts, feats...the manner in which or the efficiency with which something reacts or fulfills its intended purpose.”*

Our problem, as faculty evaluators, is knowing what “work, acts, and feats” best characterize resident or fellow clinical performance of the “intended purpose” (*i.e.*, the provision of safe, efficient, and effective anesthesia patient care) and being able to assess this without our individual biases skewing the process.

Baker uses several strategies to minimize evaluator bias, equalize the different yardsticks evaluators

use to assess performance, and validate a reliable evaluation process so we might know what work, acts, and feats best characterize performance of the intended purpose so we can educate residents ultimately to be able to provide safe, efficient, and effective anesthesia patient care. First, Baker recognizes that performance evaluation is ethnography, the study of human behavior in everyday contexts, rather than under experimental conditions. Ethnography employs an investigative strategy in which techniques such as up-close observation, interviews, and questionnaires are used to gather data that “paint portraits” and provide a narrative description of the studied population. “Ethnography enhances and widens top down views and enriches the inquiry process, taps both bottom-up insights and perspectives of powerful policy-makers ‘at the top,’ and generates new analytic insights by engaging in interactive, team exploration of often subtle arenas of human difference and similarity.”³

Second, as a part of the ethnographic study strategy, Baker cultivated faculty and resident support for and owner-

Photograph: J. P. Rathmell.

Accepted for publication June 10, 2011. The author is not supported by, nor maintains any financial interest in, any commercial activity that may be associated with the topic of this article.

In conjunction with the FAER 25th anniversary, articles and editorials in the ANESTHESIOLOGY October 2011 issue celebrate the accomplishments of FAER. For additional information visit www.FAER.org.

* Available at: <http://dictionary.reference.com/browse/performance>. Accessed March 30, 2011.

Copyright © 2011, the American Society of Anesthesiologists, Inc. Lippincott Williams & Wilkins. Anesthesiology 2011; 115:681-2

◆ This Editorial View accompanies the following article: Baker K: Determining resident clinical performance: Getting beyond the noise. ANESTHESIOLOGY 2011; 115:862-78.

ship of the performance evaluation system. Faculty and residents reviewed and commented upon an evaluation form drafted by the Education Committee, which was refined by the feedback process; consensus was developed around the performance evaluation instrument. This performance evaluation tool used the collective wisdom of the informed perspectives of the evaluators (faculty) and those evaluated (residents). The evaluation data collection process gathered ethnographic descriptors of resident performance from multiple data sources: absolute and relative-to-peers performance of the Accreditation Council for Graduate Medical Education core competencies, clinical competency committee questions, rater confidence in having a resident perform cases of increasing difficulty, and free text comment.² This variety in data collected viewed resident performance from different perspectives (how much help the resident needed relating to each competency, how the resident performed compared with other residents in the same training year, comments about specific strengths and specific areas for improvement, positive or negative affirmation of five statements relating to essential competency attributes, and an evaluator's willingness to let the resident provide independent and unsupervised care for each of eight cases of increasing difficulty), adding depth and breadth to the composite performance evaluation of each resident. The evaluation form developed a multidata source structure that, in essence, "trained" the faculty how to observe, categorize, and evaluate performance of residents.

Lastly, Baker employed a Z-score methodology to eliminate the unique attributes of individual faculty evaluators; Z-scores enable comparison of scores from different evaluators that are, therefore, from different normal distributions. Z-score, the statistical measure that quantifies the distance (standard deviation)

tions) a data point is from the mean of a data set, allowed Baker to "standardize" the resident evaluation scores of each faculty evaluator and scrub away their uniqueness; the Z-scores were a more true measure of a resident's performance and far less an indication of the evaluators' biases.

Baker makes a strong argument that the performance evaluation schema described in his study has validity with respect to anesthesia patient care by residents. His evidence is: Z-scores were (1) directly related to the Anesthesiology In-Training Examination assessment of knowledge, (2) predictors of "problem" residents identified as such and referred to the Clinical Competence Committee, (3) indicative of residents more rapidly gaining the faculty's clinical care confidence, and (4) improved when assessed after remedial intervention.² Despite the perceived difficulty with conducting valid and meaningful educational research, Baker has provided anesthesiology faculty with an ethnographic tool that has practical utility, recognizing high-performing residents and low performers who will benefit from educational interventions.

Alan Jay Schwartz, M.D., M.S.Ed., Department of Anesthesiology and Critical Care, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania. schwartz@chop.edu

References

1. Reznich CB: Designing a course, *An Introduction to Medical Teaching*. Edited by Jeffries WB, Huggett KN. New York, Springer, 2010, pp 123-41
2. Baker K: Determining resident clinical performance: Getting beyond the noise. *ANESTHESIOLOGY* 2011; 115:862-78
3. Genzuk M: A synthesis of ethnographic research, *Occasional Papers Series*. Edited by Center for Multilingual, Multicultural Research. Los Angeles, Rossier School of Education, University of Southern California, 2003