Random Thoughts . . .

HOW TO EVALUATE TEACHING

RICHARD M. FELDER, REBECCA BRENT
North Carolina State University • Raleigh, NC 27695

Student ratings of teaching get a bad rap in some academic circles. Faculty members are repeatedly and authoritatively assured that "They're just popularity contests," "High ratings go to the easy graders," and "If I get low ratings it's only because I set high standards and students don't like demanding teachers."

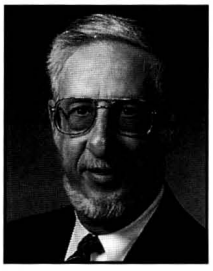
In fact, student ratings have been repeatedly shown to have a high level of validity, and those complaints about them have been debunked by research.[1–3] Students are in a better position than anyone else to judge certain aspects of teaching, such as how clear, interesting, respectful, and fair a course instructor is, and they're the only ones who can say how an instructor has influenced their attitude toward the course subject, their motivation to learn it, and their self-confidence. For these and other reasons, student ratings should be considered an essential component of faculty teaching performance evaluation.

But it makes little sense to use *only* student ratings. Few students are equipped to judge whether a course is accurate and up-to-date, the assignments and tests are appropriately challenging, and the content and learning objectives are consistent with the course's intended role in the department (for example, to serve as a prerequisite to other departmental courses or to address certain outcomes in the department's accreditation plan). Only faculty colleagues are in a position to make such judgments.

Moreover, classroom teaching may only be a small part of a faculty member's educational activities. He/she may also advise students, develop new courses and redesign old ones, adapt and develop courseware and innovative teaching strategies for use in both traditional classroom instruction and distance education, coordinate departmental preparation for accreditation, offer seminars, workshops, consulting, and mentoring to help faculty colleagues and/or graduate students improve their teaching skills, write textbooks, and conduct educational research. All of these activities can have a dramatic effect on a department's teaching quality, student retention, and chances of receiving full accreditation, but student ratings don't indicate whether and how well an instructor is doing them.

In short, a key to effective teaching evaluation is to collect data from multiple sources (*triangulation*), making sure that all education-related activities are rated by the people best qualified to rate them. Figure 1 presents a multiple-source evaluation model designed to work that way. The remainder

Richard M. Felder *is Hoechst Celanese Professor Emeritus of Chemical Engineering at North Carolina State University. He received his BChE from City College of CUNY and his PhD from Princeton. He is coauthor of the text* Elementary Principles of Chemical Processes *(Wiley, 2000) and codirector of the ASEE National Effective Teaching Institute*



Rebecca Brent *is an education consultant specializing in faculty development for effective university teaching, classroom and computer-based simulations in teacher education, and K-12 staff development in language arts and classroom management. She co-directs the ASEE National Effective Teaching Institute and has published articles on a variety of topics including writing in undergraduate courses, cooperative learning, public school reform, and effective university teaching.*

of this column briefly elaborates on the model components.

Peer Ratings

The usual form of peer evaluation, in which an observer visits a lecture and jots down whatever happens to catch his or her attention, has its own drawbacks. Most obviously, a single observed class may not be representative of someone's normal teaching. Even if it is, faculty members have widely disparate ideas of what constitutes good teaching, so that the same class could get an excellent rating from one observer and a poor rating from another. More importantly, a single class observation provides no assessment data at all on aspects of teaching performance other than lecturing.

A far more effective procedure is for two or more reviewers to use standardized checklists to rate instructional materials and at least two class observations independently and then to reconcile their ratings.[4] The checklists should consist of items taken from a list of attributes known to correlate with effective teaching,[5,6] and should be approved by the department faculty before they are used. This procedure has a high level of inter-rater reliability and includes measures to address commonly expressed concerns about peer review, including possible rater bias and excessive time demands imposed on reviewers.[4]

Student Ratings

Tested forms for student evaluation of teaching are given in a recent National Research Council publication,[7] and more information about how to make student evaluations effective is provided in that reference and by Felder.[8] Faculty performance evaluations should take into account student ratings collected over a period of several years, with relatively little

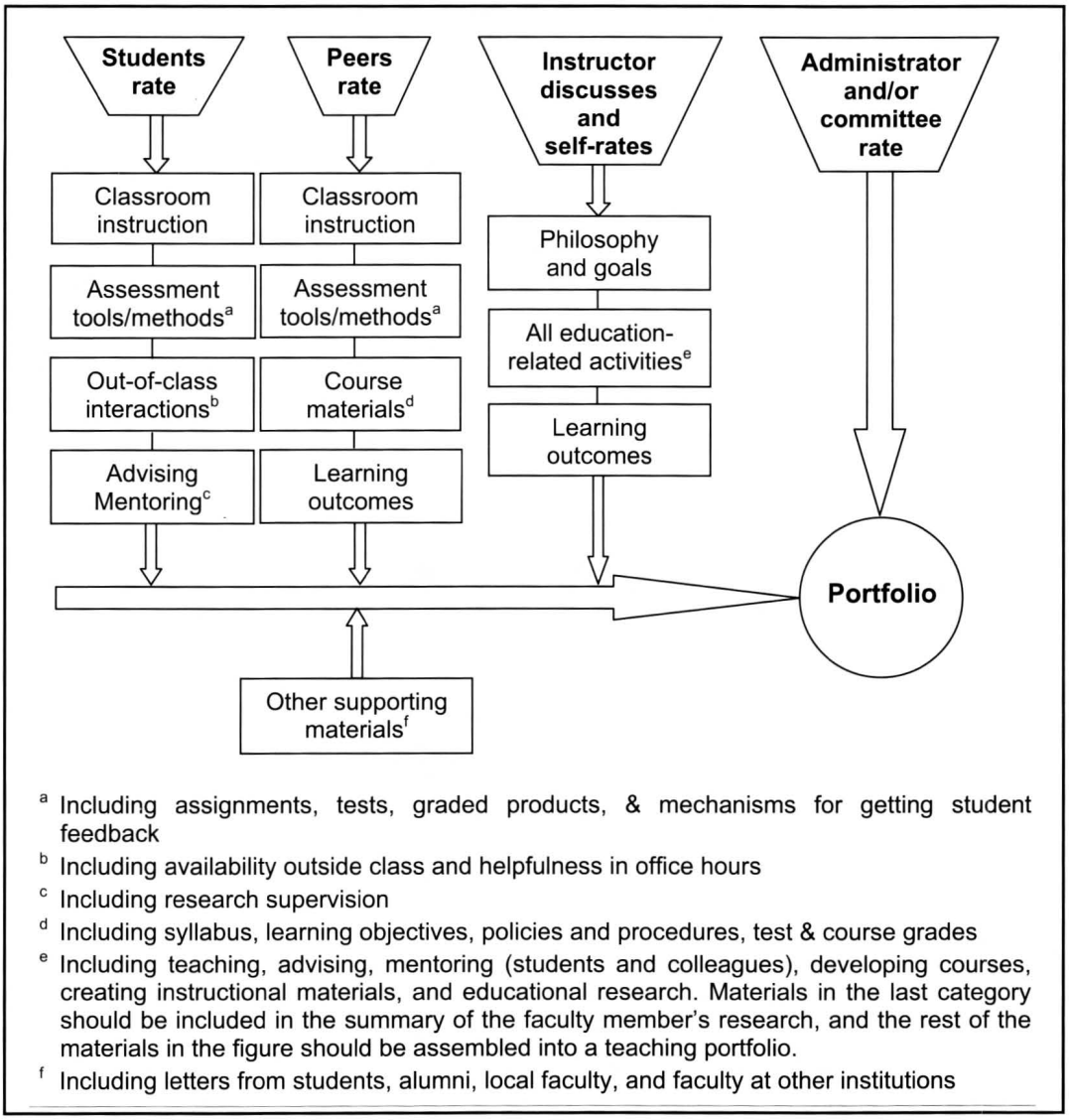


Figure 1. *Teaching performance evaluation model.*

weight being attached to ratings of someone's first semester of teaching.

The Teaching Portfolio

Just as some performance assessment data can best be provided by students and some by peers, certain important information can only be supplied by the faculty member being reviewed. Instructors should assemble materials summarizing all of their education-related activities, including developing new courses and redesigning old ones, developing and evaluating innovative instructional methods, advising and mentoring students, writing new texts and courseware, providing instructional development to faculty colleagues and graduate students, and carrying out educational research. All of these materials except those related to educational research (which we discuss in the next section) should be incorporated into a teaching portfolio, along with summaries of student ratings over the past two or three years, peer ratings, and reference letters from alumni and colleagues at other institutions who are familiar with the instructor's educational activities. The portfolio provides a solid basis for evaluating the faculty member's teaching performance and contributions to education.[9–11]

The Scholarship of Teaching and Learning

When done properly, educational research is every bit as demanding, rigorous, and important to the future of an academic discipline as traditional disciplinary research.[12] There is no legitimate reason to separate the two categories of research by making educational scholarship just another component of teaching performance, or worse, not to count it at all in faculty performance reviews. Any material related to educational research (including lists of grants, publications, presentations, and awards, along with supporting letters) should be combined with documentation of disciplinary research in faculty activity reports and in tenure and promotion dossiers, and the same high standards should be applied to the evaluation of performance in both research categories.

Consistency of Multiple-Source Ratings

For triangulation to be most effective, data from different sources should overlap to the greatest extent possible. For example, items on student rating forms related to aspects of teaching that both students and peers are equipped to evaluate (*e.g.,* the instructor's preparedness, clarity, responsiveness to questions, and respect for students) should parallel items in peer review checklists. If the two sets of ratings lead to the same conclusions, it affirms the validity of both, while if they disagree substantially it suggests that at least one of the sets is suspect and further investigation should be undertaken. For example, the department head might bring in someone from outside the department (such as a consultant from the campus center for teaching and learning) to conduct focus group interviews with students related to the issues in question.

Summative and Formative Evaluation

Evaluation of teaching may be *summative* (to provide data for use in making decisions regarding reappointment, tenure, promotion, and merit raises, and for selection of award recipients) or *formative* (to improve the teaching of the instructor being evaluated). The full procedure depicted in Figure 1 and described above should be implemented for summative evaluation. Once the portfolio is assembled, only minor effort should be required to update it in successive years. For formative evaluation, a subset of the procedure should be carried out (for example, only one peer rater may be used), and the results should be shared only with the instructor rather than being passed on to the department head or a performance review committee. Carrying out formative reviews in the first few years of a faculty member's career should substantially increase the chances that a subsequent summative review will be favorable.

REFERENCES

1. Cashin, W.E., "Student Ratings of Teaching: The Research Revisited," IDEA Paper No. 32, Kansas State University Center for Faculty Evaluation and Development, *<www.idea.ksu.edu>*, September (1995)
2. Felder, R.M., "What Do They Know, Anyway?" *Chem. Eng. Ed.,* **26**(3), 134 (1992), *<http://www.ncsu.edu/felder-public/Columns/Eval.html>*
3. McKeachie, W.J., "Student Ratings: The Validity of Use," *Amer. Psychologist,* **52**(11), 1218 (1997)
4. Brent, R., and R.M. Felder, "A Protocol for Peer Review of Teaching," *Proc. 2004 An. ASEE Meet.,* ASEE, June (2004), *<http://www.ncsu.edu/felder-public/Papers/ASEE04(Peer-Review).pdf>*
5. Chism, N. Van Note, *Peer Review of Teaching*, Anker Publishing, Bolton, MA (1999)
6. Weimer, M., J.L. Parrett, and M. Kerns, *How am I Teaching?* Magna Publications, Madison, WI (1988)
7. National Research Council, *Evaluating and Improving Undergraduate Teaching in Science, Technology, Engineering, and Mathematics*, Washington, DC, National Academies Press (2003)
8. Felder, R.M., "What Do They Know, Anyway? 2. Making Evaluations Effective," *Chem. Eng. Ed.,* **27**(1), 28 (1993), *<http://www.ncsu.edu/felder-public/Columns/Eval2.html>*
9. Seldin, P., *The Teaching Portfolio: A Practical Guide to Improved Performance and Promotion/Tenure Decisions,* 2nd ed., Anker Publishing Co., Bolton, MA (1997)
10. Edgerton, R., P. Hutchings, and K. Quinlan, *The Teaching Portfolio: Capturing the Scholarship in Teaching,* American Association for Higher Education, Washington, DC (1991)
11. Felder, R.M., "If You've Got It, Flaunt It: Uses and Abuses of Teaching Portfolios," *Chem. Eng. Ed.,* **30**(3), 188 (1996) *<http://www.ncsu.edu/felder-public/Columns/Portfolios.html>*
12. Huber, M.T., and S. Morreale, eds., *Disciplinary Styles in the Scholarship of Teaching and Learning: Exploring Common Ground.* AAHE/Carnegie Foundation for the Advancement of Teaching, Washington, (2002) ❐