



Teaching Concerns

Newsletter of the Teaching Resource Center for Faculty and Teaching Assistants

Hotel D, 24 East Range (804) 982-2815 trc-uva@virginia.edu <http://www.virginia.edu/~trc/>

Fall 1999

Evaluations Can Fit Into A Holistic Teaching Assessment Program

by Bill McAllister, TRC (History)

One of the surest ways to increase the blood pressure of faculty is to raise the issue of end-of-semester student evaluations. Introduced a generation ago in an attempt to improve teaching, student evals routinely receive criticism from the very instructors they are supposed to benefit. This essay does not argue for or against student ratings, because they have become institutionalized to the point where they are not likely to disappear. What I will attempt is to provide a context within which to view semester-end evaluations and offer some ideas about how best to make use of them. Put simply, student ratings can most profitably serve as an evaluation tool when used as part of a more comprehensive program to assess teaching proficiency.

My remarks are based on a careful reading of several key books and articles that deal with teaching evaluation issues, on discussions with faculty, and on my own observations. Readers should note that the issue of teaching evaluations in general, and end-of-semester evaluations in particular, is the most examined topic in higher education research over the last 70 years; a rich vein of information supports the works I consulted. (All titles cited below are available at the TRC library.)

Some evaluation forms are better than others. The most important distinction revolves around the differences between properly constructed student evaluations and “home-made” ratings. Carefully designed evaluation tools that incorporate psychometric principles and rigorous statistical procedures can serve as reliable and valid indicators of teaching effectiveness. Such professionally-developed forms are available; Arreola’s *Developing a Comprehensive Faculty Evaluation System*, for example, discusses the merits of ten instruments currently in use and offers advice on how to go about selecting an appropriate evaluation tool. Those interested in finding out more about such forms can contact the TRC as well. Unfortunately, “home grown” evaluation instruments often do not make use of such expertise. Most departmental committees, undergraduate groups, or administrators who develop student ratings forms do not consult with measurement specialists. Consequently, questions about reliability and validity render student rating

instruments less than fully effective. Those designing or redesigning student evaluations can secure professional advice or consult the literature themselves, taking care to make faculty aware of the extensive literature that supports the use of properly-constructed forms.

A proper understanding of the strength of student evaluations can dispel faculty reservations about ratings forms. The research literature demonstrates that well designed student evaluations belie most of the negative beliefs commonly held by faculty. One very common impression is that student ratings of instructors correspond strongly with the actual or expected grade students receive in that class. Over 400 studies of this issue indicate that no significant correlation exists between grades and properly constructed ratings. Two other popular beliefs are (1) students cannot make consistent judgments about instructor quality (they can), and (2) student ratings are simply popularity contests (students can certainly praise professors who display a pleasing classroom affect *and* criticize them for poor course design, lack of organization, inadequate knowledge, or other instructor-related hindrances to learning). Factors such as the gender of the instructor or the rater, the time of day when class is held, a student’s major, and an instructor’s rank have no impact on properly designed student ratings. Moreover, no consistent relationship exists between class size and evaluation—students do not consistently rate small classes higher than large classes simply because of the lower teacher-to-student ratio.

The best way to make sure end-of-semester forms provide beneficial information is to incorporate them as one of several gauges of teaching effectiveness. I find it useful to think of feedback in terms of a multi-legged chair or table. Self-evaluation provides an important “leg” when assessing one’s teaching. Observations from peers can provide another valuable source of information. Student evaluations can be conducted in many ways, and can provide additional “legs” to support teaching.

For example, it is possible to utilize student input to find out “how it’s going” at any time in the semester—there is no need to wait until the last days of class. Previous editions of *Teaching Concerns* exhibit the array of methods currently used by U.Va. instructors. In *Classroom Assessment Techniques*, Angelo and Cross outline many more possibilities. Many of these techniques require very little class time, and some involve constructing assignments that

do “double duty” by assessing instructor teaching as well as student learning. The TRC also offers several types of mid-semester assessment services, including Teaching Analysis Polls, videotaping, in-class observations, and assistance with interpreting one’s own evaluation tools. Some departments and schools offer such opportunities as well. The Law School encourages faculty to join in Teaching Partnerships, in which colleagues pair for a year to observe each other’s classes and talk about pedagogical issues. McIntire School of Commerce faculty routinely open their classes to visits from colleagues, who then normally offer feedback.

The research literature, including a recent study completed here at U.Va., demonstrates that combining mid-semester evaluations with consultation of the type offered by the TRC does improve teaching effectiveness. Getting multiple “looks” at one’s teaching throughout the semester eliminates the “home run or strike out” nature of semester-end evaluations; they become one indicator among many. Additional mid-semester gauges paint a fuller picture, enabling instructors to confirm what their ratings indicate or to provide alternative evidence if they believe end-of-semester assessments do not accurately portray their classroom effectiveness.

Many departments and schools are currently revising their student evaluation forms or are considering doing so. That process can provide faculty an excellent opportunity to grapple with key pedagogical questions and to discuss their teaching priorities and assumptions. What *is* “good teaching” within a particular disciplinary context? If many different types of teaching can be considered “good,” should we value certain approaches or outcomes over others? What aspects of teaching are students best able to judge? What aspects are peers or supervisors most qualified to comment upon? Some institutions have used such consensual conversations to develop clear, concise departmental teaching statements that outline what departments intend to evaluate and the manner in which that task will be accomplished.

Once faculty have established a clear idea of how they want to use the semester-end ratings, preferably as one portion of a multifaceted evaluation scheme, the process of selecting items can become less contentious. Whether a department or school chooses to use an off-the-shelf evaluation instrument, adapt questionnaires from available websites, or develop an in-house version, incorporating flexibility is a good idea. Most professionally-developed forms provide mix-and-match options, and Arreola presents a bank of 504 sample questions arranged in 24 categories. Schools or departments could agree on certain items to be included in all questionnaires, and tailor another section according to type or size of course or other categories such as type of assignments. Instructors could select additional questions that would reflect their particular teaching style and provide information about issues of individual interest.

When viewed within this holistic perspective, end-of-semester evaluations can provide valuable input. Through thoughtful consideration about the issues raised by student ratings, departments and individual instructors can discern their teaching values, discuss priorities, and set verifiable,

achievable goals. It behooves us to make use of all possible avenues to pursue the vital interest of both faculty and students—passing on knowledge as effectively as possible.

Works Cited and Consulted

- Angelo, Thomas A. and Patricia K. Cross. *Classroom Assessment Techniques: A Handbook for College Teachers* (2nd ed.). San Francisco: Jossey-Bass, 1993.
- Arreola, Raoul. *Developing a Comprehensive Faculty Evaluation System*. Bolton, MA: Anker, 1995.
- Chism, Nancy. *Peer Review of Teaching: A Sourcebook*. Bolton, MA: Anker, 1999.
- Felder, Richard. “What Do They Know Anyway?” *Chemical Engineering Education*, 26 (3), 124-135 (Summer 1992) and 27 (1), 28-29 (Winter 1993).
- Marsh, Herbert and Lawrence Roche. “Making Students’ Evaluations of Teaching Effectiveness Effective: The Critical Issues of Validity, Bias, and Utility,” *American Psychologist*, 52 (11), 1187-1197 (November 1997).
- Murray, Harry G. “Does Evaluation of Teaching Lead to Improvement of Teaching?” *International Journal for Academic Development*, 2 (1), 8-23 (May 1997).
- University of Virginia/Teaching Resource Center, Impact Self- Study, 1999. Available at the TRC.
- University of Wisconsin-Madison, Peer Review of Teaching website at <http://www.wisc.edu/MOO>.
- To view previous editions of *Teaching Concerns*, go to the TRC website at <http://www.virginia.edu/~trc/>, click on “Publications” and then on “Teaching Concerns;” or you can visit the TRC, or request copies via messenger mail.