Summative Evaluation and Formative Feedback

Introduction

As renewed interest in faculty evaluation grows due to university and state mandates for accountability, administrators and faculty increasingly find themselves in the business of creating new evaluation guidelines on which faculty will be judged. Particularly difficult is the task of devising a new system to evaluate teaching. Whereas old evaluation systems are often punitive—that is, teaching only receives attention if it is judged as poor—new systems devise fair estimates of college teaching, and, ultimately, reward good teaching. As administrators and faculty strive to create such a fair *summative* evaluation system, faculty developers labor to develop and improve teaching through *formative* feedback. Many argue that a fair summative evaluation system must build in time and support for formative feedback, so that faculty members may develop and grow as teachers before being evaluated summatively on their pedagogy. This report outlines the distinction between summative evaluation and formative feedback. Two cases illustrate the tensions and interconnectedness of these two concepts, as well as offer suggestions for both developing and evaluating teaching.

Summative evaluation and formative feedback

In a still relevant *AAHE Bulletin* article on evaluation and teaching improvement, Weimer (1987) describes summative evaluation of teaching. She writes that summative evaluation is a judgment about teaching that is used to make a decision—a decision about promotion, tenure, or, even, teaching awards. A summative evaluation of teaching attempts to *summarize* the complex phenomena of teaching. It typically consists of global ratings of teaching, thus, reducing teaching to generalities.

An example of summative evaluation is end-of-course student ratings. Student ratings instruments prompt students to rate six to ten aspects of teaching (e.g., organization, clarity, enthusiasm, etc.), and they ask students to make an overall, summarizing judgment of the instruction. Students' ratings are then averaged and given to faculty as summaries of their teaching. The overall summary rating of instruction may even be shared with students and other faculty members. Such ratings are easily compared and then used for making decisions and answering questions of accountability.

Individual faculty members have little control in the traditional summative evaluation process. They usually do not choose the instrument by which they are evaluated; they do not choose the time of the semester at which they are evaluated; nor do they choose the groups of people (e.g., students, department chairs, administrators, or fellow faculty members) to whom the results are distributed.

Formative feedback differs from summative evaluation in several ways. In their sourcebook on formative assessment of teaching, Rando and Lenze (1994) describe formative feedback as information gathered for the purpose of improving and developing teaching. This information is meant to *inform* change. The feedback itself must be specific and concrete enough to suggest actions for improvement. It is typically requested well before the end of a semester of teaching, so that the faculty member may use the information to improve teaching in the current course.

For example, consider the professor who asks her students, halfway through the semester, to write a few sentences about how the current instruction has promoted or hindered learning. This faculty member has gathered information

that she may now use to make teaching and learning more effective.

The formative feedback process entails requesting information, making sense of the information, and implementing changes based on the information collected. This process belongs to faculty members. They are responsible for making all decisions along the way. They may gather, analyze, and react to the information on their own; or they may decide to invite an instructional consultant to assist in this process. In the case of the professor whose students responded in writing to the current instruction, the faculty member chose the method of collecting data and, after reading students' responses, determined which of the issues would be addressed.

Several other examples of formative feedback exist. Faculty may opt to conduct classroom assessment for the purpose of improving teaching. They may ask an instructional consultant to videotape their teaching in order to reflect on teaching development. They may even invite trained facilitators to conduct small group "quality control" interviews with students. In all of these examples, only the faculty member who has initiated the formative feedback process need see the results.

Issues involved in their co-existence

The distinctions just drawn between these two concepts exist at the definitional level; but in practice, the two are intertwined. For example, how can faculty's teaching be judged fairly unless they are provided opportunities to improve and develop their teaching? A system that *judges* teaching without any effort to support the *development* of good teaching ultimately fails. Before successfully negotiating the co-existence of summative evaluation and formative feedback, at least two issues must be considered.

First, the two must reside in separate offices. If faculty are to trust formative feedback, formative feedback activities must be separated from summative evaluation activities. Consider the junior faculty member who is contemplating a mid-semester review of his teaching. Not sure of the value he places on students' feedback, he wants to request that a peer visit his classroom. In order to be trusted, formative feedback must be provided by someone other than his department chair or faculty members who are responsible for his tenure review. Feedback provided by these folks, no matter how helpful and well-intended at the time, could be subconsciously held against the faculty member as examples of "inadequacies" when making summative decisions. Thus, administrators must see to it that the formative activities are in some way shielded from summative evaluation.

Second, summative evaluation must evolve. If a college truly advocates formative feedback, its summative evaluation procedures must not hamper formative efforts. That is, universities cannot uphold two climates at the same time: the punitive climate often associated with summative evaluation precludes the supportive climate necessary for formative feedback. Thus, summative evaluation systems must become more humane.

Two universities have navigated the waters between formative feedback and summative evaluation. Duquesne University added a formative feedback program to an already existing summative evaluation process. The University of Massachusetts added a new summative evaluation system to an already existing formative feedback program. These two cases offer powerful insights about the co-existence of formative feedback and summative evaluation.

The Case of Duquesne University

Since 1986, Duquesne University has used end-of-semester student rating instruments for its summative evaluation of teaching. In 1989, faculty asked why there was no formal support system to assist faculty in their *development* as teachers. Furthermore, faculty wanted to know about alternate methods for reviewing teaching. Subsequently, the University Teaching Committee was formed to 1) create a faculty development center, 2) develop alternate means for reviewing teaching, and 3) revise old student evaluation instruments.

Committee members immediately began working on plans for the "Center for Teaching Excellence." They decided its mission would be to develop and improve faculty members' teaching. The Center would offer a variety of services to faculty members, including individual consultations, classroom visitations, instructional videotaping, or mid-semester formative assessment. The director of the Center was to report directly to the provost, thereby increasing the importance and visibility of the Center.

While the University Teaching Committee made plans for the Center for Teaching Excellence, committee members also created new student evaluation forms and developed a Peer Review program to supplement its traditional summative evaluation procedures. The

addition of a peer review program lessened the defensiveness that had previously surrounded summative review.

As the committee designed it, the peer review system consisted of three reviews. The first occurs in the first year of employment in order to acquaint faculty with the process and identify for faculty strong and weak aspects of their teaching. The next review takes place in the third year in time for third year contract renewals. The last review occurs in conjunction with the tenure decision.

In each review year, a team of two peers visits the assigned faculty member's classroom at one scheduled time and at one unannounced time. The committee developed their own observation instruments and their own evaluation process, beginning with an initial meeting look at teaching materials and objectives and ending with a feedback report to the faculty member.

The summative evaluation process became more fair with the addition of the Peer Review program. Faculty realized that there were many differences between teaching in traditional classrooms and teaching in the clinical setting. Thus, the Committee developed two sets of guidelines for peer review, one for classroom settings and the other set of guidelines to accommodate clinical teaching.

Even though the Center for Teaching Excellence and the Peer Review program were connected in University legislation, the Center does not have a role in the peer review process. Faculty perceive the Center as sufficiently separate from Duquesne's summative evaluation process; thus, they trust its staff to assist them in improving teaching. Of the faculty who re-

sponded to the yearly evaluation of the Center, just under 100 percent said, "Yes, I have benefited from the existence and programming of the Center for Teaching Excellence."

Duquesne stands as a good example of how administrators and faculty can work together to create a formative feedback program for a school that once had only summative evaluation. Duquesne's experience also suggests that with sincere efforts to improve and develop teaching, existing summative evaluation systems will ultimately change to be more fair and, consequently, better accepted.

The Case of The University of Massachusetts at Amherst

At the University of Massachusetts at Amherst, a thriving Center for Teaching has for seven years offered instructional consultations, videotaping services, and other resources to develop faculty as teachers. The teaching culture at this university is notably supportive. With the recent addition of several new programs for newly hired faculty members, new faculty now count on the Center's voluntary, mid-semester, formative evaluation service. (Last year, 75 percent of new faculty participated in this program.)

At the same time, there is increasing pressure from the legislature and board of directors to assess student learning and quality of teaching. One way the University has proposed to address these concerns is through end-of-course, summative evaluations. The University's recently hired Associate Director for Assessment was asked to create a new student rating instrument to assess both student learning and quality of teaching.

Staff in both the Center for Teaching and the Office of Planning and Assessment were concerned about how a new summative evaluation process might affect the existing supportive climate for teaching. Although the mission of the Center for Teaching concerns only formative assessment, Center staff acknowledged that formative assessment ultimately prepares faculty for summative evaluation, and that an unfair summative evaluation system would undermine the formative feedback process. Thus, the two offices crossed the formative-summative line in order to consult in the creation of a new, reliable and valid student rating form and a humane evaluation process. The Office of Planning and Assessment was ultimately responsible for the creation and implementation of its new rating instrument, but the director of the Center for Teaching acted as a consultant to the assessment office. Staff in both offices agreed that careful planning and development of the summative evaluation system was crucial to ensuring continued emphasis on development and support for teaching.

Currently, the new evaluation system has been piloted and well-received in several departments. Although the rating instrument has not yet been finalized, word about its clarity and its facility has spread on campus. Other department continue to ask to be on the pilot list. It seems that this University is carefully and successfully incorporating a new summative evaluation system while preserving the existing formative feedback process.

Action items

At most universities, committees, rules, standards, and guidelines already exist to guide the summative evaluation process. For exam-

ple, tenure reviews follow pre-set procedures that ultimately assist department chairs and senior faculty in making decisions about junior faculty. If summative evaluation and formative feedback are to co-exist, universities must provide an infrastructure for formative feedback as well. This infrastructure includes a commitment to developing faculty as teachers, rewards for good teaching, a process by which all faculty are given the opportunity to develop as teachers, and a support system for developing teachers. Without the infrastructure to cultivate formative feedback, efforts to foster both will fail.

Hand in hand with this first recommendation, faculty and administrators must commit to making summative evaluation systems more humane. This means standards must be outlined clearly, faculty must receive feedback along the way to the tenure decision, faculty must receive some sort of formal mentoring, and tenure decisions must not come as a surprise.

The following list suggests specific items for action.

■ Employ faculty development offices to develop and improve teaching.
⎮ Offer services to faculty on a voluntary basis.
⎮ Emphasize the development of teaching, not the repair of teachers.
⎮ Keep faculty development activities separate from summative evaluation.

■ Place value on faculty members' efforts to conduct formative assessment.
⎮ Reward faculty for commit-ment to teaching improve-ment.
⎮ Encourage teaching portfolios to showcase development as teachers.
⎮ Publish (in college newspaper) innovative efforts to improve teaching.

■ Create fair summative review procedures.
⎮ Outline the steps taken to get tenure.

⎮ List clearly the criteria for gaining tenure.
⎮ Provide checkpoints along the way to tenure.

■ Create fair summative review instruments.
⎮ Review the literature on student ratings.
⎮ Consider alternate means for evaluating teaching summatively.
⎮ Consult with those responsible for providing formative feedback.

References

Rando, William C., & Lenze, Lisa Firing. (1994). *Learning from Students: Early Term Student Feedback in Higher Education.* University Park, PA: National Center on Postsecondary Teaching, Learning, & Assessment.

Weimer, Maryellen Gleason. (1987, March/April). "Translating evaluation results into teaching improvements." *AAHE Bulletin,* pp. 8-11.

nea
NATIONAL
EDUCATION
ASSOCIATION