DOCUMENT RESUME

ED 283 444 HE 020 418

AUTHOR Weimer, Maryellen Gleason
TITLE Translating Evaluation Results into Teaching Improvements.
PUB DATE 87
NOTE 5p.
PUB TYPE Viewpoints (120) -- Journal Articles (080)
JOURNAL CIT AAHE Bulletin; v39 n8 Apr 1987

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Change Strategies; Evaluation Methods; *Evaluation Utilization; *Faculty Evaluation; *Formative Evaluation; Higher Education; *Instructional Improvement; Summative Evaluation; *Teacher Attitudes

ABSTRACT

Five problems concerning faculty evaluation are identified, along with sever proposed solutions. The problems are as follows: (1) summative and formative evaluation objectives are mixed; (2) most instructional evaluation is designed with institutional convenience in mind; (3) evaluation results are given impersonally; (4) most evaluation instruments used are homemade products that have not been empirically evaluated; and (5) evaluation results look precise, objective, and meaningful. These factors contribute negatively to faculty attitudes about evaluation and explain why evaluation results tend not to be translated into teaching improvements. Solutions may include: separating formative and summative evaluation but linking the results; allowing the individual faculty member to be in control of formative activities; de-emphasizing the evaluative and judgmental aspects of formative activities; promoting the idea that acquiring data on teaching behaviors and practices is needed to assure the integrity of the classroom; seeing formative evaluation in ways other than standardized machine-scorable forms; encouraging dialogue about formative evaluation; and rewarding faculty for instructional excellence. (SW)

TRANSLATING EVALUATION RESULTS INTO TEACHING IMPROVEMENTS

by Maryellen Gleason Weimer

My position as head of the Instructional Development Program at Penn State provides plenty of opportunities to interact with faculty about evaluations of their teaching effectiveness. These exchanges lend support to two conclusions. First, faculty have bad attitudes about instructional evaluation in general and student evaluation in particular. Quite bluntly, they do not like the activity and would not solicit the evaluations if they were not required by the institution. The data make them defensive, even in cases where the results are not all that bad.

Second, faculty have trouble translating evaluation results into action. The results give them no indication of what instructional alterations should be implemented. If they cannot answer the question "what needs improvement," they change the question and pose a much less healthy query: "How do I get these ratings up?" Now the move to change instructional practices is to benefit subsequent evaluations, not students. Maybe the change is small, but the point is important. The motivation to teach effectively should not derive principally from a desire to win the ratings game.

Where do these counterproductive responses come from? Is instruction evaluated at Penn State by some perverse system? No, Penn State evaluates instruction much like many other institutions, maybe a bit more systematically than some. Faculty members at various junctures in their academic careers are expected to provide evidence of their teaching effectiveness. The system is flexible and allows departments and colleges to make and use their own instruments. All in all, it's pretty typical.

Maryellen Gleason Weimer is head of the Instructional Development Program at The Pennsylvania State University. She edits a monthly newsletter on teaching for faculty, The Teaching Professor.

This article is based on the author's remarks at the 1987 National Conference.

From my observations, these responses tend to be quite common among faculty at all types of institutions and there seem to be five factors that cause them. In addition, I believe there are seven solutions which, if implemented, would change faculty attitudes about participation in evaluation activities and would increase the likelihood that evaluation results would be translated into teaching improvements.

PROBLEM 1: Summative and formative evaluation objectives are mixed. Summative evaluation defines teaching effectiveness with global, judgmental items and compares the instructional competence of one faculty member with that of others. These are the data institutions need to make personnel decisions. Formative evaluations describe teaching in behavioral terms and provide diagnostic input into the effectiveness and impact of a set of instructional practices. These are the data instructors need to be able to improve. Both evaluation types are legitimate and necessary.

The problem is that you cannot get both birds with the same stone. You cannot evaluate instruction summatively, give faculty members that data, and expect them to improve. Research has documented that this does not work[1] and logic tells us why. For example, suppose an item evaluates the instructor's attitude towards teaching the course and the mean is low. Those data tell the instructor there is a problem. They stop far short of identifying offensive attitudes

Faculty have bad attitudes about instructional evaluation . . . the data make them defensive, even if the results are not all bad.

If comparison data do not help faculty identify areas in need of alteration, and if no opportunities to discuss the results are provided, faculty may be motivated to become defensive, not better teachers.

or making the instructor aware of behaviors and practices that convey those attitudes. It is not surprising that faculty, given data like these, are in a quandary as to what they should do. The translating tasks confronting them are legitimately challenging.

PROBLEM 2: Most instructional evaluation is designed with institutional convenience in mind. I'm not arguing against administrative efficiency. However, this factor does contribute to negative faculty attitudes if it is the only way evaluation information is acquired. To be sure, institutions should collect data for use in personnel decisions via standard procedures. But administrative efficiency tends to be cold and impersonal. The net result is a faculty member who feels that evaluation is something done *to* him/her.

PROBLEM 3. Evaluation results (if faculty get them back) are returned via some equally impersonal, albeit efficient method. Generally, results come back to faculty via the mail—"in a plain, brown wrapper marked confidential," one told us with only a small twinkle in his eye. They came with varying amounts of statistical cybernetics to decipher and varying degrees of helpful instructions. Sometimes well meaning department heads try to add impact to the results. One we know lists all sixty faculty members by the last four digits of their social security numbers and then rank orders them from top to bottom by their overall rating of effectiveness. To be last on such a list is devastating. Being tenth from the bottom is hardly encouraging. And to what end? The comparison may indeed motivate faculty, but if the data do not help them identify specific areas in need of alteration, and if no opportunities to discuss the results are provided, faculty may be motivated to become defensive, not better teachers.

PROBLEM 4: Most evaluation instruments used are homemade products which have not been subjected to empirical evaluation. Seldin asked 616 institutions, "Has your institution conducted research on the validity of these forms?" 11.2 percent responded yes.[2] Evaluation instruments do not automatically produce good data. Sometimes the results can be misleading and unfair to the faculty member involved. Obviously, personnel decisions based on such data are implicated. Of equal concern are the impacts on the faculty member who may have felt (rightly or wrongly) that all was going well in the class. These data are a blow from behind. The faculty member's confidence is shaken. And if the data are summative then, by their very nature, they provide little in the way of clues to the problems or solutions. The point: if an institution is going to evaluate faculty, the data must be valid and reliable.

PROBLEM 5: Evaluation results look so precise, so objective, so meaningful. The calculation of means and standard deviations, placement in percentile groups, and construction of fever charts can occur quite independently of the empirical health of the instrument that has been used. Moreover, numeric representations mask the dynamic and variable milieu that is the college classroom. To be sure, the use of instructional evaluation has been and continues to be researched and some of the findings are definitive and can inform the practices. But the science is imprecise, our instruments crude. Unfortunately, the computer printout in the faculty mailbox doesn't say that. Quite the contrary, it tells a faculty member his/her overall effectiveness on a 7 point scale is 4.13 which obviously

make that faculty member better than his/her colleague with a 4.10 rating. That is what the faculty member comes to believe and that explains the need to argue so bitterly about the results.

These factors contribute negatively to faculty attitudes about the evaluation enterprise and explain why evaluation results tend not to be translated into teaching improvements. But there is good news—in the form of these seven, practical solutions.

SOLUTION 1: Separate formative and summative evaluation activities, but link the results. The two activities should run on separate tracks with points of convergence at the beginning and end. Summative evaluation constructs the comprehensive picture of instructional competency. Formative evaluation closes in on the picture, dissects the component parts, analyzes their relationship to one another, identifies what parts should be changed, and provides initial feedback on the success of those changes. Summative evaluation occurs again to create another composite picture, this time to show the differences. The connection between the two cannot be overemphasized. Formative evaluations must target appropriate areas of change. Summative assessment must reflect the impact of those changes.

SOLUTION 2: Put all formative activities under the aegis and control of the individual faculty member. The rationale is really quite simple. Faculty have what might be called ultimate instructional prerogative anyway. On any given day in class, in the case of any given alteration, they alone decide whether or not to make the change. This is no more than simple recognition of the fact that institutions and individuals cannot force a faculty member to improve. Granted, if the summative evaluation indicates the need for better teaching and the faculty member chooses not to improve, there should be consequences resulting from that decision.

> Putting faculty in charge of their improvement efforts does not imply that they should set out on the quest for good teaching unaided . . . resources and services must be available.

The faculty control advocated here extends beyond this level of responsibility to more concrete arenas as well. Putting formative activities under the aegis and control of the faculty means they get to choose the method of evaluation: open-ended questionnaires, informal interviews with students, submission of a videotaped teaching sample for analysis, classroom observation by a colleague, a program of reading, or whatever. It also means that if several areas are targeted for improvement, the faculty member selects where to work first.

All this is predicated upon the assumption that the faculty member will make good choices. In some cases that can be a tenuous proposition. But the advantages outweigh the risks. Faculty are in control. That decreases defensiveness. More importantly, it increases motivation. The faculty member acquires information about teaching via credible methods. That increases the credibility of the results, which in turn increases the chance that they will effect instructional practices.

One final note: putting faculty in charge of their improvement efforts does not imply that they should set out on the quest for good teaching unaided. If an institution is serious about its commitment to instructional excellence, resources and services must be available to support a faculty member's efforts to attain that goal.

SOLUTION 3: De-emphasize the evaluative and judgmental aspects of formative activities. Summative assessments establish an instructor's overall teaching effectiveness. This is a threatening proposition. Faculty egos being what they are, admissions of fear are unlikely, but reluctance to participate may prove the point. The evaluative aspects of summative assessment are not easily avoided, but why repeat the emphasis in the formative realm? An instructor does not need to know how he/she compares with anyone else in order to improve. Formative activities can justifiably be presented as ways of getting feedback and acquiring input. Formative evaluation lets faculty members find out what they'd like to know about their teaching. If that's the agenda and the activities are under the control of the faculty member, participation is a much more appealing proposition.

SOLUTION 4: Cultivate the notion that acquiring these kind of data are no big deal. This is not to say that the activity is unimportant; rather it is to imply that acquiring this sort of descriptive, diagnostic input is what any faculty member worth his/her salt does on a routine basis. Keeping close tabs on teaching behaviors and practices is essential if one hopes to assure the integrity of what transpires in the classroom.

SOLUTION 5: Think about formative evaluation in other ways than standardized machine-scorable forms. When the purpose of evaluation is to get a handle on what is happening in an individual class, the need for empirically rigorous instruments is much less compelling. Braskamp, Brandenburg, and Ory point out that norms are "considerably less important" when improvement is the agenda.[5]

Faculty can be encouraged to consider other kinds of forms: course material reviews, made-to-order forms where faculty (or students, provided a faculty member is basically competent and confident) construct an instrument from a collection of items, or forms that provide for reciprocal feedback where the students tell a faculty member specifically what needs improvement and the instructor returns the feedback by providing the class with the same sort of input.[6]

In the formative arena, faculty can be encouraged to think of activities that do not involve instruments. For example, faculty can review videotaped teaching samples, either their own or those belonging to others. They can visit each other's classes—not with an evaluation agenda at all, but simply to see how someone else integrates computer activities, or skeletal notes, or mini-discussions, or whatever. Students in class can be interviewed by a colleague, instructional development type, or trained student, and that information is passed on to the faculty member[5]. One iteration of this approach uses the Japanese management idea of quality circles to provide regular interaction between the faculty member and students regarding class matters of mutual interest[6]. The instructor can ask directly for students responses: "I'm going to review for the exam next Thursday. Write me a note and let me know what topics you'd like to have covered." Feedback like this gives an instructor some sense as to the areas of content about which students may be unclear or feel merit further discussion. That's valuable information about one's teaching. There are many approaches, so most faculty members will be able to find some approach of interest. Thus, participation in formative activities occurs and the faculty member receives input that constructively contributes to improvement efforts.

SOLUTION 6: Encourage dialogue about formative evaluation activities. Discussing evaluative data does make a difference in terms of subsequent evaluations." That's a dollars and cents reason to do it, but there are larger issues. Classrooms are not castles where one reigns behind the closed door in privacy. Academic freedom is not the issue. We are not talking about the *right* of the professor to make decisions about content or method. However, dialogue about teaching jeopardizes neither. Indeed, constructive analysis of challenging instructional perplexities (like the possibility of personalized teaching in large classes) and sharing ideas and answers infuses teaching with a steady source of energy that keeps its practice fresh and vigorous. Evaluation activities provide a perfect opportunity to begin the kind of dialogue about college teaching that could truly make a difference.

An instructor does not need to know how he/she compares with anyone else in order to improve.

SOLUTION 7: To the extent that engaging in formative evaluation activity represents on-going faculty commitment to instructional excellence, it ought to be recognized and rewarded. Good teaching does not happen automatically. For most faculty, it is the result of concerted efforts made throughout a teaching career. That sort of on-going commitment deserves recognition. It merits reward. When institutions provide recognition and reward, they extend to faculty powerful reasons to make that commitment. ■

Notes

[1]For example see: Cohen, P.A. "Effectiveness of Student-rating Feedback for Improving College Instruction: A Meta-analysis of Findings." *Research in Higher Education*, 13 (1980), 321–41 or Rotem, A. and Glasman, N. S. "On Effectiveness of Students' Evaluative Feedback to University Instructors." *Review of Educational Research*, 49 (Summer 1979), 497–511.

[2]Seldin, P. A. *Changing Practices in Faculty Evaluation*. San Francisco: Jossey-Bass, 1984, p. 54.

[3]Braskamp, L. A.; Brandenburg, D. C.; and Ory, J. C. *Evaluating Teaching Effectiveness*. Beverly Hills: Sage, 1984, p. 48.

[4]Samples of the Made-to-Order Form and specifically, What Needs Improvement Forms, are available upon request from The Instructional Development Program, 1 Sparks, The Pennsylvania State University, University Park, PA 16803.

[5]Redmond, M. V. and Clark, D. J. "A Practical Approach to Improving Teaching." *AAHE Bulletin*, (February 1982), 1, 9–10.

[6]Kogut, L. S. "Quality Circles: A Japanese Management Technique for the Classroom." *Improving College and University Teaching*, 32 (Summer 1984), 123–7. McKeachie, W. et al. "Using Student Ratings and Consultation to Improve Instruction." *British Journal of Educational Psychology*, 50, 168–174.