

DOCUMENT RESUME

ED 144 442

HE 009 225

AUTHOR Rose, Clare
TITLE Faculty Evaluation in an Accountable World: How Do You Do It?
PUB DATE Mar 76
NOTE 47p.; Address at the national conference of the American Association for Higher Education (Chicago, March 1976)

EDRS PRICE MF-\$0.83 HC-\$1.67 Plus Postage.
DESCRIPTORS College Faculty; Dental Schools; *Educational Quality; Evaluation Criteria; Evaluation Methods; *Faculty Development; *Faculty Evaluation; *Higher Education; *Instructional Improvement; Performance Based Teacher Education; Systems Approach; *Teacher Improvement; *Teaching Quality
IDENTIFIERS University of California Los Angeles

ABSTRACT

As the accountability movement has brought forth renewed interest in the quality of teaching and new demands for faculty evaluation, it could be expected that the relationship between the two would have created the foundation for a rigorous program of faculty evaluation. Instead, educators continued a fruitless search for qualities and behaviors that defined the perfect teacher. In the meantime, we have lost our perspective about the function of faculty and the real purpose of higher education, and similarly, about the purpose of faculty evaluation. Evaluation is not an end in itself; it is the means to improving instruction, to providing a better education for our students. It must be recognized that higher education is accountable first to the students for high quality instruction. There is a variety of sources of evaluative data, each source having certain limitations. In the UCLA School of Dentistry an ongoing, systematic, and comprehensive program of teaching and curricular evaluation has been established that uses the criterion-referenced system of instruction and measurement, supplemented informally by other, more traditional measures. An instructional development program for faculty is an integral part of the evaluation system, making available the kind of help that will enable them to maximize attainment of their contracted responsibility. Unless colleges and universities today make similarly serious commitments to improving instruction and evaluating teaching by providing substantial financial and personnel support, higher education may never recover its former importance and respect in society. (Author/MSE)

Documents acquired by ERIC include many informal unpublished materials not available from other sources. ERIC makes every effort to obtain the best copy available. Nevertheless, items of marginal reproducibility are often encountered and this affects the quality of the microfiche and hardcopy reproductions ERIC makes available via the ERIC Document Reproduction Service (EDRS). ERIC is not responsible for the quality of the original document. Reproductions supplied by EDRS are the best that can be made from original.

ED144442

FACULTY EVALUATION IN AN ACCOUNTABLE WORLD:

HOW DO YOU DO IT?*

Clare Rose
Evaluation and Training Institute

U.S. DEPARTMENT OF HEALTH
& WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRO-
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN-
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT
OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

Every year vast numbers of faculty members are "evaluated", and major decisions affecting their lives and the lives of their students are made on the basis of inaccurate, unreliable, inferential, subjective and unsystematically collected information. On paper, particularly at large universities, faculty are evaluated in three areas -- research, community service and teaching. In reality, we all know that "publish or perish" prevails in the trinity. We hear that the majority of administrators at liberal arts colleges are themselves satisfied and believe that their faculty are satisfied with the policies and practices used for evaluation. In fact, we know that this is not true. The majority of faculty evaluations at liberal arts colleges are based on chairmen and deans' subjective evaluations, and faculty look upon them much as a child forced to take cod liver oil -- with suspicion and distaste.

If the audience detects a note of dissatisfaction in these statements, they are correct. I am appalled at the shoddy practices that characterize the "state-of-the-art". I am dismayed by the fact that many of the very same critics of faculty evaluation who argue that teaching cannot be assessed because it

Invited address at the National Conference of the Association of American Higher Education, Chicago, Illinois, March 7-10, 1976

HE 009 202 5-

is too complex a process and that "really important learning" cannot be measured (an idea that survives without a shred of evidence), then accept evaluation procedures based on a generally haphazard collection of opinions from students, colleagues and administrators; and even more, accept almost uncritically important decisions based on inferences drawn from these opinions.

Certainly, I am not alone in my concern about faculty evaluation. Several of education's formidable gadflies have been arguing for systematic faculty evaluation procedures for over thirty years. In my mind, however, the progress we have made would astonish even the most cautious tortoise. During the last few years, as the accountability movement brought forth renewed interest in the quality of teaching and new demands for faculty evaluation, one would have thought that the rather obvious relationship between the two factors would have emerged as the foundation for a rigorous program of faculty evaluation. No such logic prevailed; and people continued their fruitless search for qualities and behaviors that define the perfect teacher. Hundreds of studies were designed to isolate an index of instructional skills, and almost without exception, all failed to yield even a hint.

In the meantime, we somehow lost our perspective about the function of our faculty and the real purpose of higher education. We have lost our perspective about the dual purposes of faculty evaluation: 1) to render judgments about an individual teacher's instructional effectiveness as part of his competence as a faculty member; but, even more important, 2) to provide faculty members

with the means to improve their teaching. Faculty evaluation is not an end in itself. It is the means to improving instruction, to providing a better education for our students.

Let us suppose for a moment that a college or university was genuinely interested in establishing a judicious and effective faculty evaluation system. What should they do? Congruent with the title of this symposium, "Faculty evaluation in an accountable world: How do you do it?", we must first ask to whom is higher education accountable . . . and for what? For my part, there is little doubt. We are accountable to the students! And for what? Unquestionably, high quality instruction.

If we accept that a high quality education is the basic mission of higher education, then the procedures used for evaluating faculty should at least include some evidence of a faculty member's ability to promote learning. If learning is the goal, then a crucial element in the teacher evaluation process is the means whereby both student and teacher can determine whether the student has in fact learned what he was supposed to have learned. Can the student analyze the problem, formulate hypotheses, solve the equation, translate the sentence, describe the imagery? Does the student demonstrate awareness and sensitivity, use critical thinking, love learning for learning's sake? If not, why not? What are the errors and why do they exist? The answers to these questions provide direction for the teacher during the instructional process and a more appropriate base for evaluating a person's teaching performance.

found that faculty tend to overrate themselves with respect to their overall effectiveness. At the same time, because they have received no training on how to focus on the really relevant aspects of their instruction, they criticize superficialities such as appearance, mannerisms and voice tone -- none of which bear any significant relationship to either student satisfaction or achievement. Relationships investigated between self-evaluation ratings and other criteria, such as student ratings and measures of student gain, have been found to be negligible.

3. Student ratings. Without question, the evaluation measure most widely employed today is the ubiquitous student rating questionnaire. Student ratings do have some merit. They are relatively easy to collect, provide a comparatively large data base, and even more important, provide a means for student input into the educational system (however superficial that may be).

That they are valid and reliable is supported by the research. But so is the fact that they are affected by such factors as class size, academic discipline, course content and faculty experience. To make promotion and tenure decisions on the basis of student opinion would result in serious injustices to the teachers of large classes or required courses in "tough" disciplines. Even more serious is the fact that there is simply no evidence to suggest that student ratings are effective aids for improving teaching.

To discover post hoc that students felt you didn't "stimulate them to high intellectual effort" gives little constructive feedback which can help you detect why you didn't and how you could do

so in the future. On the other hand, finding that you "encouraged students to express themselves freely and openly" or "spoke with expressiveness and variety in your tone of voice" might lead you to join Actor's Equity, but tells you nothing about whether the students learned as a result of your vibrant tones.

Consumer satisfaction is certainly an important consideration in any evaluation scheme, and student ratings are reliable and valid indicators of students' feelings of satisfaction. But, as the Dr. Fox experiments amply demonstrated, student satisfaction with learning represents little more than the illusion of having learned. Inclusion of student ratings in an evaluation system must be done with this reality in mind.

4. Criterion-referenced measurement. Measuring change in student behavior toward explicitly defined educational objectives is a relatively new and most promising approach to faculty evaluation. Commonly referred to as criterion-referenced measurement, this system focuses on the knowledge, skills or attitudes students have acquired as a result of their instruction. Based on the assumption that effective teaching means promoting changes in student learning, criterion-referenced evaluation directly measures the teacher's impact on student learning.

Instead of identifying a student's relative status in the class (as in the traditional norm-referenced measurement system), attention is focused on the individual student and his changes in behavior -- in other words, changes in his learning. The precise boundaries of the behavior to be assessed are defined, and criteria for judging the adequacy of the student responses are identified prior

to instruction. Thus, an absolute rather than a relative standard is used to measure learning.

Criterion levels can be established at any point during instruction where information concerning the level of an individual's performance is desired. The term "criterion" does not necessarily refer to final, end-of-the-course behavior. The point is that specific behaviors and attitudes implied at each level of proficiency are identified and used to describe the specific tasks a student must be capable of performing.

One of the chief problems with using criterion-referenced measurement, or its outgrowth (contract plans) for evaluating teaching, is that since different teachers set different goals within the same discipline, and often, even within the same course, it is difficult to compare the relative proficiency of teachers who are pursuing different goals with different groups of students. The adoption of such a process also requires the education of faculty and administrators so that they both know how to define explicit teaching goals, evaluate the worth of the goals, and to construct the difficult, but absolutely necessary, appropriate criterion-referenced tests. Such an educational program is both expensive and time consuming.

The advantages of the system to my mind, however, outweigh its disadvantages. While the process is difficult and time-consuming, the payoff is well worth the effort. First of all, by defining specific objectives prior to instruction, the teacher can more likely employ relevant instructional materials that will help students attain the objectives. The process of specifically defining objectives forces instructors to carefully think through exactly what they really want to teach, and it is then much more

likely that they will head for truly worthwhile goals. Secondly, students' progress toward attainment of the objectives can be continuously monitored so that instruction can be continuously improved while the course is still in progress, not after the fact. Finally, the criterion-referenced measurement system provides for a fairer evaluation of a faculty member's competence AS A TEACHER. There is a mass of evidence that no one style or method of teaching produces superior learning. The criterion-referenced measurement system takes this well-established but often forgotten fact into consideration. The faculty member whose students make progress toward and achieve educational goals is an effective teacher regardless of the methods he uses.

The evaluation of teaching is a far too serious and complicated process to be based solely on the personal assessment of administrators, the judgment of visiting peers, an examination of course syllabi or teaching methods, or especially student opinions. Each of these is useful, but none is sufficient by itself. They are not even that valuable in total. We must recognize the deficiencies in these traditional approaches, continue to explore and refine more direct measures of learner growth, and, until the perfect evaluation system is found, if ever, opt for multiple indicators of assessing faculty performance.

With this last statement in mind, let me now describe a teaching evaluation program that my associate, Glenn Nyre, and I have recommended be implemented at the UCLA School of Dentistry next year. The proposed design of the program is based on many of the ideas contained in this paper, and takes into consideration the advantages and disadvantages of the various data forms previously described.

THE UCLA SCHOOL OF DENTISTRY TEACHER EVALUATION PROGRAM

First, let me say that the climate is right! For the last two years, we have worked with the administrators and faculty of the School of Dentistry to establish the foundation for an ongoing, systematic and comprehensive program of teaching and curricular evaluation.* With almost complete and direct involvement of the entire faculty and a major portion of the student body, explicit, measurable goals were defined for the School and its 14 curricular sections. In the process of clarifying these goals, faculty saw the need to become more explicit about their own course objectives -- making sure that what they were teaching in class corresponded to and contributed to the larger and more long-range school goals. This process led to a renewed focus on teaching, and an excitement was generated that resulted in the faculty asking for faculty development workshops on everything from the evaluation of instruction to the development of a self-paced instructional program.

A comprehensive education program for the faculty was thus developed at their request which included teaching them how to specify instructional objectives, to analyze the instructional

*A more comprehensive description of the renewal project at the UCLA School of Dentistry can be obtained by writing to the Evaluation and Training Institute, 11110 Ohio Avenue, Suite #202, Los Angeles, California 90025

tasks implicit in the target goals, to incorporate into their instruction empirically derived instructional practices that increase students' chances of attaining objectives, to construct valid and reliable criterion tests and criterion checks to monitor students' progress toward the objectives and how to measure student achievement and attitudes at the end of the course. In sum, faculty were taught the criterion-referenced system of instruction and measurement.

Currently, faculty are working with us on the development of a comprehensive questionnaire which will be used to survey recent graduates of the School to determine their attitudes and behaviors in their practices as they relate to the goals which were established. In this manner, the goals themselves will be evaluated to determine their relevance to the world of the practicing dentist, and thus, their worth and viability for future graduates. The curriculum will be revised accordingly, if necessary, and the cycle of evaluation and change will continue.

It should be clear that this group of faculty is ready for a program of systematic teaching evaluation, and their cooperation is insured. They have learned the rubrics of criterion-referenced

measurement and have used the system in a variety of ways. As far as their instruction is concerned, they are committed to the measurement of outcomes as indicators of student learning and the logical next step for them is to implement an outcomes-based teacher evaluation system.

Colleague visitation and student ratings will continue to be used, but colleague visits will be made on an informal basis only, at the discretion of the faculty member, and will not be included in the evaluation profile. Student ratings will be used as indicators of satisfaction only, and included with them will be data concerning the size of the class, the nature of the course and whether or not the course is required -- factors that have been shown to influence student ratings. The main thrust of the evaluation profile will be a combination of concrete evidence of direct teaching effectiveness within each course.

The evaluation program will begin before classes start in September with a meeting of each faculty member and the section chairman. At this time, the objectives for each course taught by the faculty member will be reviewed, as well as the evidence that will be collected from students to determine their degree of attainment. Levels of performance will be established for both individual students and the class, and the conditions under which the students' behaviors are to be demonstrated will be specified. Although each faculty member will be expected to continuously monitor class progress throughout the course, for the first year the faculty member and section chairman will not meet again until the end of the quarter (unless the faculty member

requests such a meeting), at which time they will review the level of class achievement, individual students achievement and the objectives themselves. As a result of this analysis, objectives may be reformulated for the next class and expectations regarding student and class levels of achievement may be revised.

At the end of the academic year, the section chairman will meet with faculty both individually and as a group to determine the overall efficacy and interrelationship of the course objectives and the general level of student achievement. By reviewing an individual's course evaluations from three quarters, an assessment will be made of each faculty member's ability to promote student learning generally, which objectives proved too difficult to accomplish, what methods proved effective or ineffective, and so on. Student satisfaction based on student rating forms administered to each class each quarter will also be reviewed and areas of dissatisfaction analyzed. It should be noted, however, that the purpose and emphasis of this program is diagnostic --- to assess each faculty member's current level of teaching effectiveness, identify problem areas, and establish the foundation for improvement.

An integral and parallel part of the evaluation program will necessarily be an on-going instructional development program to which faculty can come for help any time they wish. We believe that it is impractical and foolhardy to establish a teaching evaluation program without an accompanying improvement program. By having available to faculty an instructional improvement program congruent with the ingredients of the evaluation program, they will receive the kind of help that will enable them to maximize attainment of their contracted responsibility. Nothing

succeeds like success itself. Thus, for the next year, and possibly even the second year, the focus of the evaluation program will be on teaching improvement and not on promotion or tenure. Once the faculty have had a chance to improve their teaching effectiveness, to demonstrate their ability to bring about student learning, and to practice the process of collecting outcome data, the end-of-year chairman/faculty meeting should ultimately result in an evaluation profile that can be used (along with traditional evidence of scholarly research and community service) in arriving at reappointment, promotion and tenure decisions. Teaching will no longer be the ugly step-sister of the group.

The experimental teaching evaluation program I have just described (as well as our entire change and renewal project at the School of Dentistry) is based on three major assumptions: 1) that the goal of curricular development and organizational change is the improvement of teaching and learning; 2) that any really meaningful changes in the curriculum and, ultimately, improvement in the teaching/learning process must be based upon a rigorous evaluation program; and 3) the focus of such an evaluation program must be on outcomes -- outcomes in terms of student achievement and satisfaction; faculty development and satisfaction; the responsiveness of the curriculum; and finally, outcomes in terms of the total school and organizational environment.

These are the conditions necessary for a viable faculty evaluation program. Accountability means more than just instructional accountability, and institutional cost-effectiveness. It means the entire system of higher education is accountable for

providing students with an effective, responsive and satisfying education. It means that the goals of each institution must be defined in order to determine what constitutes a "good" education for the students at that institution. It means that institutional changes must be made in the environment and particularly the reward structure so that faculty members who are truly effective teachers are recognized and encouraged in their efforts. It means that institutions must make a serious commitment to improving instruction and evaluating teaching by providing substantial financial resources and personnel support. I am convinced that unless colleges and universities today make such a commitment, higher education may never recover its former importance and respect in society. That would be a sad ending indeed.

REFERENCES

Centra, J.A. Colleagues as raters of classroom instruction. Journal of Higher Education, 1, May/June, 1975, 327.

Cohen, A.M., Trent, J.W. and Rose, C. Evaluation of teaching. In Travers, R.M.W. (Ed.) Second Handbook of Research on Teaching. Chicago: Rand McNally, 1973, 1041-1071.

Eash, M.J. Systems design and the evaluation of educational programs. Paper presented at American Educational Research Association annual meeting, Washington, D.C., March, 1975.

Eble, K.E. Evaluation and the improvement of teaching. Teaching in Higher Education, II, Ohio Board of Regents. Delaware, Ohio: 1975, 822-835.

Garman, N. A mastery model of evaluation as a strategy for developing professional teaching competencies. Paper presented at American Educational Research Association annual meeting, Washington, D.C., March, 1975.

Glasman, N.S., Killait, B.R. and Gmelch, W. Evaluation of Instructors in Higher Education: Concepts, Research and Development. Santa Barbara, Ca: Regents of the University of California, 1974.

Grasha, T. Principles and models for assessing faculty performance. Teaching-Learning Monograph Series, I, (2), Cincinnati, Ohio: IRTE, 1972.

McNeil, J.D. Toward Accountable Teachers: Their Appraisal and Improvement. New York: Holt, Rinehart and Winston, 1971.

McNeil, J.D. and Popham, W.J. The assessment of teacher competence. In Travers, R.M.W. (Ed.) Second Handbook of Research on Teaching. Chicago: Rand McNally, 1973, 218.

Marcus, J.S. A discussion of a proposed system for rewarding and improving instructional effectiveness. Paper presented at American Educational Research Association annual meeting, Washington, D.C., March, 1975.

Naftulin, D.H., Ware, Jr., J.E. and Donnelly, F.A. The Doctor Fox Lecture: A Paradigm of Educational Seduction. Teaching in Higher Education, II, Ohio Board of Regents, Delaware, Ohio: June, 1975, 860-863.

Parlett, M. and Hamilton, D. Evaluation as illumination: a new approach to the study of innovatory programs. Centre for Research in the Educational Sciences, University of Edinburgh: Occasional paper, October, 1972.

Piercé, H.B. and Schroeder, L.L. An objectives-based participatory evaluation plan for teaching faculty. Educational Technology, August, 1974, 28-32.

Popham, W.J. Educational Evaluation. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1975.

Popham, W.J. (Ed.) Criterion-referenced Measurement, An Introduction. Englewood Cliffs, New Jersey: Educational Technology Publications, 1971.

Seldin, P. How Colleges Evaluate Professors. New York: Blythe-Pennington, Ltd., 1975.

Smock, H.R. The practice of evaluation as a purposeful activity. Paper presented at American Educational Research Association annual meeting, Washington, D.C., March, 1975.