

## Evaluating Instructional Effectiveness Can Be a Rewarding Experience

LAWRENCE M. ALEAMONI, Office of Instructional Research and Development, University of Arizona, Tucson 85721

More than anything else, an institution of higher education needs visible ways of convincing its faculty that: 1) good instructors are desired, 2) good instruction is respected, 3) the improvement of instruction is a major concern, and 4) there are specific ways of recognizing and rewarding good instruction. Every institution is already committing a major portion of its budget to maintain its instructional program: faculty salaries, classrooms, libraries, laboratories, instructional services, sabbatical leaves, and the like. Interestingly enough, however, when a new graduate prepares to embark on a professional career at any one of the institutions of higher education in the United States or abroad, he or she is told by the department head or dean that the institution embraces the three general objectives of excellence in teaching, research, and service and that rewards are based on satisfactory to excellent performance in any one or a combination of those objectives. The sad fact is that after a short period of time at the institution, the new faculty member realizes that, although the three objectives of teaching, research, and service are appropriate for any institution of higher education, most of the institutions reward faculty primarily for their performance in the research function (11,23).

This is a disturbing state of affairs because it indicates, basically, that most institutions are only interested in supporting and encouraging excellence in research. Its consequences are obvious in that faculty, regardless of their interest, may neglect their teaching and service activities in order to attain the professional recognition required to remain and succeed at their institution. Students are perhaps the most unfortunate pawns in such a game because they are forced to take courses from faculty who are not able or willing to spend the time to prepare and organize their courses or to do such things as spend time outside class discussing problems and concerns that would help most students learn material better.

Some take the position that there is no inconsistency in this type of skewed reward system because excellent researchers are, in fact, the best teachers (12,14). The research evidence (2,10,15,16,18,20,24), however, does not support this point and shows that, in general, there is no correlation between scholarly productivity and effective teaching.

It is further suggested that the evaluation of scholarly productivity is much easier, more valid, and more reliable than is the evaluation of instructional effectiveness. This attitude has generally resulted in a policy (dictated by practice) indicating that, regardless of the quality of the evaluative teaching evidence, published research will still take precedence in the reward considerations. Under this policy, the relationship between teaching and scholarly productivity is believed to be high and positive even though the research evidence to date does not support this belief.

The criteria used to judge scholarly achievement rest basically on the belief that judgments by colleagues (synonymous with "the academic community at large") provide the final evidence. These judgments typically include: 1) obtaining research grants, 2) publishing in journals where expert evaluation is required for acceptance, 3) receiving favorable reviews of books, 4) receiving appointments or

awards that require evaluation of professional competence, 5) being elected to office in learned societies, and 6) receiving fellowships. When one looks closely at these judgments, it becomes obvious that each is very subjective; in fact, it would be very difficult to find a set of objective criteria on which many colleagues could agree when judging journal publications, books, etc. Those who have read several advisory editors' reviews of their articles or experienced the acceptance of a once-rejected article by another equally reputable journal could attest to this subjectivity.

Where, then, is the rationale for maintaining that the evaluation of teaching based on the testimony of department heads and deans, the comments of colleagues who are well acquainted with the teaching performance of the instructor, the achievements of students, the quality of teaching materials prepared by the instructor, and the judgments of students exposed to the instructor is more subjective, less reliable, and less accurate than the judgments of scholarly achievement? Apparently, the answer lies in the fact that the methods and techniques used in evaluating scholarly achievement have remained unchanged for so long that they have become accepted as standard without question or evaluation. In contrast, the evaluation of teaching has been subject to a wide variety of approaches with no commonly accepted methods or techniques except for student judgments gathered via rating forms, some of which have not, admittedly, been professionally designed.

A study by Ladd and Lipset (19) further emphasizes the need for institutions formally recognizing and rewarding instructional efforts. They conducted a survey of U.S. faculty members in all disciplines and found that as a group they are more interested in teaching than in research and scholarship. In fact, for every professor strongly devoted to research they found nine equally devoted to teaching.

Scriven (22) states that the evaluation of instruction must be part of a system that also has an appropriate process for evaluating research and service and a specific commitment to their relative weighting. The instructional evaluation component must itself be a system with several components.

If institutions of higher education (and their departments) are serious about encouraging excellence in teaching, they must be prepared to reward such excellence in a manner similar to that for excellence in scholarly achievement. One approach toward reaching this goal would be to have departments define what constitutes effective instruction and then establish a system of evaluation that would have as its two major outcomes guidance for improvement and reward for success. An integral element would be an agreement (or contract) between the department head and the faculty member outlining the percentage of time commitment to teaching, research, and service. This article outlines one possible system of evaluation for encouraging instructional and faculty development (3).

### Methods of assessing instructional effectiveness

Evaluation of student attitudes about instructors and instruction is perhaps the most commonly used method of assessing instructional effectiveness. The rationale for this is that students are assumed to be the only ones who are constantly exposed to the elements of the course (e.g.,

instructor, textbook, homework, course content, method of instruction, etc.) and, therefore, are the most logical evaluators of the quality and effectiveness of those elements. In addition, student attitudes can indicate areas of rapport, degrees of communication, and the existence of problems and thereby help instructors as well as instructional development personnel describe and define the learning environment more concretely and objectively than they could through other types of measurements (5,8,9).

Self-evaluation on the part of the faculty member is necessary to determine the rationale for the various elements of the course(s) taught. Such self-evaluation could deal with: 1) course content, 2) method of instruction, 3) ability to keep student interest and attention high, 4) ability to promote student learning, 5) course and instructional objectives, and 6) course and instructional organization.

The quality of student learning is another important variable to consider and could be used to determine if the course objectives are being met. It would be possible to get a measure of student achievement in a course by employing course or departmentally agreed on examinations that were designed to fit the instructional objectives. Nationally standardized examinations could also be used if they were judged to accurately measure the satisfaction of course and/or departmental objectives. Each of these could then be used either as final examinations or as a part of the final examination. Other measures could be obtained by follow-up assessments of student performance in succeeding courses or by following the students after they have left the institution and recording their performances at other institutions of higher education or in their jobs or professions. Such follow-up information would be appropriate if it could be related to specific courses.

Faculty (colleague) and department head attitudes and judgments represent the most comprehensive method of assessment. For example, these judgments could be based on an evaluation of: 1) the procedural adequacy of the course(s), 2) the appropriateness of the course content, 3) the appropriateness of course objectives, 4) the adequacy and quality of the course examinations, 5) the stability of an instructor's teaching effectiveness over the years, 6) the student attitudinal report, 7) the instructor's self-report, 8) the quality of student learning, and 9) other course materials or comments that the instructor wishes to submit. Such evaluations ideally should be made by three-person departmental review committees using well-structured guidelines and criteria of excellence and, where needed, instructional development personnel expertise.

In order to accurately measure and evaluate instructional effectiveness, departmental faculty need to agree on a set of criteria and then guidelines for that evaluation. Once a set of criteria has been agreed on, the departmental faculty would need to establish the guidelines (standards) necessary to evaluate them. For example, the departmental standard may require the peer evaluation committee to verify that at least 25% of the content of a particular course be based on information generated in the last 5 years (7). The peer rating item could be constructed as follows: "What percent of material in syllabus is current (based on information generated in the last 5 years)? 1) 100-75% (definitely needs no improvement), 2) 74-50% (needs very little improvement), 3) 49-25% (needs some improvement), 4) 24-0% (definitely needs improvement)."

Standards can be established by having each departmental faculty member submit information pertaining to the agreed on criteria. A departmental committee (either elected or appointed) would then organize the material in a qualitative (and in some cases quantitative) order to establish the initial standards to be used in evaluating future material submissions. With the agreed on criteria and standards, the departmental peer review committee could conduct an evaluation of a specific faculty member in approximately 1-3 hours. Such evaluations would normally be conducted once a year. For examples of the effectiveness of such systems see Harrell (17) and O'Connell and Smartt (21).

## Feedback to the instructor

All assessment results should be provided to the faculty member in a form that allows easy and accurate identification of instructional problems as well as suggestions for solving such problems. After the results have been carefully analyzed and possibly used by the faculty member to design and implement solutions to instructional problems, summary results could be submitted to the department head and weighted appropriately in a teaching + research + service equation for use in rank, pay, and tenure considerations.

The student attitude or rating data should be presented to the instructor in an easily readable and interpretable format. Student responses to the questionnaire items should be reported in summary form for each course, and appropriate comparisons should be made between the mean scores in the instructor's course (by each item and groups of items) and the means for similar instructors teaching similar courses in the same department, college, and university. For examples of appropriate normative comparisons see Costin et al (13) and Aleamoni (1,4,6).

The instructor should be provided with enough interpretive material to understand the summary information as well as with the possibility of consultation with the instructional development personnel responsible for providing the student attitudinal information. Any additional evaluative information and resources needed in this area should also be discussed with the instructional development personnel. All interpretive material and consultation should concentrate on identifying instructional strengths and weaknesses and then agreeing on what should be done to remedy the weaknesses.

The quality of student learning data gathered should also be presented to the instructor in an easily readable and interpretable report. This report, however, should concentrate on the achievement of course, department, college, and university instructional objectives.

Faculty (colleague) and department head attitudes and judgments should be presented to the faculty member in an easily readable and interpretable report. The report should outline the committee and individual judgments made on each element of the instructional setting considered and identify the strengths and weaknesses observed. Constructive criticisms that would allow the instructor to plan a strategy for improvement in the areas of weakness should appear in the report. Personal consultation should take place between the department head and the faculty member concerning all aspects of this report. The department head should explain the options or resources available to instructors to help them in their attempts to remedy their instructional weaknesses.

The department head should hold an individual conference with the instructor concerning all the evaluations that have been conducted and what they mean to the instructor. (This, of course, assumes that the department head understands the evaluations made.) The department head should emphasize what the instructor needs to do in order to resolve the weaknesses or problems uncovered in the various evaluations made. The department head should also clarify how such evaluations and efforts can be used in the rank, pay, and tenure considerations of the individual faculty member.

## Options for the instructor

If all the assessments of instructional effectiveness are highly positive, the instructor could request that the summary results (e.g., course content category evaluation rather than evaluations made on specific items such as textbook, homework, etc.) be placed in his or her departmental file to be used for rank, pay, and tenure considerations. If, on the other hand, all the assessments of instructional effectiveness are not highly positive, the instructor should identify (with the possible aid of instructional development personnel as well as the department head) what needs to be done to remedy the instructional weaknesses. The instructor could then take the

necessary steps to rectify those weaknesses. If additional resources are needed to plan and implement an instructional improvement strategy, however, a proposal could be generated along the following lines:

1. The instructor drafts a proposal outlining what needs to be done, what instructional resources will be needed, how long it will take, how much it will cost, and how it will be evaluated.
2. The instructor discusses the proposal with the department head and obtains any necessary approval and/or additional resources.
3. The instructor also discusses the proposal with the appropriate instructional development personnel to determine what services are available in trying to implement the proposal.
4. Finally, the instructor modifies the proposal according to the suggestions made in discussions with the department head and instructional development personnel.

### **Implementation of the instructional improvement proposal and the project's final report**

Once a final draft of the proposal is generated and all approvals obtained and resources identified, the instructor should initiate the timetable needed to implement the proposal.

After the instructional improvement project has been completed and all evaluations have been made, the instructor should prepare a comprehensive final report. Instructional development personnel could be used at this point, if needed. The final report (containing the results of the instructional improvement effort) and the proposal should then be submitted to the department head and placed in the instructor's departmental file for use in rank, pay, and tenure considerations.

### **Administrative procedure for considering instruction in the institutional reward system**

All of the instructional effectiveness assessments should be evaluated and considered relative to the teaching load required of the faculty member. If an instructional improvement proposal and final report are in the instructor's departmental file, they should be analyzed carefully and evaluated by a departmental review committee utilizing methods and techniques developed by the department. This evidence should be considered proportionate to the teaching load required of the faculty member and to the time needed to generate the proposal, complete the project, and write up the final report. That evidence should be weighted accordingly in the teaching + research + service equation in arriving at rank, pay, and tenure recommendations.

In conclusion, the high-level administration of the institution (including deans and department heads) should encourage the development and use of such a system as an integral part of the institutional reward scheme. When such "systems" have been developed, the satisfaction on the part of the participant faculty

and administrators appears to be very high (17,21). It therefore seems logical to encourage the development of comprehensive systems of instructional evaluation that have guidance for improvement and reward for success as their two major outcomes. Not to do so may suggest a lack of interest in and a lack of concern with instruction and instructional improvement.

### **LITERATURE CITED**

1. Aleamoni, L. M. 1972. Illinois Course Evaluation Questionnaire (CIEQ) Results Interpretation Manual Form 66 and Form 32. Res. Rep. 331. Measurement and Research Division, Office of Instructional Resources, University of Illinois, Urbana.
2. Aleamoni, L. M. 1972. Response to Professor W. Edward Deming's "memorandum on teaching." *Am. Stat.* 26(4):54.
3. Aleamoni, L. M. 1976. Proposed system for rewarding and improving instructional effectiveness. *Coll. Univ.* 51(3):330-338.
4. Aleamoni, L. M. 1976. Typical faculty concerns about student evaluation of instruction. *Natl. Assoc. Coll. Teach. Agric. J.* 20(1):16-21.
5. Aleamoni, L. M. 1978. Development and factorial validation of the Arizona course/instructor evaluation questionnaire. *Educ. Psych. Meas.* 38(4):1063-1067.
6. Aleamoni, L. M. 1979. Arizona Course/Instructor Evaluation Questionnaire (CIEQ) Results Interpretation Manual Form 76. Office of Instructional Research and Development, University of Arizona, Tucson.
7. Aleamoni, L. M. 1981. Standards for evaluation of instruction. Note to the Faculty No. 11. Office of Instructional Research and Development, University of Arizona, Tucson.
8. Aleamoni, L. M. 1981. Student evaluations of instruction. Pages 110-145 in: *Handbook of Teacher Evaluation*. J. Millman, ed. Sage Publications, Beverly Hills, CA.
9. Aleamoni, L. M., and Spencer, R. E. 1973. The Illinois course evaluation questionnaire: A description of its development and a report of its results. *Educ. Psychol. Meas.* 33:669-684.
10. Aleamoni, L. M., and Yimer, M. 1973. An investigation of the relationship between colleague rating, student rating, research productivity, and academic rank in rating instructional effectiveness. *J. Educ. Psychol.* 64:274-277.
11. Astin, A. W., and Lee, C. B. T. 1966. Current practices in the evaluation and training of college teachers. *Educ. Rec.* 47:361-375.
12. Borgatta, E. F. 1970. Student ratings of faculty. *AAUP Bull.* 56:6-7.
13. Costin, R., Greenough, W. T., and Menges, R. J. 1971. Student ratings of college teaching: Reliability, validity, and usefulness. *Rev. Educ. Res.* 41:511-535.
14. Deming, W. E. 1972. Memorandum on teaching. *Am. Stat.* 26(1):47.
15. Guthrie, E. R. 1949. The evaluation of teaching. *Educ. Rec.* 30:109-115.
16. Guthrie, E. R. 1954. *The Evaluation of Teaching: A Progress Report*. University of Washington, Seattle.
17. Harrell, R. A. 1979. The development of a comprehensive faculty evaluation system. *Natl. Counc. Meas. Educ. meeting*, San Francisco.
18. Hayes, J. R. 1971. Research, teaching and faculty fate. *Science* 172:227-230.
19. Ladd, E. C., Jr., and Lipset, S. M. 1975. How professors spend their time. *Chron. Higher Educ.* 11(5):2.
20. Linsky, A. S., and Straus, M. A. 1975. Student evaluations, research productivity, and eminence of college faculty. *J. Higher Educ.* 46(1):89-102.
21. O'Connell, W. R., Jr., and Smartt, S. S. 1979. Improving faculty evaluation: A trial in strategy. *Rep. SREB Faculty Eval. Proj.* Southern Regional Education Board, Atlanta, GA.
22. Scriven, M. 1980. The evaluation of college teaching. *Natl. Counc. Prof. Inservice Educ. Newsl.*
23. Thorne, G. L., Scott, C. S., and Beaird, J. H. 1976. Assessing faculty performance: Final project report. Teaching Research Division, Oregon State System of Higher Education, Monmouth.
24. Voeks, V. W. 1962. Publications and teaching effectiveness. *J. Higher Educ.* 33:212.