University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

POD Network Conference Materials

Professional and Organizational Development Network in Higher Education

10-1-1993

How Colleges Evaluate Professors, 1983 v. 1993

Peter Seldin

Follow this and additional works at: http://digitalcommons.unl.edu/podconference

Part of the Higher Education Administration Commons

Seldin, Peter, "How Colleges Evaluate Professors, 1983 v. 1993" (1993). *POD Network Conference Materials.* Paper 110.
http://digitalcommons.unl.edu/podconference/110

This Article is brought to you for free and open access by the Professional and Organizational Development Network in Higher Education at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in POD Network Conference Materials by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

AAHE Bulletin Oct 1993

Results from a survey of liberal arts campuses.

HOW COLLEGES EVALUATE PROFESSORS

1983 v. 1993

by Peter Seldin

Years ago, the process of faculty evaluation carried few or none of the sudden-death implications that characterize contemporary evaluation practices. But now, as the few to be chosen for promotion and tenure become fewer, and faculty mobility decreases, the decision to promote or grant tenure (or not) can have a stunning impact on a professor's career. Administrators, for their part, are under growing pressure to make fiscally sound decisions in the face of higher operating costs, fund shortages, and bold competition from giant corporations, some of which are moving aggressively into higher education. These conflicting pressures have prompted college professors to question their institution's evaluation criteria and academic deans to reexamine the validity and legality of their personnel decisions.

METHODOLOGY

In that context, in early 1993 a survey of faculty evaluation policies and practices was undertaken. Questionnaires were sent to the academic deans in all of the accredited, four-year, undergraduate, liberal arts colleges listed in the *Higher Education Directory.* (University-related liberal arts colleges were excluded to make the population more manageable.)

Of 658 deans surveyed, 501 (76%) responded — an unusually high response rate. Many of the deans added their comments and attached committee reports and sample evaluation forms used at their college. All this material was read carefully, and the impressions thus gained are included here.

Peter Seldin is distinguished professor of management at Pace University, Pleasantville, NY 10570.

The survey also sought to uncover changes in institutional policies and practices since 1983, when a similar survey was conducted (see the March 1989 *AAHE Bulletin*). The base data for both surveys are identical. Although some institutions had since been accredited and others had closed or merged, the effect of the few differences was negligible.

RESULTS

The survey (and its 1983 counterpart) was designed to gather information on the policies and procedures that guide institutions in evaluating faculty performance for decisions on retention, promotion in rank, and tenure. Reported here are the most significant findings, those on changes in the evaluation of *overall* faculty performance and *classroom teaching* performance.

When Evaluating Overall Performance

When considering a professor for promotion, tenure, or retention, institutions select and weigh a wide range of factors. The questionnaire listed thirteen criteria for consideration by the deans, and they were asked to rate each criterion as a "major factor," "minor factor," "not a factor," or "not applicable." **Table 1** summarizes the relative importance of those criteria as "major factors" in 1983 and 1993.

Examining the data reveals scant change in ten years. In fact, of the thirteen criteria, only the importance of two — student advising and personal attributes — changed by as much as 3 percent.

Classroom teaching continues to be by far the most often reported "major factor" in evaluating *overall* faculty performance. Many deans also give high regard to the other traditional measures of academic repute — research, publication, and activity in professional societies.

The high visibility of published research and professional society membership clearly is a by-product of the economic stress

being experienced by many colleges. As an Illinois dean said, "Our budget comes directly from the state legislature. They want scholarship with high visibility. So our faculty must publish, publish, publish." A New Jersey dean wrote, "The only way our college can stay in the public eye is if our professors do research, publish journal articles, and present papers at professional meetings." The latter remark lends credence to the oft-repeated observation that professors are paid to teach but are rewarded for their research and publication.

The importance of "staying in the public eye" probably is reflected also in the consideration given by some deans to public service by faculty members. At the same time, deans seem to expect faculty to involve themselves in on-campus activities such as committee work and student advising. The emphasis on campus committee work seems to reflect a trend of decentralization and broader sharing of the institution's non-teaching load. In their emphasis on student advising, colleges seem to recognize its value as an outreach effort to keep students content and in school.

Length of service in rank still merits major importance in a professor's overall evaluation at about half the colleges surveyed. Colleges relying on this factor presumably would argue that a positive correlation exists between the number of years in rank and the faculty member's overall contribution to the college. That argument is open to challenge by younger faculty with fewer years of service but rapidly expanding reputations.

Personal attributes, a conveniently elusive criterion used for years by some deans and department chairs to ease targeted faculty out of a job or to deny them promotion or tenure, has declined significantly in importance. This suggests that fewer faculty are being punished today for having the "wrong" personality or practicing the "wrong" politics. A Texas dean wrote, "'Fitting in' today means doing your fair share of teaching and research and doing it effectively. It's no big deal if a professor is from a different mold." A California dean agreed, "Diversity is the name of the game today."

T-test results. To assess change since 1983 in the overall importance deans give these various criteria in evaluating faculty performance, t-tests of differences in mean scores were performed.

First, each of the four possible responses to the criteria was assigned a numerical weight: "major factor"-1, "minor factor"-2, "not a factor"-3, and "not applicable"-3. Next, to determine the mean score of each criterion, its weights were added and that sum was divided by the number of deans reporting. The resulting value was that criterion's overall importance. This ranking process, used by the American Council on Education in an earlier study, simplifies the identification of important factors.

As can be seen in **Table 2,** the mean scores of only two criteria changed significantly over the ten years: personal attributes and supervision of graduate study. Each had a significantly higher mean score in 1993 compared with 1983, indicating a decline in overall importance.

When Evaluating Teaching Performance

Liberal arts institutions have long taken pride in the high caliber of teaching offered by their faculties, a fact supported by the deans' almost unanimous citing of classroom teaching as a "major factor" in evaluating *overall* faculty performance. But how is that

Table 1. Percentage of liberal arts colleges that consider each criterion a "major factor" in evaluating *overall* faculty performance.

Criterion[a]	1983 (N=616)	1993 (N=501)
Classroom teaching	98.7 (1)[b]	98.7
Student advising	61.7 (2)	58.6
Campus committee work	52.6 (3)	53.4
Length of service in rank	46.8 (4)	46.2
Research	33.4 (5)	33.8
Publication	29.2 (6)	29.1
Activity in professional societies	24.5 (8)	23.0
Personal attributes	28.6 (7)	21.9
Public service	17.4 (9)	19.3
Supervision of graduate study	3.7 (10)	2.8
Competing job offers	1.8 (13)	2.4
Consultation (govt., business)	2.4 (11)	2.4
Supervision of honors program	1.9 (12)	1.7

a. In descending order by 1993 scores. b. Rank in 1983.

Table 2. T-tests of differences in mean scores of criteria considered in liberal arts colleges in evaluating *overall* faculty performance.

Criterion[a]	1983 (N=616)	1993 (N=501)	t
Classroom teaching	1.01 (1)[b]	1.00	1.77
Student advising	1.40 (2)	1.41	-0.52
Campus committee work	1.49 (3)	1.46	0.62
Length of service in rank	1.63 (4)	1.64	-0.30
Research	1.71 (5)	1.71	0.05
Publication	1.76 (6)	1.75	0.30
Activity in professional societies	1.80 (7)	1.80	0.10
Public service	1.92 (9)	1.88	1.12
Personal attributes	1.86 (8)	2.01	-3.35**
Supervision of graduate study	2.14 (10)	2.31	-2.31*
Supervision of honors program	2.36 (13)	2.32	0.84
Consultation (govt., business)	2.36 (12)	2.38	-0.48
Competing job offers	2.72 (11)	2.76	-1.37

Note: Test was a t-test for differences in independent proportions.
a. In descending order by 1993 scores. b. Rank in 1983.
* Significant at a 0.05 level of confidence. ** Significant at a 0.01 level of confidence.

teaching itself assessed?

The survey asked the deans to indicate the frequency with which the fifteen sources of information were used to evaluate teaching performance. The deans had four possible responses, and a numerical weight was assigned to each response: "always used"-1, "usually used"-2, "seldom used"-3, and "never used"-3. **Table 3** presents the survey results for both 1983 and 1993.

It is evident that some significant changes are occurring in the ways liberal arts colleges evaluate teaching performance. Over the ten-year period, six of the sources changed in frequency by at least 5 percent; more significantly, all but one of those (dean evaluation) is used *more* widely today. It would seem that the information-gathering process is becoming more structured and systematic, and that many colleges are reexamining and diversifying their approach to evaluating classroom teaching.

The predominate sources of information continue to be ratings by students, the department chair, and the academic dean. However, their relative importance has shifted considerably since 1983.

Student ratings have become the most widely used source of information to assess teaching. A dean in California wrote, "No doubt, students are the most accurate judge of teaching effectiveness." Remarked a dean in Massachusetts, "I would not want to promote or tenure a faculty member without giving student views top priority." Although student ratings are enjoying unprecedented popularity, some dissenters disparage them. Said a dean in North Carolina, "Student ratings have led directly to grade inflation."

Since evaluations from chairs and deans continue to have a major impact, one might ask how sound are the judgments of those chairs and deans? What method do they use? These are questions with no easy answers. In defense, some cite the analogy of clinical medicine, where experienced physicians can make correct diagnoses from obscure symptoms but would be at a loss to explain how they do so.

What "symptoms," then, do administrators look for? Beyond student ratings, they rely to varying degrees on the other criteria listed in Table 3. Interestingly, significantly fewer deans consider research/publication a major factor, suggesting a growing skepticism that a professor's record of scholarly research/publication provides a reliable indicator of teaching competency.

T-test results. The shifts in emphasis over the ten-year period are highlighted by the results of t-tests of differences in mean scores of the sources of information, as shown in **Table 4.** (See above for the t-test methodology used.)

The overall importance of seven sources of information showed statistically significant changes since 1983. Five increased in importance (student ratings, classroom visits, course syllabi/exams, alumni opinions, and self-evaluation/report), and two decreased in importance (dean evaluation and scholarly research/publication).

The increased overall importance of classroom visits to the evaluation of teaching performance served to intensify the conflict over the value of such

(continued on page 12)

Table 3. Percentage of liberal arts colleges that "always used" the source of information in evaluating faculty *teaching* performance.

Information source[a]	1983 (N=616)	1993 (N=501)
Systematic student ratings	67.5 (3)[b]	85.7
Evaluation by department chair	81.3 (1)	78.7
Evaluation by dean	75.0 (2)	67.9
Self-evaluation or report	41.9 (6)	56.0
Committee evaluation	46.1 (4)	48.6
Colleagues' opinions	43.3 (5)	48.6
Classroom visits	19.8 (9)	33.4
Course syllabi and exams	20.1 (8)	29.1
Scholarly research/publication	27.3 (7)	23.9
Informal student opinions	11.5 (10)	13.9
Alumni opinions	3.9 (12)	8.7
Long-term follow-up of students	3.4 (14)	6.1
Grade distribution	4.5 (11)	4.6
Student examination performance	3.6 (13)	3.0
Enrollment in elective courses	1.1 (15)	2.2

a. In descending order by 1993 scores. b. Rank in 1983.

Table 4. T-tests of differences in mean scores of sources of information used in evaluating faculty *teaching* performance.

Information source[a]	1983 (N=616)	1993 (N=501)	t
Systematic student ratings	1.44 (3)[b]	1.19	6.25**
Evaluation by department chair	1.26 (1)	1.30	-0.99
Evaluation by dean	1.36 (2)	1.51	-3.02**
Self-evaluation or report	1.96 (5)	1.65	5.19**
Committee evaluation	2.06 (6)	1.99	0.88
Colleagues' opinions	1.71 (4)	1.68	0.60
Classroom visits	2.42 (10)	2.05	6.76**
Course syllabi and exams	2.22 (7)	2.00	4.23**
Scholarly research/publication	2.23 (8)	2.37	-2.21*
Informal student opinions	2.41 (9)	2.41	-0.27
Alumni opinions	3.08 (13)	2.94	-2.88**
Long-term follow-up of students	3.15 (15)	3.04	2.03
Grade distribution	3.07 (12)	3.07	-0.04
Student examination performance	3.03 (11)	3.14	-2.24
Enrollment in elective courses	3.12 (14)	3.19	-1.46

Note: Test was a t-test for differences in independent proportions.
a. In descending order by 1993 scores. b. Rank in 1983.
* Significant at a 0.05 level of confidence. ** Significant at a 0.01 level of confidence.