Senate Committee on Faculty Affairs

Evaluating Teaching Effectiveness for Promotion and Tenure

(Advisory and Consultative)

Introduction

A university faculty bears much of the responsibility for the quality of education offered at its institution. The faculty carries out this responsibility by setting curricula and by establishing and maintaining standards of excellence in teaching, research and service. The faculty of the Pennsylvania State University recognizes that systematic evaluation of teaching effectiveness, by both peers and students, is essential to maintaining high quality teaching and a vital faculty.

Thus PS-23 requires both peer and student input for the evaluation of teaching effectiveness for promotion and tenure. This report addresses both peer and student evaluation of teaching.

Background

In its final report issued in January of 1980, the Joint Faculty/Administrative Commission to Review and Make Recommendations on University Policies Governing Promotion and Tenure recommended that evaluation of teaching should include both peer and student input and that "the methods of evaluation to be used within a unit shall be selected by the faculty of that unit." Senate legislation based on the Commission's recommendation was passed in February, 1982, but was not accepted by former President Oswald a year later.

Subsequent Senate and Faculty Affairs Committee activity, including a forensic session at which time Dr. John Centra from Educational Testing Services and Dr. Frederick Gottheil, a faculty member at the University of Illinois-Urbana-Champaign, made presentations, led to passage of recommendations concerning the Evaluation of Teaching Effectiveness for promotion and tenure purposes. This legislation differed from earlier legislation in four ways. First, a standard set of four classifications was proposed to rate teaching effectiveness. Second, Commonwealth campus peer reviews were specified in addition to departmental peer reviews. Third, two global questions that every unit should ask were presented. Fourth, the legislation included a set of guidelines for the administration of student evaluations.

However, speaking to the Senate in September, 1984, President Jordan stated that "neither the legislation in recommendation #1, nor the nine guidelines in recommendation #4 should be incorporated into the policy itself, or into the current guidelines..." His reasons were basically that leaving each unit to develop its own methods of student evaluation of teaching effectiveness could create inconsistencies. He suggested that the faculty should consider at least a partially uniform system of evaluation across all units of the University. The recommendation that the Office of the Executive Vice President and Provost of the University act to facilitate
the exchange of methods and ideas concerning teaching evaluation was accepted. Dr. Carol Cartwright, Vice Provost was asked to assume leadership of the effort to review policies and procedures with the Faculty Affairs Committee. Our joint efforts have resulted in this report.

Summary of Research on Student Evaluation of Teaching Effectiveness*

There is an abundance of research on all aspects of student evaluation of teaching effectiveness. The consensus in the literature is that while student evaluations are the most common strategy of evaluation, by themselves they are not sufficient to provide a complete evaluation of teaching. Students, however, are in a unique position to make evaluations and are an appropriate source of information when they are judging student-instructor relationships, organization of course, their views of the instructor's professional and ethical behavior, their work load, what they have learned in the course, fairness of grading and instructor's ability to communicate. They are not good sources from which to judge relevance and recency of course content and knowledge and scholarship of instructor.

Items found on student rating forms are based on commonly identified characteristics of effective teaching and generally fall into three groups: 1) organization, structure or clarity of course and course material; 2) teacher-student interaction and 3) teaching skill. Other subjects of evaluation include evaluation of workload in the course, grading and examinations, student outcomes and global questions. For promotion and tenure purposes, the global or general questions have been found to be most reliable. Reliability of student evaluations is increased by ensuring that an appropriate number of raters per class and number of classes are included for each faculty member. Several factors have been found to have some relation to student ratings: class size, subject matter and expected grade. Whether a course is in a student's major, is being used to fulfill a requirement outside the major, or is an elective has also been found to have some relation to student ratings. Although beginning faculty tend to receive poorer ratings than experienced faculty, faculty characteristics that are generally considered irrelevant to teaching effectiveness (e.g., rank, gender, teaching load, research productivity) have not been found to influence ratings. Student rating forms have limitations. Because student evaluations commonly elicit numerical responses, it is easy to assign them a precision that they do not possess. When used for personnel decisions the possibility of faculty influencing the ratings must be taken into consideration. Standardized and systematic procedures for administering student evaluations are essential to ensuring the usefulness of ratings. Furthermore, student evaluations alone are not sufficient for either personnel decisions or for improvement of teaching.


Principles

The Faculty Affairs Committee bases its current recommendations on the following principles.

1. Student evaluation of teaching shall be in the form of in-class rating surveys and may be supplemented by other forms of student evaluation at the discretion of the unit.

(In this legislation "unit" shall be used to mean department or other corresponding and appropriate academic entity.)

2. A computer-based, "cafeteria" system of student evaluation is appropriate for a university as diverse as Penn State. A cafeteria system is based on a fixed pool of items from which departments and individual faculty can select items most appropriate for their courses.

3. At least a portion of the student evaluation shall be universal across all faculties in all units and on all campuses.

4. A cafeteria system of student evaluation is complex and in order to be used to its maximum benefit, an office must be designated to coordinate the development and administration of the system in consultation with the individual units.

Recommendations

Recommendation 1. The following shall be added to section III of PS-23 after the third paragraph of that section:

Evaluation of Teaching

The evaluation of teaching effectiveness shall be based on both peer and student input. For Commonwealth Campus faculty, this shall include input from peers on the candidate's campus as well as peers from the candidate's unit. (In this section, "unit" refers to the department or similar academic entity.) Specific procedures shall conform to the Statement of Practices for the Evaluation of Teaching Effectiveness for Promotion and Tenure.
Recommendation 2. The following shall be adopted as the Statement of Practices for the Evaluation of Teaching Effectiveness for Promotion and Tenure

Statement of Practices for the Evaluation of Teaching Effectiveness for Promotion and Tenure

Introduction

PS-23 requires that the evaluation of teaching effectiveness for purposes of promotion and tenure be based on both peer and student input. This statement outlines the procedures for obtaining and reporting that input.

A. Student evaluations

1. All units shall use in-class rating surveys for student evaluation of teaching. These surveys may be supplemented by other forms of student evaluation at the discretion of the faculty of the unit.

2. The in-class survey system shall be a computer based "cafeteria" system with a fixed pool of items from which departments and individual faculty members can select items most appropriate for their courses.

3. Each survey form shall consist of three sets of questions -- a University core, a departmental core, and individual faculty items.

   a. The University core shall consist of two global questions that shall be included in all survey forms asking students to give the overall rating of the course and the overall rating of the instructor.

   b. The departmental core. Each department shall select from 5-15 additional items from the pool to constitute a departmental core for the survey form. These items will reflect the nature of the discipline, type of class, and other factors the department deems appropriate. Therefore each department may actually have several different departmental forms; e.g., one for discussion courses, one for lecture courses, one for lab courses, etc. The faculty of each unit shall be responsible for selecting the items that will constitute the departmental cores, subject to the approval of the appropriate academic officer.

   c. Individual faculty members may add up to five additional items to supplement the two global questions and the departmental core.

4. The Office of the Executive Vice President and Provost of the University shall be responsible for coordinating and administering the student evaluation system in consultation with individual units.
5. A pool of items (questions about teaching) shall be developed by the Office of the Executive Vice President and Provost of the University in consultation with appropriate persons from each unit and the Faculty Affairs Committee of the Senate. This pool will include both general and specific items centered around the areas of organization, structure or clarity of the course or course material, teacher-student interaction, teaching skills, instructional environment and specific instructional settings. This pool should be as inclusive as possible rather than restrictive.

6. A set of demographic questions and information shall be developed by the coordinating office to facilitate the proper interpretation of survey results. Included should be such information as class size, percent of students completing survey, whether the course is in a student's major, is a requirement, or an elective, expected grade, etc.

7. All items will be ranked on a seven point scale, and appropriate descriptors shall be provided for the end point and the mid-point of the scale. faculty of the

8. Student Evaluation surveys shall belong to the unit which administers them. The faculty member shall be furnished with a copy of all survey results. Results shall be included in promotion and tenure documentation.

9. Reporting of Results

(See sample in Appendix)

a. Demographic information

Appropriate demographic information shall be reported for each class completing the survey.

b. Evaluation items

(1) The reporting of results of the surveys shall include the following information:
percent of students selecting each response category;
number of students selecting each response category;
mean for each item;
median for each item;
standard deviation for each item;
appropriate normative data (i.e., means or medians for norm groups)

(2) The method of reporting shall indicate the potential degree of uncertainty (confidence interval).
(3) Normative data

Normative information shall be supplied in accordance with the following guidelines.

(a) for the University-wide core: University norms, College norms, departmental norms, and CES norms for CES faculty

(b) for the departmental core: departmental norms for each departmental form

(c) for individually selected items: norms for all faculty who used the item

(d) Other norm groups may be used at the discretion of the unit (e.g., by instructional setting).

(e) There shall be no ranking of faculty in terms of percentiles.

c. The coordinating office shall report the survey results to the appropriate academic unit.

d. Appropriate controls for confidentiality of information shall be implemented.

10. Administration of the in-class student survey system

Implementation guidelines for the use of the in-class survey system shall be developed by the coordinating office based on the guidelines listed below. Specific procedures for the administration and collection of surveys shall be developed by the faculty of each unit.

a. Responses to evaluation questions must remain anonymous.

b. Directions to the students shall be uniform across administrations.

c. The candidate shall not participate in the administration, collection, or compilation of the survey results.

d. Evaluations dependent on samples should be conducted in accord with standard sampling procedures.

 e. At least two-thirds of the students must be present for the evaluation to be administered.

f. The candidate shall not be present during the administration of the evaluation.
g. The responses shall be collected and returned to the appropriate unit office by the person administering the evaluation.

h. Evaluations shall be administered during the next to last week of the semester on a day when no examination is scheduled.

11. Frequency of reviews

The coordinating office shall develop guidelines for determining the frequency of reviews. The specific procedures for determining the frequency of reviews for the faculty members within a unit shall be determined by the unit. Surveys shall also be conducted when requested by the faculty member. The following principles are suggested concerning the frequency of reviews:

a. Where possible, evaluations should be conducted over a period of years and in a variety of courses.

b. Units should refrain from having every course evaluated every time it is taught.

B. Peer Review of Teaching

In addition to student evaluation of teaching, there shall also be evaluation of a candidate's teaching by peers from the candidate's unit and campus. For CES faculty, there shall also be evaluations by peers from the candidate's campus. The same individuals can fulfill both roles.

The methods of peer evaluation to be used by a unit or a campus, as well as the manner in which the results are presented in the dossier, shall be selected by the faculty of the unit or the campus. The procedures must be developed by or selected by the faculty of the unit (or campus) for purposes of evaluating teaching for promotion and tenure. The procedures must be approved by the appropriate dean. The procedures for each campus must be approved by the Vice President and Dean of the Commonwealth Educational System.

While the procedures are to be selected by the unit, the following procedures are suggested for consideration: class visitation, interviews with the candidate, and review of the course material. The unit may adopt a standard evaluation reporting form if it chooses.

C. Review Committee Reports

It is the responsibility of the departmental review committee to make a judgment of the candidate's teaching effectiveness based on both peer and student reviews in terms of the following classification: excellent, very good, satisfactory, and unsatisfactory. Reviewers should understand that unsatisfactory carries a negative connotation; satisfactory should convey a neutral evaluation; very good, a positive one; and excellent, a highly positive evaluation. The committee must provide clear documentation for its judgment.
For Commonwealth campus faculty, the campus review committee shall also make a judgment of the candidate's teaching effectiveness in terms of the four category classification. It is the responsibility of the campus review committee to provide appropriate documentation of its judgment.

D. The section of this document titled "Summary of Research on Student Evaluation of Teaching Effectiveness" shall be used as an appendix in the Statement of Practices for the Evaluation of Teaching Effectiveness for Promotion and Tenure.

Recommendation 3. This legislation shall supersede all previous University Faculty Senate Legislation concerning the evaluation of teaching for purposes of promotion and tenure. The Senate hereby repeals all other such legislation.

Subcommittee
Leonard J. Berkowitz, Chair
Danielle K. Boal
Carol A. Cartwright
Alfred Engel
Zachary Irwin
Elizabeth Walters

Senate Committee on Faculty Affairs
C. Gregory Knight, Chairperson
Leonard J. Berkowitz, Vice Chairperson
Alfred Engel
George W. Franz
Diane Garner
George Heitmann
Zachary Irwin
Philip A. Klein
Harvey Manbeck
Shirley Marchalonis
Ellen L. Perry
William C. Richardson
Irwin Richman
William J. Ross
Robert A. Schlegel
Laura Sims
Paul T. Sindelar
Elizabeth Walters
### Appendix

**SAMPLE FORM: REPORTING OF RESULTS**

FSU Instructor Effectiveness Report  
Fall Semester, 1987

Instructor: Smith, J.  
Dept: Paleobotany  
College: Edulas  
Location: Any Campus  
Course: PALEO  
Number: 932  
Section: 23  
Section Size: 60  
% Responding: 83  
Required in Major? No

<table>
<thead>
<tr>
<th>University Core</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>Dept. Mean</td>
<td>Univ Mean</td>
<td>Low Aver</td>
</tr>
<tr>
<td>e.g., Rate the Course</td>
<td>3.8</td>
<td>3.6</td>
<td>3.3</td>
</tr>
<tr>
<td>N = 5</td>
<td>8</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>Z = 10</td>
<td>16</td>
<td>22</td>
<td>30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item 2</th>
<th>Dept. Mean</th>
<th>Univ Mean</th>
<th>Low Aver</th>
<th>Aver</th>
<th>High Aver</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>e.g., Rate the Instructor</td>
<td>6.1</td>
<td>3.8</td>
<td>3.4</td>
<td>1.</td>
<td>000X000</td>
<td></td>
</tr>
<tr>
<td>N = 0</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>8</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>Z = 0</td>
<td>4</td>
<td>8</td>
<td>8</td>
<td>16</td>
<td>32</td>
<td>28</td>
</tr>
</tbody>
</table>

**Department Core - Form A**

**Item 21**

<table>
<thead>
<tr>
<th>Description of Item</th>
<th>Dept. Mean</th>
<th>Univ Mean</th>
<th>Low Aver</th>
<th>Aver</th>
<th>High Aver</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>N = 0</td>
<td>3</td>
<td>6</td>
<td>9</td>
<td>18</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Z = 0</td>
<td>6</td>
<td>12</td>
<td>18</td>
<td>36</td>
<td>18</td>
<td>10</td>
</tr>
</tbody>
</table>

**Item 35**

<table>
<thead>
<tr>
<th>Description of Item</th>
<th>Dept. Mean</th>
<th>Univ Mean</th>
<th>Low Aver</th>
<th>Aver</th>
<th>High Aver</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>N = 8</td>
<td>10</td>
<td>20</td>
<td>10</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Z = 16</td>
<td>20</td>
<td>40</td>
<td>20</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

(etc.)

**Individual Core**

**Item 28**

<table>
<thead>
<tr>
<th>Description of Item</th>
<th>Dept. Mean</th>
<th>Univ Mean</th>
<th>Low Aver</th>
<th>Aver</th>
<th>High Aver</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>N = 1</td>
<td>3</td>
<td>5</td>
<td>10</td>
<td>19</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>Z = 0</td>
<td>6</td>
<td>12</td>
<td>18</td>
<td>36</td>
<td>18</td>
<td>10</td>
</tr>
</tbody>
</table>

(etc.)