UNIVERSITY OF CALIFORNIA, BERKELEY

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO   SANTA BARBARA • SANTA CRUZ

BERKELEY, CALIFORNIA 94720-1500

OFFICE OF THE VICE PROVOST FOR
TEACHING & LEARNING
228 CALIFORNIA HALL #1500

September 30, 2009

GEORGE BRESLAUER,
Executive Vice Chancellor and Provost

CHRISTOPHER KUTZ,
Chair, Berkeley Division of the Academic Senate

Subject: Final Report from the Joint Senate-Administration Task Force on Teaching Evaluation

Dear George and Chris,

As you know, a task force on teaching evaluation was formed in the Spring Semester of 2008. The Task Force was charged with reviewing the procedures by which teaching is assessed on campus and to recommend changes as to how the evaluation and promotion of teaching could be improved. Of additional interest was the possibility of implementing a campus-wide online system.

In the attached report you will find that the Task Force has collected and summarized the existing guidelines on teaching evaluation for our campus and reviewed the wide range of systems used to evaluate teaching. Based on these findings, the Task Force proposes substantial changes in the evaluation of teaching on the Berkeley campus. In essence, we recommend streamlining the end-of-term evaluations, making these evaluations more uniform across campus, placing increased emphasis on peer evaluation of teaching, encouraging faculty to use mid-semester evaluations, introducing a separate student-oriented teaching evaluation system, and pursuing the implementation of an online or hybrid evaluation system.

While we realize that many of the changes we are recommending will not come without costs, we hope that the report will provide a blueprint for the changes that should be implemented once funding becomes available. We also hope that the report does justice to extraordinary work that is performed on evaluating and improving teaching each year on our campus.

Respectfully submitted,

Christina

Christina Maslach
Vice Provost for Teaching and Learning
Co-Chair Task Force on Teaching Evaluation

Oliver

Oliver M. O'Reilly
Professor of Mechanical Engineering
Co-Chair Task Force on Teaching Evaluation

Attachment: Teaching Evaluation Task Force Final Report

Task Force on Teaching Evaluation Final Report

September 2009

Summary of Recommendations

The Task Force proposes substantial changes in the evaluation of teaching on the Berkeley campus. In essence, we recommend streamlining the end-of-term evaluations, making these evaluations more uniform across campus, placing increased emphasis on peer evaluation of teaching, encouraging faculty to use mid-semester evaluations, introducing a separate student-oriented teaching evaluation system, and pursuing the implementation of an online or hybrid evaluation system.

Recommendation 1: End-of-term evaluations completed by students should meet the following criteria:

- Uniformity: a common subset of data for all instructors;
- Transparency: forms should provide valid and reliable information on instructor effectiveness;
- Flexibility: departments should be able to tailor their forms according to their own needs and practices, adding questions of interest as appropriate.

Recommendation 2: The following information should be gathered in the end-of-term evaluation process for every course in every department:

- Numerical evaluation for instructor's overall teaching effectiveness, as per memorandum from Vice Provost de Vries on 11/19/02;
- Numerical evaluation of how much the student learned in the course;
- Self-reported level, major, and GPA of the student completing the form.
- Departmental averages and standard deviations for overall teaching effectiveness;
- Grade distributions for the course and the average grade distributions for comparable courses;
- Student written comments.

Recommendation 3: Information other than student ratings should be provided for evaluating teaching in merit and promotion cases:

- A self-statement that is included in every case for merit and promotion, and that specifically addresses teaching;
- A statement from the department chair that comments on the teaching effectiveness of the faculty member. The chair's statement should contain information and observations beyond what is found in the end-of-semester evaluation forms;

- Syllabi for the courses that show the content, grading philosophy, and assignments;
- For mid-career appraisals, peer observation of a faculty member's teaching.

Recommendation 4: The campus (through the Office of Educational Development) should develop a database of good end-of-term evaluation questions that can be made available to all departments; departments should be encouraged to revise and streamline their forms, as appropriate, and the campus should provide departments with consultation and support in revising their current forms.

Recommendation 5: The Administration, Academic Senate, and departments should strongly and consistently urge faculty to use mid-semester evaluations and other methods of early feedback in their classes.

Recommendation 6: The campus should fully support efforts to pilot a separate system that will provide more relevant course evaluation information to students.

Recommendation 7: The campus should pursue and fund the development and implementation of an online or hybrid evaluation system. To accomplish this, the Task Force recommends appointing a small working group with technical and pedagogical expertise to recommend the selection of a centralized system. Such an evaluation would include addressing policy, legal, and procedural issues.

Recommendation 8: The campus should provide departments with information and consultation to facilitate the recommended changes in evaluation.

.

Task Force on Teaching Evaluation

Final Report

September 2009

Table of Contents

I. The Charge and Organization of the Task Force

The Task Force on Teaching Evaluation (TFTE) (see Attachment A for the list of members) was convened in the Spring Semester of 2008 at the request of the Administration and the Divisional Council of the Academic Senate. Their request was in response to a proposal coauthored by the chairs of the Committee on Budget and Interdepartmental Relations (BIR) and the Committee on Teaching (COT) (see Attachment B for both the proposal and the official charge letter). The general charge for the TFTE was to review the procedures by which teaching is assessed on campus and to recommend changes, if any, as to how the evaluation and promotion of teaching could be improved. The purview of the task force included teaching evaluation forms, teaching evaluation and promotion policies, peer and external review, and the pros and cons of an online teaching evaluation system.

II. Goals for Evaluating Teaching at Berkeley

1. At the campus-wide level:

   Informing academic personnel decisions regarding merit and promotion by collecting information required by the BIR and the Vice Provost for Academic Affairs and Faculty Welfare.

2. At the department level:

   Helping departments evaluate the overall teaching effectiveness of their faculty, both on a regular basis and during formal program reviews. Helping departments to evaluate program-level learning goals. Alleviating administrative burdens associated with a manual and distributed system.

3. At the individual instructor level:

   Helping all instructors, including Senate and non-Senate faculty, and graduate student instructors (GSIs), to improve their teaching and evaluate course requirements and workload.

4. At the student level:

   Helping students to choose classes by providing them with peers' evaluations of instructors and courses.

III. The Current Status of Evaluation of Teaching on the Berkeley Campus

The primary functions of the existing evaluation systems are to improve teaching on campus and to provide quantitative and qualitative data on an instructor's teaching for use in merit and promotion cases. For faculty, the end-of-semester teaching evaluation is intended to satisfy, in part, the Academic Personnel Manual (APM) section on "Criteria for Appointment, Promotion, and Appraisal," particularly section APM-210(d), which covers the criteria that are intended to serve as guidelines for minimum standards in judging the candidate on teaching (section (1) of APM-210(d)). Information from the APM is reproduced in Attachment C. The COT policy on evaluating teaching is shown in Attachment D.

There is much to commend about the existing systems and the dedication of staff and faculty to the promotion of excellent teaching on the Berkeley campus. Below we describe how teaching is currently evaluated on campus.

a. *Administration and processing:* Each semester on the Berkeley campus over 3000 courses are offered and over 150,000 teaching evaluation forms are collected at the end of the term to evaluate instructors and GSIs. In terms of work-hours, we estimate that over 5000 hours are spent each semester processing and compiling all the teaching evaluation data. This estimate is based on a survey conducted by the TFTE and distributed to MSOs throughout campus.

In contrast to many of our peer institutions, where there exists some sort of centralized system of course evaluation, the Berkeley campus has a decentralized environment where the design of forms and the collection and analysis of teaching evaluation data are at the discretion of the individual departments, schools, and colleges. A survey of these forms conducted by the TFTE revealed that there are over 50 different end-of-term evaluation forms used throughout the campus. There are wide disparities in the range and scope of questions asked on these forms.

The procedures for administering and collecting teaching evaluation forms were formulated in a memorandum by the COT and discussed in a directive from Provost King in September 1987. In compliance with Campus Policy on Teaching Evaluations (see Attachment E: Vice Provost de Vries's memorandum of November 2002), all end-of-term teaching evaluation forms include a question rating the teaching effectiveness of the instructor on a scale from 1-7. In addition, all of the evaluation forms on campus feature requests for narrative comments on the instructor and the course. There is wide-spread consensus that the written comments are an invaluable aid for evaluating and improving teaching.

Over 50 departments and schools on campus use written forms whose quantitative responses must be tabulated by hand. This includes some programs with large undergraduate teaching responsibilities, such as the Department of Physics. In contrast, the Boalt School of Law and the School of Education have developed and implemented their own in-house online evaluation systems. In terms of processing the data from evaluation forms, the College of Engineering, the Graduate School of Public Policy, the College of Chemistry, the Haas Business School and a handful of departments in other Colleges use Scantron-based systems to compile the quantitative responses. The Department of Plant and Microbial Biology uses a hybrid system connected to the Our Unit management system in the College of Natural Resources, which is also being used, or being considered for use, by Chemistry, ESPM, the School of Information Studies, Psychology, and Integrative Biology.

b. *Content and format of current evaluation forms*. In terms of content, the forms used on campus have a wide range of formats and questions. For instance, the form used in the College of Engineering has nineteen questions on topics ranging from assignments and exams to teaching effectiveness; the Department of Astronomy's form has six distinct questions on pedagogy and instructor effectiveness. The form for the Department of English has two questions: answering the first question provides students with an opportunity to comment on the instructor's effectiveness, helpfulness, and availablity, and the effectiveness of the course; the second question features the quantitative teaching effectiveness question. Examination of current forms reveal questions that are hard to answer, that have two-pronged formats, and that may not solicit the information intended. With help from the Office of Educational Development (OED), many of the forms could be updated and improved to take into account recent research in the area of teaching evaluation. For example, the interdisciplinary undergraduate major in Media Studies (formerly Mass Communications) recently redesigned their form.

c. *Access to data*: The data from teaching evaluation forms is normally kept confidential. One of the exceptions to this is the Haas School of Business which has an online database of quantitative data accessible to students and faculty in the School. Reasons for preserving the confidentiality of the evaluation data include compliance with the Family Educational Rights and Privacy Act (FERPA) and the California Public Records Act (CPRA). The overall absence of access to teaching evaluation data has the unfortunate effect of alienating and frustrating students on campus. In response, the Graduate Assembly passed a resolution requesting that the campus release data from teaching evaluations. The ASUC partnered with a commercial online evaluation system, pickaprof.com. Pickaprof's system also lists the grading distributions of each class on campus supplied, under court order, by the Registrar's Office. Recently Pickaprof has rebranded itself as MyEdu and expanded its services. For a fee in certain cases, it will now provide evaluation information about courses, as it always has, but also information about degree requirements and a selection of course planning tools.

In the College of Engineering, the honor society Tau Beta Pi has an online evaluation system,[1] and in Electrical Engineering and Computer Sciences (EECS) the honor society Eta Kappa Nu has posted the departmental teaching effectiveness and course worth data online.[2] However, professors and GSIs in EECS may choose to opt out of this system, thereby keeping their evaluation data confidential. The Tau Beta Pi system uses specially developed evaluation forms, not departmentally administered questionnaires.

Our review of the current practices and policies for evaluating teaching reveal that the collection of data is inconsistent, haphazard, and time-consuming. Further, data needed for decisions about merit and promotion are not always forwarded in a format that can be easily used by the BIR. We also note that end-of-term evaluations are not the best way to improve teaching; often data come too late to instructors, and the information may not necessarily be what a faculty member needs to modify aspects of his or her teaching. Finally, students lack access to evaluation data that they could use to make informed choices about classes.

[1] http://tbp.berkeley.edu/chapter/coursesurveys/
[2] http://hkn.eecs.berkeley.edu/student/CourseSurvey/

IV. Requirements for an Ideal Teaching Evaluation System

Although it is arguable if any teaching evaluation system can effectively accommodate the requirements of all its stakeholders, it is useful to benchmark the ideal teaching evaluation system for the Berkeley campus. In what follows, we outline such a system without regard to the costs or capabilities of existing systems. Our discussion of such a system builds on the Policy for the Evaluation of Teaching for Advancement and Promotion. This policy, which is reproduced as Attachment D to this report, was developed by the COT in April 1987 and adapted thereafter by the campus. The procedures to conduct teaching evaluations are discussed in this policy and are not repeated here.

The ideal teaching evaluation system would be user friendly—students would find it easy to fill out, and faculty (and other designated parties) would find it easy to access the information. It would be adaptable to the wide variety of course permutations on campus (e.g., courses taught by three instructors, such as Bio 1A; GSI-led sections; labs). It would mesh easily with other campus systems. It would be cost-effective: it would require less staff time than is now used to tabulate and transcribe data from evaluations. It would be robust: at the end of each semester, large amounts of data would have to be processed within a short period.

Just as citations and peer reviews are parts of a researcher's portfolio, we view the teaching evaluations as a component of an instructor's teaching portfolio, and much of our discussion is framed in this context.

The ideal teaching evaluation system would cater to all the needs of the parties vested in the system:

The Instructor. The ideal system would provide a means of evaluating instructor teaching efforts and encouraging course improvement. It would provide a confidential venue for students to voice their opinions, praises, and complaints on a course and its instruction, and to hear informed comments from other students on the course.

Traditionally, a primary function of teaching evaluation for the instructor is course improvement; however, the data collected at the end of the semester can only be used to improve subsequent offerings of the course. Further, as most of the written comments are no longer transcribed, convenient access to this resource is not available. A better source of useful information for current course improvement is a mid-semester evaluation process (in either online or paper format), so a teaching evaluation system should ideally include both a mid-semester and end-of-term component.

Department Chair, Dean, BIR, and Vice Provost for Academic Affairs and Faculty Welfare. The ideal system would provide an efficient and economic means for evaluating an instructor's teaching portfolio. Such an evaluation would include comparisons with other instructors within a department, analysis of efforts made to improve courses and pedagogy, and ready access to course materials and teaching evaluation data. A set of questions common across the campus would allow for the possibility of more accurate and helpful comparisons between and among departments. This would be of great benefit not only to the BIR, but to the campus in general, allowing us to make more concrete statements about teaching across the campus than we have been able to make to date.

Students. The ideal system would be standardized so that students are not answering randomly different questions on different forms. The system would provide fewer questions. It would be clear how the student responses would be used (too frequently, students comment that they have no idea what use is made of end-of-term evaluations). It would allow students to respond to other questions that would be made available to their peers to help them select courses.

Only students who attended classes during the course could submit mid-semester and end-of-term evaluations. While a 100% rate-of-return on evaluations is unrealistic, an expectation of 70% or higher should be normative. The evaluations would be collected in a manner which ensures that the confidentiality of the students' comments is safeguarded, and the integrity of the evaluation process is maintained. In analyzing student comments, it would be possible to correlate their qualitative comments with their numerical evaluation of the course, number of hours per week they worked on the course, and their expected grade for the course. For students considering taking the course in the future, it would be of interest to have a separate set of questions on workload, course difficulty, and course worth. Because use of a single number or a pair of numbers to evaluate the teaching of an instructor can easily be misleading, any evaluation system should include ample opportunity for the students to submit written comments on the course and the instructor.

A truly comprehensive, ideal evaluation system would not be limited to end-of-term and mid-semester student evaluations. It would also include:

- Periodic peer evaluation of the instructor and course (e.g., randomly scheduled classroom visits), to be included in the instructor's teaching portfolio;
- Comments from Cal alumni who took the course in earlier years. This type of evaluation is used when considering nominees for the campus' distinguished teaching award and in promotion cases;
- Instructors’ responses to concerns raised about their teaching in the alumni, peer, and student evaluations;
- Syllabus, electronic course notes, examinations, grade distribution data, and grading policy for all of the instructor’s courses;
- An analysis of the role played by the instructor's courses in the curriculum, and how these quantitative evaluations compare to similar courses taught across department, college, and campus. Per Provost King's memo from 1987 (see Attachment D), these data at a department level are already required for merit and promotion cases. Having campus-wide benchmarks for these data could be helpful and informative for all instructors.

The campus and its departments would be required to periodically examine their teaching evaluation systems. First, the questionnaires distributed to the students would be updated to ensure that they are in compliance with the latest research in the field of teaching evaluation. Second, information on advances in the systems used to assemble, compile, and distribute the evaluation data would be made available to departments.

V. Summary of Methods for Gathering Evaluation Data

The Task Force surveyed several existing methods for gathering and processing teaching evaluation data. Five of these methods are currently in use by departments on campus. Listed below are brief advantages and disadvantages of the nine principal methods that we found.

A. Paper
   Pros: familiarity, allows for detailed qualitative comments, good control of response rate because evaluations are distributed during class
   Cons: labor intensive to transcribe comments and to calculate mean and standard deviations, error-prone, possibility of selective use of comments

B. Scantron[3]
   Pros: good control of response rate because done during class; can get mean and standard deviations
   Cons: requires available Scantron machine; still must transcribe qualitative comments

C. Class Climate[4]
   Pros: can be handled either in class or online (i.e., hybrid); can get mean and standard deviations
   Cons: if done in class, qualitative comments are only scanned; complex system which is difficult for staff to learn and modify for a department's needs. Limited integration with other campus systems.

D. Other commercial online systems
   Pros: often "out of the box"; most companies are eager to tailor to the needs of a campus.
   Cons: no assurance that a given company will remain in business; security of data and student identifiers; costs may be prohibitive.

E. Individually developed online systems[5]
   Pros: can be tailored to the specific needs of the department
   Cons: expensive to develop individual systems; no common way to gather/upload data for use by the administration. Investments made across campus with no oversight and increased risk for failure.

F. Sakai system
   Pros: integrate with the campus course sites in bSpace and other campus databases; use of student ID assures that the proper person is filling out the evaluations. Integrated instructional support services would help with staff and faculty training.

---

[3] The Haas School of Business processes 18,500 evaluations each year using a Scantron system (https://acadadmin.haas.berkeley.edu/TIES/). Written comments are only made available to instructors, and GSI evaluations are kept confidential. The College of Engineering also uses Scantron forms for evaluations.

[4] The College of Chemistry recently abandoned this system in favor of the home-grown Our Unit.

[5] The School of Education has developed its own system which has been in use for several years.

Cons: central campus investment in time and money; care would be required to design the system to avoid system overload at peak evaluation and exam times.

G. The evaluation component of Our Unit[6]
Pros: already in use in some departments, home-grown, hybrid, adaptable, inexpensive.
Cons: not based on a widely used platform; data security issues. Who will oversee it and maintain it, given budget cuts, is unclear. Roadmap and future on campus also unclear.

H. For-profit online systems for students (e.g. Ratemyprofessors and MyEdu)
Pros: Ease of use, can rate at any time
Cons: no control over who evaluates a class (no way to know whether it is really a student in a class or not); no control over irrelevant or libelous evaluations; reluctance on the part of faculty to rebut student comments; no control over whether evaluations are current; response rate low and most likely bi-modal.

I. CourseRank online system for students[7]
Pros: has many features that students can use; can be linked to SIDs; can be made to interface with campus systems
Cons: Unclear as yet how much commentary or how many questions students can answer; response rate lower—students do not have to use the system. System is currently in flux, moving from Stanford-incubated to private.

VI. Recommendations of the Task Force

A. At the Campus-wide Level:

Goal: Informing academic personnel decisions regarding promotion and tenure by collecting information required by the BIR and the Vice Provost for Academic Affairs and Faculty Welfare.

The end-of-term teaching evaluation is intended in part to satisfy the APM's section on "Criteria for Appointment, Promotion, and Appraisal," particularly section APM-210(d), which covers the criteria that are intended to serve as guidelines for minimum standards in judging the candidate on teaching (section (1)). This section of the APM is reproduced in Attachment C.

Recommendation 1: End-of-term evaluations should meet the following criteria:

- Uniformity: a common subset of data for all instructors (see Recommendation 2 for specifics);

[6] This system is currently being used by Chemistry, ESPM, PMB, the School of Information Studies, Psychology, and Integrative Biology.
[7] CourseRank is an evaluation and course planning system developed by and in use at Stanford University: http://courserank.com

- Transparency: forms should provide valid and reliable information on instructor effectiveness;

- Flexibility: departments should be able to tailor their forms according to their own needs and practices, as long as they also include the campus-wide information from Recommendation 2, below;

- Brevity: beyond including the common set of questions, we recommend that departments examine the questions they currently ask students with the goal of including only those questions that are reliable and tested. (We note that some evaluation forms have laundry lists of questions). With shorter forms, not only will the information be more concise, but students will not suffer from "evaluation burn-out" and the department staff will have a reduced processing workload.

An end-of-term form would begin with a set of campus-wide questions, followed by those the school, college, and/or department wishes to add: hence the requirements for both uniformity and flexibility. The Task Force sees the uniform set of questions as being important beyond their use by the BIR. The common set of data for all instructors is absolutely essential if the campus is to present anything like a coherent picture of the quality of teaching on campus. The single question now in use is fine, but only begins, in the most general way, to provide a snapshot of teaching on campus. In addition, the campus-wide summaries from a set of such questions would provide OED, as well as other offices charged with supporting faculty, with excellent guidelines about what workshops and other support to offer in general to the campus.

We expect, too, that the process of filling out evaluations will be less onerous for students if the various forms bear more similarity to each other than they do now. With so many different kinds of forms and questions, students are more apt to see their education as a series of separate courses rather than as a unified whole. In addition, we hope that they might give more informed answers to the questions if they are answering the same questions for every course.

Recommendation 2: The following information should be gathered in the end-of term evaluation process for every course in every department:

- Numerical student evaluation for instructor's overall teaching effectiveness, as per memorandum from Vice Provost de Vries on 11/19/02 (see Attachment E);

- Numerical student evaluation of how much the student learned in the course;

- Self-reported level, major, and GPA of the student completing the form;

- Student written comments: Currently, copies of the original evaluation forms are sent forward with academic personnel cases only when enrollment is fewer than ten students. In the new system, we would wish to be able to pull up student written comments from all courses, according to such criteria as the rating the student awarded, the student's major or level, and so forth.

In terms of response rates, we encourage departments and instructors to achieve a minimum response rate of 70%.

- Grade distributions for the course and the average grade distributions for comparable courses;

- Departmental averages and standard deviations for overall teaching effectiveness.

Current policy, articulated in the de Vries memorandum of 11/19/02, is that the department should provide departmental averages (later specified as mean and median) for all courses taught in the same category in the same semester. Former Vice Provost de Vries gives "upper division lecture course" as an example. Other helpful categories would be: required vs. elective course, courses of similar enrollment, scores for the same course if different instructors have taught it, and so forth. If the database is electronic, it would be helpful if reviewers could query it in the largest possible number of ways. More generally, increasing the number of different averages will increase the usefulness of the data.

Recommendation 3: Information other than student ratings should be provided for evaluating teaching for merit and promotion cases.

It is important that the campus continue to nurture a culture where good teaching is encouraged, fostered, and rewarded. Although the current end-of-term teaching evaluation forms used by many departments provide a measure of the effectiveness of an instructor, they by no means present a complete picture. Other means of evaluation should be encouraged and documented, such as:

- A self-statement that is included in every case for merit and promotion and that specifically addresses teaching. It is helpful if faculty provide a written statement of their teaching approach, including the learning goals of specific courses and their choices of teaching strategies. They may also comment about their efforts to improve their own instruction, as well as to improve the instruction of the GSIs they supervise, and to respond to criticisms of their teaching performance made by the department chair or by students on prior end-of-term evaluations. Faculty may wish to comment on the courses, indicating which are new or team-taught. They may also want to describe their role in directing senior theses, masters and doctoral studies, and postdoctoral scholars. The number of graduate students successfully completing degree work could also be included. Although these categories are currently included in the Academic Bio-bibliography, we mention them as a reminder of their importance.

- A statement from the department chair that comments on the teaching effectiveness of the faculty member. The chair's statement should contain information and observations beyond what is found in the end-of-term evaluation forms.

- Syllabi for the courses, which show the content, grading philosophy, and assignments.

- For mid-career appraisals, faculty observation of a faculty member's teaching, which provides a measure of peer review of teaching effectiveness.

Much of the above information is outlined in the University of California Academic Personnel Manual, section ii, Appointment and Promotion, d. Criteria for appointment, promotion, and appraisal; and in the Berkeley Campus' 1987 "Policy for the Evaluation of Teaching for Advancement and Promotion." These two documents are reproduced in Attachments C and D.

B. At the department level:

Goals: Helping departments evaluate the overall teaching effectiveness of their faculty, both on a regular basis and during formal program reviews, and assisting departments to evaluate program-level student learning goals, while also alleviating administrative burden of a manual and distributed system.

Recommendation 4: The campus should develop a database of good end-of-term evaluation questions that can be made available to all departments; departments should be encouraged to revise their forms, and the campus should provide departments with consultation and support in revising their current forms.

Many campus forms suffer from being overly-comprehensive, and research has shown that a small number of precisely aimed questions do a better job than a laundry list of many questions. That is, more questions do not necessarily produce more added value. In addition, there are many two-pronged questions that cannot be answered easily, e.g., "has a grading system that is clearly defined and equitable." Obviously, a grading system might be clearly defined, but not equitable.

With a slightly larger common set of campus-wide questions (see Recommendation 2, above) and a shorter list of good questions, most campus evaluations would be much more useful, as well as less onerous for students to fill out.

In Attachment F, we provide some examples of questions that could be asked in an end-of-term form. We also present a brief sample questionnaire that could be used by departments. Of course, departments might want to add questions—judiciously—that address certain specific areas. For instance, seminar courses may want more specific questions about class participation while studio courses may want to focus on development of skills. Catalogs of valid and reliable questions have already been developed at the University of Michigan and other campuses, which UC Berkeley could use.

C. At the individual instructor level:

Goal: Helping instructors (including Senate and non-Senate faculty, and GSIs) to improve their teaching.

Recommendation 5: The Administration, Senate, and departments should strongly and consistently urge faculty to use mid-semester evaluations and other methods of early feedback in their classes.

The most widely used method for evaluating teaching is the end-of-term evaluation completed by students. These evaluations arrive too late, however, to benefit the current students doing the evaluation. Nor do the evaluations typically encourage students to give the specific comments an instructor might need either to spot weaknesses in classroom organization, pacing, and work load or to identify how well students have understood the material.

Much more effective are early feedback activities that take place during the term. Early feedback activities can elicit the specific comments and constructive criticism that faculty need to improve students' understanding of the material and their subsequent performance on exams and projects. In addition, considerable research shows that gathering feedback throughout the term allows an instructor to improve even very challenging classes (see References).

Attachment G describes some strategies that faculty can use to get feedback on their teaching. In addition, faculty can use online surveys, through bSpace or as a stand-alone, to solicit input from students during the semester.

We note that the term "mid-semester" evaluation is not technically applicable in all cases, such as those classes in which a sequence of three instructors teach a class. In such cases, however, each instructor can certainly solicit the same information found on a mid-semester form.

We are not recommending that mid-semester evaluations be mandated. However, we note that efforts by faculty members to improve their classes based on feedback from mid-semester evaluations are easy to document in, and can become a regular part of, a faculty member's self-review for merit and promotion cases.

D. At the student level:

Goal: Helping students to choose classes by providing them with peers' evaluations of faculty.

Recommendation 6: The campus should fully support efforts to pilot a separate system that will provide course evaluation information to students.

Although students fill out teaching evaluations at the end of each semester, these results are retained for use in faculty personnel cases and are not normally shared with students. This means that a natural motivation for filling out an evaluation form, namely to aid other students, is absent.

A separate system such as CourseRank not only allows students to provide their peers with evaluations of courses, but also provides students with a host of other tools for planning their schedules. Unlike for-profit systems, e.g., MyEdu.com, this system can be linked to SIDs, thereby assuring that only registered students are providing evaluations.

Such a system would not include the questions needed for instructor merit and promotion cases. Rather, it would feature questions tailored to the students' interests, such as work load, grading, quality of instruction, and appropriateness of the prerequisites (see Attachment F). Consistent with its purpose of helping students choose courses and plan their schedules, it could be linked to the Telebears system.

Because the data for such a system would not be used in instructional personnel cases, rate-of-return issues would not be so critical. In the long run, the student system should interface with a potential campus-wide online evaluation system, so that students could fill out a single evaluation form, and the results would be distributed appropriately to the various stakeholders.

During the course of writing this report, the TFTE became informed that a group from Students Systems at Berkeley, led by David Scronce, was evaluating both the CourseRank and MyEdu systems.

E. Moving toward an online or hybrid evaluation system

With the advent of an online personnel system in the Office of the Vice Provost for Academic Affairs and Faculty Welfare (currently known as the Berkeley Achievement and Review System, or BARS), we will have the opportunity to gather in one place comprehensive teaching and research dossiers for every faculty member. In order for this to be successful, departments will need to be able to provide data from teaching evaluations in an electronic format, something that only a few departments can currently do.

Recommendation 7: The campus should pursue the development of an online or hybrid evaluation system. To accomplish this, the TFTE recommends appointing a small working group with technical and pedagogical expertise to evaluate available hybrid or online systems. Such an evaluation would include addressing policy, legal, and procedural issues.

Clearly, an online teaching evaluation system —or a hybrid one (i.e., which uses paper evaluations that can be scanned, and the data recorded and tabulated)— offers many advantages, because it would allow importation of evaluation data into the dossiers of faculty members. However, given the wide range of forms now in use, the various methods departments employ to administer and analyze evaluation data, and the lack of infrastructure to manage centralized data collection and analysis, it does not seem prudent to mandate an online system without substantial investments of time, effort, and money to make sure it will be successful. Attachment H describes issues in the implementation of an online teaching evaluation system. In our experience, the presentation of such a system can look excellent and the system easy to use, but there can be large architecture and interface issues that are beyond the expertise of the TFTE, not to mention hidden and upfront costs. Although the TFTE was able to assess the major pros and cons for the various options that are currently available, it did not have the necessary technological expertise to make specific selection recommendations, for either a single centralized system, or a variety of possible systems that could be utilized by individual departments. However,we strongly recommend that a small working group that has the necessary technological and pedagogical expertise should be charged with this specific task.

It should be noted that other universities, such as the University of Michigan,[8] who have recently introduced online teaching evaluation systems, have pre-existing campus-wide offices dedicated to developing databases of questions, distributing forms, and collecting and analyzing data. These evaluation offices are usually distinct from an office like OED, which provides direct consultation and other services to faculty. Thus, the issue of how a selected centralized online system will be operated and maintained at Berkeley is a necessary one to include in the overall assessment.

Until the campus decides to invest in a centralized system for everyone, the only thing that can be done now is to stay decentralized, and allow departments/schools to continue to choose their own system. Any system will be satisfactory, as long as the department can upload the appropriate end-of-term evaluation data into the electronic personnel system.   The findings from the TFTE suggest a path forward that would aim to streamline and digitize the workflow for departmental staff and the Office of Academic Affairs and Faculty Welfare, while continuing to support the distribution of the day-to-day oversight within the departments. This approach would involve an end-of-term evaluation template to be integrated into a bSpace site, which could be exported in a standard format for migration into a departmental database, or re-collated and uploaded to the BARS for faculty personnel cases. This approach is not without costs (see Attachment H), but would minimize the initial investment, bring some relief to overburdened departmental staff, and increase the ability for the campus administration to manage the data into the future. This assumes that the BARS will act as the campus-wide data store for this information.

F. <u>General recommendation for facilitating improvments in evaluation forms, assessment of teaching, and teaching effectiveness</u>:

Recommendation 8: The campus should provide departments with information and consultation to facilitate the recommended changes in evaluation.

The Office of Educational Development (OED) can consult with departments about changes in the forms for end-of-term evalutions. The Campus should also sponsor regular workshops on methods to improve teaching effectiveness and to perform mid-semester teaching evaluations. A section of the OED website should be devoted to teaching evaluation practices and systems on campus. In addition, a campuswide forum, open both to staff who are involved in evaluations in their departments and to all interested faculty, could provide:

- Opportunities for staff members to discuss existing teaching evaluation systems used on campus;

[8] http://www.umich.edu/~eande/ (Office of Evaluations and Examinations at the University of Michigan).

- An occasion for the Vice Provost for Academic Affairs and Faculty Welfare and BIR to comment on the evaluation of teaching in merit and promotion cases;

- A venue for information dissemination on the best practices for mid-semester and end-of-term teaching evaluations;

- Improved communication between various groups on campus, which should also have the effect of improving the efficiency and cost-effectiveness of existing teaching evaluation systems.

VII. Next Steps

In the short term:

1. The TFTE recommendations should be presented to the Academic Senate for their comments and approval. We anticipate that the Academic Senate Chair will communicate the TFTE report to BIR and other relevant Senate Committees. Once the TFTE recommendations have been approved, they should be disseminated to departments.

2. The Berkeley Division of the Academic Senate and the relevant administrative officers should provide Deans, Chairs, and all faculty with information (perhaps via CalMessages) on the value of midsemester evaluations, a sample form, and instructions on how to use them. This can be followed by an announcement on the teach-net maillist and presentations at departmental meetings by OED staff. Much of this information is currently available at the OED website, http://teaching.berkeley.edu.

In the intermediate term:

3. OED should proceed to create a database of good evaluation questions and make them available to departments, along with help in revising and streamlining departmental questionnaires. The OED website, http://teaching.berkeley.edu, should develop a section devoted to these questions. The Berkeley Division of the Academic Senate should announce this information/service and encourage departments to take advantage of it.

4. A small working group with the appropriate technical expertise should be convened to examine and recommend online and hybrid systems for end-of-term evaluations. The primary function would be to determine which system(s) would be cost-effective, user-friendly, and robust, and would work with other campus systems. The group should also consider any legal issues surrounding end-of-term evaluations, and the pedagogical value of any system. The group should be composed of representatives from ETS, IST, and OED, with consultation from campus legal authorities, who can clarify which types of evaluation data can be made accessible to the campus community while remaining compliant with FERPA and the California Public Records Act.

Acknowledgements

The TFTE wishes to express their gratitude to the following individuals for their comments and opinions on the teaching evaluation systems currently in use on the Berkeley campus: Donna Craig and Shareena Sampson from the Department of Mechanical Engineering, Denise Harris from the College of Chemistry, Paget Harris from the Departments of Comparative Literature and French, Devin Jones from Boalt Law School, Professor Lisa Kala from the School of Education, Rachel Richardson from the Haas School of Business, Nancy Schimmelman from the Department of Plant and Microbial Biology, Tony Christopher from the Office of Student Affairs, and Johanna Metzgar from the Office of the Registrar.

References

Cohen, P. "Effectiveness of Student Rating Feedback for Improving College Instruction: A Meta-Analysis of Findings." *Research in Higher Education*, 1980, 13, 321-341.

Davis, B.G. *Tools for Teaching*, Second Edition. San Francisco: Jossey-Bass, 2009.

Fabry, V. J., Eisenbach, R., Curry, R. R., and Golich, V. L. "Thank You for Asking: Classroom Assessment Techniques and Students' Perceptions of Learning." *Journal on Excellence in College Teaching*, 1997, 8(1), 3-21.

Hamilton, D. M., Pritchard, R. E., Welsh, C. N., Potter, G. C., and Saccucci, M. S. "The Effects of Using In-Class Focus Groups on Student Course Evaluations." *Journal of Education for Business*, 2002, 77(6), 329-333.

Hampton, S. E., and Reiser, R. A. "Effects of a Theory-Based Feedback and Consultation Process on Instruction and Learning in College Classrooms." *Research in Higher Education*, 2004, 45(5), 497-527.

Kreutzer, C. S. "Midterm Evaluation of Teaching Provides Helpful Feedback to Instructors." *Teaching of Psychology*, 1993, 20(4), 238-240.

L'Hommedieu, R., Menges, R., and Brinko, K. "Methodological Explanations for the Modest Effects of Feedback." *Journal of Educational Psychology*, 1990, 82(2), 232-241.

Marincovich, M. “Using Student Feedback to Improve Teaching.” In P. Seldin and Associates (Eds.), *Changing Practices in Evaluating Teaching*. Bolton, MA: Anker, 1999.

Overall, J. U., and Marsh, H. W. "Midterm Feedback From Students: Its Relationship to Instructional Improvement and Students' Cognitive And Affective Outcomes." *Journal of Educational Psychology*, 1979, 71(6), 856-865.

Rando, W. C., and Lenze, L. F. Learning from Students: Early Term Feedback in Higher Education. University Park, PA: *National Center on Postsecondary Teaching, Learning and Assessment Publications*, Pennsylvania State University, 1994.

Svinicki, M. D. "Encouraging Your Students to Give Feedback." *New Directions for Teaching and Learning*, 2001, no. 87, 17-24

Attachment A

Members of the Task Force on Teaching Evaluation

Christina Maslach (Co-Chair)
Professor of Psychology
Vice Provost for Teaching and Learning

Oliver M. O'Reilly (Co-Chair)
Professor of Mechanical Engineering
Former Chair of the Senate Committee on Teaching

Paul Albertus
Graduate Student

Barbara Gross Davis
Assistant Vice Chancellor for Equity and Inclusion

Mara Hancock
Director, Educational Technology Services

Kushal Kadakia
Undergraduate Student

Tsu-Jae King Liu
Professor of Electrical Engineering and Computer Sciences
Member of the Graduate Council

John Lindow
Professor of Scandinavian
Member of the Budget and Interdepartmental Relations Committee

Steve Tollefson
Lecturer for College Writing
Director, Office of Educational Development

Mark Wilson
Professor of Education
Member of Committee on Educational Policy

Sheldon Zedeck
Professor of Psychology
Vice Provost for Academic Affairs and Faculty Welfare

Attachment B

Proposal to DIVCO from COT and BIR on the Formation of a Task Force, followed by the Charge Letter

January 21, 2008

To: William Drummond, Chair, Academic Senate, 2007-08

From: Carla Hesse, Chair, Budget and Interdepartmental Relations Committee, 2007-08
Kevis Goodman and Oliver M. O'Reilly, Co-Chairs, Committee on Teaching, 2007-08

Re: Evaluation of Teaching

Need for Review of Teaching Evaluation Processes and Procedures

The last comprehensive examination of the teaching evaluation process on the Berkeley campus took place over 20 years ago. One outcome of that review was the "Policy for Evaluation of Teaching (for Advancement and Promotion)," which was prepared by the Committee on Teaching, and first implemented in the Fall 1987. Since then, the only significant change in the evaluation process has been the inclusion of a seven-point scale for overall teaching effectiveness; this requirement became campus policy in Fall 2002.

One year later (August 2003), the Chair of the Committee on Teaching, Deborah Nolan, together with Christina Maslach, Vice Provost of Undergraduate Education, felt the need for a new examination of teaching evaluation practices on campus, more wide-ranging than the simple insertion of the seven-point scale. In a written request to the Chair of the Academic Senate, Professor Nolan and Vice Provost Maslach requested "the creation by DIVCO of a Senate working committee to undertake just an examination." The request was not granted at the time. We believe that the need still persists and that it is now time to renew aspects of the appeal.

The existing evaluation process is far from ideal:

- While reviewing nominations for the Distinguished Teaching Award, the COT has noted that, among the many different evaluation forms used across campus, some are poorly designed, others are unnecessarily complicated or cumbersome, while others barely leave room for student commentary or explanation of the numbers circled.

- Both the COT and the BIR have been struck by the frequently poor rate of return in a number of courses, regularily as low as 15% of enrolled students; these circumstances make fair, informed, and representative assessment impossible.

- For many departments, information from completed evaluations is not available to students. Partly as a result, many students have turned to commercial sites (Ratemyprofessor.com., Pickaprof.com), whose results are selective at best. The ASUC itself has now partnered with Pickaprof.com, and at least two student groups (Eta Kappa Nu and Tau Beta Pi) have constructed their own websites.[9]

• Other questions have emerged in both inter- and intra-departmental discussions as well, including the authority that student evaluations should (or should not) command in personnel procedures; the possible, supplementary role of "peer review;" the appropriate time(s) during the semester to have students evaluate a course; and the best conditions for their doing so.

• The last question touches on the issue of on-line evaluation, also raised by the 2003 request to DIVCO. There are a host of questions and concerns surrounding this issue (e.g.: Will the response rate be high enough? Will the process be secure? What will happen to the data? Should the site be open for posting anytime?), and, especially given the variety of distinctive disciplines taught on campus, their suitability for individual departments on the Berkeley campus remains an open issue.[10] Nonetheless, given the problems listed above, and also the high cost that some departments sustain while tallying results—especially in labor time for already burdened staff employees—we feel that it is time for a reexaminination of the pros and the cons of on-line evaluation, combined with a more thorough survey of the considerable range of systems available and already in use at peer institutions across the country.

Given all of the above, we hereby request the creation by DIVCO of a joint Academic Senate-Administration task force to undertake an examination of how teaching is evaluated on the Berkeley campus and to make recommendations, if any, as to how the measures involved might be improved.

Charge to the Committee

The general charge to the committee would be to review the procedures by which teaching is assessed on campus and to recommend changes, if any, as to how the evaluation and promotion of teaching could be improved. More specifically, the committee's purview would include:

• Teaching Evaluation Forms -- Review current evaluation forms on campus that students use to evaluate not only their instructors but also their own performance (how much they have learned, how hard they have worked, etc.) in their courses. Examine the content and organization of the different forms, the way in which the results are disseminated and used (e.g., advancement and promotion) both inside the home departments and outside. Identify best practices and make recommendations for change, whether of forms, policy, or both, as necessary.

• Teaching Evaluation and Promotion Policies -- Review current policies on the evaluation of teaching, the collection of data, and the uses of information collected. Identify best practices and make recommendations for change as necessary.

[9] See http://hkn.eecs.berkeley.edu/student/CourseSurvey/ and http://tbp.berkeley.edu/chapter/coursesurveys/ respectively.

[10] We do have some results from three pilot projects conducted at U. C. Berkeley on a volunteer basis during the spring and fall terms of 2005; these are available at http://hrweb.berkeley.edu/ldp/ldp.htm. It would be up to the committee we are proposing, as part of its charge, to interpret those results and to decide if they warrant action, or whether further trials are now necessary.

- Peer and External Review -- Explore whether "peer review" can and should be part of the evaluation of teaching during merit and other promotion considerations. Examine what is being done in this regard across departments on campus and at comparable universities across the country (e.g., classroom observation, creation of teaching portfolios, research on the scholarship of teaching). Make recommendations as necessary.

- Online Evaluation -- Investigate the pros and cons of gathering student evaluation data electronically. Should there be some type of centrally-supported online system that departments could choose to use? Or should it be left to those individual departments seeking an online environment in the near future to choose their own vendors or design their own systems (i.e., the current practice)?

This last aspect of the charge touches upon an important issue. We wish to be clear that we are not advocating either a single campus form or uniform procedures for the collection of responses. Uniformity is neither possible nor desirable in a University committed to instruction in multiple and diverse disciplines.

Composition of Committee

A committee that would convene panels of relevant experts related to the specific aspects of the charge is optimal. We recommend that the committee include representatives from the Budget & Interdepartmental Relations (BIR) Committee, the Committee on Teaching (COT), the Committee on Educational Policy (CEP), and the Graduate Council; campus faculty with expertise in methodology and evaluation; representatives from the Division of Undergraduate Education (Barbara Gross Davis and Steve Tollefson); and at least one undergraduate and one graduate student.

Staff support for the committee will be provided by the OED.

February 21, 2008

CHRISTINA MASLACH
Vice Provost – Undergraduate Education

SHELDON ZEDECK
Vice Provost – Academic Affairs and Faculty Welfare

Subject: Formation of a joint Senate administration task force on teaching evaluation

Dear Christina and Shelly,

As you know, at its February 21, 2008 meeting, Divisional Council (DIVCO) authorized the formation of a joint Senate-administration task force on teaching evaluation. I write to invite your respective offices to participate in this important effort. Specifically, we would appreciate representation from your offices on the task force. In addition, we would like to request that the Office of Educational Development staff the task force. Attached is the revised proposal for your consideration. As you can see from the proposed timeline for the task force to report back to DIVCO, we hope to move forward expeditiously. We would appreciate receiving your response by Friday, February 29, 2008.

Sincerely,

William Drummond
Chair, Berkeley Division of the Academic Senate

Cc: Carla Hesse, Chair, Committee on Budget and Interdepartmental Relations
Kevis Goodman, Co-chair, Committee on Teaching
Oliver O'Reilly, Co-chair, Committee on Teaching
Donald Mastronarde, Chair, Committee on Committees
Aimee Larsen, Manager, Committee on Budget and Interdepartmental Relations
Steve Tollefson, Director – Office of Educational Development staffing the Committee on Teaching
Anita Ross, Senate Analyst, Committee on Committees

Attachment C

Academic Personnel Manual Section on Evaluating Teaching

Section APM-210-d:

> "In judging the effectiveness of a candidate's teaching, the committee should consider such points as the following: the candidate's command of the subject; continuous growth in the subject field; ability to organize material and to present it with force and logic; capacity to awaken in students an awareness of the relationship of the subject to other fields of knowledge; fostering of student independence and capability to reason; spirit and enthusiasm which vitalize the candidate's learning and teaching; ability to arouse curiosity in beginning students, to encourage high standards, and to stimulate advanced students to creative work; personal attributes as they affect teaching and students; extent and skill of the candidate's participation in the general guidance, mentoring, and advising of students; effectiveness in creating an academic environment that is open and encouraging to all students, including development of particularly effective strategies for the educational advancement of students in various underrepresented groups."
>
> …
>
> "It is the responsibility of the department chair to submit meaningful statements, accompanied by evidence, of the candidate's teaching effectiveness at lower-division, upper-division, and graduate levels of instruction. More than one kind of evidence shall accompany each review file. Among significant types of evidence of teaching effectiveness are the following: (a) opinions of other faculty members knowledgeable in the candidate's field, particularly if based on class visitations, on attendance at public lectures or lectures before professional societies given by the candidate, or on the performance of students in courses taught by the candidate that are prerequisite to those of the informant; (b) opinions of students; (c) opinions of graduates who have achieved notable professional success since leaving the University;(d) number and caliber of students guided in research by the candidate and of those attracted to the campus by the candidate's repute as a teacher; and (e) development of new and effective techniques of instruction, including techniques that meet the needs of students from groups that are underrepresented in the field of instruction.
>
> All cases for advancement and promotion normally will include: (a) evaluations and comments solicited from students for most, if not all, courses taught since the candidate's last review; (b) a quarter-by-quarter or semester-by-semester enumeration of the number and types of courses and tutorials taught since the candidate's last review; (c) their level; (d) their enrollments; (e) the percentage of students represented by student course evaluations for each course; (f) brief explanations for abnormal course loads; (g) identification of any new courses taught or of old courses when there was substantial reorganization of approach or content; (h) notice of any awards or formal mentions for distinguished teaching; (i) when the faculty member under review wishes, a self evaluation of his or her teaching; and (j)

evaluation by other faculty members of teaching effectiveness. When any of the information specified in this paragraph is not provided, the department chair will include an explanation for that omission in the candidate's dossier. If such information is not included with the letter of recommendation and its absence is not adequately accounted for, it is the review committee chair's responsibility to request it through the Chancellor."

Attachment D

Policy for the Evaluation of Teaching (For Advancement and Promotion)
http://apo.chance.berkeley.edu/evaluation.html

Berkeley: OFFICE OF THE CHANCELLOR

September 18, 1987

DEANS, PROFESSIONAL SCHOOLS AND COLLEGES:

Re: Policy for Evaluation of Teaching (for Advancement and Promotion)

Last year the Committee on Budget and Interdepartmental Relations asked the Senate Committee on Teaching to formulate a comprehensive policy with guidelines for how the assessment of teaching should be presented in recommendations for faculty promotion or advancement. The Committee on Teaching conducted a review, which included among other things invoking the aid of faculty who have received the Distinguished Teaching Awards.

Their statement, "Policy for the Evaluation of Teaching (for Advancement and Promotion)" and its attachment, "Recommendations for Administering and Analyzing Student Course Evaluations," are enclosed. Please distribute these documents to your respective teaching units. The Policy has the full endorsement of the Committee on Budget and Interdepartmental Relations and unit heads are urged to follow it as closely as possible.

C. Judson King
Provost, Professional Schools and Colleges

Attachment

POLICY FOR THE EVALUATION OF TEACHING
(FOR ADVANCEMENT AND PROMOTION)

Committee on Teaching
April, 1987

I. Criteria for Advancement Promotion

The University community believes that excellence in teaching and excellence in research go hand in hand, and as a matter of policy teaching and research are both essential duties of every faculty member. Promotion depends upon the demonstration of excellence in both areas.

The essential question in the evaluation of teaching is whether the candidate contributes in an effective, creative, and appropriate way to the teaching mission of the department. Attention should be paid to the varieties of demands placed on instructors and the range of teaching activities called for in various disciplines and at various levels. It is imperative that clear documentation of ability in teaching be included in all advancement and promotion cases. Incomplete advancement or promotion cases will be returned to the originating department. While no two cases will be alike, there are several recurring themes which may be addressed in the preparation of the teaching component and several useful techniques for verifying performance in these areas.

II. Aspects of Teaching to be Evaluated

Teaching plays a major role in decisions regarding advancement and promotion. Therefore, an explicit discussion of the teaching performance of a faculty member is essential. The following components of teaching may be evaluated in a personnel review decision.

Design and redesign of courses. Does the course "work"? Are the course objectives reasonable? Are course requirements clearly stated and communicated to students? Is the course continuously updated to reflect recent developments in the field?

Presentation of material. Does the teacher convey enthusiasm for the subject matter? Does the teacher present material with logic and force, arousing curiosity in beginning students and stimulating advanced students to creative work?

Command of the subject. Is the instructor knowledgeable in the subject matter of the courses he or she teaches? Does the instructor engage in reading or research in the subject matter of the course in order to keep up to date with current research developments?

Contributions to curriculum and instruction. Has the teacher developed instructional materials, such as textbooks, videotapes, computer courseware, slides, publications related to teaching, or the like?

In what ways has the teacher participated in departmental or campus curriculum design or development efforts?

Directing student research. How active is the teacher in guiding the research projects of graduate and undergraduate students?

Advising. What formal advising duties or informal advising has the teacher undertaken? How much time does the teacher spend consulting with students?

Guiding and supervising Graduate Student Instructors. To what extent has the teacher prepared, trained, and supervised graduate student instructors?

III. Sources and Methods for Evaluating Teaching

Each department or teaching unit can be expected to have a unique culture which supports and encourages teaching excellence and which will have its own traditions of teaching evaluation that serve to define and reinforce that culture. Departments should work to improve their evaluation methods and seek to make them supportive and encouraging rather than investigative or punitive. The following is a list of sources of information for evaluating teaching; departments will choose those most appropriate for the particular case.

The candidate's faculty colleagues who have appropriate expertise in the discipline are best able to evaluate the scholarship that informs the design and organization of courses and curriculum, the choice or development of texts and other instructional materials (syllabus, handouts, etc.), the nature of examinations and assignments, and so on.

Current students can comment on an instructor's ability to communicate clearly, the extent and level of the instructor's course preparation, whether the instructor makes effective use of class time, how sensitive and responsive the instructor is to difficulties students may be having in the course, the workload, and so on. Students should not be used to judge the adequacy, relevance, and timeliness of the course content nor the breadth of the instructor's knowledge and scholarship.

Former students can comment on the long-term effectiveness of the candidate's teaching: for example, the role of the instructor's courses in preparing the student for advanced study or work in the discipline.

If the candidate teaches with Graduate Student Instructors, these students can often comment on the teachers role and effectiveness in the classroom and in preparing, training, supervising and evaluating GSIs.

Self-evaluations can be both descriptive and evaluative and may address such issues as teaching goals and philosophy, involvement in curriculum projects, efforts to improve teaching, and so on.

Various methods can be used to gather data from these sources: rating forms or detailed questionnaires, written appraisals (letters or responses to open-ended questions on rating forms), interviews, observations, and so on. Combining sources and methods, it is possible to collect a variety of information about a faculty member's teaching. For example, colleagues can evaluate instructional materials or observe an instructor's classroom teaching. Students can complete evaluation forms at the end of a course, participate in individual or group interviews, or fill out surveys when they graduate.

IV. The Teaching Dossier

A teaching dossier should include the following items:

Departmental letter summarizing the candidate's teaching. A carefully prepared Chair's letter is an essential part of a dossier. An effective letter from the chair will describe departmental teaching evaluation procedures, the nature and quality of a candidate's teaching, and the evidence upon which this evaluation is based.

Departmental ad hoc committee report for mid-career review, promotion to tenure, and promotion to professor. For these types of reviews, departments are encouraged to convene an ad hoc committee (two or more faculty members) to examine evaluation data and assess a candidate's teaching performance. The ad hoc committee's report should be included in the dossier.

Candidate's statement. It is helpful if candidates provide a written statement of their teaching approach, including the goals of specific courses and choices of teaching strategies. They may also comment about their efforts to improve instruction and respond to criticisms of their teaching performance made by the department chair and by students on end-of-course evaluations.

Description of courses taught. A list of courses by course number and enrollment should be included. The candidate may wish to comment on the courses indicating which are new, team taught, and so on.

Description of student research directed. Candidates may want to describe their role in directing senior theses, masters and doctoral studies, and postdoctoral scholars. The number of graduate students successfully completing degree work could also be included (indicate the date each student began graduate school).

Peer evaluation. Reports or letters about the candidate's teaching performance from faculty colleagues familiar with the content could be included in the dossier. The letters should cite the basis and evidence for judgments made (observation, review of instructional materials, and so on).

Student evaluation. Some form of student evaluation data (e.g., end-of-semester student ratings) for each different course taught in the period under review should be presented. The data should include both summaries of student evaluations of teaching and sufficient "raw" data (i.e., representative student comments) to allow the reviewers to see the candidate's teaching from the students' point of view. In addition, the dossier can include letters from current students or summaries of interviews.

Alumni evaluation. Former students, as well as Graduate Student Instructors (if the candidate teaches with GSIs), can provide information about the candidate's teaching performance. Dossiers may include letters, results of group interviews, or summaries of surveys of alumni that specifically address the candidate's teaching.

V. Resources for Evaluating Teaching

The Office of Educational Development, under the auspices of the Committee on Teaching, has prepared a handbook that suggests a range of useful methods and procedures for evaluating teaching.

Attachment E

DATE: **11/19/02**

TO: DEANS AND DIRECTORS

FROM: **Jan de Vries**
Vice Provost for Academic Affairs and Faculty Welfare

SUBJECT: Policy on Teaching Evaluations

Over the past year, committees of the Academic Senate have urged that a uniform question, with a uniform scale of quantitative evaluation, be included in all departmental forms used for the evaluation of teaching on the Berkeley campus. In March of this year I solicited opinions from all deans and department chairs on this question. The responses were made available to the Academic Senate and its committees, and they have recently confirmed their support for the campus-wide administration of a uniform question, and the uniform reporting of the quantitative responses to this question.

Consequently, upon the recommendation of the Committee on Budget and Interdepartmental Relations, with the support of the Committee on Teaching and the endorsement of the Divisional Council of the Academic Senate, all departments are to include in their teaching evaluation forms the following question:

Considering both the limitations and possibilities of the subject matter and course, how would you rate the overall teaching effectiveness of this instructor?

1	2	3	4	5	6	7
not at all			moderately effective			extremely effective

This question, scored using a seven-point scale, should be included in all teaching evaluation forms as soon as practically possible, but no later than the fall semester of 2003.

In conformity with the desires of the Academic Senate and its committees, departments should also calculate and report the mean and the median score for the question for each course taught by an individual faculty member in all merit and promotion cases. Finally, the department is to supply departmental averages for all courses taught in the same category in the same semester. For example, if Professor X receives a mean score of 4.5 and a median of 4.6 in an upper division lecture course, the department should also supply the mean and median scores for all upper division lecture courses offered by the department in the same semester.

In implementing this policy, please bear in mind the following:

1. If you do not currently use this question, its introduction does not require that you in any way change the existing questions or scoring conventions used by your department. You may wish to make additional changes, but this question is envisioned as in addition to rather than in the place of the rest of your evaluation form.

2. The measurement of teaching performance provided by this question is intended to supplement other information on teaching performance. It is not intended to replace other types of information, let alone become the sole indicator used in merit and promotion evaluations. Consequently, it is important that your evaluation forms ask additional questions to the uniform question cited above and/or solicit narrative responses from students. Chairs' letters and other departmental documents should summarize these other sources of information, and not dwell exclusively on the descriptive statistics of the uniform question in their characterizations of a faculty member's teaching.

3. Student evaluations of teaching are a necessary source of information on teaching performance, but they should not be regarded as a sufficient source. Departments are encouraged to solicit other types of information, including class observation, comments by GSIs and co-instructors, course materials, and self-evaluations supplied by the faculty member. The obligation to make use of the uniform question does not negate the importance of drawing on a variety of sources in evaluating teaching effectiveness.

4. Finally, several technical matters: For courses with fewer than ten students, the original evaluation forms should be submitted for merit/promotion review; there is no need to calculate summary statistics. For all courses, the number of evaluation forms submitted should be reported alongside the course enrollment. Large discrepancies between the two numbers should be explained. In many cases, the median and mean scores provide an adequate description of the underlying raw data, but when the scores do not approximate a normal distribution, a frequency distribution should be provided to assist in the interpretation of the data. Departments are free to define the categories of their courses (upper division, lower division, lectures, seminars, requirements, electives, etc.), but should be consistent in the categorization they select. Departments that currently score this question using a 5 point scale should put the new scores and averages in perspective or develop a methodology to translate the old averages to the new scale. Whichever approach is chosen should be fully described in the review file.

Jan de Vries
Vice Provost for Academic Affairs and Faculty Welfare

Attachment F

Examples of Questions for Evaluating Teaching and Sample End-of-Semester Evaluation Form

We present here topics of interest to different stakeholders and within each topic, examples of the kinds of questions that could be posed. Most likely only one question would be asked under each topic.

Students

1. Overall rating of instructor and/or course

    a. " Overall, this was an excellent course"
    b. " Overall, the instructor did an excellent job teaching this course"
    c. "This instructor is one of the best UC Berkeley teachers I have known"

2. Workload

    a. "Relative to other courses with the same number of units, this course required a lot of work"
    b. "On average, how many hours per week have you spent on this course, including attending classes, doing readings, reviewing notes, completing assignments and any other course related work?
    c. "How would you rate the difficulty of this course?" (from easy to very difficult)

3. Grade

    a. "What grade to you expect to earn in this course (or did earn in this course)?"

4. Learning

    a. "To what extent has this course developed your intellectual skills and interests?"
    b. "To what extent has the instructor inspired you to learn (or stimulated your interest in the course subject)"
    c. "How much did the following aspects of the course help you in learning?" (Examples might be course lectures, labs, readings, assignments, tests, etc.)
    d. "As a result of your work in this class, how much did you learn about the following?" (Instructors might list concepts and/or skills most important to the course)
    e. "What helped your learning the most in this class?" "What hindered your learning the most in this class?"

5. Instructor skills

    a. "The instructor was well-prepared for class"
    b. "The instructor explained clearly"
    c. "The instructor responded respectfully to students' questions and viewpoints"
    d. The instructor showed a genuine interest in students and their learning"

e. “This course was well-organized”
f. “Course grading procedures were fair”
g. “Evaluations of students’ work (e.g., exams, graded assignments and activities) were good measures of what students learned in the course”

6. Open-ended questions

   a. “Please identify what you perceive to be the greatest strengths and weaknesses of this instructor’s teaching/ of this course”
   b. “What are the one or two tips you would give to students who enroll in this course next semester?”
   c. “How would you summarize this course for a fellow student?”

Students are also interested in other information that can be obtained from other sources, such as the course grade distribution, the size of the course (enrollment), the percentage of students responding to the questions.

<u>BIR Committee</u>

1. Numerical overall ratings

   a. “Considering both the limitations and possibilities of the subject matter and course, how would you rate the overall effectiveness of this instructor?”
   b. “Considering both the limitations and possibilities of the subject matter and course, how would you rate the overall effectiveness of this course?”

2. Student demographic information

   a. year in school
   b. major
   c. self-reported GPA

3. Open-ended questions

   a. “Please identify what you perceive to be the greatest strengths and weaknesses of this instructor’s teaching/ of this course”

The BIR would also like the departmental averages and standard deviations for the overall ratings for like courses, required versus elective courses, courses of similar enrollment, scores for the same course if different instructors have taught it; the grade distribution for the course and the average grade distribution for comparable courses; and the response rate. -

<u>Departments/Deans/Individual Faculty</u>

1. learning (“This course helped me develop intellectual skills, such as critical thinking, analytical reasoning, and integration of knowledge”)
2. organization (“The instructor seemed well prepared for each class”)
3. rapport (“The instructor showed a genuine interest in students and their learning”)
4. workload (“The amount of work required was appropriate for the unit credit received”)
5. grading (“Grades were assigned fairly and impartially”)
6. group interaction (“Group activities in this class contributed to my learning”)
7. student level of effort (“Overall, to what extent did you attend class, prepare for each class, actively participate in each class, seek help at office hours”)

Sample End-of-Semester Evaluation Form

The instructor:	Strongly Disagree		Neither Agree or Disagree			Strongly Agree		
1. was well-prepared for class.	1	2	3	4	5	6	7	NA
2. explained clearly.	1	2	3	4	5	6	7	NA
3. responded respectfully to students' questions and viewpoints.	1	2	3	4	5	6	7	NA
4. was well-organized.	1	2	3	4	5	6	7	NA
5. assigned grades fairly and impartially.	1	2	3	4	5	6	7	NA
6. helped me develop intellectual skills such as critical thinking, analytical reasoning, and integration of knowledge.	1	2	3	4	5	6	7	NA
7. showed a genuine interest in students' learning the material.	1	2	3	4	5	6	7	NA

Overall, to what extent:	Not at All		Somewhat			Very Much So	
8. did you learn new knowledge, skills, concepts, or ways of thinking?	1	2	3	4	5	6	7
9. did you attend class, do the readings and assignments, spend time studying, actively participate in class?	1	2	3	4	5	6	7

10. Considering both the limitations and possibilities of the subject matter and course, how would you rate the overall effectiveness of this <u>instructor</u>?

1	2	3	4	5	6	7
not at all effective			somewhat effective			very effective

11. Considering both the limitations and possibilities of the subject matter and course, how would you rate the overall effectiveness of this <u>course</u>?

1	2	3	4	5	6	7
not at all effective			somewhat effective			very effective

12. Year in school (circle one): 1st 2nd 3rd 4th 5th graduate

13. Major: ____

14. Grade Point Average at UC Berkeley: _______________

15. Please identify what you perceive to be the greatest strengths and weaknesses of this instructor's teaching of this course (use back of page if necessary):

Attachment G

Examples of Early Feedback Strategies for Instructors

(Adapted from Davis, 2009, pp. 461-471)

Distribute index cards during the first or last five or ten minutes of class, if your class size is less than 100.

Pass out three-by-five cards to students and ask them to respond anonymously to two questions, one on the front of the card, the other on the back. Explain that the purpose is for you to get feedback. You can pose general questions about what is going well in the course and what needs to be improved or changed. Other general questions include these: What do you want more of? Less of? How are you finding the course? Any suggestions for improving the course? Any problems? What do you need before the end of the term? How can I help you learn better? What is helping you learn in this course? What is making learning difficult? If you could change on thing about this course, what would it be?

Leave the room while students write their comments. Ask a student volunteer to collect the cards in a manila envelope and return them to your office.

Ask students to complete a brief questionnaire in class.

During the first or last few minutes of class, distribute a short questionnaire to students or to a random sample of students in a large enrollment class. Keep the number of items small—no more than ten or twelve. You can use relevant items from your campus's or department's end-of-term questionnaire, create your own questions, or select them from existing instruments.

The questions you select should be ones you can respond to during the term; otherwise your students may develop false expectations about the remainder of the course. Include questions that help students focus on their own learning (How do you know when you have learned the material?), or probe what students themselves could do in class (What is the one thing you want the other students to do to improve this course? What do *you* need to do to improve your learning in this course?). You might ask about the level of difficulty of course content, the quality and quantity of assignments, the use of class time, things the instructor does that enhance or inhibit student learning, the nature of student preparation outside of class, or the pace of the class.

Ask students to write *to* you rather than about you. To assure that students respond candidly, leave the room while they complete the questionnaire anonymously, and ask a student volunteer to collect the forms in a manila envelope and return them to you or place in your department mailbox.

Ask students to complete a brief questionnaire online.

Some faculty give modest extra credit if students complete the online evaluation form.

Break students into small groups for Stop/Start/Continue.

Ask small groups to identify teaching activities and behaviors that they wish you would stop doing, start doing, and continue doing, and to supply a reason or explanation for each request. A variation is to focus on the class instead of the instructor: "What is the class doing that frustrates or blocks your learning?" What could we be doing as a class that would help you learn?" Have the groups report back to the class and share their lists.

Arrange for your students to be interviewed.

Invite a colleague or staff member from the Office of Educational Development to conduct an oral evaluation with your students during the last ten or fifteen minutes of class. The evaluator asks students to cluster into groups of five or six and to take several minutes to do the following:

- Select a spokesperson who will also write down the groups' comments.
- Name something in the course that they find helpful or worthwhile and that has helped their learning.
- Name something that has hindered their learning and that they would like to see changed; identify how the change would be implemented.
- Name something that students can do themselves to improve the course.

Convene focus groups.

Ask for volunteers to meet with you to talk about the course and improvements in organization, structure, assignments, and the like. Begin by posing an issue, concern, or dilemma and ask the group to help you understand it. This is more effective than asking, "What improvements are needed in this course?" Faculty who use this technique report animated discussions that contain helpful suggestions. Instructors who convene focus groups also tend to receive higher student ratings at the end of the term.

Name a set of "spotters" for each class session.

Spotters are students who arrive in class a few minutes early and report to you on troublesome aspects or challenges in the previous assignment, readings, or class session. Plan to address the spotters' issues immediately, toward the start of class. In large courses, the spotting system will also give you the opportunity to meet more students.

Establish a student liaison committee.

Ask two to four students to meet with you periodically outside of class to discuss difficulties or dissatisfactions with the course. Membership on the committee may be rotated from a list of volunteers, as long as the entire class always knows who the liaisons are and how and why to contact them. If you teach a course with a large number of discussion sections, invite each section to select a delegate. Encourage the liaison students to circulate and seek out information formally or informally from other students. Report back to the class about your meetings with the liaison committee.

Form a student management team.

After the first three or four weeks of class, ask for student volunteers to serve as the management or resource team for the course. The team meets regularly to identify problem areas and propose changes, and you visit their meetings to hear their suggestions. Faculty who use this strategy have found it effective for improving course quality and are enthusiastic about the results.

Encourage students to form study groups.

Invite representatives of the study groups to meet with you to discuss any difficulties with the subject matter. Study groups provide students with opportunities to learn from one another, and some students may find it easier to seek assistance as a group rather than as individuals. While this technique seems workable in small classes, it can be particularly effective in large-enrollment courses, where students may feel less connection with their peers.

Attachment H
Implementing an Online Teaching Evaluation System

One of the aims of the TFTE was to investigate the possibility of introducing an online teaching evaluation system to the Berkeley campus. Although such a system offers potential cost savings and increased efficiency to date, as documented by the online resource onset.byu.edu, fewer than one hundred universities have implemented online teaching evaluation systems campus-wide. The most recent example is the University of Michigan.[11]

To be effective, an online teaching evaluation system to support end-of-term (EOT) evaluation needs a robust infrastructure. Currently the primary support structure for EOT evaluation comes from the departments. An online system that can cover the entire campus will require centralized resources to support the information technology (IT) systems, end-user support for faculty and departmental staff, and data analysis and reporting support. In order to minimize additional costs to the university, leveraging and augmenting existing organizational and technical structures wherever possible is recommended.

Regardless of the final solution, at minimum an online course evaluation system should achieve the following:

1. An overall decrease in the campus' expenditures on the evaluation process.
2. Efficient and less error-prone collection, storing, and reporting of data for use in campus administrative offices and systems (such as student systems, departments, and faculty achievement databases).
3. Quick and comprehensive feedback to instructors, departments, and administration on teaching effectiveness.

There are many systems, both vendor-based and open source, which can help the campus achieve these goals. Colleagues at institutions housing online course evaluation programs shared a common theme, namely that establishing an effective online course evaluation program faces various non-technical challenges and issues that need to be addressed:

1. Steps to insure that student response rates match current levels or above.
2. Carefully conceived hierarchical administrative structures based on departmental roles and distributed oversight.
3. A phased roll-out to address possible issues of faculty skepticism and resistance.
4. Additional incentives and reminders for students to complete the online end-of-term evaluation.

In particular, concerns about low response rates have been addressed by various strategies, some of which individual instructors can undertake, and some of which require campus wide changes. These strategies include (adapted from Davis, 2009):

- Making completion of the online evaluation form (or going to the web site and opting out) an ungraded class requirement.

[11] The website http://www.umich.edu/~eande/tq/faq.htm discusses the pros and cons of online teaching evaluation and the reasons for its implementation at the University of Michigan at Ann Arbor.

- Telling students why they are using an online form and making sure they know how to access it.
- Demonstrating in class how to complete the online form.
- Reminding students about the form during the evaluation period, and sending them an e-mail reminder a day or two before the survey period ends.
- Emphasizing the importance of students' opinions and the fact that the results will be used to improve instruction.
- Offering an incentive such as extra credit points for participation, with the proviso that 80–90 percent of the class has to complete the form for any student to receive extra credit.
- Using class time to have students complete the evaluation forms on their laptops or mobile learning devices.
- Giving students access to the results of selected items for choosing future courses.
- Giving students early access to their grades if they complete an online evaluation form (or opt out from completing the form at the web site).
- Entering students in a raffle with a meaningful prize (for example, bookstore coupons, early course registration).
- Providing a positive incentive such as contributions to a charity for each form completed.
- Explaining how the system preserves students' anonymity and reminding them that instructors will not see the results until after grades have been submitted.

Online teaching evaluation systems also face several technical challenges:

1. The online system should integrate with well-utilized front-end systems (such as bSpace) as well as with back-end systems such as the new Berkeley Achievement and Review System (BARS) for faculty personnel cases or other reporting systems. This often means a good deal of customization for both open source and vendor solutions.
2. The online system should provide reporting for a variety of campus roles, which are not currently managed through an identity management system.
3. Most online systems are focused on a centralized management model and require significant investment for ongoing oversight and analysis, integration, and customization.
4. General system performance challenges can occur with the large uptake of courses during a short period at the end of the term.

The findings from this Task Force suggest a path forward that would aim to streamline and digitize the workflow for departmental staff and the Office of Academic Affairs and Faculty Welfare, while continuing to support the distribution of the day to day oversight within the departments. This approach would support an end-of-term template to be integrated into a bSpace site which could be exported in a standard format for migration into a departmental database, or re-collated and uploaded to the forthcoming BARS system. This approach is not without costs. Estimates for the initial implementation of such a system range from $400,000-$500,000, with ongoing costs ranging from $100,000 to $300,000 annually depending on the support requirements. This estimate includes both program support in departments and systems support. Such a system would minimize the initial investment, bring some relief to overburdened departmental staff, and increase the ability for the campus administration to manage the data into the future. This assumes that the BARS system will act as the campus-wide data store for this information.