December 12, 2012

To: Christina Maslach, Chair, UC Berkeley Division of the Academic Senate

Re: **Online Evaluation of Courses Initiative**

I am pleased to send you a mid-term progress report for the Online Evaluation of Courses Initiative for distribution to key Academic Senate committees. The report details the results to date from the two-semester pilot conducted in academic year 2011-12. A more complete report, which will include the results of the Fall 2012 pilot, will be available in the spring.

I hope you will disseminate this report to key committees of the Academic Senate and share any feedback you may have. I would also be happy to meet with you to discuss any issues or recommendations related to this initiative to ensure that it meets the needs of the Berkeley faculty. We will also continue to consult closely with the representatives of the Academic Senate through the Online Evaluation of Courses Steering Committee, which will convene again in Spring 2013 to review the results of the Fall 2012 pilot.

With my warm regards,

Catherine P. Koshland
Vice Provost, Teaching, Learning,
Academic Planning & Facilities
Wood-Calvert Professor in Engineering
Professor, Environmental Health
Sciences, and Energy and Resources

Attachment

cc: Executive Vice Chancellor and Provost George Breslauer
        Assistant Vice Provost Cynthia Schrager
        ACIO, Academic Engagement & Interim Director Ben Hubbard
        Manager ETS Lisa Rothrauff

2011-12 Mid-Term Progress Report

Online Evaluation of Courses Initiative

The following report is a mid-term progress report of the multi-semester online evaluation of courses pilot. A fuller report, which will include results of the Fall 2012 pilot, will be available in Spring 2013.

Summary of Fall 2011 and Spring 2012 pilots

During the 2011-2012 academic year, the Online Evaluation of Courses (OEC) Initiative piloted its two chief deliverables: 1) a new questionnaire based on instructional format and, 2) a new online evaluation system that automatically distributes end-of-term evaluations. The goals of the pilots were as follows:

- Establish baseline response rate for participating courses using existing formats/protocol
- Assess response completeness
- Pilot test new questions and assess quality and usefulness of responses to faculty

And specific to the Spring 2012 pilot where we tested an online system:

- Adapt system to UC Berkeley system and environment
- Assess resources required for running and administering system

Summary results: The pilots met their stated goals. In so doing, they also established for the campus baseline average and median response rates for the current protocols (largely paper based system) and for the online system. Those response rates were:

	Fall 2011(Paper Pilot)	Spring 2012 (Online Pilot
Average responses	70%	64%
Median responses	81%	65%
Response Completeness	80%	91%

The average response rate for the paper-based evaluations was not as high as many had anticipated. Still, the Fall 2011 pilot established useful data on our existing protocols, and as a result, we now have a basis of comparison as we make the move to online distribution. The response rate that we achieved in Spring 2012 was lower than that with paper—64%, but it was higher than what most of our peers experienced in their initial move online, and more important, it comes close to meeting the 66% response rate required by the Academic Senate. In addition, this baseline online response rate was achieved without use of incentives. In terms of response completeness, the percentage of students completing all questions was significantly higher in the online pilot.

The Spring 2012 pilot also provided us with preliminary data on the quality of feedback garnered through online course evaluations, which we'll share below.

In this report, we state the goals and scope of each of the pilots as well as offer preliminary data on the quality and quantity of responses they achieved. We then explain challenges the project team faced in adapting online system and articulate plans for the Fall 2012 pilot.

Fall 2011 (Paper-based Pilot) Summary

The goal of this pilot was to test a newly developed questionnaire for courses delivered by paper; the questionnaires were administered by the departments following the same protocol as before. The questionnaire—of which there are 13 versions each based on an instructional format—were designed in 2010-2011 in collaboration with faculty.

Scope:

- 10 departments, 49 courses, 10 instructional formats (see addendum for list of participating departments and instructional formats)
- 3038 student enrollments

Goals:

- Establish baseline response rate for participating courses using existing formats/protocol
- Assess response completeness
- Pilot test new questions and assess quality and usefulness of responses to faculty
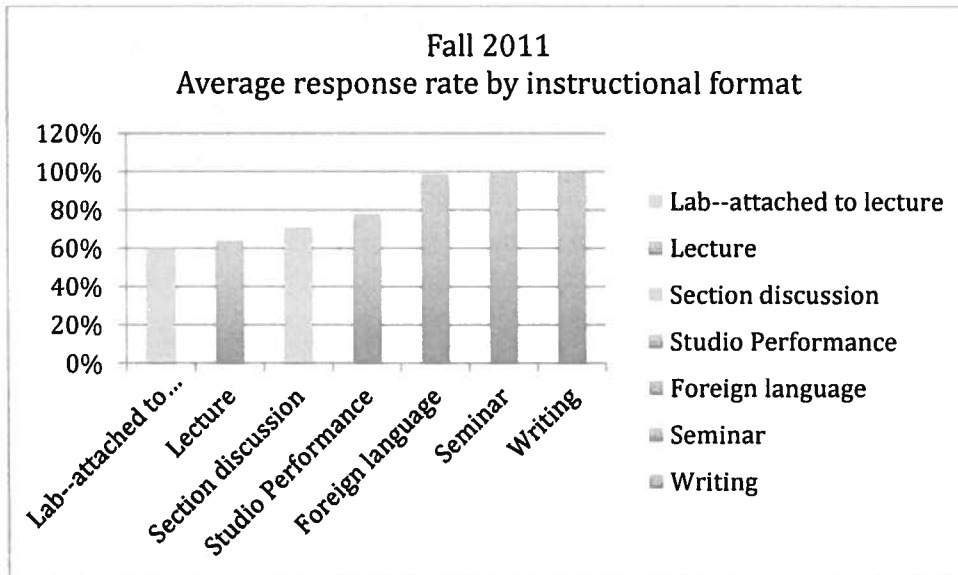
Results

Eight out of ten departments, and a total of 42 sections (primary and secondary combined) submitted results from this pilot. The project team compiled the results, calculating average and median response rates; average response rate by instructional formation; and degrees of completion. These calculations were driven by the goals initially set out by the team.

Response and completion rates

Response rate	Average response rate	70 %
	Median response rate (in range from 2% to 100%)[1]	81 %
	Response rate by instructional format	See chart below
Completion rate for primary sections	Responded to all questions	80 %
	Responded to close ended, quantitative questions only	8 %
	Responded to all close ended, quantitative questions and to some open-ended, qualitative questions	12 %

[1] The 2% comes from one class with just 2 evaluations out of 118 turned in. Staff from the participating department confirm that the evaluations were distributed but only 2 were returned (and the whereabouts of the others are not known). The next response rate on the median scale is 13%. In that case, the participating department confirmed that the lab sections for this course generated a very low response rate (and that this rate of return is typical for this course's lab sections). Finally, it is important to note that 11 sections (out of 42) achieved a response rate of 100%.

Fall 2011
Average response rate by instructional format

Legend:
- Lab--attached to lecture
- Lecture
- Section discussion
- Studio Performance
- Foreign language
- Seminar
- Writing

The campus does not have accurate central data on response rates under the current system. The 70% the average response rate for paper-based evaluation were not as high as anecdotal reports had led the project team to expect. That said, overall, the average response rate, does meet the Academic Senate's requirement of 67%. There are significant differences, however, by instructional format, with classes with smaller enrollments achieving higher response rates as expected.

As shown in the table above, the project team also sought to determine **completion rates** of the submitted evaluations. In our preliminary analysis, we counted three levels of completeness in primary sections: 100% complete (all closed- and open-ended questions completed); all closed-ended and some open-ended questions completed; only closed-ended questions completed. Our reason for measuring completion across the two pilots in this way was that many of our peers who achieve a 99% rate require only one response in order for the evaluation to be completed; yet in order to achieve not only high response rates but also high-quality responses, our campus seeks to elicit thorough and thoughtful responses from students.

Quality of the responses and the user experience in Fall 2011

In addition to establishing response and completion rates, the project team sought to measure the quality of the feedback elicited by the new questions and questionnaires. Overall, open-ended responses were pertinent and appropriate. There were no responses that could be read as disrespectful. Key repeated themes in open-ended responses include general appreciation of instructors; level of organization; and use and allocation of course materials. Below are some representative responses to two **open-ended questions directed at the course instructor**:

1. Please identify what you consider to be the strengths of the course.

"The course was interesting and provided a lot of knowledge. The structure was clear."

"Thorough and comprehensive."

"The organization, lectures/presentations were the course strengths."

"X. is an incredible instructor. His lectures are extremely engaging."

"I like how organized the class was. The lectures were uniform and easy to follow. I especially liked the illustrated outline to the class."

2. Please identify area(s) where you think the course could be improved.

"If the homework solution could be posted on bSpace, it would be more useful."

"Give homework every week or two instead of a big project at the end."

"More recapping; explain labs; provide more practice example problems"

"The text."

We also reviewed the quality of responses in the **student-to-student open-ended question**, used for the first time in this new set of questionnaires. In response to this question, which asked students to advise peers considering enrolling in this class, key repeated themes included time management; course planning and sequencing; and effective use of instructor and GSI office hours. Below are some representative responses:

"Wake up earlier. Do homework and get help from your GSI."

"Attend section and actually start the problem sets early; [they're] very long."

"My advice is to take your time on the homework and do your readings."

"Be prepared to engage with the readings. Writing assignments are not difficult but require critical thinking and organization of ideas and research."

"Study early and study hard; the assignments tend to pile up and if the readings are not studied early, the homework becomes painfully difficult. Attending lectures and discussion is also imperative."

"Take it because you love the material. It takes more time commitment than other 3-unit courses but is totally worth it."

Faculty Experience

The team set out to measure the value of the responses elicited from the new questionnaires. To that end, we sent out a survey to faculty in February 2012, asking whether the responses captured by the new questionnaires provided valuable feedback. Out of 40 instructors evaluated, only 5 responded: two found the new forms helpful, 1 did not, and two had not yet received the results of the evaluations from their departments due to the long processing time associated with paper-based evaluations.

Spring 2012 (Online Pilot) Summary

The overarching goal of this pilot was to implement an online course evaluation system. The implementation involved absorbing campus data, delivering evaluations, and processing evaluation results.

Another objective was to achieve an overall response rate that would match or better the response rate achieved through existing practices and to elicit feedback on teaching that would match or better that gained through current practices.

Scope:

- 8 departments, 22 courses, 8 instructional formats (see addendum for list of participating departments and instructional format types)
- 2394 projected enrollments

Goals:

- Establish baseline response rate for participating courses using online system *without the use of incentives*
- Assess response completeness
- Continue to pilot test new questions and assess quality and usefulness of responses to faculty
- Adapt system to UC Berkeley system and environment
- Assess resources required for running and administering system

Results

The project team had access to and reviewed results of 71 sections in this pilot. The project team compiled those results, again calculating average and median response rates; average response rate by instructional formation; and degrees of completion. These calculations were driven by the goals initially set out by the team. We also sought to measure quality of the responses and of the student experience with a new system, as explained below.

Response and completion rates

Response rate for primary sections	Average response rate	64 %
	Median response rate (in range from 25% to 92%)[2]	65 %
	Response rate by instructional format	See chart below.
Completion rate for primary sections	Responded to all questions	91 %
	Responded to close ended, quantitative questions only	5 %
	Responded to all close ended, quantitative questions and to some open-ended, qualitative questions	4 %
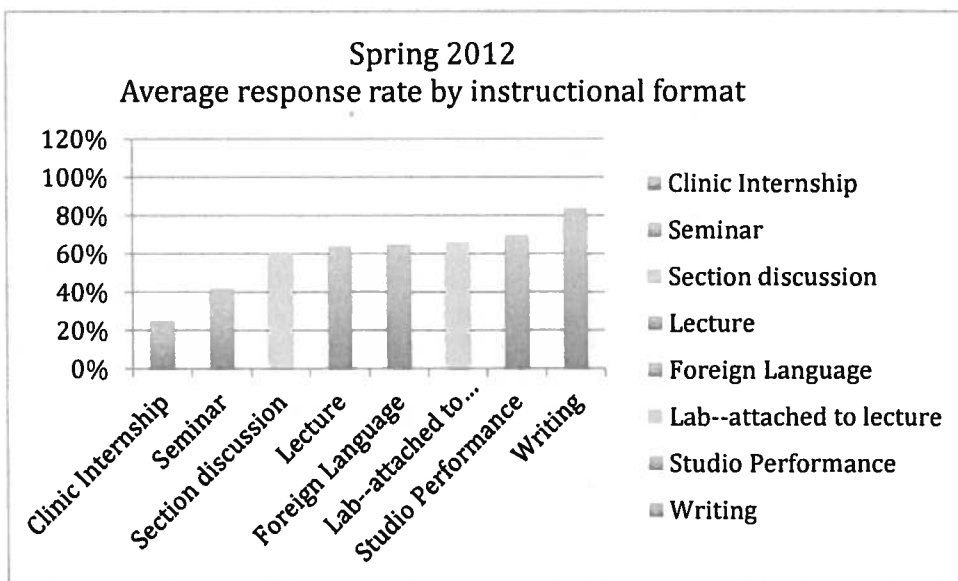
One of the key priorities for the Spring pilot was to determine what happens to **response rates** when an online system is used. Surprisingly, the response rate for the Spring pilot was higher than expected—64%. This result is better than what many of our peers experienced when they first moved online. More important, it is very close to achieving the Academic Senate policy minimum requirement of 67% response rate, as well.

It is important to note that no incentives, aside from an instructor's encouragement, were used to stimulate student responses. The project team at Berkeley agreed to first determine how and whether the encouragement of instructors was sufficient for getting students to complete their evaluations. Now, however, we need to see rates move higher. We plan to implement some tangible incentives for students in our Fall 2012 pilot.

---

[2] While there were 6 sections whose response rates fell below 50%, the vast majority of response rates in this pilot fell between 50-88%, with the top being 92%.

Related to completion rates, in Spring 2012, more courses saw a 100% completion rate than in the Fall (paper-based) pilot. Yet as noted in the section on completion rates for the Fall 2011 pilot, completion can be achieved by filling in each question with just one word (whereas many of our peers with response and completeness rates in the 90% or above count any evaluation with at least one question answered complete). With additional resource coming onto our project in August, we plan to further investigate the connection between completion and quality.

Quality of the responses and the user experience in Fall 2011

Research into peer institutions told us that when course evaluations are moved online, response rates typically drop (significantly), but the quality of the feedback goes up. To determine whether this was the case on our campus, the project team reviewed evaluation results and found the same level of maturity and relevance in the open-ended answers in both Fall 2011 and Spring 2012 evaluations.

As well, a review of the open-ended responses indicate that responses tended to be just as long if not longer than in the Fall 2012 pilot. Key repeated themes in open-ended responses included appreciation for instructors; level of organization; usefulness of assignments; lecture style; use and allocation of course materials. Below are representative responses to 2 questions:

1. Please identify what you consider to be the strengths of the course.

"X. was very prepared and always had good study guides and activities for us to do during section. It also felt like a safe and friendly environment to have discussions in, so we had some very insightful comments."

"I thought it was helpful to compare across readings at the end of each unit in order to help students think critically about the material."
"The small papers assigned in section were a helpful balance to the exam-only assessments in lecture."

"Extremely organized lectures. Review sessions were very helpful. Tests were challenging but very fair."

"The material was presented in a neat and consistent manner, and was given due attention during classes. In particular, the pacing of the course was excellent, something I find to be rare."

2. Please identify area(s) where you think the course could be improved
"Vary the structure of section a little more - sometimes small group work, sometimes have the class form a big circle and have one big discussion, etc."

"I thought it would have been more helpful to go over key concepts for each reading in discussion sections on a regular basis rather than having class be mostly small group work. Identifying key points from each reading to help guide the discussion was helpful, but I would have liked it if each section incorporated this more."

"The handouts were generally not very helpful, with the exception of the one for the final exam, reviewing all the semester's readings according to topic area."

"Topic sheets/study sheets outlining general lecture material/things to know for midterm/final."

In response to the question where students were asked to advise other students on the course, students tended to comment on these aspects: Time management; resources. Below are representative responses to this question:

"Stay on top of readings and try to do the writing assignments early. The readings are incredibly interesting- I only wish I had more time to fully engage myself."

"Make sure to read and attend lecture. In some classes that is not necessary but that is essential for success in this class."

"Setting up some method or schedule for regular studying, rather than cramming, makes life in this class much easier. At the very least, this will help with understanding the basic concepts behind any problem that they are likely to ask."

"Find as many midterms from years past as you can to see how the material will be presented on the tests, as it is often different from what you find in lectures and in the book."

Faculty experience: The project team sought to understand whether our faculty saw an improvement in the quality of the feedback elicited from the online questionnaires.

Survey: In June 2012, we surveyed instructors about the quality of the feedback received from the new questionnaires. This is a question we have pursued both semesters, but where we have had minimal response.

That said, in the June survey, we asked a series of questions about the value of the responses elicited from the new questionnaires. Of the instructors polled in June 2012, only two responded:

- One thought the immediate access to the online feedback helpful.
- The other asked that we reformat reports so that each student's set of responses could be read in its entirety so that he could have the context for the individual responses.

Interviews: Our team conducted four in-depth interviews with faculty who participated in the Spring online pilot. We asked them to comment on four main areas: 1) Response rate, 2) The questions in the new instrument, 3) The value of the feedback they received. 4) Their overall impression of taking evaluations online. Below are their most relevant comments:

Response rate

- Instructor has been teaching in the program for 21 years, and the response rate has been very consistent between online/offline evaluations.
- 67% seems high for online. Instructor's impression is that online completion rates can be as low as 20%. Response rate for this class in the past has ranged between 50% and 75%.
- If Instructor had done on paper, she could have ensured higher response by giving out when a homework is due; and normally does at beginning of class. So with online, this [71%] seems a little lower.

New instrument/questions

- Likes question: How satisfied were you with your effort in this course?
- Would have liked a more pointed question about instructor, especially about GSIs so we can determine whether to nominate them for GSI teaching award.

Value of the feedback

- Better spelling, but also slightly more coherent responses. Instructor thinks responses were largely similar in terms of quality, but perhaps slightly better due to students' not having to rush to finish writing at the end of a class.
- Instructor described the quality as "mixed." Some people wrote a lot, some were brief. Ones that are detailed seem to be more detailed than before, (i.e. with paper). Feels like the comments are longer, more descriptive. Old form had a box that was filled in by hand. Limited input.
- Many of the responses were similar.
- Did you have text-based responses before? Yes. And responses seem to be in the same range in terms of length, language register, and overall reactions/ratings.
- In past, we have liked ability to see comments in conjunction with quant. Questions for context. Can't do that here. Prose comments are sometimes illuminated by answers to other questions.

Pros and cons of going online

- Instructor thought it was nice to get the responses nearly instantly; there was no waiting for paper results to be collated. This is particularly relevant considering the department has been without an MSO for many months, and to her knowledge, the paper results from spring term are still sitting in a box waiting to be addressed.
- Instructor also thought it was more beneficial, this method of evaluating meets students' expectations to do more activities online. It's more efficient, and it does not take away from class time to conduct evaluations, which takes the pressure off the instructor to incentivize/reward students (e.g., returning assignments on evaluation day, etc.).
- Instructor is a "big fan" of online evaluations. Maybe get a higher rate with paper, but don't know that you'd get better answers, better quality. These seem more detailed. Also appreciates the time savings for staff. Not a good use of anyone's time.

Student experience: As the move online represented a new dimension of this initiative and a new process for students, the project team sought to understand what motivated the students to complete the evaluations, whether email notifications were helpful, and what we could improve about the online process.

Of the approximately 2300 students surveyed over the summer, 63 responded as follows:

- Students who remember completing the evaluation—63%
- Students who had completed the evaluations in their dorm room or apartment— 85%

Email notifications:

- 71% thought there were the right number
- 93% thought the messages were clear and understandable
- 74% thought the timing of the notifications were right
- 75% thought the notifications helpful

The project team plans to follow up with students who volunteered to give additional feedback in Fall 2012.

Additional information about the 2011-2012 pilots

Implementation and adaptation of EvalSys, our online evaluation system

With the Spring pilot, the project team implemented the Sakai EvalSys system. Configuring and administering the system to meet the pedagogical and administrative needs of the campus required significant resource from ETS and IS&T. One of the goals of the Spring pilot was to test whether we could adapt the system to the campus infrastructure. We faced two challenges in this work:

1) While the system worked well, its administration required a centralized model where the project staff created, distributed and reported out on all evaluations for individual departments.

2) The system also required manual input of data from ETS's Operations, Instructional Design and User Experience teams. This was due in part to the system itself, but mostly due to the data sources (e.g. instructional formats and course codes; teaching models; enrollment data) and business practices on campus (enrollment changes) that need to inform the evaluation system.

This centralized and manual work was manageable but presented challenges for scalability. In Spring 2012, we will investigate the system's ability to scale given the challenges we mention above.

Resources required for running and administering system

Another goal of this pilot was to make sure that the resources identified for the project were appropriate for its ongoing operation. We found that the resource budget scoped out was sufficient, but that the amount of work performed by team members was somewhat different from what we expected: we needed far more time from our UX designer because EvalSys required significant work on the user interface in order to be usable. We have since re-allocated resources to account for that additional UX time, while remaining within our budget.

The project team included the following staff, all funded by the project budget and listed here with the amount of their time dedicated to the project:

- User experience designer—80%
- Instructional designer—50%
- Operations and database manager—30%
- Developer—20%
- Technical project manager—30%

Plans for the Fall 2012 pilot

Scope: The fall pilot is a continuation of the OEC project, the purpose of which is to assess suitability of the Sakai online evaluation solution for long-term implementation and to determine the fit for meeting overall campus expectations.

The project team will accomplish this by administering the evaluations of one or two departments. Our criteria include: covering a range of instructional formats; co-taught courses;[3] courses with primary and secondary sections.

In taking on a department, we seek to better understand whether the online system can meet the pedagogical and administrative needs of an entire department. We can also start our work in defining the reporting needs for the system and for the EDW, which will ultimately produce the reports on teaching that the campus requires.

Specific Goals:

- Identify a scalable and sustainable evaluation model that supports campus wide adoption and provides a level of flexibility to support UCB growth/changes; specifically:
  - Investigate and determine what is required to enable the automation of key processes, i.e. course and evaluation creation
  - Assess system's ability to evaluate courses that are team taught in parallel and in sequence
  - Serve a larger user base (full departments or small college)
- Determine staff effort required to support manual processes and developer effort to automate key processes
- Test impact of incentives on response rate
- Continue to use and assess the quality of feedback received in new questionnaires.
- Provide faculty with feedback on their teaching that matches or exceeds the quality and amount of feedback received in the spring pilot
- Provide the campus community information and access to high-level summary of evaluation pilot results.

---

[3] To better understand current practice of course evaluation in co- or team-taught courses, we interviewed departments who according to the Office of Planning and Analysis offer co-taught courses regularly. The pattern on campus now is to provide one form per instructor, and therefore, students may complete multiple forms for one class. We seek to understand the impact of this practice (when evaluations are moved online) on response and completion rates and overall quality of the feedback.

Addendum

Participating departments/schools in Fall and/or Spring pilots:

- College of Electrical Engineering and Computer Science (EECS)
- College of Environmental Design
- College of Natural Resources
- Department of Chemistry
- Department of Linguistics
- Department of Music
- Department of Political Science
- Department of Spanish and Portuguese
- Haas School of Business
- School of Public Health
- Statistics

Instructional Formats tested (in questionnaires)

- Clinic with supervision
- Foreign language
- Lecture
- Lab_attached to lecture
- Lab_standalone
- Project based
- Section_discussion_GSI
- Seminar
- Studio_performance
- Writing
- Activity*
- Clinic without Supervision*
- Section-problem solving*

Instructional formats for which there exist questionnaires but that were not tested in these pilots