Pilot Study for Assessing the Viability of Using Online Course Evaluations at California State University
Sacramento

A Report from the Electronic Course Evaluation Task Force

Fall 2010
Sacramento State conducts a majority of their student teaching evaluations using paper and pencil. The practice of collecting teaching evaluations using paper and pencil has come under question. In particular, it is argued that collecting teaching evaluations in this manner is costly (Miller, 1987), demands large amounts of staff time to collect and process the forms (Kronholm, Wisher, Curnow, & Poker, 1999), delays feedback to the faculty (Layne, DeCristoforo, & McGinty, 1999), consumes large quantities of storage space (Donmeyer, Baum, Hanna, & Chapman, 2004), and is environmentally unfriendly (Anderson, McCain, & Bird, 2005). One alternative to conducting paper and pencil student teaching evaluations is to use a system in which evaluations are collected online.

In an effort to explore the implications and practical effects of shifting student evaluations from the traditional paper and pencil to an online format at Sacramento State, an eleven-person Electronic Course Evaluation Task Force (ECETF), representing the relevant university groups related to the collection of teaching evaluations, was formed (see Table 1 for a full list of ECETF members). This document identifies the relevant literature, reports on the findings of a pilot study, and concludes with recommendations regarding the use online evaluations at Sacramento State.

Table 1 - Electronic Course Evaluation Task Force Members

<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kimo Ah Yun</td>
<td>Director, Center for Teaching and Learning, Co-Chair ECETF</td>
</tr>
<tr>
<td>Jean-Pierre Bayard</td>
<td>AVP for Academic Affairs Technology Initiatives and Director, Academic Technology and Creative Services, Co-Chair ECETF</td>
</tr>
<tr>
<td>Wendy Cunningham</td>
<td>Faculty Policies Committee</td>
</tr>
<tr>
<td>Doug Jackson</td>
<td>AVP for Information Resources and Technology</td>
</tr>
<tr>
<td>Kathryn Kay (replaced in Summer 09 by Jesse Cuevas)</td>
<td>ASI VP for Academic Affairs</td>
</tr>
<tr>
<td>David Lang (replaced in Spring 10 by Brett Holland)</td>
<td>Curriculum Policies Committee</td>
</tr>
<tr>
<td>Raymond Pina</td>
<td>Manager, AIT, Academic Technology and Creative Services</td>
</tr>
<tr>
<td>Ann Stoltz</td>
<td>Chair, Nursing</td>
</tr>
<tr>
<td>Harry Theodorides</td>
<td>Faculty, Kinesiology and Health Science</td>
</tr>
<tr>
<td>Jennifer Underhill</td>
<td>Student, Department of Sociology</td>
</tr>
</tbody>
</table>
Literature Review

Benefits, Drawbacks, and Questions of Electronic Teacher Evaluations

Literature on shifting teaching evaluations from face-to-face to online delivery identifies the potential benefits and drawbacks of such an endeavor. Benefits include (a) timely student feedback to instructors, (b) reduced use of class time devoted to conducting evaluations, (c) increased time for students to complete evaluations, (d) greater number of written comments from students, (e) lower costs associated with the process and (f) increased opportunities for students to complete evaluations. In contrast, disadvantages include (a) lowered student response rates, (b) need for computer access to complete the evaluations, and (c) the risk that electronic evaluations will be less secure than paper evaluation forms (Anderson, McCain, & Bird, 2005; Lieberman, Bowers, and Moore, 2001).

Given that student response rates tend to decrease when teaching evaluations are shifted from a face-to-face process to an online one, implementation research has focused on indentifying which strategies maximize student response. For universities that value full student participation, they have created policies that require students to complete their teacher evaluations in order to enjoy continued access to their Learning Management System (LMS) or to register for classes. The strategy to block LMS access or class registration has been extremely successful at obtaining student compliance, but some of the collected data have been called into question as some students may simply “click buttons” to access the LMS or register for classes without paying attention to what they are doing. In contrast, other universities have taken a less rigid approach. For example, some universities simply make students aware that the teaching evaluations are online and students are required to log into a central site to
complete their evaluations. Overall, when employing strategies that are less rigid, research reveals that response rates drop by approximately twenty percent (Anderson, McCain, & Bird, 2005).

Most college and university implementation strategies fall between the most and least strict models, as described above. It has been found that, when students receive e-mail messages with a link to conduct their teaching evaluations, are reminded on several occasions by their teachers to complete the teaching evaluation, are informed by the faculty member that their teaching evaluation feedback is valuable to the teaching process, and are reminded by e-mail to complete the teaching evaluation, that the response rate nears those of paper and pencil rates. Further, when incentives, such as slight increases in their grade, or opportunities to win prizes in a raffle for their completion of teaching evaluations, are made available, the response rate tends to be slightly higher than rates of current paper and pencil formats (Donmeyer, Baum, and Hanna, 2002).

In light of the current fiscal constraints Sacramento State faces, if online teaching evaluations are more cost effective while providing equally good or better data from students, then such a move should receive careful consideration. With respect to costs for paper and pencil evaluations at Sacramento State, the current process requires the printing of thousands of pages of evaluation forms, purchasing of scantron forms and pencils, human hours to collect and send the data to Information Resources & Technology (IRT), and IRT resources to process, calculate, and mail teaching evaluation reports to each faculty member and department.

To date, one study attempted to quantify cost savings of shifting from paper and pencil to online evaluations (Kronholm, Wisher, Curnow, and Poker, 1999). In this study, it was estimated that paper evaluations were approximately 30 times more costly than online evaluations. Given the obvious potential cost saving of moving from paper and pencil to online teaching evaluations, Sacramento State could save significant dollars.
Based on a careful reading of the literature, taking into account information received from the Senate Executive Committee, the University Appointment, Retention, Tenure and Promotion (UARTP) Committee, Faculty Senate, and discussions within the ECEFT, the following list of pros, cons, and issues that were deemed particularly relevant at Sacramento State was composed (see Table 2).

**Table 2. The Pros and cons of online course evaluations in comparison to the traditional face-to-face course evaluations**

<table>
<thead>
<tr>
<th>Pros</th>
<th>Cons</th>
<th>Issues/Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Security of the submissions with students being authenticated in the campus' LMS (SacCT)</td>
<td>• Lower return rate.</td>
<td>• Getting students to submit their evaluations (return rate) outside class time</td>
</tr>
<tr>
<td>• Security of the data given that departments get a copy of the electronic data, with the original submission remaining on the server</td>
<td>• Students could complete the evaluations together, which may impact how they answer.</td>
<td>• How will the data be analyzed, and who will analyze the data?</td>
</tr>
<tr>
<td>• Complete anonymity - No possibility of instructors recognizing students' handwriting</td>
<td>• System technical reliability.</td>
<td>• The need for faculty buy-in</td>
</tr>
<tr>
<td>• Access to the evaluation data very quickly, as soon as course grades are submitted, and in time to impact next semester's course design</td>
<td></td>
<td>• Responding to student concerns for privacy and anonymity</td>
</tr>
<tr>
<td>• Cost reduction for departments and the University in the administration and processing of the evaluations</td>
<td></td>
<td>• The need to change the culture to support online student evaluation processes</td>
</tr>
<tr>
<td>• Save on storage space – No need to store evaluations</td>
<td></td>
<td>• Verifications of submissions?</td>
</tr>
<tr>
<td>• GREEN – Saves paper</td>
<td></td>
<td>• Having all students potentially submit an evaluation, including those who have not been attending class.</td>
</tr>
<tr>
<td>• Flexibility – allows for students to complete the evaluation at their leisure. Even if they miss class they can participate.</td>
<td></td>
<td>• Faculty choosing when to administer the survey.</td>
</tr>
</tbody>
</table>
This list was used to guide the design of the pilot study presented here in order to minimize any potential problems.

**Using Online Teaching Evaluations**

The use of online teaching evaluations is spreading in colleges and universities throughout the country. Cited as being at the forefront of collecting teaching evaluations, Drexel University shifted from paper and pencil to online course evaluations in the 1998-1999 academic year. Initial research on the impact of such a move revealed that the response rate decreased. Follow-up research suggests that colleges and universities need not sacrifice dramatic decreases in their response rates. For example, Columbia University in their initial test of implementing online teaching evaluations achieved response rates of 85%, which match most current universities paper and pencil response rates (Sax, Gilmartan, & Bryant, 2003).

Although the Drexel and Columbia University data provide a starting point to consider the impact of shifting teaching evaluations to an online format, their data do not provide sufficient comparison groups between face-to-face and online teaching evaluations. Two studies helped to illuminate the expected effect on completion rates of traditional versus online evaluations. In a study conducted at Georgia State University ($N = 2,453$) in which students were randomly assigned to either a traditional paper or pencil versus online evaluation condition, it was revealed that the response rates for face-to-face participants (60.6%) were higher than those in the online format (47.8%). However, most important, the average ratings for faculty members did not differ across the selected mediums (Layne, DeCristiforo, & McGinty, 1999). This study also revealed an unanticipated effect. Their data revealed that students in the online evaluation condition provided a substantially greater number of written comments about their educational experience than those in the paper and pencil condition.
Another typical study on the impact of shifting from a paper and pencil to an electronic format was conducted at California State University, Northridge (Donmeyer, Baum, & Hanna, 2002). In this study, 16 faculty members were assigned to have one of his/her sections evaluated in-class and the other evaluated online ($N = 696$). In the online evaluation condition, faculty members were also assigned to one of four treatment conditions, which included: (1) A very modest grade incentive (one-quarter of a percent), (2) demonstration by teacher of how to complete the online evaluation, (3) ability to receive early final grade feedback from the instructor, and (4) no information or incentive. Results revealed that the grade condition (87%) had the highest response rate, followed by demonstration (53%), feedback (51%), and the no information or incentive condition (28%). Overall, the results indicated a response rate difference from the paper and pencil format (75%). However, consistent with most research comparing teacher ratings between those conducted using paper and pencil versus an online format there were no differences in teacher rating scores amongst any of the online treatments groups compared to the paper and pencil control group.

While a change from a paper and pencil format to an online one at Sacramento State deserves careful consideration, general data trends suggest that a shift would be unlikely to yield detrimental effects on faculty members, or the university. Within the CSU system, there are a number of instances in which individual faculty members and departments have already shifted to online evaluations without major difficulty. In fact, San Diego State has already shifted all of their teaching evaluations to an online format.

On the Sacramento State campus, online evaluations are currently used. For example, the Division of Nursing shifted exclusively to the use of online teacher evaluations several years ago. Findings from this department reveal a drop in response rates, no differences with respect to individual teacher evaluation ratings, and an increase in student comments.
While Nursing is the only Sacramento State department to fully make use of online evaluations, there are numerous faculty members from other university units, such as the College of Education, College of Arts and Sciences, and The College of Business, who currently use some form of online evaluations. In these other units, the process continues to be utilized and no major issues have been noted. Although comparison data are not readily available for these units in their use of electronic evaluations, anecdotal data suggest that such a system is viable for use at Sacramento State.

**Pilot Study**

The ECETF pilot study was designed using feedback from Human Resources, Faculty Senate Executive Committee, Faculty Senate, UARTP Committee, CFA (state and local), and ECETF discussions. The design was such that:

1. No electronic evaluations used in this pilot study became part of any faculty member’s file.
2. Maximization of representation of faculty from as many colleges was sought.
3. Faculty participation was voluntary.
4. Faculty participation was disallowed if their RTP document precluded the use of non-paper and pencil evaluations, or mandated the inclusion of all course evaluations in their RTP file.
5. A sufficiently large enough sample was obtained to address issues of generalizability.

**Sample and Methodology**

Recruitment of faculty was solicited through an e-mail to department chairs asking for faculty volunteer participants. This initial request yielded approximated 40 faculty volunteer members. After applying the volunteer requirements some faculty volunteers were deemed ineligible. To complete the targeted 40 faculty participants, volunteer participant recruitment was undertaken by ECETF members to guarantee a representative cross section of participants. In the recruitment stage, a total of 30 classes
met the participation criteria and electronic evaluation data were collected on these 30 classes. Unfortunately, some of the data for the 30 classes could not be used. For example, in some cases the department did not keep the hard copies of the evaluations for comparison purposes. The final participation count consisted of 25 courses with online evaluation, and 25 courses with traditional evaluation. The 50 courses were paired by the course contents and taught by the same instructors and included about 1,600 student participants. Courses were distributed across the university as follows, Arts and Letters \((n = 12)\), Business \((n = 12)\), Education \((n = 4)\), Health and Human Services \((n = 6)\), Natural Sciences and Mathematics \((n = 4)\), Social Science and Interdisciplinary Studies \((n = 12)\). No data were collected from Engineering and Computer Science faculty as the College requires faculty to collect student teaching evaluations from all of their courses.

All comparative analyses utilized Paired Samples T-Test. According to the rule of thumb for mean comparison, the smallest group should not be less than \(1/3\)rd the size of the largest group. Therefore, all classes were divided into two groups: one was a large class group with 43 or more students in each class; another was a small class group with less than 43 students in each class. One pair of classes was excluded from the group analyses as their sizes were mismatched \((58 \text{ vs. } 25)\).

All incomplete evaluations (such as a student answering all but one of the questions) were included when calculating the mean score for each course. The incomplete evaluations were a very small portion of each teaching evaluation so they had little impact on the average score of individual courses.

**Method of Deployment**

This pilot study was built upon the experiences and practices of delivering electronic course evaluations at Sac State using the software, Flashlight, which is used for many of the online courses. Flashlight uses a Secure Sockets Layer (SSL), which is the current commercial standard to protect data by
encryption over the internet. Access to the survey databases was limited to an authorized person only by means of a username and password. The data center that physically contains the survey database is physically secured and data are backed up nightly: should the need arise to restore information. Student anonymity is maintained as the system does not associate the student identifiers with the raw data. The evaluation delivery process was as follows:

1. A list of faculty participants and selected courses was provided by the Task Force to ATCS.
2. Existing course evaluation formats were converted to an electronic format.
3. Respondent pools were created for each survey, including student e-mail addresses.
4. A template message (describing how the evaluation was to be delivered) was provided to each faculty member that could be sent to each student or posted in SacCT.
5. Course evaluations were sent by e-mail message to each student and the message included a unique url. The evaluations were available for a two-week period.
6. Midway through the evaluation period, an e-mail reminder was sent to each student who had not completed the evaluation.
7. The evaluation data were sent to the authorized department contact. Evaluation data were also provided to the OIR for the purposes of this report. No evaluation data were provided to faculty prior to them submitting their final grades.

**Results and Analyses**

<table>
<thead>
<tr>
<th>Class Size</th>
<th>Online</th>
<th>Traditional</th>
<th>Comparison</th>
<th>Significance (p&lt;.001)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Large</td>
<td>Small</td>
<td>Total</td>
<td>Large</td>
</tr>
<tr>
<td>Course selected</td>
<td>10</td>
<td>14</td>
<td>24</td>
<td>10</td>
</tr>
<tr>
<td>Mean Score</td>
<td>4.42</td>
<td>4.33</td>
<td>4.38</td>
<td>4.50</td>
</tr>
<tr>
<td>Response Rate</td>
<td>45.7%</td>
<td>55.4%</td>
<td>50.3%</td>
<td>62.1%</td>
</tr>
</tbody>
</table>

*Note: Comparison is based on Paired Samples T-Test. Higher values are highlighted in yellow.*

1. There was no significant difference between online evaluation and traditional evaluation among large classes in terms of response rates and mean scores.
2. The response rate for online evaluation compared to traditional evaluation for the small classes was statistically significant. The mean scores of both evaluations showed no significant difference among small classes.

3. Among the paired 25 courses, the mean scores of online evaluation were slightly lower than the mean score of traditional evaluation but the difference was not statistically significant. However, the response rate of online evaluation was significantly lower than the rate of traditional evaluation.

It is worth mentioning that the traditional evaluations were selected from different semesters, ranging from fall 1998 to summer 2009. However, the online evaluations were completed in spring 2010.
According to the results, the time difference did not impact the ratings that students gave to their instructors.

In conclusion, students’ ratings for each of the paired courses were consistent regardless of implementing online evaluation or traditional evaluation. However, the response rate of online evaluation was significantly lower than traditional evaluation for small classes as well as for overall comparison. In general, a low response rate means less representativeness of the evaluation and the margin of error has to be added when interpreting the mean score of each class.

Discussion

Results of this pilot study mirror findings from other universities. Shifting from paper and pencil to online evaluations reduced the overall percentage of students opting to complete the teacher evaluation, but the teacher rating scores were not statistically different between the two forms.

While these data confirm differences in mean scores that are not statistically significant, the actual numerical differences may be interpreted by RTP committees as indicative of a pattern in decreasing scores. Caution must be used by faculty and the in the event that faculty rating scores change slightly.

Finally, one might also ask whether the lowered response rate impacts the quality and quantity of the qualitative data through student written responses. Universities that have studied the effect of shifting the medium of collecting on teaching evaluations have found no negative impact on the comments. Although no statistical analysis of the qualitative data were undertaken, a review of the qualitative data of those evaluations collected in this study and with informal discussions with faculty participants, the total number of student comments either remained the same or increased and the comments reflected greater depth.
Recommendations

This pilot study reveals that online evaluations will yield somewhat similar results to traditional paper and pencil evaluations at Sacramento State. Given that online evaluations are substantially more cost effective, less labor intensive, and can provide more timely feedback to faculty members, the following recommendation is proposed:

Recommendation 1: Faculty members should be provided with the option to conduct their teaching evaluations electronically with approval from their department.

Because some departments may require discussion about their policies regarding the use of electronic evaluation, it is further recommended that departments engage in discussions about how best to implement and allow electronic evaluations. As such, the following recommendation is proposed:

Recommendation 2: Academic Affairs ask each of the Deans to communicate with individual departments that policy regarding electronic evaluations be revisited to accommodate, where appropriate, the use of electronic teaching evaluations.

Academic Technology and Creative Services successfully served as the unit responsible for coordinating and working with faculty members to collect electronic evaluations for this pilot study. Given their prior experience with working with faculty within and outside of the pilot study to collect electronic teaching evaluations, it seems reasonable to charge this unit with future oversight in this area. As such, the following recommendation is proposed:

Recommendation 3: Academic Technology and Creative Services be charged with creating a formal campus procedure for collecting and distributing electronic course evaluations that is secure and expedient.
Recommendation 4: Academic Technology and Creative Services assume all duties related to the collection of electronic course evaluations at Sacramento State.

Because the California Faculty Association plays a role in representing faculty and that use of online teaching evaluations has been a topic of interest to the California Faculty Association, the following recommendation is made:

Recommendation 5: Human Resources and CFA should confer with one another to find an acceptable approach to supporting the use of online teaching evaluations at Sacramento State.
References


Sax, L. J., Gilmartin, S. K., and Bryant, A. N. Assessing Response Rates and nonresponse bias in web and