Report to the Senate:

Spring 2013 Online Evaluations Pilot at Humboldt State University

Richard W. Bruce

Humboldt State University

September 2013

Table of Contents

Introduction 3
Methods 3
Results 5
Cost Savings 5
Response Rates 5
Likert Scales 8
Analysis 9
References 10
Appendix A: University-Wide Semester Descriptive Data 12

Introduction

In spring semester of 2012, the College of Arts, Humanities & Social Sciences (AHSS) at Humboldt State University (HSU) performed a pilot study of online evaluations (Bruce, 2013). The faculty collective bargaining agreement allowed for evaluations to be delivered electronically. Additionally, the faculty contract required instructors to evaluate all courses, instead of two per year for tenured faculty. To alleviate some of the increased workload now required by evaluating more courses, the University Senate approved a university-wide online evaluation pilot to be held that semester (spring 2013) with a report presented to the Senate by October 2013.

The AHSS pilot resulted in expected outcomes of lower response rates (Avery, Bryant, Mathios, Kang & Bell, 2006, p. 29; Donovan, Mader & Shinsky, 2006, p. 287; Norris & Conn, 2005, p. 16; Nowell, Gale & Handley, 2010, p. 466; Stowell, Addison & Smith, 2012, p. 469), faculty ratings on a Likert scale (lower but not significantly, per Wang, 2010, p. 10), and an increase in the length of student comments to open-ended questions (Hardy, 2003, p. 35; Johnson, 2003, p. 54; Layne, DeCristoforo & McGinty, 1999, p. 229). The AHSS pilot did not show a significant increase in the number of student comments. The literature review of this analysis informed the implementation of the university-wide pilot performed in spring 2013.

Methods

In the past, each college had implemented evaluations differently: one college asked departments for a list of classes for evaluation, generated course evaluations in the ClassClimate software, printed evaluations from a template, collated evaluations, distributed evaluations to departments for proctoring, received evaluations back, scanned evaluations into ClassClimate, and distributed the reports; another left most of the process to departments, but the college

scanned evaluations for all the departments within that college; the third completely decentralized the evaluation process, leaving everything to the departments. The University Senate's mandate was to create a task force "to oversee the implementation…to ensure the highest response rate possible" (University Senate, 2012). Information from the 2012 AHSS pilot (Bruce, 2013) was used as guidelines for getting the highest response rate: consistency in practices across campus and advertising were most important.

For a university-wide and consistent approach (during the AHSS pilot in spring 2012, evaluations were set up inconsistently, some with no student reminders), the three colleges were each responsible for evaluations within the college. Departments received a list of potential courses and requested confirmation of which courses were to be evaluated. The colleges set up the evaluations, including every-other-day reminders to students beginning on April 22, 2013, and mid-way notification to faculty with response rates lower than 70% in each course on May 1, 2013. Evaluations were open for three weeks, closing on May 10, 2013.

For publicity, a website was created (http://www.humboldt.edu/cebs) with a sample evaluation and information for students and faculty. With consultation from the campus internal communications specialist, students were sent an email from "Humboldt State University" notifying them of electronic evaluations the week before evaluations opened with a link to the website, and faculty received a similar email from the Provost encouraging faculty to talk to their students about the importance of teaching evaluations. A notice went in the weekly University Notices email to the campus community each Friday that evaluations were active. The Provost's Office also funded a half-page ad in the campus newspaper, the *Lumberjack*, during the last two weeks of the evaluations. Finally, one of the *Lumberjack* staff wrote an article on online evaluations.

Results

Cost Savings

From the AHSS report from spring 2012, cost savings were disaggregated into space/storage costs, printing costs, and staff time (Bruce, 2013, p.11). Estimates from the college pilot predicted a university-wide savings of approximately $25,000 between staff time and printing costs. Production costs are easily identifiable, but personnel time is not traceable.

In spring 2012, there were 31,904 students enrolled in the 993 sections evaluated. Each enrolled student received a double-sided paper evaluation and, as most courses had a cover sheet, the campus printed 65 reams of 60-weight paper for a cost of $585. Printing costs would be an additional $540 at $0.0081 per copy. Total printing costs from spring 2012 should be $1,125 per semester, or $2,250 per year.

For the spring 2013 semester, 38,223 student evaluations were generated for all the courses across campus in 1320 sections. Campus saved 80 reams of paper ($720) and duplication costs of $646. This semester's production costs would have been $1,366, or $2,732 per year.

Because this was a first run, costs to personnel time were redistributed to the college office staff who worked on advertising and addressing questions, and to the task force members who discussed implementation issues for the new campus process. Staff from across campus departments did not spend hours scanning evaluations into ClassClimate as students submitted evaluation feedback directly into the software, but much of personnel costs were reduced and concentrated in fewer people overall.

Response Rates

The average response rate for paper evaluations in spring 2012 was 78.07% (N=993) while the spring 2013 online-only evaluations average response rate was 72.14% (N=1320), a

difference of 5.93% lower in the electronic delivery but with 327 more sections evaluated. The difference in the number of sections between above and the data in table 1 can be explained simply: the table includes both paper and online evaluations for the spring 2012 semester; the spring 2013 semester below is lower because the data does not include courses evaluated before the last three weeks of the semester (i.e., part-of-term courses that ended before the last week of the semester were evaluated earlier). From this point, "spring 2013" implies "electronic" and "spring 2012" implies "paper." For spring 2013, 60% of courses received a response rate of 70% or higher for the electronic evaluations while 60% of spring 2012 courses received 74% response rates or higher.

Table 1

Descriptive Statistics for Paper and Electronic Delivery of Evaluations

	N	Minimum	Maximum	Mean	Std. Deviation
Spring 2012 (Paper) Response Rates	1027	3	100	74.88	18.546
Spring 2013 (Electronic) Response Rates	1297	19	100	72.13	12.846

Response rates for courses were slightly lower, but with a smaller standard deviation. Spring 2012 had a bell curve skewed to the right, with a strong peak at the 100% response rate. Electronic evaluations resulted in a curve a little more to the left and with a lower peak at 100% (figures 1 and 2).

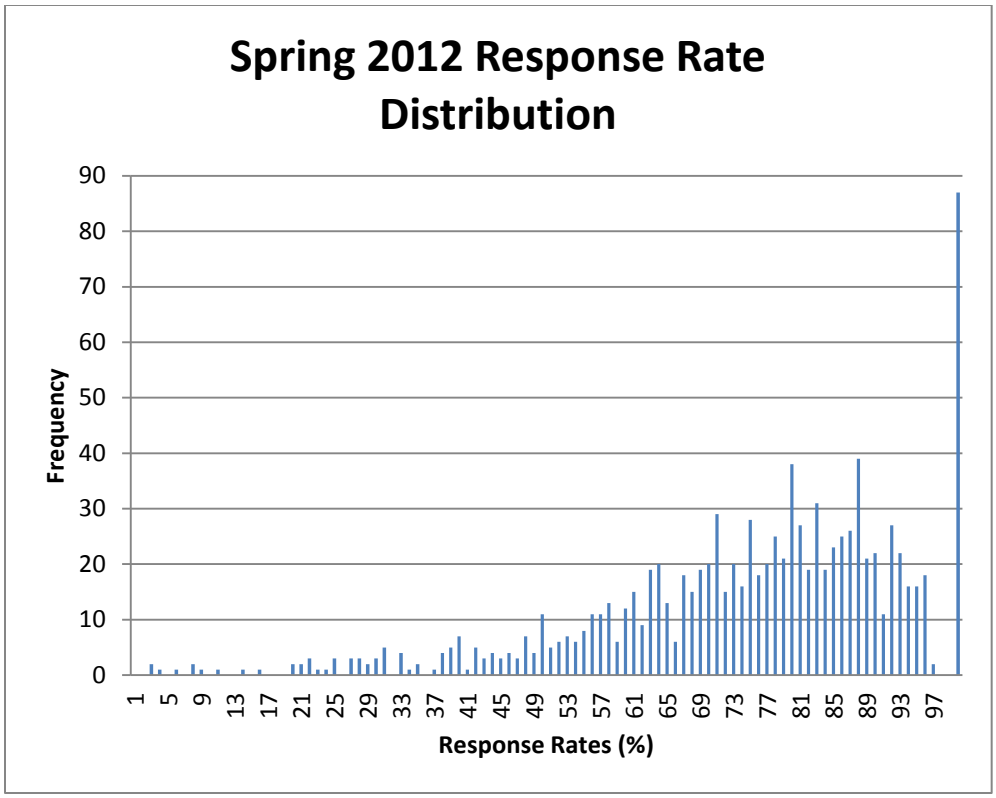

Figure 1. Distribution of course response rates for spring 2012 paper evaluations.

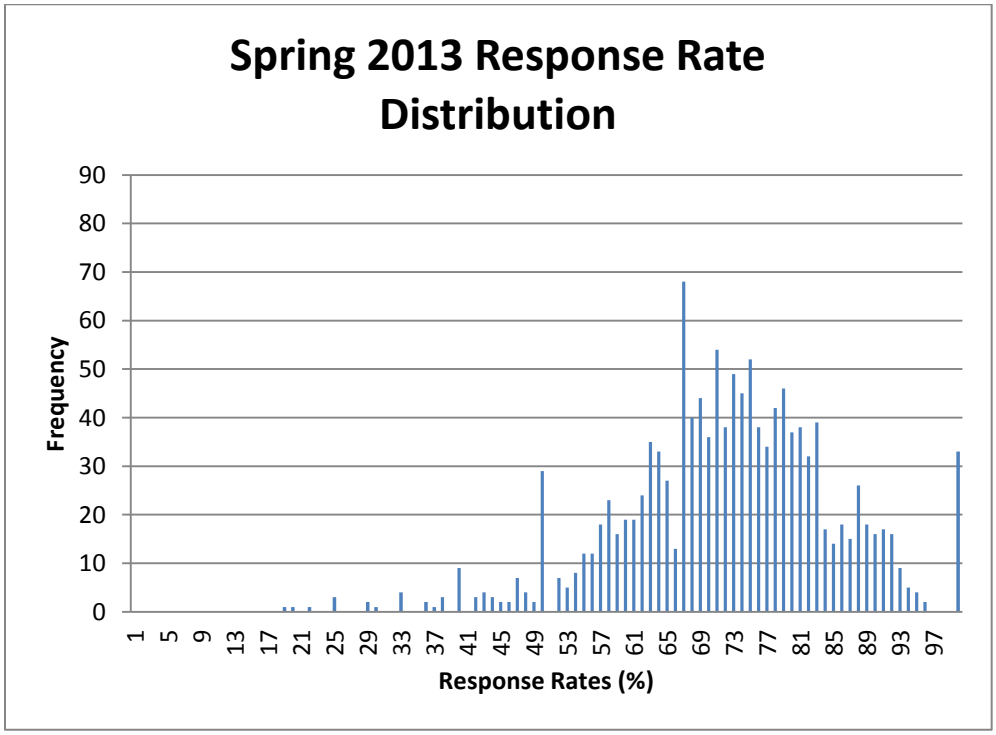

Figure 2. Distribution of course response rates for spring 2013 electronic evaluations.

Another perspective for response rates is to look at the total number of students enrolled in all sections evaluated each term and to look at the number of responses across the university. In the spring 2012 semester, there were 31,904 students enrolled in the 993 sections evaluated and 22,814 students responded to evaluations (a response rate of 71.51%). The following spring had 38,223 students registered and 27,602 responses, a university-wide response rate of 72.21%. Overall, more students responded to the number of evaluations, but the individual response rates of courses was lower.

Likert Scales

The standard university questions consist of ten opportunities for students to rate professors on a scale of 1-5 on teaching ability, availability and the classroom environment. The data from these scales is discrete: spring 2012 data are for paper evaluations, spring 2013 data are for online evaluations. Each question is independently averaged and compared across campus for each term/delivery, as well as a campus-wide overall mean for each term. Overall, the means for spring 2013 were statistically significantly lower, meaning there is a difference between the paper evaluations and the online evaluations. The difference between the averages, however, is small. The greatest difference is in question 10: “The instructor's overall teaching effectiveness in this course was:” while the least difference was in question 3: “The instructor's explanation of the grading system was:”

Table 2
University Averages for Evaluation Questions

Question #	Spring 2012	Spring 2013	Difference in Means	% Difference
1	4.3095	4.1970	-0.1125	-2.25%
2	4.2600	4.1916	-0.0684	-1.37%
3	4.1591	4.1144	-0.0446	-0.89%
4	4.2794	4.1474	-0.1319	-2.64%
5	4.3447	4.2313	-0.1134	-2.27%
6	4.4613	4.3546	-0.1067	-2.13%
7	4.5822	4.4875	-0.0947	-1.89%
8	4.5979	4.4735	-0.1244	-2.49%
9	4.4194	4.2864	-0.1330	-2.66%
10	4.2921	4.1513	-0.1408	-2.82%
Overall	4.3706	4.2635	-0.1071	-2.14%

Analysis

Of the data gathered, nothing is particularly surprising. Production costs for running evaluations are estimated to be lower. Response rates were slightly lower as well as the instructor ratings. The available information to date should be sufficient to demonstrate the influence of online evaluations for faculty file reviews. To date, there has been no analysis of the frequency, content, or length of student comments. Based on data from the AHSS pilot, the evaluations should have comparable results for open-ended questions.

References

Avery, R.J, Bryant, W.K., Mathios, A., Kang, H., & Bell, D. (2006 Winter). Electronic course evaluations: Does an online delivery system influence student evaluations? *Journal of Economic Education, 37*(1), 21-37.

Bruce, R.W. (2013, February). An Online Evaluation Pilot Study in the College of Arts, Humanities & Social Sciences. Retrieved from https://www.humboldt.edu/cebs/docs/AHSS%20Online%20Eval%20Pilot%20Program2012.pdf

Donovan, J., Mader, C.E., & Shinsky, J. (2006, Winter). Constructive student feedback: Online vs. traditional course evaluations. *Journal of Interactive Online Learning, 5*(3), 283-296. Retrieved from http://www.ncolr.org/jiol/issues/pdf/5.3.5.pdf

Hardy, N. (2003). Online ratings: Fact and fiction. *New Directions for Teaching and Learning, 2003*(96), 31-38.

Johnson, T.D. (2003). Online student ratings: Will students respond?. *New Directions for Teaching and Learning, 2003*(96), 49-59.

University Senate. (2012, March**).** Resolution on the Pilot of Electronic Anonymous Student Evaluations of Teaching. Retrieved from https://www.humboldt.edu/senate/resolution-docs/2012-2013/Resolution27-12-13-FAC2ndReadingPASSEDApproved.pdf**.**

Layne, B.H., DeCristoforo, J.R., & McGinty, D. (1999). Electronic versus traditional student ratings of instruction. *Research in Higher Education*, 40, 221-232.

Norris, J. & Conn, C. (2005). Investigating strategies for increasing student response rates to online-delivered course evaluations. *The Quarterly Review of Distance Education, 6*(1), 13-29.

Nowell, C., Gale, L., & Handley, B. (2010). Assessing faculty performance using student evaluations of teaching in an uncontrolled setting. *Assessment & Evaluation in Higher Education*, *35*(4), 463-475.

Stowell, J.R., Addison, W.E., & Smith, J.L. (2012). Comparison of online and classroom-based student evaluations of instruction. *Assessment & Evaluation in Higher Education*, *37*(4), 465-473. doi: 10.1080/02602938.2010.545869

Wang, J. (2010, Fall). 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. Retrieved from http://www.csus.edu/acse/archive/1011/ECETF%20Distributed%20Report.docx

Appendix A

University-Wide Semester Descriptive Data

	Term	N	Mean	Std. Deviation	Std. Error Mean
1. The instructor's contribution to my understanding of concepts/ideas was:	S2012	21251	4.309	.8668	.0059
	S2013	27568	4.197	1.0052	.0061
2. The instructor's accessibility/availability for consultation outside of class (office hours, by appointment, email, phone) was:	S2012	21104	4.260	.9023	.0062
	S2013	27478	4.192	.9842	.0059
3. The instructor's explanation of the grading system was:	S2012	21112	4.159	.9608	.0066
	S2013	27498	4.114	1.0382	.0063
4. The instructor's ability to present information clearly was:	S2012	21183	4.279	.9317	.0064
	S2013	27503	4.147	1.0668	.0064
5. The instructor's ability to challenge me was:	S2012	21124	4.345	.8761	.0060
	S2013	27516	4.231	1.0002	.0060
6. The instructor's ability to create an atmosphere where students were comfortable asking questions and/or engaging in discussion was:	S2012	21180	4.461	.8707	.0060
	S2013	27469	4.355	.9868	.0060
7. The instructor's ability to create a classroom environment that was respectful of diversity (ethnicity, socio-economic background, sexual-orientation, nationality, age, ability, religion, gender) was:	S2012	21117	4.582	.7461	.0051
	S2013	27472	4.488	.8568	.0052
8. The instructor's ability to demonstrate knowledge of the subject matter was:	S2012	21176	4.598	.7187	.0049
	S2013	27474	4.474	.8708	.0053
9. The quality of the instructor's overall preparation for class was:	S2012	21109	4.419	.8502	.0059
	S2013	27472	4.286	1.0039	.0061
10. The instructor's overall teaching effectiveness in this course was:	S2012	21076	4.292	.9379	.0065
	S2013	27425	4.151	1.0952	.0066