INVESTIGATING STRATEGIES FOR INCREASING STUDENT RESPONSE RATES TO ONLINE-DELIVERED COURSE EVALUATIONS

John Norris  
University of Hawaii

Cynthia Conn  
Northern Arizona University

Student evaluations of college instruction provide one critical source of information for the improvement of courses, curriculum, and practitioners' pedagogic efforts. Increasingly, online-delivered course evaluations are being put to use for eliciting student feedback in Web-based as well as face-to-face courses. Two studies were conducted in order to, first, understand the level of student responsiveness to existing online-delivered course evaluations at Northern Arizona University, and, second, to investigate strategies for improving evaluation response rates in a subset of Web-based and face-to-face courses. Findings indicated that a combination of simple and easily implemented strategies was associated with considerably higher student response rates to online-delivered course evaluations than had been attained in existing online-delivered evaluations.

INTRODUCTION

At Northern Arizona University (NAU), like other institutions of higher education, end-of-semester student evaluations of courses provide a ubiquitous source of potentially useful information on instructional features, instructor effectiveness, and student learning. Certainly, a student's subjective evaluation of course features is only one means of gathering information for understanding and improving instruction (Angelo, 1998; Angelo & Cross, 1994; Messick, 1999), but it is a critical one. In addition, student course evaluations are frequently used as a primary source of data for faculty performance evaluations—at NAU, Arizona Board of Regents and Faculty Senate policies require that student feedback be collected and considered during faculty tenure and promotion reviews. Finally, especially where they are otherwise required, student course evaluations can also provide one
readily available means for collecting students’ self-assessments of their own learning outcomes, a valuable source of information for accountability, improvement, and related assessment purposes.

Despite their potential value, the effective implementation of student course evaluations may be constrained as traditional forms of college instructional and administrative practice evolve to incorporate more varied teaching and learning contexts, delivery mechanisms, and demands on efficiency and quality. One potential solution to apparent constraints has come in the form of online-delivered course evaluations. Clearly, within electronically-mediated instruction, the delivery of course evaluations via a Web-based application is mandatory. In addition, many colleges and universities are also considering or already implementing course evaluations via an online format for face-to-face courses, in order to lower costs, increase timeliness of feedback, and facilitate record-keeping and data analysis. However, a key preliminary concern with using online course evaluations for both types of course is whether student response rates will prove dramatically less than has been the case for paper-and-pencil evaluations administered to students in face-to-face classes.

Several limitations of online-delivered survey instruments in general have been found to result in low rates of response from online students and other participants. Some of these limitations include technical problems with the online tool (Schmidt, 1997), difficulty accessing open computers in campus laboratories, and students’ relative levels of technological literacy (Handwerk, Carson, & Blackwell, 2000). Others findings have shown that students’ use of multiple e-mail addresses and the frequent changing of e-mail addresses (Bradley, 1999), as well as slow Internet connection rates (Crawford, Couper, & Lamias, 2001) may also decrease response rates.

Although such challenges may reduce their utility, a number of benefits may be realized through the use of online-delivered course evaluations and survey instruments. These benefits include time and cost savings, quick return of responses, longer and more in-depth responses to open-ended questions, and simple convenience for participants who are becoming increasingly technologically literate. Using the Internet to deliver surveys or evaluations eliminates time-consuming and expensive printing processes on the front-end, and it is also an environmentally friendly alternative. In addition, direct input of data into Web-based forms by respondents is much more efficient and less expensive than secondary manual input of responses or even the scanning of paper-based responses (Franceschini, 2000; Mertler, 2002; Schmidt, 1997; Tse, 1998). Of course, costs can increase if substantial time and money is spent building a custom tool. However, if those costs are spread over a number of projects and several years, the initial upfront development may prove less expensive than using a commercial product or the more traditional, paper-based course evaluation alternatives.

In order to pursue these potential benefits for various institutional constituents, NAU’s Office of Academic Assessment (OAA) decided to confront the potential limitations of online-delivered course evaluations head-on, by investigating both the nature of low student response rates and the effectiveness of strategies for improving them in online-delivered course evaluations. The purpose of this work was to provide initial empirical evidence for informing policies and practices, rather than simply accepting the common perception that response rates for online-delivered course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. Accordingly, the OAA undertook two studies. The first study was intended to determine how low response rates actually were for existing online course evaluations and surveys will always be low. According to the OAA, the second study explored the effectiveness of a combination of very simple strategies for increasing response rates to online-delivered evaluations in both Web-based and face-to-face courses.
STUDY 1

Study 1 sought to establish a baseline of student response rates, and factors that might contribute to higher or lower rates, for online-delivered course evaluations at NAU. This study tapped available data from an existing online course evaluation system, Evalajack, which has been in use since the Spring semester of 2000. Evalajack is a proprietary Web-based tool that was developed by NAU’s Information Technology Services (ITS) department to provide an immediate, if not comprehensive, solution to the burgeoning demands for online-delivered evaluations in conjunction with rapidly increasing Web-based course offerings. Between Spring 2000 and Fall 2002, Evalajack forms were made available for over 1,000 Web-based course sections at NAU. In this study, we first analyzed existing response-rate data for these 2 years of online course evaluations; we then followed up with a questionnaire to a subset of online instructors in order to identify variables which might be associated with higher or lower student response rates. Unfortunately, we were not able to collect students’ insights into the variables which may or may not encourage their responses to online-delivered evaluations, and future investigations have been planned in order to tap this additional source of information.

Method

The Evalajack database was queried by an ITS administrator in order to produce a complete set of anonymous course evaluation data for the six semesters from Spring 2000–Fall 2002. These data included: (a) semester/year designation; (b) course/section identification number; (c) course classification (graduate, undergraduate); (d) number of students enrolled; and (e) number of students who completed the evaluation. Instructor information was not queried. Data were subsequently pruned by OAA staff for course repetitions, false (e.g., cancelled) courses, and other potential anomalies, resulting in a final set of course evaluation data consisting of 1,108 distinct course sections. Courses represented all disciplines and all colleges/schools at NAU in which Web-based courses were offered during the 2000–2002 period. Enrollments averaged 27 students per section, while a standard deviation of 34 students indicated considerable variability in section sizes. Using the individual course section as the unit of analysis, descriptive statistics were calculated for this final data set in order to establish baseline student response rates for all online course evaluations as well as for the distinct sub-sets of undergraduate versus graduate online course evaluations.

Subsequently, in order to identify factors which might have been associated with higher or lower student response rates within these Evalajack data, a questionnaire was administered in early 2003 to all instructors who had taught Web-based courses during the Fall 2002 semester. This group of instructors was selected due to the likelihood that they would be able to recall and report on the actions that they had taken in administering course evaluations to their online students at the end of the preceding semester. A voluntary Web-based questionnaire was sent to 85 instructors, 50 of whom responded (59% response rate). The questionnaire (see Appendix A) asked instructors to provide information regarding: (a) when and how the course evaluation was announced to students; (b) whether/how the value of evaluations was discussed; (c) whether/how students were reminded to complete the evaluation; (d) whether students were given extra credit or were required to complete the evaluation; (e) whether a mid-semester evaluation had been administered; and (f) what additional strategies, if any, were employed to encourage student responses. Instructors were asked to respond to all questions for a single course (which they identified), and to repeat the questionnaire for additional courses if desired. They were ensured that their responses would remain anonymous and would be used only for the purposes of recom-
mending potential strategies for increasing student response rates. Finally, patterns in instructors’ answers to the questionnaire were compared with response-rate data for their Fall 2002 online courses in order to identify any associations with higher or lower student response rates.

Findings

Average Online Evaluation Response Rates

Analysis of existing data for Evalajack online evaluations indicated generally low student response rates. Table 1 shows descriptive statistics for response rates, with course section taken as the unit of analysis. Based on data for 1,108 course sections, an average of only 31% of students per course responded to end-of-semester evaluations, although a standard deviation of 28% also indicated considerable variability from course to course. Furthermore, on average, 11% more students responded to evaluations in graduate courses than did in undergraduate courses, while variability was equally high in each of these subsets of data.

In order to better understand the extent to which low student response rates were found for courses using the Evalajack system, the frequency of course sections was plotted graphically for various response levels. Figure 1 shows the frequency of course sections with response rates falling at 5% increments, from 0% to 100%. The positive skew in the frequency distribution clearly reflects the low average rates identified previously. However, note also the very high frequency of courses (over 200) which reported zero student responses to online evaluations. For these cases of zero response, it is unclear whether students may or may not have been provided the opportunity to complete the course evaluation or whether they simply elected not to, encountered technological incompatibilities, etc. Likewise, it is important to note that, while considerably fewer, a number of courses did exhibit much higher student response rates. Further analyses revealed that graduate and undergraduate course sections were relatively evenly distributed across the range of response rates, while the averages for these two groups obviously fell at different points; thus, while the graduate/undergraduate distinction accounted for some systematic difference in response rates, the equal distribution of both types of courses across the entire range suggested that other factors contributed more substantially to observed differences. Finally, in order to check whether increasing course size

<table>
<thead>
<tr>
<th>Course Sections</th>
<th>Statistic</th>
<th>Response Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined (N = 1,108)</td>
<td>Mean</td>
<td>31%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>28%</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>100%</td>
</tr>
<tr>
<td>Graduate (N = 625)</td>
<td>Mean</td>
<td>36%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>28%</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>100%</td>
</tr>
<tr>
<td>Undergraduate (N = 483)</td>
<td>Mean</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>27%</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>100%</td>
</tr>
</tbody>
</table>
was associated with decreasing student response rates (a common assumption), a Pearson product-moment correlation was calculated between the number of students enrolled per course section and the associated response rates; the resulting correlation (-.02) indicated no relationship.

Of course, these findings should not be overinterpreted. Given the anonymity and multiyear nature of the data, it was impossible to confirm with individual instructors whether Evalajack forms had been made available to students in every course section, or whether other problems were encountered in accessing and completing the forms. In addition, it may well have been the case with a subset of the data that alternative evaluations (e.g., instructor-created forms) were used in lieu of, or in addition to, the Evalajack forms. Nevertheless, given the fact that evaluations were created and made available online, and that instructors were requested to announce the evaluations to their students, it is reasonable to interpret the current findings as evidence that student response rates to Evalajack online course evaluations have been, on average, relatively low. However, a sub-set of courses/instructors did manage to elicit reasonable to good levels of responsiveness from students, and this finding was not largely attributable to class size or undergraduate/graduate classification of the course.

Instructor Strategies Potentially Affecting Response Rates

Results from the Fall 2002 instructor questionnaire helped to identify factors potentially associated with differences in response rates to Evalajack online evaluations. Obviously, given the low number of instructors who were queried (85) and who responded (50), defini-
tive associations could not be drawn between specific variables or combinations of variables and increases in response rates; multivariate inferential statistical assumptions (cf. Tabachnick & Fidell, 1996) were simply not met by the low number of participants in this study. However, patterns in their answers to the questionnaire did emerge, and tentative relationships identified with increases in response rates were used to inform further investigations.

First, none of the 50 respondents reported having required students to complete the evaluation (e.g., in order to receive a final course grade), and only two provided extra credit to students for completing the evaluation. Thus, coercion/reward strategies were not found to play a role in encouraging or discouraging responsiveness. However, these two questions did elicit a number of open-ended comments from instructors, who argued that requiring or rewarding evaluation completion would only lead to overly negative or overly positive instructor ratings from students, and therefore should not be pursued. Second, exactly half of the respondents indicated that they had discussed the value of student evaluation of courses with their students prior to making the Evalajack online evaluation available; however, the degree and manner in which instructors did so varied considerably, with most having provided some kind of short statement on e-mail about the usefulness of student feedback on the course. Average student response rates were virtually identical between these courses and the courses in which instructors had made no explicit statement regarding the value of course evaluations, so the role to be played by values statements remained ambiguous. Third, only two instructors reported that students had completed mid-semester evaluations of their courses, so the potential relationship between mid- and end-of-semester evaluations also remained uncertain.

Instructors’ answers to questions about announcing the evaluation and reminding students to complete the evaluation produced the most apparent associations with increasing student response rates. First, timing of the initial announcement of the online evaluation seemed to have played an important role. When instructors announced the evaluation more than 3 weeks prior to the end of the semester (typically, by simply listing it in the syllabus or on the Website at the beginning of the course), average student response rates were only 25%. However, when instructors announced the evaluation between 2 and 3 weeks prior to the end of the semester, average response rates were 48%; this figure dropped to 44% when the evaluation was not announced until the final week of the semester. Second, location and frequency of announcement postings may have influenced student responsiveness. Figure 2 shows average student response rates associated with the posting of evaluations in various locations (clearly there may have been interactions among these variables which could not be discerned due to the low N; with a greater number of respondents a regression analysis could be used to predict whether combinations of these variables were associated with differences in response rates). Posting the evaluation announcement to a class discussion board was associated with the highest response rates (nearly 50%), and posting the announcement in at least two locations was associated with higher rates than a single posting. It should be noted that twice as many instructors reported using e-mail and Website postings versus discussion boards or online assignments.

Figure 3 shows the average response rates found for several types and frequencies of reminder strategies. Note that of the 50 instructor respondents, approximately one third reported that they did not remind students to complete the evaluation after the initial announcement. Interestingly, average response rates for these courses were virtually identical to those for courses which did include a reminder; however, it was impossible to disentangle the differential effects of multiple initial announcements versus announcement plus reminders. Nevertheless, within the subset of courses that did use intentional reminder strat-
Investigating Strategies for Increasing Student Response Rates to Online-Delivered Course Evaluations

Figure 2
Average student response rates for evaluation announcement strategies

Figure 3
Average student response rates for evaluation reminder strategies
egies, posting a reminder to the discussion board was associated once again with the highest average response rates, and multiple reminders resulted in higher responsiveness than did single reminders. By far the most common reminder strategy was an e-mail message to students.

Despite the lack of definitive relationships in these questionnaire data, patterns did indicate several potential strategies for increasing student response rates to online evaluations, including the timing and location of evaluation announcements/reminders and the use of multiple postings. Instructors’ open-ended feedback also served to further contextualize these possibilities. First, several instructors commented on the incompatibility of Evalajack online forms with certain Web browsers, and they stated that a stable and accessible evaluation platform was a prerequisite for enabling students to respond online. Second, a number of instructors noted that they were worried about “bombarding” students with too many reminders to complete an evaluation, as well as with too many evaluations (e.g., from the distributed learning office as well as from their college), and they were concerned that students might react negatively within evaluations as a result. Finally, several instructors commented that low response rates to online course evaluations posed a serious problem for tenure and promotion review, and they added that evaluations should ask questions that are directly related to the course and comparable to evaluations of other courses within the college/school, in order to be meaningful to student respondents as well as to faculty and administrative users. All of the findings above, and in particular the insightful open-ended comments by instructors, informed the development of strategies investigated within the subsequent study.

STUDY 2

Study 2 investigated the effectiveness of a set of combined strategies intended to increase student response rates to online-delivered course evaluations. Based on factors observed in study 1, simple recommendations were compiled for instructors to use in announcing, discussing, and reminding students to complete the online course evaluations. Several constraints influenced the strategies selected. First, given the potential positive or negative outcomes for instructors, penalty/reward strategies were not recommended. Second, in order to minimize potential technological influences on response rates, an alternative and reliable online evaluation system was utilized in lieu of the proprietary Evalajack system. Third, in response to instructors’ concerns with over-burdening and annoying students, it was decided that only a minimal set of strategies would be operationalized based on (a) ease of implementation across widely varying individual course contexts (e.g., wherein different Web-based course management tools were used) and (b) likelihood of increasing response rates. Finally, because large numbers of faculty/instructors were not available for participation in the study, it was decided that individual strategies could not feasibly be investigated in isolation. In addition, because the course evaluation process and results would be taken seriously and used for various purposes by students, instructors, and administrators, ethical concerns dictated that all courses should be treated equally and that the study intervention should attempt to maximize student response rates. Accordingly, only a single set of combined strategies was operationalized in all courses in the current study.

Method

Two distinct subsets of courses/instructors were included in the study. First, voluntary participation in the study was solicited from NAU faculty and instructors teaching fully Web-based courses during the Spring or Summer 2003 semesters. Participants were solicited from those colleges/schools/departments in which Evalajack was the only online evaluation tool in operational use at the time (sev-
eral schools had implemented other in-house evaluation tools in order to meet unique requirements). Twenty-one instructors and faculty, teaching a total of 39 distinct course sections (22 graduate, 17 undergraduate), volunteered from a variety of disciplines in four colleges/schools: Arts and Sciences, Business Administration, Education, and Social and Behavioral Sciences. The number of students enrolled in these online courses ranged from 3 to 49, with an average of 24 students per section. Courses were all taught entirely online during 5-week, 10-week, or 16-week semesters. Participants agreed to the following conditions of the study: (a) the OAA would design and host course evaluations; (b) evaluations would include all items from the required college/school forms plus any additional items about the course or student learning outcomes that the instructor requested; (c) participants would follow a set of basic strategies, provided by the OAA, for increasing student response rates; and (d) anonymous results would be reported to the instructor and other required parties (e.g., department chairs) within one week following the grades due date for the semester.

In addition to these Web-based courses, instructors for University Colloquium 101 (UC101), a required freshman-level course that is taught in small face-to-face sections, were invited to participate in the response rates study during the Spring 2003 semester. Given the very large number of UC101 sections offered each semester at NAU, the university administration and the OAA had decided to investigate the possibility of migrating the end-of-semester course evaluations to an online rather than paper-and-pencil format, in order to facilitate gathering, analyzing, and reporting evaluation data. Of the 25 sections offered during Spring 2003, instructors of 21 sections agreed to participate in the study; enrollment in these sections ranged from 19 to 30 students, with an average of 23 per section. For comparison purposes, additional response-rate data were collected for paper-and-pencil evaluations in the UC101 course for both Fall 2002 and Spring 2003 semesters, and response rates for pilot online evaluations from the Fall 2002 semester (when no strategies for increasing responsiveness had been implemented) were also included.

In order to ensure that technology problems (e.g., incompatibility between evaluation software and certain Web browsers) would not influence response rates, a stable Web-based evaluation system was identified in the commercial survey-design and -hosting software, SurveyMonkey (http://www.surveymonkey.com). The OAA had pilot-tested several online course evaluations using this software during Fall 2002, and instructors/students had reported very few difficulties in accessing or completing the evaluations on the SurveyMonkey Website. Therefore, end-of-semester evaluations for all participating courses were designed and hosted on this stable platform.

After evaluation forms had been designed and uploaded, a unique URL was provided to each instructor for each Web-based course section. Participants were then asked to implement the following strategies (exact instructions are provided in Appendix B). First, the evaluation URL was to be posted, and clearly identified, in a prominent location on the course management home page 2 weeks prior to the last day of class for the 16-week semester (times were adjusted for the shorter summer semesters). Second, students were to be notified simultaneously of the availability of the course evaluation and the URL for completing it via postings in all locations appropriate to a given class (including e-mail, the online course syllabus or course calendar, and online course discussion boards). Participating instructors were also asked to include in these announcements a brief statement regarding the value of completing course evaluations, as well as instructions to students for doing so and a date by which evaluations should be completed. Finally, one week prior to the final day of classes, instructors were asked to send an e-mail message reminding students to complete the evaluation and including the URL and a “complete-by” date. A single e-mail
The reminder strategy was selected over other possible reminders because all instructors utilized e-mail for regular communication with their online students, but not necessarily other Web-based tools (e.g., discussion boards). In addition, given the proximity of the reminder to the initial announcement (1 week later), multiple reminders were not recommended in order to avoid potential negative reactions from students.

The implementation of strategies in the UC101 sections differed in several ways due to the face-to-face course format. Instructors were provided with an evaluation instruction handout to distribute to students two and three weeks prior to the last day of classes. These instructions informed students about the value of completing course evaluations, provided them with the online evaluation URL and instructions for accessing it outside of class, and stated the date by which forms should be completed. Instructors were asked to announce the online evaluation in class and to spend a few minutes discussing the instructions. One week prior to the last day of classes, the OAA sent an e-mail reminder message to all students enrolled in UC101; this message requested that students complete the evaluation if they had not already done so, and it provided a "clickable" URL link to the form for each course section.

After the semester ended, a questionnaire was sent to all instructors to determine whether the intended strategies were followed. Responses confirmed that the strategies were adhered to with very few deviations, in both the Web-based and the UC101 courses.

**Findings**

Overall, considerably higher average student response rates to online evaluations were found for courses implementing the simple strategies outlined above than had been observed for Evalajack online evaluations over the previous 2 years. Table 2 shows response rate data for the subset of Web-based courses included in the study, both combined and broken down according to graduate/undergraduate classification. In order to provide an accurate indication of the extent to which strat-

<table>
<thead>
<tr>
<th>Courses</th>
<th>Statistic</th>
<th>No Strategies (Evalajack)*</th>
<th>Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined</td>
<td>N</td>
<td>37</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>42%</td>
<td>74%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>22%</td>
<td>18%</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>6%</td>
<td>33%</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>81%</td>
<td>100%</td>
</tr>
<tr>
<td>Graduate</td>
<td>N</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>44%</td>
<td>73%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>22%</td>
<td>18%</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>6%</td>
<td>38%</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>77%</td>
<td>96%</td>
</tr>
<tr>
<td>Undergraduate</td>
<td>N</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>39%</td>
<td>74%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>23%</td>
<td>18%</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>9%</td>
<td>33%</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>81%</td>
<td>100%</td>
</tr>
</tbody>
</table>

*No strategies* statistics were calculated based on response rate data available from all previous administrations of Evalajack online evaluations for the subset of participating courses; data were unavailable for two of the participating courses which had not previously administered online evaluations.
egies may have influenced student responsiveness, existing response rate data from previous administrations of Evalajack were also analyzed for the same subset of courses on their own (i.e., rather than comparing the overall Evalajack average response rate of 31% with this subset). For these Web-based classes, average response rates with the strategies were found to be 32% higher than without, up from 42% to 74%. Likewise, average response rates for graduate courses were 29% higher with strategies and for undergraduate courses, 35% higher with strategies. Minimum and maximum response rates for individual sections were also consistently higher under the strategies condition.

Figure 4 shows clearly the large difference between average student response rates for courses under the two conditions (without strategies on the left and with strategies on the right). The average response rate is indicated by the dash in the middle of the two boxes (42% for no strategies, 74% for strategies). Each box indicates the size of a 95% confidence interval within which the observed average may be trusted. These confidence intervals were calculated according to the number of courses and the variability of response rates within each condition (see Wilkinson & Task Force on Statistical Inference, 1999); conceptually, each box reflects two standard errors above and below the average response rate observed for each condition. Thus, the average response rate of 42%, found for the courses without strategies, could be trusted with 95% certainty to fall between 34.5% and 49.5%. Similarly, the average response rate of 74% for the courses with strategies could be trusted with 95% certainty to fall between 68.5% and 79.5%.

In addition, the confidence intervals shown in Figure 4 provide the information necessary for a test of the statistical significance of the differences between the average response rates found under the two conditions. If the confidence intervals surrounding the two averages were to overlap (vertically), that would indicate that an observed difference between the averages could not be trusted with 95% certainty—in other words, the difference would not be found statistically significant and would fall within the realm of a chance occurrence. It is obvious in Figure 4 that the two boxes do not overlap vertically and hence that the observed difference between the average response rates for the two conditions could not be attributed
to chance (i.e., the difference was found to be statistically significant). Of course, statistical significance aside, the important finding here was the substantial difference in the magnitude of response rates between the two conditions, and the fact that an average of 74% of the students within the strategies condition completed the online evaluations.

Similar patterns were observed for response rates in the UC101 face-to-face courses. Table 3 shows response-rate analyses for three groups of UC101 sections: (a) the Fall 2002/Spring 2003 sections that administered paper-and-pencil evaluations in class (Paper 02-03); (b) the Fall 2002 sections that administered online evaluations prior to the strategies study (E 02); and (c) the Spring 2003 sections that administered online evaluations and implemented the recommended strategies (E 03). Once again, the average student response rate for sections that implemented the strategies for online evaluations was found to be much higher than for the prestrategies sections (up from 34% to 67%). Note also that the minimum response rate for any section using the strategies was observed to be 41% (compared with 0% for the prestrategies sections), and the maximum was 96%. While the average response rate for paper-and-pencil sections was found to be 83% (16% higher than the online evaluations with strategies), this difference had been anticipated, given the fact that none of the UC101 course sections was taught in a Web-based format. In other words, an average of 67% of the freshmen students in this face-to-face course responded to the request that they complete an online-delivered evaluation of the course at the end of the semester outside of class time, a rather remarkable finding. Of additional interest is the fact that 100% response rates were not found for the UC101 sections that elected to administer paper-and-pencil evaluations in class; indeed, a number of the online plus strategies sections received higher response rates than did several of the paper-and-pencil sections. Clearly, absenteeism in the paper-and-pencil sections played a role in this observation. Finally, it should be noted that three additional UC101 sections that had elected to administer paper-and-pencil evaluations failed to do so, for a response rate of zero (data for these sections were not included in the calculation of average response rates).

Figure 5 provides a graphic depiction of the average evaluation response rates and 95% confidence intervals for the three sub-sets of UC101 course sections. These confidence

<table>
<thead>
<tr>
<th>Table 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Response Rate Descriptive Statistics for UC101 Evaluations</strong></td>
</tr>
<tr>
<td><strong>UC101 Sections</strong></td>
</tr>
<tr>
<td><strong>Paper 02-03 (N = 11)</strong></td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>S</td>
</tr>
<tr>
<td>Min</td>
</tr>
<tr>
<td>Max</td>
</tr>
<tr>
<td><strong>E 02, no strategies (N = 48)</strong></td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>S</td>
</tr>
<tr>
<td>Min</td>
</tr>
<tr>
<td>Max</td>
</tr>
<tr>
<td><strong>E 03, strategies (N = 21)</strong></td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>S</td>
</tr>
<tr>
<td>Min</td>
</tr>
<tr>
<td>Max</td>
</tr>
</tbody>
</table>

*Three additional sections elected to administer paper evaluations but did not return any completed forms (i.e., 0% response rates). Data from these sections are not included in the current response rate analyses.
Intervals were calculated and should be interpreted as described above for Figure 4. Thus, the average response rate of 83% for the paper-and-pencil condition could be trusted with 95% certainty to fall between 74.5% and 91.5%; the average of 34% for online evaluations without strategies fell between 29.5% and 39.5%; and the average of 67% for online evaluations with strategies fell between 59.5% and 74.5%. Once again, as described for Figure 4, non-overlapping confidence intervals indicated that average differences between all of the three groups were probabilistically rare, that is, statistically significant. Note, however, that the lower end of the confidence interval for the paper-and-pencil condition fell at exactly the same value as the upper end of the confidence interval for the online evaluations with strategies, reflecting the similar magnitude and certainty of response rates observed for these two conditions.

It should be obvious that study 2 was not designed to inform causal interpretations about the relative effectiveness of individual or combined intervention strategies at increasing student response rates; such interpretations would only be warranted through the operationalization of individual strategies in a large number of carefully sampled courses and the control of other potentially intervening variables. In addition, a variety of unexamined factors might have influenced the different response rates observed within the courses in the current study, including: (a) overall consistent increases in students’ online literacy and familiarity from semester to semester; (b) the variable interactiveness and engagement of students and instructors in different online courses; (c) differential perceived worth of the course evaluation process among students; (d) interaction effects between user-side technology and online evaluation technology; and many others.

However, despite these obvious limitations in the current study, findings clearly indicated that students responded with relatively high average frequency to online course evaluations that were administered via a stable platform and in conjunction with a very simple set of strategies to encourage their responsiveness. It is also clear that response rates for online courses that employed these strategies were
substantially higher than for other online courses that did not, as well as for the same courses prior to the use of such strategies, and this pattern was replicated for online-delivered course evaluations used in a set of face-to-face classes that had no online instruction whatsoever. Perhaps most encouraging was that observation that, on average, 74% of the students enrolled in online courses responded to the course evaluation when simple strategies were employed, a response rate that was only 9% less than that observed for paper-and-pencil evaluations administered in the face-to-face UC101 sections. Furthermore, it should not be overlooked that average response rates for the UC101 online evaluations without strategies (34%) were very similar to average response rates found for Evalajack overall (31%), this despite the fact that a stable online platform was used to host the UC101 evaluations. This finding lent considerable support to the interpretation that strategies were the primary contributor to the observed differences in response rates (i.e., rather than the simple reduction of potential technological problems).

CONCLUSION

Student evaluations of college instruction provide one critical source of information for the improvement of courses, curriculum, and practitioners’ pedagogic efforts, and their use as a component of faculty review is a well-established tradition in higher education not likely to disappear any time soon. Of course, many other sources of information should be tapped via a variety of evaluative methods in order to provide comprehensive and trustworthy evidence for purposes as important as tenure and promotion, course improvement, and others. However, where student evaluations are to be employed in this regard, steps need to be taken to ensure that useful data are gathered and thoughtful student feedback is facilitated. Many institutions are currently considering or implementing the use of online course evaluations, both for Web-based courses, in which other alternatives are few, and for face-to-face courses, in which issues of expense, timeliness of feedback, ease of record-keeping and analysis, and others may argue in favor of electronically-mediated evaluation solutions. A key preliminary concern with using online course evaluations for both types of course is whether student response rates will prove dramatically less than has been the case for paper-and-pencil evaluations administered to students in face-to-face classes.

In the series of studies reported above, we found that a combination of simple and easily implemented strategies were associated with considerably higher student response rates to online-delivered course evaluations than had been the case for prior online evaluations at NAU. Average positive differences of 32% were found for online courses implementing these strategies, and the resulting average response rate of 74% approximated closely the rate found for paper-and-pencil evaluations administered in multiple face-to-face sections of a freshman-level class (83%). Similar positive differences were also found for the implementation of these strategies with online evaluations used in a face-to-face course.

Thus, while the mere creation and posting of an online course evaluation is probably insufficient for ensuring that students will respond to it—as clearly demonstrated in the very low average response rates found over several years of online evaluation use at NAU—it is not necessarily the case that this situation cannot be rather easily ameliorated. An obvious prerequisite in this respect is the use of a stable and accessible online platform for designing and hosting course evaluations, such that students encounter minimal technological impediments to the evaluation process. In addition, it would seem that a few simple efforts may go a long way toward successfully encouraging considerable proportions of students to provide their critical feedback, including: (a) explicitly announcing the availability and location of the evaluation within a few weeks of the end of the course; (b) explaining the value of the course evaluation process and
student feedback; and (c) reminding students to complete the evaluation. It may be, as was found in the current study, that these minimal efforts (similar, by the way, to the most fundamental actions we take as instructors in assigning homework or administering an online quiz) will prove sufficient for gathering adequate levels of student response data via online course evaluations, without the addition of dubious coercion or reward methods (likely to have systematic negative or positive effects on evaluation outcomes). Of course, additional investigations of students’ perspectives on such strategies will shed important light on what they consider the most effective means for eliciting their evaluative feedback on courses. Likewise, attention will need to be paid to changes in students’ online behaviors over the coming years, as online instruction becomes increasingly common within all reaches of college teaching.

In the end, it is our hope that paying due attention to basic concerns, such as those investigated here, will result in improvements in the increasing use of online-delivered course evaluations. Just as effective online instruction requires more than simply posting a syllabus and course readings to the Web, so too will effective online course evaluation require something more than simply posting the evaluation URL to a course management Web site. Finally, once we are able to move beyond concerns with student response rates, the real work of developing useful course evaluation instruments that will inform reliable and valid interpretations about instruction can begin.

**APPENDIX A**

*Questionnaire: Online Course Evaluation Response Rates*

Please provide us with the following information about your end-of-semester course evaluations for Fall, 2002, online courses.

Please list the course names and section numbers for those courses to which your answers on the following questions apply:

1. When was the end-of-semester course evaluation first announced to students?
   a. 1 week before the last day of the semester
   b. 2 weeks before the last day of the semester
   c. 3 weeks before the last day of the semester
   d. More than 3 weeks before the last day of the semester
   e. Other (please specify):

2. How was the end-of-semester course evaluation first announced to students? (Check all that apply)
   a. In class verbal announcement (hybrid class)
   b. Paper handout (hybrid class)
   c. E-mail announcement
   d. Posted to class Website (online or hybrid class)
   e. Posted to class assignments (online or hybrid class)
   f. Posted to class discussion board (online or hybrid class)
   g. Other (please specify)

3. Did you discuss the value of end-of-semester course evaluations with your students? If yes please explain when and how:

4. Did you use any reminder strategies to encourage students to respond to the evaluation? (Check all that apply)
   a. In class verbal reminder (hybrid class)
   b. Paper handout (hybrid class)
   c. E-mail reminder
   d. Posted reminder to class Website (online or hybrid class)
   e. Posted to class discussion board (online or hybrid class)
   f. No reminder strategy used
   g. Other (please specify)
5. Did you offer students credit (participation points, extra points, etc.) for completing the evaluation?
If yes please explain what kind of credit was offered:

6. Did you require students to complete the evaluation (e.g., in order to access the last assignment)?
If yes please describe the requirement:

7. Did students complete a mid-semester evaluation of your course?

8. Please comment on any additional strategies used for encouraging student responses to online course evaluations and please add any other insights into this issue.

APPENDIX B

Strategies for Increasing Response Rates to Online Course Evaluations

1. Posting course evaluation link: For WebCT courses, post the link (url) for the course evaluation directly to the course home page, with the heading “End-of-semester Student Evaluation of Course”. Make the link available at least 2 weeks prior to the last day of class. Likewise, for courses offered via other course management systems, post the link for the course evaluation in a prominent location on the course home page. DO NOT post the link to the old Evalajack evaluation. (We will also inform Statewide/Distributed Learning not to post the Evalajack evaluation for your course.)

2. Announcement of course evaluation availability: Simultaneous with posting the evaluation link, announce to students the date of initial availability of the course evaluation in each of the following locations, where applicable to your course: (a) syllabus; (b) discussion board; (c) e-mail message; (d) course calendar.

Include a statement of value for the course evaluation process and instructions for completion in the initial announcement. Make sure to also include a “complete by” date in the initial announcement. [NOTE: Course evaluations should always be completed prior to the final day of the course (before grades are turned in).] The following example may be used and adapted/personalized for announcement purposes:

Course evaluations are very important at NAU. They enable students to provide instructors and the university with feedback about the effectiveness, quality, and value of courses. Your feedback on this course evaluation is essential for revising and improving this course and how it is taught. These improvements will be of direct benefit to future students. Improvements will also benefit you, by raising the overall quality of courses at NAU and increasing the reputation and value of an NAU degree. Finally, I greatly value your feedback, and I appreciate your attention to the evaluation of this course.

By clicking on the link below, you will access the anonymous end-of-semester evaluation for this course. Please contact d-oaa@jan.ucc.nau.edu (the online evaluation administrator) if you run into any problems with completing the evaluation online.

Please complete the course evaluation by [PROVIDE THE DATE FOR THE LAST DAY OF CLASSES]

3. Reminder to complete course evaluation:
One week prior to the last day of class, send an e-mail reminder to students to complete the course evaluation if they have not already done so. Include the evaluation url link in the e-mail message.
and remind students of the “complete by” date.

Acknowledgments: We would like to thank the faculty members and instructors who participated in the research reported here, as well as the many students who responded to their online course evaluations. Thanks also to Brian Sawert at NAU Information Technology Services for his assistance in querying the Evalajack database. Finally, we appreciate Paul Rowland’s support, encouragement, and feedback.

REFERENCES


