Online Student Evaluation of Instruction: An Investigation of Non-Response Bias

Stephen W. Thorpe
Assistant Provost
Institutional Research, Planning, & Assessment
Drexel University

ABSTRACT

One concern for online student course evaluation over in-class paper methods is the potential for low student response and the likelihood of non-response bias. This paper provides results of an analysis of respondents to an online versus in-class paper method of conducting student course evaluations.

INTRODUCTION

The advent of web-based surveys has become an established practice at many institutions of higher education. The advantages of online methodologies over paper-based methods include the reduced costs of online research (personnel, mailing, printing, etc), the ease of reaching representative samples of a population, the ability to validate data during collection, and rapid dissemination of results (Couper, 2000; Dillman, 2000).

Published research studies have found that the results of web-based versus paper surveys are not significantly different (Handwerk, Carson, and Blackwell, 2000; Matz, 1999; Sax et al, 2002), although demographic differences have been found in comparing respondents to non-respondents of online surveys (Underwood, Kim, & Matier, 2000; Tomsic, Hendel, & Matross, 2000).   Additional research suggests several factors that might influence an individual's decision to complete an online survey, such as familiarity with the Internet, the ease of completing the survey, and concerns for privacy and confidentiality (Dillman, 2000; Handwerk, Carson, and Blackwell, 2000).

Unlike student satisfaction surveys, the results of student evaluation of instruction questionnaires are frequently used (at least in part) to make critical faculty personnel decisions. Faculty are naturally concerned about the potential effects of non-response bias if a limited number of students complete an online course evaluation instrument. The option of conducting course evaluations using in-class paper methods are unlikely to suffer the effects of non-response bias since most students are assumed to be in attendance when in-class evaluations are conducted.

The use of online course evaluation systems is relatively limited in higher education. In one study of 200 institutions ranked as most "wired" by Yahoo!, only 2 institutions reported institution-wide use of web-based evaluation systems (Hmieleski, 2000). Low response rates for web-based surveys was suggested as the primary issue that limits institutions to in-class, paper-based course evaluation instruments.

Cummings and Ballatyne (1999) describe the experiences of Murdoch University in piloting a web-based course evaluation process. The response rates were lower for the web-based system, 31% compared to 65% for the in-class paper-based method. However, an analysis of the ratings received from the previous term that was paper-based and the rating results from the web-based evaluation system revealed no pattern of response bias.

STUDY DESIGN

The purpose of this study was to examine whether significant differences exist in student responses to a course evaluation instrument based on the delivery method, in-class paper versus web-based. The question of significant differences is important since prior research suggests that student response rates are likely to be lower in the web-based evaluation process versus the in-class, paper method. Additionally, this study will determine whether non-response bias exists within the web-based method of course evaluation.

The study used two large classes (more than 80 students) with two distinct course sections from the Fall 2001 academic term where the same professor was assigned as the primary instructor for both sections. An additional class with two large sections taught by the same professor was selected from the Winter 2002 academic term. The courses selected for this study were in computer science, math, and statistics. One section for each course was designated to receive the in-class, paper evaluation form while the other section received the web-based evaluation form.

The student evaluation instrument, which consists of 23 Likert-scale questions and three open-ended questions, was recently developed to incorporate student-learning outcomes in the course evaluation process (Thorpe, 2001). The Likert-scale items include statements about the student, about the course, about the instructor, and overall summary items (the survey items are shown in Table 4). The items regarding the student, course, and instructor are rated on a scale from 1-never, 2-seldom, 3-sometimes, 4-usually, and 5-always. The summary questions are rated on a Likert-scale from 1-extremely low, 2-low, 3-adequate, 4-high, and 5-extremely high.

Students from web-based sections of each course received an email two weeks prior to the end of classes that asked them to complete an online course evaluation instrument. The email solicitation included a URL link to access the online survey. Students were required to enter their university ID number to validate their participation in the course

evaluation survey. Students who had not responded within three days received a follow-up email reminder.

During the last week of classes, students from the other section of each course completed the same course evaluation instrument through an in-class paper evaluation process.

RESULTS OF STUDY

Table 1 provides the response rates by course section and evaluation method. For the Fall 2001 courses, the web-based evaluation generated a higher response rate than the section using the in-class, paper-based methodology. For the CS course, 45% of the students in the section that was evaluated through the web completed the course evaluation compared to 37% of the students in the paper-based evaluation section. Likewise, half of the students enrolled in the MATH course section that used the web-based form completed the course evaluation compared to 44% of the students enrolled in the section that utilized the in-class, paper-based evaluation method.

Table 1
Completed Course Evaluations by Course Section

Course Section	Enrollment	Responses	Response Rate
CS (in class)	126	46	37%
CS (web)	163	73	45%
Math (in class)	152	67	44%
Math (web)	186	93	50%
Stat (in class)	136	95	70%
Stat (web)	81	30	37%

The STAT course that was added from the Winter term generated response rates that might be expected when comparing in-class versus web-based evaluation methods. In this case, about 70% of the students enrolled in the in-class, paper-based evaluation section completed the course evaluation survey compared to 37% of the students enrolled in the web-based course evaluation section.

Comparison of Respondents vs Non-Respondents to Web-Based Evaluation

Unlike traditional paper-based evaluation methods, conducting a web-based course evaluation process provides demographic data to compare respondents and non-respondents.   Although the actual evaluations remain anonymous, a web-based system provides a method to identify students who respond to the survey. In this way, a determination can be made regarding potential non-response bias.

Table 2 provides a comparison of the respondents and non-respondents to the web-based course evaluation sections on the basis of student sex, minority status, the actual grade received for the class, and the cumulative GPA of the student.   Other variables, such as class standing and college of study, were not included because of the nature of the courses selected for this study. That is, students in these courses tended to be enrolled in a particular discipline and were virtually all at the same class standing, so comparisons on these demographic variables would not be informative.

Table 2
Comparison of Web-Based Evaluation Response Rates

	CS Class		Math Class		Stat Class		Combined	
	N	% Resp	N	% Resp	N	% Resp	N	% Resp
Sex								
Female	70	56%*	55	65%*	30	50%	155	58%*
Male	93	37%	131	44%	51	29%	275	39%
Total	163	45%	186	50%	81	37%	430	46%
Minority Status								
Minority	57	42%	40	45%	24	29%	121	40%
Non-Minority	84	44%	116	52%	46	37%	246	46%
Course Grade								
A	63	51%	56	61%*	9	56%	128	55%*
B	64	48%	35	57%	8	38%	107	50%
C	24	33%	48	48%	11	73%	83	47%
D	6	17%	29	45%	23	35%	58	38%
F	5	20%	17	18%	24	25%	46	22%
Other	1	0%	1	0%	6	0%	8	0%
Cumulative GPA								
Less than 1.0	0		3	0%*	1	100%	4	25%*
1.0-1.9	9	22%	33	21%	7	29%	49	22%
2.0-2.9	40	33%	61	51%	46	30%	147	39%
3.0-above	114	51%	89	62%	27	48%	230	55%

* - Chi-square significant difference, p<.05

Within the CS course section that used the online evaluation process, women were significantly more likely than men to complete the evaluation, 56% compared to 37%. Likewise, women were more likely than men to complete the course evaluation process for the MATH class.   The difference in response rates in the STAT class were not significant.

No significant differences were found in the response rates between minority and non-minority students in the any of the class sections.

In looking at the actual grades received by the students for each course, students who earned higher grades in the MATH course were significantly more likely to complete the online course evaluation. Over 60% of the students who earned an "A" in the course completed the online evaluation compared to 45% of the students who earned a "D", and just 18% of the students who earned an "F". No significant differences were found in the completion rates of the CS and STAT class based on the grade earned in the course.

Students in the MATH class with higher cumulative grade point averages (including the term in which the course was taken) were significantly more likely to complete the course evaluation process. No significant differences were found in the completion rates by student cumulative GPA in the CS and STAT classes.

At the aggregate level, that is by combining the respondents and non-respondents across all three courses, female students were significantly more likely than men to respond to the evaluation process, 58% and 39% respectively. Academic performance of the student was found to be related to response rates. At the aggregate level, students who earned higher grades in the course were more likely to complete the web-based evaluation process. Over half of the students (55%) that earned an "A" in the course completed the online evaluation, compared to less than 40% of the students who earned a "D" in the course. Just 25% of the students earning an "F" completed the online evaluation process.

Likewise, significant differences were found in the completion rates of students based on their cumulative grade point average at the university. Over half of the students (55%) with a cumulative GPA of 3.0 or higher completed the online course evaluation process, compared to just 22% of the students with a cumulative GPA below a 2.0.

Comparison of Student Responses: Paper vs Web-Based Methods

The course evaluation instrument includes two questions that ask the student to provide their expected grade for the course and whether the course was a requirement for their major. The results of these two questions are shown in Table 3.

For the CS and MATH classes, the proportion of students responding by expected course grade was not significantly different. The respondents were significantly different in the case of the STAT class. About 40% of the students completing the in-class evaluation instrument expected to receive an "A" for the course, compared to just 14% of the student who completed the web-based evaluation form.

In terms of whether the course was required for the student's major, the results from the three classes were mixed. In the case of the CS class, a higher proportion of students completing the web-based instrument indicated that the course was required for their major, 94% compared to 77% of the students completing the in-class paper form. The MATH class shows just the opposite where a higher proportion of students completing

the paper-based evaluation instrument indicated that the class was a requirement for their major. The proportional difference between students enrolled in the paper-based versus web-based course evaluation sections of the STAT class was not signficant.

Table 3
Expected Grade and Required Status by Course and Type

	CS		MATH		STAT	
	Paper %	Web %	Paper %	Web %	Paper %	Web %
Expected Grade in Course						
A	39%	44%	33%	41%	40%*	14%*
B	44%	47%	34%	29%	37%	31%
C or below	17%	8%	33%	30%	23%	55%
Is course required for major?						
Yes	77%*	94%*	95%*	83%*	98%	90%
No	23%	6%	5%	17%	2%	10%

* - Chi-square significant difference between course sections, $p<.05$

For analysis of the course evaluation items, several statistical techniques were applied. The t-test was used to evaluate the differences in average ratings for each survey item between the course sections.   Chi-square analysis and the Mann-Whitney test were also used to determine if significant differences existed in the student ratings between course sections.   Essentially each of these techniques resulted in substantively similar conclusions regarding differences in student ratings between the in-class and web-based course evaluation processes.

Table 4 provides the average (mean) course evaluation response for each survey item and course section.   Although arguably Likert-scale data is not interval level, average ratings are nevertheless often used in reporting student course evaluation results. For this reason, Table 4 provides the average rating by item so the reader can discern what the "reported" difference in ratings might be in comparing course sections. Statistically significant differences shown in Table 4 are based on the t-test.

For the CS course, the average rating of students using the online evaluation process was slightly higher on most items than the students completing the in-class paper-based form. For example, the average rating on the first item regarding preparation for class was 3.85 for students completing the online evaluation process compared to 3.74 for the students that completed the in-class form.   Nevertheless, only three statistically significant differences were found in the average course ratings between the in-class and web-based evaluation processes.   These items included "I invested enough time and energy to meet/exceed course requirements" ($p<.01$); "the course objectives and requirements were clearly communicated" ($p<.05$); and "the instructor treated students, their ideas and opinions, with respect" ($p<.05$). In each case, the average rating from the students completing the web-based course evaluation process was significantly higher than for those students completing the in-class evaluation process.

Table 4

Mean Course Evaluation Scores by Course and Survey Type

	CS		MATH		STAT	
	Paper	Web	Paper	Web	Paper	Web
Students Completing Evaluations	46	73	67	93	95	30
About the Student						
I was well-prepared for each class session	3.74	3.85	3.61	3.83	3.84	4.03
I asked the instructor for help/feedback when I needed it	2.54	2.97	2.95	2.80	3.21	3.20
I invested enough time and energy to meet/exceed course requirements	**3.53****	**4.11****	3.76	3.72	3.67	3.77
I participated actively and contributed thoughtfully during class sessions	2.61	2.76	**3.36***	**2.99***	2.98	3.14
I attended class sessions (and related, required meetings)	4.24	4.50	**4.61***	**4.35***	4.49	4.37
I practiced self-directed learning, including self-assessment and reflection	3.39	3.60	3.63	3.43	3.75	3.63
About the Course						
The course provided the opportunity to work in teams and team projects	2.02	1.72	2.85	2.58	1.77	1.77
The course objectives and requirements were clearly communicated	**2.96***	**3.38***	4.03	4.07	4.01	3.70
The course increased my desire to continue learning about this material	2.26	2.45	3.18	2.99	2.68	2.23
The course provided the opportunity to learn/practice oral comm skills	1.93	1.93	2.18	2.03	2.05	1.90
The course provided the opportunity to learn/practice written comm skills	2.35	2.19	2.37	2.16	2.46	2.20
The course provided opportunities to identify problems/formulate solutions	3.24	3.33	3.65	3.72	3.77	3.60
The course allowed oppt to exhibit an understanding of how solutions impact society	2.44	2.36	2.58	2.23	3.16	2.80
About the Instructor						
The instructor provided me clear, useful, and timely feedback	3.37	3.33	3.68	3.55	3.89	3.67
The instructor inspired interest/excitement in course material	3.28	3.31	3.10	3.12	3.69	3.43
The instructor was available and helpful when asked	3.61	3.96	4.05	3.75	3.95	3.80
The instructor communicated ideas and information clearly/effectively	3.39	3.57	3.90	3.73	3.96	3.60
The instructor treated students, their ideas and opinions, with respect	**3.85***	**4.29***	4.21	4.26	4.23	4.33
The instructor organized and used class time effectively to promote learning	3.57	3.88	4.27	4.30	4.36	4.10
Summary Questions						
Before the course began, my level of interest in this course/topic was	2.48	2.78	3.33	3.24	2.25	2.33
Overall, I would rate the quality of my work in/for this course	3.39	3.60	3.55	3.35	3.57	3.40
Overall, I would rate the effectiveness of the instructor	3.22	3.33	3.71	3.44	**3.86***	**3.33***
Overall, I would rate the amount I learned in this course	2.85	3.04	3.33	3.00	3.56	3.17

Scale for Student, Course, Instructor sections: 1-Never, 2-Seldom, 3-Sometimes, 4-Usually, 5-Always

Scale for Summary questions: 1-Extremely low, 2-Low, 3-Adequate, 4-High, 5-Extremely high

* - Significantly different, p<.05

** - Significantly different, p<.01

In comparing the course evaluation results of the two sections of the MATH course, just two significant differences in the evaluation results were found among the 23 items. The students evaluating the course through the web-based format were significantly less likely than students completing the in-class evaluation to participate actively and contribute thoughtfully during class sessions and to attend class sessions (p<.05).

Like the MATH course sections, the average ratings between the STAT course sections were sometimes slightly higher from the in-class section and at other times slightly higher from the web-based course evaluation section. Only one statistically significant difference was found in the average evaluation ratings between the two course sections. Students evaluating the course using the in-class paper-based form provided significantly higher ratings in their evaluation of the overall effectiveness of the instructor, 3.86 compared to 3.33 for the students who used the web-based course evaluation process.

SUMMARY

The purpose of this study was to determine if significant differences exist in evaluation ratings obtained from conducting student course evaluations through an in-class, paper-based method versus an online web-based approach. Second, this study intended to measure the potential for non-response bias in conducting web-based course evaluation surveys.

Using web-based survey techniques for course evaluation provides several potential benefits. A principal advantage of web-based surveys over paper methods is the reduced costs for personnel, printing, and administration of the survey. With course evaluation models, the web-based approach allows for timely and flexible feedback of evaluation results to faculty and administrators. In addition, no time is required to re-type open-ended questions to protect student identity and valuable class time is not lost for course evaluations to be conducted during the class. Because students can complete the course evaluation on their own time, the potential for more thoughtful responses is possible since the student is not rushed to complete the evaluation in class, often in the last five to ten minutes. Recent studies have documented improved quality and quantity of student comments through web-based surveys.

The primary concern for using web-based course evaluation processes is the fear that low response rates may affect the overall results of the course evaluation. It is interesting to note that in two of the three courses included in this study, a higher response rate was achieved through the web-based course evaluation process over the traditional in-class process.

The results of this study suggest that the evaluation responses between the in-class, paper-based method and the web-based evaluation process are not significantly different in the vast majority of survey items. While a few significant differences were

found in the average evaluation ratings between the sections of each course, no pattern was evident to suggest that the web-based course evaluation process would generate substantially different course evaluation results from the in-class, paper-based method. This finding is consistent with prior research studies that have found no significant differences in student survey responses based on survey delivery (web versus paper). Even in the case of the STAT course where response rates varied widely, just one statistical difference was found among the average ratings of the 23 survey items.

Similar to other studies of respondents and non-respondents to online surveys, this study also found that some students are more likely to respond to an online course evaluation survey than others. For this institution, female students were significantly more likely to respond to the web-based course evaluation survey than men. More importantly, students performing poorly academically were less likely to respond to the online course evaluation process. However, it should be noted that these students may not complete in-class evaluation instruments either. In the CS and MATH course sections, the distribution of anticipated course grades between students completing the in-class versus the web-based evaluation process were not significantly different. Even in the case of the STAT course where a significant difference was found, the average ratings on the course evaluation items were not significantly different.

In conclusion, the results of this study suggest that concerns regarding low response rates and the potential for non-response bias in web-based course evaluation methods may not be warranted. Nevertheless, faculty and administrators considering the move to web-based course evaluation models would be well advised to replicate this study on their own campuses to determine the potential effects on course evaluation results from online survey methods.

REFERENCES

Couper, M. (2000). Web surveys: A review of issues and approaches. Public Opinion Quarterly, (64), pp 464-494.

Cummings, R. and Ballatyne, C. (1999). Student feedback on teaching: Online! On target? Paper presented at the Australisian Society Annual Conference, October, 1999. (Available online at http://cleo.murdoch.edu.au/evaluation/pubs/confs/aes99.htm)

Dillman, D. (2000). Mail and Internet Surveys: The Tailored Design Method. New York: John Wiley & Sons.

Handwerk, P., Carson, C., and Blackwell, K. (2000). On-line vs. paper-and-pencil surveying of students: A case study. Paper presented at the 40th Annual Meeting of the Association of Institutional Research, May 2000.

Hmieleski, K. (2000). Barriers to online evaluation: Surveying the nation's top 200 most wired colleges. Report prepared by the Interactive and Distance Education Assessment Laboratory at Rensselaer Polytechnic Institute, Troy, NY.

Matz, C. (1999). Administration of web versus paper surveys: Mode effects and response rates. Masters Research Paper, University of North Carolina at Chapel Hill. (ERIC Document Reproduction Service ED 439 694).

Sax, L., Gilmartin, S., Keup, J., Bryant, A., and Plecha, M. (2002). Findings from the 2001 pilot administration of Your First College Year (YFCY): National norms. Higher Education Research Institute, University of California. (Available online at http://www.gseis.ucla.edu/heri/yfcy/yfcy_report_02.pdf)

Tomsic, M., Hendel, D., Matross, R. (2000). A world wide web response to student satisfaction surveys: Comparisons using paper and Internet formats. Paper presented at the 40th Annual Meeting of the Association of Institutional Research, May 2000.

Thorpe, S. (2001). Linking learning outcomes to student course evaluation. Paper presented at the 28th Annual Conference of the Northeast Association for Institutional Research, Boston MA, November, 2001.

Underwood, D., Kim, H., and Matier, M. (2000). To mail or to web: Comparisons of survey response rates and respondent characteristics. Paper presented at the 40th Annual Meeting of the Association of Institutional Research, May 2000.